Distinguishing Random and Black Hole Microstates

Jonah Kudler-Flam,[1,*] Vladimir Narovlansky,[2,†] and Shinsei Ryu[2,3,‡]

[1]*Kadanoff Center for Theoretical Physics, University of Chicago, Chicago, Illinois 60637, USA*

[2]*Princeton Center for Theoretical Science, Princeton University, Princeton, New Jersey 08544, USA*

[3]*Department of Physics, Princeton University, Princeton, New Jersey 08540, USA*

In this work, we study the distinguishability of random states drawn from the Wishart ensemble as well as black hole microstates. We compute the relative entropy and many generalizations, including the Petz Rényi relative entropy, sandwiched Rényi relative entropy, fidelities, and trace distances. These generalized quantities are able to teach us about new structures in the space of random states and black hole microstates where the von Neumann and relative entropies were insufficient. We further generalize to generic random tensor networks where new phenomena arise due to the locality in the networks. These phenomena sharpen the relationship between holographic states and random tensor networks. We discuss the implications of our results on the black hole information problem using replica wormholes, specifically the state dependence (hair) in Hawking radiation. Understanding the differences between Hawking radiation of distinct evaporating black holes is an important piece of the information problem that was not addressed by entropy calculations using the island formula. We interpret our results in the language of quantum hypothesis testing and the subsystem eigenstate thermalization hypothesis (ETH), deriving that chaotic (including holographic) systems obey subsystem ETH for all subsystems less than half the total system size.

I. INTRODUCTION AND SUMMARY OF RESULTS

A unifying idea spanning quantum information theory, quantum chaos and thermalization, and black hole physics is that of (in)distinguishability of quantum states. In quantum information theory, we would like to understand what the space of quantum states are. In particular, how can we characterize which states are close or far away and endow the Hilbert space with a geometry? This notion of distinguishability is critical for storing and processing quantum information.

Quantum chaos and thermalization is all about distinguishibility. A natural *definition* of quantum thermalization is that the state is indistinguishable (up to a certain error) from a completely thermal, e.g., Gibbs state. It is then important to characterize which systems thermalize and the mechanism for thermalization to occur. We can begin with two states that are easily distinguishable, e.g., the "all spin up" and "all spin down" states of a quantum spin chain. If we evolve these states with a thermalizing Hamiltonian, the states will become indistinguishable using "simple" measurements.

Similarly, the black hole information problem is most naturally framed in terms of thermalization and indistinguishability. Black holes can be formed in many different ways. Moreover, they have an extraordinary number of microstates [1,2]. Even so, using semiclassical calculations, Hawking showed that all black holes with identical thermodynamic quantities (mass, charge, and angular momentum) will radiate *thermal* radiation [2]. This means that at late times, after the black hole has evaporated, all of these microstates are completely indistinguishable, which is in sharp tension with the unitarity of quantum mechanics. To resolve this apparent paradox, different black hole microstates must be made distinguishable directly from the radiation.

Remarkably, all three of these broad problems may be addressed using random matrix theory calculations of distinguishability measures. The purpose of this paper is to make this statement precise and elucidate the intricate and surprising connections, extending the analysis of Ref. [3].

The techniques of random matrix theory have become ubiquitous across far-ranging fields of physics. Originally

*jkudlerflam@uchicago.edu

†narovlansky@princeton.edu

‡shinseir@princeton.edu

used to characterize the spectra of heavy nuclei [4], random matrix theory has flourished in its applications in quantum information theory [5], quantum chaos and thermalization [6], and black hole physics [7,8]. What is more is that these fields are now understood to be deeply related to one another and, to some extent, inseparable.

In Sec. II, we lay the foundation by reviewing the precise definitions of distinguishability in quantum information theory. In particular, we review various distinguishability measures that are used to characterize how well different states can be discriminated between. This is formalized by the operational tasks of *quantum hypothesis testing* and *state discrimination*. These are the most fundamental information processing tasks and make precise the operational meaning of our subsequent results.

In Sec. III, we undertake our main technical computations. We introduce the ensemble of random mixed states (Wishart ensemble) and a diagrammatic approach in evaluating moments of the reduced density matrices. We exactly compute the relative entropy, Petz Rényi relative entropies, sandwiched Rényi relative entropies, fidelities, and trace distance of random states in the limit of large Hilbert space dimensions. This characterizes the space of generic quantum states. In particular, we find that, when the logarithm of the dimension of the sub-Hilbert space that we consider is less than half of the total Hilbert space, then generic states are indistinguishable up to exponentially small terms in the system size. When the sub-Hilbert space is larger, we find that the states are completely distinguishable up to exponentially small terms. We find interesting $O(1)$ crossover behavior. We compare these "large-N" results to finite size numerics and find precise agreement.

In Sec. IV, we begin to interpret the random matrix theory results in the language of gravity. First, we show that in AdS/CFT, if one considers two different black hole microstates, the evaluation of distinguishability measures between the black hole microstates when an observer only has access to a subregion of the boundary is formally identical to the formulas present in Sec. III. This occurs in special states called "fixed-area states" [9,10]. With this realization, we characterize the distinguishability of black hole microstates in anti–de Sitter (AdS) space or, equivalently, high-energy states in conformal field theories (CFTs). We subsequently apply this formalism to a toy model of an evaporating black hole [11]. We conclude that before the Page time, an observer of an evaporating black hole is only able to distinguish different black hole microstates if one has an $O(e^{1/G_N})$ number of copies of the radiation, meaning that the microstates are nearly completely indistinguishable. After the Page time, an observer of an evaporating black hole can easily distinguish microstates with a single copy of the radiation, though the needed measurement will be quite complex. In these calculations, replica wormholes play a central role.

In Sec. V, we generalize our computations to random tensor networks. These represent new ensembles of random matrix theory and introduce a notion of locality into the quantum state. We find qualitatively new features in distinguishability with larger tensor networks being more distinguishable than smaller tensor networks and the Haar random states of Sec. III.

In Sec. VI, we discuss thermalization in chaotic quantum many-body systems. Using an ansatz for the structure of high-energy eigenstates in chaotic systems [12–14], we evaluate the various distinguishability measures. We interpret these results using the subsystem eigenstate thermalization hypothesis [15], a very strong version of thermalization. We determine that chaotic systems obeying the ansatz also obey subsystem eigenstate thermalization hypothesis (ETH) for subsystem sizes that are less than half the total system, after which subsystem ETH is violated. Furthermore, we find similar structures in gravity, generalizing the calculations of Sec. IV to black hole microstates without fixed areas. Analogous conclusions apply and we conclude that holographic CFTs in generic dimensions obey subsystem ETH.

We relegate certain details and extensions of our work to Appendices A through E. We highlight alternative derivations using free probability theory in Appendix A.

II. DISTINGUISHABILITY MEASURES AND THEIR USE

A. Review of distinguishibility measures

In this section, we review various distinguishability measures commonly used in quantum information theory. Each measure has an operational meaning and there are various relations between the measures. Readers familiar with distinguishability measures and hypothesis testing may move to Sec. III as there are no new results in this section.

1. Relative entropy

We begin with the quantum relative entropy that is arguably the most important quantity in quantum information theory as many of the deepest results in the field are directly derivable from its fundamental properties. The classical relative entropy or Kullback-Leibler divergence is defined as

$$D_{\mathrm{KL}}(P||Q) := \log\left[\sum_{x \in X} P(x) \log\left(\frac{P(x)}{Q(x)}\right)\right], \quad (1)$$

where P and Q are classical probability distributions over a set X. The quantum relative entropy is the noncommutative analog defined for two density matrices, ρ and

σ, as

$$D(\rho||\sigma) := \mathrm{Tr}(\rho \log[\rho] - \rho \log[\sigma]). \qquad (2)$$

This is only well defined when the support of ρ is contained within the support of σ. Otherwise, the relative entropy is infinite.

The relative entropy acts as a distinguishability measure, as can be seen from its basic properties. The first is positivity, $D(\rho||\sigma) \geq 0$, with the inequality saturated if and only if $\rho = \sigma$. The second is referred to as the data processing inequality or monotonicity of relative entropy [16], which states that the relative entropy is nonincreasing under completely-positive trace-preserving (CPTP) quantum channels, $\mathcal{N}$ [17]:

$$D[\mathcal{N}(\rho)||\mathcal{N}(\sigma)] \leq D(\rho||\sigma). \qquad (3)$$

This property is crucial for a distinguishability measure because it asserts that if you are given two quantum states, after performing operations on them, they can never become easier to distinguish.

A particularly important quantum channel is the partial trace operation on a bipartite Hilbert space

$$\mathcal{N} : \mathcal{H}_A \otimes \mathcal{H}_B \to \mathcal{H}_A, \qquad (4)$$

$$\rho \mapsto \mathrm{Tr}_B[\rho] := \rho_A. \qquad (5)$$

Under the partial trace, we lose all information about region B, making ρ harder to distinguish from other states that look similar on A. The partial trace will play a central role throughout the rest of the paper because we are generally interested in how to distinguish states when only having access to a subregion.

While the relative entropy characterizes the structure of the space of quantum states, importantly, it is not a metric. This is most obviously seen from the definition which is not symmetric under exchange of ρ and σ. This is a feature and not a bug, as can be seen by its operational meaning that we will soon explore.

The relative entropy is a parent quantity to many other central information-theoretic quantities, such as the von Neumann entropy

$$S_{vN}(\rho_A) = \log[d_A] - D\left(\rho_A \,\middle\|\, \frac{\mathbb{1}}{d_A}\right), \qquad (6)$$

where d_A is the Hilbert space dimension, the mutual information

$$I(A, B) = D(\rho_{AB}||\rho_A \otimes \rho_B), \qquad (7)$$

and conditional entropy

$$S(B|A) = \log[d_B] - D\left(\rho_{AB} \,\middle\|\, \rho_A \otimes \frac{\mathbb{1}}{d_B}\right). \qquad (8)$$

In these terms, the strong subadditivity of von Neumann entropy,

$$S_{vN}(\rho_B) + S_{vN}(\rho_{ABC}) \leq S_{vN}(\rho_{AB}) + S_{vN}(\rho_{BC}), \qquad (9)$$

is a straightforward consequence of the data processing inequality

$$D(\mathrm{Tr}_C[\rho_{ABC}]||\mathrm{Tr}_C[\rho_A \otimes \rho_{BC}]) \leq D(\rho_{ABC}||\rho_A \otimes \rho_{BC}). \qquad (10)$$

2. Rényi relative entropies

Like the Kullback-Leibler divergence, the relative entropy can be generalized into Rényi relative entropies. However, because of the noncommutativity of density matrices, there are many inequivalent ways to generalize the relative entropy such that it reduces to the classical α-Rényi divergences, the unique set of quantities satisfying the five axioms of a generalized divergence [18]

$$D_{\mathrm{KL},\alpha}(P||Q) := \frac{1}{\alpha - 1} \log\left[\sum_{x \in X} P(x)^\alpha Q(x)^{1-\alpha}\right], \qquad (11)$$

where α is a positive semidefinite real variable. We study two complementary families that have served the most uses in quantum information theory.

The first is the most obvious quantum analog of Eq. (11) and is referred to as the Petz Rényi relative entropy (PRRE) [19]

$$D_\alpha(\rho||\sigma) := \frac{1}{\alpha - 1} \log(\mathrm{Tr}[\rho^\alpha \sigma^{1-\alpha}]). \qquad (12)$$

The PRRE satisfies various nice properties, such as reduction to the von Neumann relative entropy when $\alpha \to 1$. For $\alpha \in [0, 1)$, the PRRE is finite even when the support of ρ is larger than the support of σ. Most importantly, the PRRE satisfies the data processing inequality when $\alpha \in [0, 2]$ [19–21]. One particularly useful case is at $\alpha = 1/2$, which defines what has been called Holevo's "just-as-good fidelity" [22] or affinity [23]

$$F_H(\rho||\sigma) := (\mathrm{Tr}[\sqrt{\rho}\sqrt{\sigma}])^2 = e^{-D_{1/2}(\rho||\sigma)}, \qquad (13)$$

which, for most purposes, is just as (if not more) useful as the more widely used Uhlmann fidelity

$$F(\rho||\sigma) := \left(\mathrm{Tr}\left[\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right]\right)^2. \qquad (14)$$

Both satisfy all of Jozsa's axioms for distinguishability measures [24] and define metrics on the space of quantum

states

$$D_B(\rho||\sigma) := \sqrt{2\left[1 - \sqrt{F(\rho||\sigma)}\right]}, \qquad (15a)$$

$$D_A(\rho||\sigma) := \arccos\left[\sqrt{F(\rho||\sigma)}\right], \qquad (15b)$$

$$D_H(\rho||\sigma) := 2\left[1 - \sqrt{F_H(\rho||\sigma)}\right], \qquad (15c)$$

called the Bures distance, Bures angle, and Hellinger distance, respectively.

The other quantum generalization of (11) we will study is the sandwiched Rényi relative entropy [137]

$$\tilde{D}_\alpha(\rho||\sigma) := \frac{1}{\alpha - 1}\log\{\mathrm{Tr}[(\sigma^{(1-\alpha)/2\alpha}\rho\sigma^{(1-\alpha)/2\alpha})^\alpha]\}. \qquad (16)$$

It is clear that this is equivalent to the PRRE when ρ and σ commute and reduces to the Uhlmann fidelity at $\alpha = 1/2$. Like the PRRE, the SRRE reduces to the von Neumann relative entropy in the $\alpha \to 1$ limit and is only finite if either $\alpha \in [0, 1)$ or the support of ρ is contained within the support of σ. The most important property of SRRE is that it satisfies the data-processing inequality for $\alpha \in [1/2, \infty)$. In this way, it is complementary to the PRRE. Similar formulas for Rényi analogs of entropy, mutual information, and conditional entropies can be written in terms of the Rényi relative entropies.

3. Trace distance

The final distinguishability measure that we study is the trace distance, defined as

$$T(\rho||\sigma) := \tfrac{1}{2}|\rho - \sigma|_1, \qquad (17)$$

where $|\cdot|_1$ is the trace norm. The trace distance defines a metric on the space of quantum states and takes values between zero and one. However, unlike Holevo's just-as-good and Uhlmann fidelities, it does not descend from a relative entropy. The trace distance is monotonically decreasing under quantum operations. It will play a central role in our discussion of eigenstate thermalization in Sec. VI.

There are various useful relations between the above distinguishability measures that we now list. First, we note that both PRRE and SRRE are monotonic in α, i.e.,

$$D_{\alpha_1}(\rho||\sigma) \leq D_{\alpha_2}(\rho||\sigma),$$

$$\tilde{D}_{\alpha_1}(\rho||\sigma) \leq \tilde{D}_{\alpha_2}(\rho||\sigma), \quad \alpha_1 < \alpha_2, \qquad (18)$$

while the SRRE lower bounds the PRRE,

$$\tilde{D}_\alpha(\rho||\sigma) \leq D_\alpha(\rho||\sigma), \quad \alpha \geq 0. \qquad (19)$$

By Pinsker's inequality, the von Neumann relative entropy upper bounds the trace distance [28],

$$\tfrac{1}{2}T(\rho||\sigma)^2 \leq D(\rho||\sigma), \qquad (20)$$

while the Fuchs–van de Graaf inequalities assert that both fidelities place upper and lower bounds on the trace distance [29],

$$1 - \sqrt{F(\rho||\sigma)} \leq 1 - \sqrt{F_H(\rho||\sigma)} \leq T(\rho||\sigma)$$
$$\leq \sqrt{1 - F(\rho||\sigma)} \leq \sqrt{1 - F_H(\rho||\sigma)}. \qquad (21)$$

These are strong results that we use throughout the paper due to the difficulty in directly computing the trace distance. They are also important, nontrivial consistency checks of our results.

B. Operational interpretations in hypothesis testing

The most fundamental information processing processes are quantum state discrimination (QSD) and hypothesis testing (QHT). It should then be no surprise that this is where the most fundamental quantities, relative entropy and trace distance, find their operational meanings. In this section, we make precise what it means for states to be distinguishable by first introducing QSD and QHT, then stating what the distinguishability measures say about our ability to perform these tasks. For more details, we refer the reader to the literature, e.g., Refs. [30,31].

The general setup is that we are given a state on $\mathcal{H}$ that is either ρ or σ and we wish to determine which state we were given. We are allowed to use any positive operator-valued measure (POVM) that is a collection of positive semidefinite operators, $\{M_i\}$, that sum to the identity operator on $\mathcal{H}$. Each subscript, i, corresponds to a measurement outcome. Because we are looking for a binary outcome (is our state ρ or σ?), we can consolidate the M_i into just two elements. For outcomes $i \in \mathcal{A}$, we conclude that the state is ρ, while for outcomes $i \notin \mathcal{A}$, we conclude that the state is σ. Our POVM is then $\{A, \mathbb{1} - A\}$, where $A := \sum_{i\in\mathcal{A}} M_i$. There are many choices for A and we want to optimize this choice so as to have the least error in our conclusions. There are two types of error. The probability of mistakenly concluding that we have σ when we were really given ρ is given by

$$\alpha(A) := \mathrm{Tr}[(\mathbb{1} - A)\rho], \qquad (22)$$

while the probability of mistakenly concluding that we have ρ when we were really given σ is given

by

$$\beta(A) := \text{Tr}[A\sigma]. \qquad (23)$$

These are referred to as the error probabilities of the first and second kind, respectively (or types I and II).

There are various ways of optimizing these errors [32]. The symmetric way is called state discrimination. The smallest combined error is given by the trace distance between the states [33,34]

$$\min_A [\alpha(A) + \beta(A)] = 1 - T(\rho||\sigma), \qquad (24)$$

where the optimization is taken over all POVM. If the trace distance is very large (close to one), we are able to choose a POVM that has very small error probabilities. If the trace distance is small (close to zero) then the combined error is close to one, the maximal optimized error that can be saturated by taking $A = \mathbb{1}$. Likewise, the probability that we correctly discriminate, P_+, is also given by the trace distance

$$P_+(A) := \tfrac{1}{2} \max_A \{\text{Tr}[A\rho] + \text{Tr}[(\mathbb{1} - A)\sigma]\}$$
$$= \tfrac{1}{2}[1 + T(\rho||\sigma)]. \qquad (25)$$

State discrimination can be made easier if instead of given one copy of the state, we are given multiple, n, copies. This is the topic of asymptotic state discrimination. With these n copies, we can ask what is the optimal POVM on $\mathcal{H}^{\otimes n}$. The error probabilities are generalized in the obvious way

$$\alpha_n(A) := \text{Tr}[(\mathbb{1} - A)\rho^{\otimes n}], \quad \beta_n(A) := \text{Tr}[A\sigma^{\otimes n}]. \qquad (26)$$

The sum of the errors can be shown to be bounded above by Holevo's just-as-good fidelity

$$\min_A [\alpha_n(A) + \beta_n(A)] \leq F_H(\rho||\sigma)^{n/2}. \qquad (27)$$

Unless the states are identical ($F_H = 1$), the error rate exponentially decays to zero as we are given a large number of copies. If the fidelity is small, we may need only one copy (or very few) to confidently discriminate the states. Asymptotically ($n \to \infty$), this is strengthened to an equality by the *quantum Chernoff bound* [35, 36]

$$\lim_{n\to\infty} -\frac{\log\{\min_A[\alpha_n(A) + \beta_n(A)]\}}{n}$$
$$= \max_{\alpha\in(0,1)} (1 - \alpha) D_\alpha(\rho||\sigma) := \xi(\rho||\sigma). \qquad (28)$$

The quantity on the right-hand side of this equation is called the *quantum Chernoff distance*.

We progress to the asymmetric treatment of this problem, quantum hypothesis testing. The asymmetric optimization is the task of minimizing one of the errors while keeping the other error below some fixed, finite threshold ϵ. We define

$$\alpha_n^*(\epsilon) := \min_{\beta_n(A)\leq\epsilon} [\alpha_n(A)], \quad \beta_n^*(\epsilon) := \min_{\alpha_n(A)\leq\epsilon} [\beta_n(A)]. \qquad (29)$$

Quantum Stein's lemma [37,38] asserts that, for *any* $\epsilon \in (0, 1)$, the type-II error decreases exponentially with the rate given by the relative entropy

$$\lim_{n\to\infty} -\frac{\log[\beta_n^*(\epsilon)]}{n} = D(\rho||\sigma). \qquad (30)$$

Quantum Stein's lemma can be further refined to optimize the error of the first kind, assuming that the error of the second kind decays exponentially. Defining

$$\tilde{\alpha}_{n,r} := \min_{\beta_n(A)\leq e^{-nr}} [\alpha_n(A)], \qquad (31)$$

the PRRE determines this error rate if $r < D(\rho||\sigma)$ [39–41],

$$\lim_{n\to\infty} -\frac{\log[\tilde{\alpha}_{n,r}]}{n} = \max_{\alpha\in(0,1)} \left[\frac{\alpha - 1}{\alpha}[r - D_\alpha(\rho||\sigma)]\right], \quad (32)$$

while the SRRE determines this error rate if $r > D(\rho||\sigma)$ [42],

$$\lim_{n\to\infty} -\frac{\log[1 - \tilde{\alpha}_{n,r}]}{n} = \max_{\alpha\in(1,\infty)} \left[\frac{\alpha - 1}{\alpha}[r - \tilde{D}_\alpha(\rho||\sigma)]\right]. \qquad (33)$$

With the above review, we have a thorough understanding of how to quantify the ability to discriminate between *two* quantum states. It would be desirable to generalize this to an arbitrary, finite number of states $\{\rho_i\}$. As we will see in Sec. IV, this is particularly important for the black hole information problem. In the case that we are discriminating many states, we no longer consolidate the M_i into A and $\mathbb{1} - A$. Rather, each measurement outcome i can lead us to conclude that we have state ρ_i. If we are given the state ρ_i with probability p_i, the error probability is given by

$$P_{\text{err}}(\rho_i, M_i) := \sum_{i=1} \text{Tr}[p_i\rho_i(\mathbb{1} - M_i)], \qquad (34)$$

whose optimized value we define as

$$P_{\text{err}}^*(\rho_i) := \min_{\{M_i\}} [P_{\text{err}}(\rho_i, M_i)]. \qquad (35)$$

Rather remarkably, building on the work of Refs. [43–46], the quantum Chernoff bound was generalized in the multiple state case, referred to as the *multiple quantum Chernoff*

bound [47,48]

$$\lim_{n \to \infty} - \frac{\log[P^*_{\text{err}}(\rho_i)]}{n} = \min_{i \neq j} \left[\max_{\alpha \in (0,1)} (1 - \alpha) D_\alpha(\rho_i || \rho_j) \right]. \tag{36}$$

The value on the right-hand side is referred to as the *multiple quantum Chernoff distance*. When comparing to Eq. (28), it is surprising that, when discriminating between arbitrarily many more states, all one needs to do is apply a global minimum.

In the one-shot case, bounds can be placed on $P^*_{\text{err}}(\rho_i)$, though, to our knowledge, an equality is not known. If we take the spectral decompositions of our POVM as $M_i := \sum_i^{T_i} \lambda_{ik} Q_{ik}$, an upper and bound is given by [47]

$$\frac{\sum_{i<j} \sum_{k,l} \min[\lambda_{ik}, \lambda_{jl}] \text{Tr}[Q_{ik} Q_{jl}]}{2(r-1)}$$
$$\leq P^*_{\text{err}}(\rho_i)$$
$$\leq 10(r-1)^2 T^2 \sum_{i<j} \sum_{k,l} \min[\lambda_{ik}, \lambda_{jl}] \text{Tr}[Q_{ik} Q_{jl}], \tag{37}$$

where r is the total number of states and $T := \max[T_i]$. The upper bound can be made more intuitive, though generally weaker, by noting that

$$\sum_{k,l} \min[\lambda_{ik}, \lambda_{jl}] \text{Tr}[Q_{ik} Q_{jl}] \leq \min_{\alpha \in (0,1)} (1 - \alpha) D_\alpha(\rho_i || \rho_j), \tag{38}$$

leading to

$$P^*_{\text{err}}(\rho_i) \leq 10(r-1)^2 T^2 \sum_{i<j} \min_{\alpha \in (0,1)} (1 - \alpha) D_\alpha(\rho_i || \rho_j)$$
$$\leq 5(r-1)^3 r T^2 \max_{i \neq j} \left[\min_{\alpha \in (0,1)} (1 - \alpha) D_\alpha(\rho_i || \rho_j) \right], \tag{39}$$

where in the second line we have removed the remaining sum to mimic the form of Eq. (36) even though this formula is strictly weaker when setting $r = 2$. In passing, we note that determining P^*_{err} is a computation that can be formulated as a semidefinite program [31,49], which means that it may be efficiently evaluated.

III. DISTINGUISHING RANDOM STATES

In this section, we consider Haar random states. This ensemble can be described in several ways. Perhaps the simplest is to consider an arbitrary reference state $|0\rangle \in \mathcal{H}$ and act with a random unitary matrix drawn from the Haar measure, the unique left-right invariant measure over $U(\dim \mathcal{H}) : |\Psi\rangle = U|0\rangle$. This ensemble is particularly nice

because the averages over the α copies of Haar random states are sums of permutations, τ, of the α copies

$$\overline{|\Psi\rangle\langle\Psi|^{\otimes \alpha}} = \frac{\sum_{\tau \in S_\alpha} g_\tau}{\sum_{\tau \in S_\alpha} \text{Tr}[g_\tau]}, \tag{40}$$

where g_τ is the matrix representation of τ and the denominator ensures that the state has unit norm. We are generally interested in ensembles for mixed states that are induced from taking a partial trace over a sub-Hilbert space. If $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, the reduced density matrix on A is given by

$$\rho_A^{\otimes \alpha} := \text{Tr}_B \left[\overline{|\Psi\rangle\langle\Psi|^{\otimes \alpha}} \right] = \frac{\sum_{\tau \in S_\alpha} g_{\tau_A} \text{Tr}[g_{\tau_B}]}{\sum_{\tau \in S_\alpha} \text{Tr}[g_\tau]}, \tag{41}$$

where the subscript on the permutation elements mean that they only permute within a sub-Hilbert space. The trace of a permutation element is straightforward to work out, equaling the dimension of the Hilbert space, $d_A d_B$, to the number of cycles in the permutation, $C(\tau)$. The denominator can then be written as

$$\sum_{\tau \in S_\alpha} \text{Tr}[g_\tau] = \sum_{\tau \in S_\alpha} (d_A d_B)^{C(\tau)}, \tag{42}$$

which can be summed exactly because we know that the number of permutations of α elements with k cycles is given by the Stirling number of the first kind. However, we can easily avoid this technical point because we will be interested in the regime where the Hilbert space dimension is large. Therefore, only the permutations that maximize $C(\tau)$ will contribute at leading order. The unique permutation that maximizes $C(\tau)$ is the identity permutation that has α cycles, so throughout this paper, we approximate the denominator as $(d_A d_B)^\alpha$.

There is an alternative description of the same induced ensemble of density matrices that will be useful for us when generalizing to tensor networks. Rather than starting with fiducial state $|0\rangle$, we begin with complete bases, $|i\rangle_A$ and $|J\rangle_B$, on $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. The Haar random state is then represented as

$$|\Psi\rangle = \mathcal{N} \sum_{iJ} X_{iJ} |i\rangle_A \otimes |J\rangle_B, \tag{43}$$

where the X_{iJ} are complex Gaussian independent and identically distributed matrix elements of $d_A \times d_B$ matrix X with (unnormalized) joint probability distribution

$$P(\{X_{iJ}\}) \propto \exp[-d_A d_B \text{Tr}(XX^\dagger)], \tag{44}$$

and $\mathcal{N}$ is a normalization constant. The reduced density matrix on $\mathcal{H}_A$ is then [50–52]

$$\rho_A = \frac{XX^\dagger}{\text{Tr}[XX^\dagger]}. \tag{45}$$

Ensemble averages over n copies are given by the same formula in terms of permutations, so at large dimensions, $\rho_A \simeq XX^\dagger$. This is the famous Wishart-Laguerre ensemble and is equivalent to the previously introduced Haar random states [53,54]. The advantage of working with random Gaussian states instead of Haar random states is due to "Wick calculus" being simpler than "Weingarten calculus." The difference will appear for random tensor networks although the ensembles will still be equivalent at large Hilbert space dimension [55]. Moreover, the class of random tensor networks used for holography involving projected Haar random states [56] precisely correspond to the states we study even at finite Hilbert space dimension, as explained in Appendix C.

We now introduce a diagrammatic approach for computations of certain moments of the Wishart ensemble involving multiple states, building on Refs. [3,57–59]. This will prove invaluable in the following calculations.

We represent the elements of the random global pure state as two vertical lines

$$|\Psi\rangle_{iJ} = X_{iJ} := \overset{i\,J}{\begin{vmatrix} \vdots \end{vmatrix}} \ , \tag{46}$$

where the solid line represents $\mathcal{H}_A$ and the dashed line $\mathcal{H}_B$. To form the density matrix, we take the outer product

$$[|\Psi\rangle\langle\Psi|]_{iJ,jK} = X_{iJ}X^*_{jK} = \overset{i\,J}{\begin{vmatrix} \vdots \end{vmatrix}}\ \overset{Kj}{\begin{vmatrix} \vdots \end{vmatrix}} \ . \tag{47}$$

We usually drop the index labeling of the lines to avoid cumbersome notation. All matrix manipulations are done on the lower ends of the lines. For example, we can take a partial trace over $\mathcal{H}_B$ by connecting the dashed line,

$$[\rho_A]_{i,j} = \sum_{J=1}^{d_B} X_{iJ}X^*_{jJ} := \overset{i\,J\qquad Jj}{\begin{vmatrix} \mathrel{\vrule height0pt}{\text{-----}}\end{vmatrix}} \ , \tag{48}$$

square the matrix by taking two copies and connecting the *bra* of the first matrix with the *ket* of the second,

$$[\rho_A^2]_{i,j} = \begin{vmatrix} \text{-----}\end{vmatrix}\ \sqcup\ \begin{vmatrix} \text{-----}\end{vmatrix} \ , \tag{49}$$

then take a trace by connecting the remaining solid lines to determine the purity,

$$\mathrm{Tr}\left[\rho_A^2\right] = \underline{\begin{vmatrix} \text{-----}\end{vmatrix}\ \sqcup\ \begin{vmatrix} \text{-----}\end{vmatrix}} \ . \tag{50}$$

For every insertion of the density matrix, we include a factor of $(d_A d_B)^{-1}$. This will give the normalization factor that we computed from Eq. (42).

The ensemble averaging of the states are done on the upper ends of the lines. The rule here is that we must add

up all diagrams contracting any *bra* with any *ket*. For α insertions of the density matrix, there will be $\alpha!$ diagrams, corresponding to the $\alpha!$ allowed permutations. Within each diagram, we count the number of loops with each loop giving a factor of the Hilbert space dimension. One can see that this diagrammatic sum is precisely the numerator of Eq. (40).

We can now practice by taking the ensemble averaged purity. There are two (2!) diagrams descending from Eq. (50), i.e.,

$$\overline{\mathrm{Tr}\left[\rho_A^2\right]} = \ \widehat{\begin{matrix}\frown\ \frown\end{matrix}}\ +\ \widehat{\begin{matrix}\frown\end{matrix}} \ , \tag{51}$$

immediately leading to $d_A^{-1} + d_B^{-1}$.

Because we are interested in distinguishing density matrices that are *independently* sampled from the ensemble, we must extend the diagrammatic technique. We do this by introducing different colors for different density matrices. When ensemble averaging, *bras* of one color can only contract with *kets* of the same color. For example, the overlap between independent induced states ρ_A (black) and σ_A (red) looks similar to the purity,

$$\mathrm{Tr}\left[\rho_A\sigma_A\right] = \begin{vmatrix} \text{-----}\end{vmatrix}\ \sqcup\ \begin{vmatrix} \text{\textcolor{red}{-----}}\end{vmatrix} \ , \tag{52}$$

but the ensemble averaging will only include a single diagram,

$$\overline{\mathrm{Tr}\left[\rho_A\sigma_A\right]} = \ \widehat{\begin{matrix}\frown\ \textcolor{red}{\frown}\end{matrix}}\ = \frac{1}{d_A} \tag{53}$$

because the second diagram would have connected the black and red indices, which is disallowed. With this formalism, we are now ready to compute each distinguishability measure using a replica trick.

A. Relative entropy

We begin with the von Neumann relative entropy, the topic of Ref. [3], both because it is the most fundamental quantity and the simplest to compute using our techniques. This will illustrate our strategy that will be used throughout. The relative entropy may be computed using a replica trick. That is, we first compute a certain series of moments of the ensemble and then analytically continue to arrive at the desired quantity. The replica trick for the relative

entropy is given by [60]

$$D(\rho_A || \sigma_A) = \lim_{\alpha \to 1} \frac{1}{\alpha - 1} \{ \log(\text{Tr}[\rho_A^\alpha]) - \log(\text{Tr}[\rho_A \sigma_A^{\alpha-1}]) \}. \tag{54}$$

We compute the ensemble average of the two terms separately. The first term is the Rényi entropy and, as a diagram, looks like

$$\text{Tr}[\rho_A^\alpha] = \underbrace{\llcorner\text{- - -}\lrcorner\,\sqcup\,\llcorner\text{- - -}\lrcorner\,\llcorner\quad\overset{\cdots}{\quad}\quad\lrcorner\llcorner\text{- - -}\lrcorner}. \tag{55}$$

While we make the dimensions of the sub-Hilbert spaces large, $d_A, d_B \propto N \to \infty$, we keep their relative sizes, d_A/d_B, finite. The leading diagrams maximize the total number of loops. These are the planar diagrams as this double line notation corresponds to the standard large-N topological expansion. Planar diagrams correspond to the *noncrossing permutations*, NC_α, a well-studied object in enumerative combinatorics and probability theory [61,62]. The ensemble averaged Rényi purity is then given by

$$\overline{\text{Tr}[\rho_A^\alpha]} = \frac{1}{(d_A d_B)^\alpha} \sum_{\tau \in \text{NC}_\alpha} d_A^{C(\eta^{-1} \circ \tau)} d_B^{C(\tau)}, \tag{56}$$

where η is the cyclic permutation, spawning from the matrix multiplication and trace in Eq. (55). The noncrossing permutations maximize the total exponent as $C(\eta^{-1} \circ \tau) + C(\tau) = \alpha + 1$ [61,62]. A more refined statement is that the number of noncrossing permutations with $C(\eta^{-1} \circ \tau) = k$ [and therefore $C(\tau) = \alpha + 1 - k$] is given by the Narayana number [61,62]

$$N_{\alpha,k} := \frac{1}{\alpha} \binom{\alpha}{k} \binom{\alpha}{k-1}. \tag{57}$$

With this information, we can reorganize Eq. (56) as a sum over k instead of a sum over permutations to obtain

$$\overline{\text{Tr}[\rho_A^\alpha]} = \frac{1}{(d_A d_B)^\alpha} \sum_{k=1}^\alpha N_{\alpha,k} d_A^k d_B^{\alpha+1-k}, \tag{58}$$

which can be rewritten again as a hypergeometric

function [63] as

$$\overline{\text{Tr}[\rho_A^\alpha]} = \begin{cases} d_A^{1-\alpha} \, _2F_1\!\left(1 - \alpha, -\alpha; 2; \dfrac{d_A}{d_B}\right), & d_A < d_B, \\[2ex] d_B^{1-\alpha} \, _2F_1\!\left(1 - \alpha, -\alpha; 2; \dfrac{d_B}{d_A}\right), & d_A > d_B. \end{cases} \tag{59}$$

The $A \leftrightarrow B$ symmetry of Rényi entropies of bipartite pure states is manifest.

Taking the logarithm and analytically continuing to $\alpha = 1$, we obtain Page's formula [64]

$$\lim_{\alpha \to 1} \frac{1}{1-\alpha} \overline{\log(\text{Tr}[\rho_A^\alpha])} = \begin{cases} \log[d_A] - \dfrac{d_A}{2d_B}, & d_A < d_B, \\[2ex] \log[d_B] - \dfrac{d_B}{2d_A}, & d_B < d_A. \end{cases} \tag{60}$$

In writing this formula, we have assumed that the logarithm and ensemble average commute. In Appendix B, we explain why this is true when the Hilbert space dimensions are large.

The second term in Eq. (54) involves both ρ_A and σ_A [65]:

$$\text{Tr}[\rho_A \sigma_A^{1-\alpha}] = \underbrace{\llcorner\text{- - -}\lrcorner\,\sqcup\,\llcorner\text{- - -}\lrcorner\,\llcorner\quad\overset{\cdots}{\quad}\quad\lrcorner\llcorner\text{- - -}\lrcorner}. \tag{61}$$

Because there is only a single copy of ρ_A, when ensemble averaging, we must contract the first density matrix with itself. There are no constraints on how to contract the red lines. This means that the S_α permutations are broken down to $\mathbb{1} \times S_{\alpha-1}$:

$$\overline{\text{Tr}[\rho_A \sigma_A^{1-\alpha}]} = \frac{1}{(d_A d_B)^\alpha} \sum_{\tau \in \mathbb{1} \times S_{\alpha-1}} d_A^{C(\eta^{-1} \circ \tau)} d_B^{C(\tau)}. \tag{62}$$

We still need to maximize the exponent by choosing noncrossing permutations, though many such permutations are disallowed by the identity factor on the first matrix. The diagrams are topological, so we have

$$\overset{\frown}{}\,\llcorner\text{- - -}\lrcorner\,\llcorner\quad\overset{\cdots}{\quad}\quad\lrcorner\llcorner\text{- - -}\lrcorner = \overset{\frown}{}\;\llcorner\text{- - -}\lrcorner\,\llcorner\quad\overset{\cdots}{\quad}\quad\lrcorner\llcorner\text{- - -}\lrcorner. \tag{63}$$

From this diagram, it is clear that the cardinality of the intersection of NC_α and $\mathbb{1} \times S_{\alpha-1}$ is given by the cardinality of $\text{NC}_{\alpha-1}$, and the number of such noncrossing permutations with $C(\eta^{-1} \circ \tau) = k$ is given by the Narayana

number $N_{\alpha-1,k}$; hence,

$$\overline{\text{Tr}[\rho_A \sigma_A^{1-\alpha}]} = \frac{1}{(d_A d_B)^\alpha} \sum_{k=1}^\alpha N_{\alpha-1,k} d_A^k d_B^{\alpha+1-k}. \tag{64}$$

This can also be represented by a hypergeometric function as

$$
\overline{\mathrm{Tr}[\rho_A \sigma_A^{\alpha-1}]}
$$
$$
= \begin{cases}
d_A^{1-\alpha}\, {}_2F_1\!\left(1-\alpha, 2-\alpha; 2; \dfrac{d_A}{d_B}\right), & d_A < d_B, \\[2ex]
d_B^{2-\alpha} d_A^{-1}\, {}_2F_1\!\left(1-\alpha, 2-\alpha; 2; \dfrac{d_B}{d_A}\right), & d_A > d_B.
\end{cases}
$$
(65)

Taking the $\alpha \to 1$ limit, we have

$$
\lim_{\alpha \to 1} \frac{1}{1-\alpha} \overline{\log(\mathrm{Tr}[\rho_A \sigma_A^{\alpha-1}])}
$$
$$
= \begin{cases}
\log[d_A] + 1 + \left(\dfrac{d_B}{d_A} - 1\right)\log\!\left[1 - \dfrac{d_A}{d_B}\right], & d_A < d_B, \\[2ex]
\infty, & d_B < d_A.
\end{cases}
$$
(66)

Therefore, we find the ensemble average of the relative entropy to be

$$
\overline{D(\rho_A || \sigma_A)}
$$
$$
= \begin{cases}
1 + \dfrac{d_A}{2d_B} + \left(\dfrac{d_B}{d_A} - 1\right)\log\!\left[1 - \dfrac{d_A}{d_B}\right], & d_A < d_B, \\[2ex]
\infty, & d_A > d_B.
\end{cases}
$$
(67)

This is a satisfying, simple answer. For small d_A/d_B, the relative entropy is given by d_A/d_B. If we think in terms of the "number of qubits," N_A and N_B, this is exponentially small in the difference $(N_B - N_A)$, meaning that the states will be very difficult to distinguish whenever we

have access to a few qubits less than half the system; the asymptotic error rate, $\beta_n^*(\epsilon)$, is very small, meaning that we need exponentially (in N_A) many copies of the state to identify it with confidence. Equation (67) is also monotonically increasing in d_A/d_B, a consequence of the data processing inequality when we take the partial trace as the quantum channel. When $d_A \to d_B$, the relative entropy approaches the curious value of $3/2$. This value of $3/2$ was also determined in Ref. [66] using very different techniques that serves as an additional consistency check of our results.

When $d_A > d_B$, every reduced state on $\mathcal{H}_A$ in the ensemble will be rank deficient with $d_A - d_B$ zero eigenvalues. This is because the Wishart ensemble has rank at most $\min(d_A, d_B)$. It is therefore overwhelmingly unlikely that two independent states, ρ_A and σ_A, will have the same support. In particular, the support of ρ_A will not be contained within the support of σ_A. This is the reason why the relative entropy becomes infinite in this regime; there will be a measurement we can choose that easily distinguishes ρ_A and σ_A.

B. Petz Rényi relative entropy and Holevo's just-as-good fidelity

To understand more sophisticated structures in Haar random states, we progress to the computation of the PRRE. The PRRE has a tricky $1 - \alpha$ exponent for σ_A, so we use the following replica trick with two replica parameters, α and m:

$$
D_\alpha(\rho_A || \sigma_A) = \lim_{m \to 1-\alpha} \frac{1}{\alpha - 1} \log(\mathrm{Tr}[\rho_A^\alpha \sigma_A^m]).
$$
(68)

We compute this for $\alpha, m \in \mathbb{Z}^+$, only taking the limit to $\alpha, m \in \mathbb{R}$ at the end of the calculation. The positive integer moments in diagrammatic form are

$$
\mathrm{Tr}[\rho_A^\alpha \sigma_A^m] = \boxed{},
$$
(69)

where there are α black density matrices and m red density matrices. When ensemble averaging, we are only able to contract using the subgroup $S_\alpha \times S_m \subset S_{\alpha+m}$, leading to the sum over permutations

$$
\overline{\mathrm{Tr}[\rho_A^\alpha \sigma_A^m]} = \frac{1}{(d_A d_B)^{\alpha+m}} \sum_{\tau \in S_\alpha \times S_m} d_A^{C(\eta^{-1}\circ\tau)} d_B^{C(\tau)}.
$$
(70)

As can be seen by the diagram, even with the restricted sum, there are many ways to contract the lines that are

noncrossing, hence maximizing the exponents. These are precisely the noncrossing permutations acting independently on the black and red indices, so the combinatorial factor will be given by the product of two Narayana numbers:

$$
\overline{\mathrm{Tr}[\rho_A^\alpha \sigma_A^m]} = \frac{1}{(d_A d_B)^{\alpha+m}} \sum_{k=1}^{\alpha} \sum_{j=1}^{m} N_{\alpha,k} N_{m,j} d_A^{k+j-1}
$$
$$
\times d_B^{2+\alpha+m-k-j}.
$$
(71)

The reason why there is an additional "-1" in the exponent of d_A is that the black and red lines are connected at the bottom of the diagram due to the matrix multiplication. Note that this expression is a generalization of the replica trick used in the previous section for the relative entropy, Eq. (64), if we set $\alpha = 1$ and $m = \alpha - 1$. As before, the double sum can be expressed in terms of hypergeometric functions as

$$
\overline{\text{Tr}[\rho_A^\alpha \sigma_A^m]} =
\begin{cases}
d_A^{-m-\alpha+1} \, {}_2F_1\left(1-m,-m;2;\dfrac{d_A}{d_B}\right) {}_2F_1\left(1-\alpha,-\alpha;2;\dfrac{d_A}{d_B}\right), & d_A < d_B, \\[2ex]
\dfrac{d_B^{-m-\alpha+2}}{d_A} \, {}_2F_1\left(1-m,-m;2;\dfrac{d_B}{d_A}\right) {}_2F_1\left(1-\alpha,-\alpha;2;\dfrac{d_B}{d_A}\right), & d_A > d_B.
\end{cases}
\tag{72}
$$

Now that the sum that required m to be an integer is complete; it is safe to take the $m \to 1 - \alpha$ limit to obtain

$$
\overline{\text{Tr}[\rho_A^\alpha \sigma_A^{1-\alpha}]} =
\begin{cases}
{}_2F_1\left(1-\alpha,-\alpha;2;\dfrac{d_A}{d_B}\right) {}_2F_1\left(\alpha-1,\alpha;2;\dfrac{d_A}{d_B}\right), & d_A < d_B, \\[2ex]
\dfrac{d_B}{d_A} \, {}_2F_1\left(1-\alpha,-\alpha;2;\dfrac{d_B}{d_A}\right) {}_2F_1\left(\alpha-1,\alpha;2;\dfrac{d_B}{d_A}\right), & d_A > d_B.
\end{cases}
\tag{73}
$$

When $d_A = d_B$, this precisely agrees with a formula from Ref. [66]. Taking the logarithm leads to an exact closed-form expression for the PRRE in the large Hilbert space dimension limit:

$$
\overline{D_\alpha(\rho_A||\sigma_A)} = \frac{1}{\alpha-1}
\begin{cases}
\log\left[{}_2F_1\left(1-\alpha,-\alpha;2;\dfrac{d_A}{d_B}\right) {}_2F_1\left(\alpha-1,\alpha;2;\dfrac{d_A}{d_B}\right) \right], & d_A < d_B, \\[2ex]
\log\left[\dfrac{d_B}{d_A} \, {}_2F_1\left(1-\alpha,-\alpha;2;\dfrac{d_B}{d_A}\right) {}_2F_1\left(\alpha-1,\alpha;2;\dfrac{d_B}{d_A}\right) \right], & d_A > d_B.
\end{cases}
\tag{74}
$$

This is a rare instance where we have an exact closed-form solution for relative entropies and can be thought of as the "Page formula" for PRRE. Importantly, this equation contains much more information about random quantum states than Eq. (67). A highlight is the finiteness of Eq. (74) for $\alpha < 1$ in the $d_A > d_B$ regime. This explains the approach of random quantum states to complete distinguishability. There are a few consistency checks that we can readily verify. Namely, we note that Eq. (74) reduces to Eq. (67) if we send $\alpha \to 1$, and Eq. (74) is monotonically increasing in d_A/d_B (data processing inequality) and monotonically increasing in α.

   An additional desirable property of Eq. (74) is that it is simple enough that we can perform the optimization needed to compute the quantum Chernoff distance

$$
\overline{\xi(\rho_A||\sigma_A)} =
\begin{cases}
-2\log\left[{}_2F_1\left(\dfrac{1}{2},-\dfrac{1}{2};2;\dfrac{d_A}{d_B}\right) \right], & d_A < d_B, \\[2ex]
-\log\left[\dfrac{d_B}{d_A} \, {}_2F_1\left(\dfrac{1}{2},-\dfrac{1}{2};2;\dfrac{d_B}{d_A}\right)^2 \right], & d_A > d_B,
\end{cases}
\tag{75}
$$

where the optimal value of α in Eq. (28) is found to be $1/2$. This definitively establishes the error rate in quantum state discrimination for a measure one set of quantum states. Because $\alpha = 1/2$ is the optimal value, this adds to the usefulness of Holevo's just-as-good fidelity, which is given by

$$
\overline{F_H(\rho_A||\sigma_A)} =
\begin{cases}
{}_2F_1\left(\dfrac{1}{2},-\dfrac{1}{2};2;\dfrac{d_A}{d_B}\right)^4, & d_A < d_B, \\[2ex]
\dfrac{d_B^2}{d_A^2} \, {}_2F_1\left(\dfrac{1}{2},-\dfrac{1}{2};2;\dfrac{d_B}{d_A}\right)^4, & d_A > d_B.
\end{cases}
\tag{76}
$$

In order to evaluate the quantum multiple Chernoff distance, we need to characterize the fluctuations in the PRRE. To compute the variance, we must compute

$$
\text{Tr}\left[\rho_A^\alpha \sigma_A^m\right]^2 = \underbrace{} \quad \underbrace{}.
\tag{77}
$$

Only the diagrams that connect the two blocks will contribute to the variance because the disconnected diagrams are subtracted. For small d_A/d_B, these contributions will be $O(d_A^{1-2\alpha-2m}d_B^{-1})$ or $O(d_A^{-1}d_B^{-1})$ after taking the relevant limit. We may use a Taylor expansion of the logarithm to determine that the variance of the PRRE, σ^2, will be the same order. The higher degree central moments, and therefore higher cumulants, will be subleading because, in general, the nth central moment will be $O(d_A^{1-n\alpha-nm}d_B^{-1})$. Therefore, the PRRE will follow a normal distribution at subleading order.

For a normal distribution, the probability of random variable X being r standard deviations, σ, below the mean, μ, is

$$\Pr(X - \mu \leq -r\sigma) = \int_{-\infty}^{-k\sigma} dx \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}$$
$$= \frac{1}{2}\text{erfc}\left(\frac{r}{\sqrt{2}}\right). \quad (78)$$

Therefore, if we have W independent samplings of ρ_A and σ_A, the probability that the minimum relative entropy will be at most r standard deviations from the mean is

$$\Pr\{\min[D_\alpha(\rho_A||\sigma_A)] \geq \mu - r\sigma\} = \left[1 - \frac{1}{2}\text{erfc}\left(\frac{r}{\sqrt{2}}\right)\right]^W. \quad (79)$$

If we are discriminating between W states, the quantum multiple Chernoff distance will be

$$\xi_W(\rho_A||\sigma_A) \geq \frac{d_A}{4d_B} - \sqrt{2}\text{erfc}^{-1}[2 - 2(1-\epsilon_1)^{1/W}]\sigma \quad (80)$$

with probability $1 - \epsilon_1$. In order to be confident in the state discrimination ($P_{\text{err}}^* < \epsilon_2$), we need

$$n \simeq \frac{\log[\epsilon_2^{-1}]}{d_A/4d_B - \sqrt{2}\text{erfc}^{-1}[2 - 2(1-\epsilon_1)^{1/W}]\sigma} \quad (81)$$

copies of the state. Because of σ being suppressed in the total Hilbert space dimension, this formula only mildly depends on W even when W is of the order of the Hilbert space dimension. Thus, the multiple Chernoff bound is essentially just as tight as the two-state Chernoff bound.

C. Sandwiched Rényi relative entropy and Uhlmann fidelity

Continuing our progression in difficulty, we now compute the SRRE using the following new replica trick requiring two replica indices:

$$\tilde{D}_\alpha(\rho_A||\sigma_A) := \lim_{m\to[(1-\alpha)/2\alpha]} \frac{1}{\alpha - 1} \log\{\text{Tr}[(\sigma_A^m \rho_A \sigma_A^m)^\alpha]\}. \quad (82)$$

The associated diagrams are more complicated because there the red and black lines are not cleanly partitioned,

$$\text{Tr}\left[(\sigma_A^m \rho_A \sigma_A^m)^\alpha\right] =$$

$$(83)$$

There are still many ways to contract the above diagram without crossing lines. We take two steps. First, we need to have the black lines contract with themselves in a noncrossing manner. For example, we may have
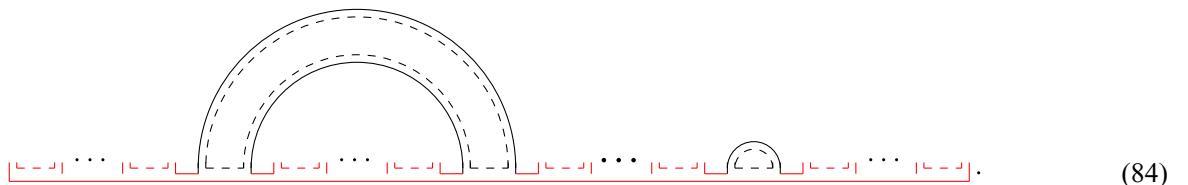
$$(84)$$

This gives a factor of $d_B^{C(\tau)}$, where τ is the noncrossing permutation of the α black lines. We can see from this diagram that depending on how the black lines are contracted, this restricts the allowed permutations for the red lines. This is why this computation is more complicated than for the PRRE where the black and red permutations simply factorized as $\text{NC}_\alpha \times \text{NC}_m$. In order for the global permutation to be noncrossing, the red permutations must be noncrossing within each block partitioned off by the black permutations. In the above example diagram, the black "rainbow" restricts the red permutation to be of the form $\text{NC}_{2m} \times \cdots$. The identity permutation on the black density matrix on the right places no additional restrictions. In terms of equations, the diagrams may be summed as

$$\overline{\mathrm{Tr}[(\sigma_A^m \rho_A \sigma_A^m)^\alpha]} = \frac{1}{(d_A d_B)^{\alpha(2m+1)}} \sum_{\tau \in \mathrm{NC}_\alpha} d_B^{C(\tau)} \prod_{\zeta \in \mathrm{cyc}(\eta^{-1}\circ\tau)} \sum_{\gamma \in \mathrm{NC}_{2m|\zeta|}} d_A^{C(\eta^{-1}\circ\gamma)} d_B^{C(\gamma)}, \tag{85}$$

where the product is over the cycles of $\eta^{-1} \circ \tau$ and $|\cdot|$ represents the length of the cycle.

We first focus on the $d_A < d_B$ regime where the inner sum may be computed as before, i.e.,

$$\overline{\mathrm{Tr}[(\sigma_A^m \rho_A \sigma_A^m)^\alpha]} = \frac{\sum_{\tau \in \mathrm{NC}_\alpha} d_B^{C(\tau)} \prod_{\zeta \in \mathrm{cyc}(\eta^{-1}\circ\tau)} d_A d_B^{2|\zeta|m} \, _2F_1(-2|\zeta|m, 1-2|\zeta|m; 2; d_A/d_B)}{(d_A d_B)^{\alpha(2m+1)}}$$

$$= \frac{\sum_{\tau \in \mathrm{NC}_\alpha} d_A^{\alpha+1-C(\tau)} d_B^{C(\tau)+2m\alpha} \prod_{\zeta \in \mathrm{cyc}(\eta^{-1}\circ\tau)} \, _2F_1(-2|\zeta|m, 1-2|\zeta|m; 2; d_A/d_B)}{(d_A d_B)^{\alpha(2m+1)}}. \tag{86}$$

In the second line we have pulled out the factors of d_A and d_B from the product by enforcing the global permutation to be noncrossing. This formula does not need m to be an integer, so it is now safe to take the $m \to (1-\alpha)/2\alpha$ limit to obtain

$$\overline{\mathrm{Tr}[(\sigma_A^{(1-\alpha)/2\alpha} \rho_A \sigma_A^{(1-\alpha)/2\alpha})^\alpha]} = \sum_{\tau \in \mathrm{NC}_\alpha} \left(\frac{d_A}{d_B}\right)^{\alpha-C(\tau)} \prod_{\zeta \in \mathrm{cyc}(\eta^{-1}\circ\tau)} \, _2F_1\left(\frac{|\zeta|(\alpha-1)}{\alpha}, 1+|\zeta|-\frac{|\zeta|}{\alpha}; 2; \frac{d_A}{d_B}\right). \tag{87}$$

The product over cycle structures makes this formula still very difficult. Fortunately, Kreweras solved exactly this combinatorial problem about cycle structure in his landmark paper on noncrossing partitions [61]. He found that the number of noncrossing permutations of $\{1, 2, \ldots, \alpha\}$ with cycle structure [67] $\{m_i\}$ is given by, what we will call, the Kreweras number [61,62]

$$K_{\{m_i\}} := \frac{\alpha!}{(\alpha-b+1)! \, m_1! \cdots m_\alpha!}, \quad b := \sum_i m_i. \tag{88}$$

Therefore, we can reorganize the sum such that there are no more references to permutations, only natural numbers,

$$\overline{\mathrm{Tr}[(\sigma_A^{(1-\alpha)/2\alpha} \rho_A \sigma_A^{(1-\alpha)/2\alpha})^\alpha]} = \sum_{m_1,\ldots,m_\alpha=0}^{\alpha} K_{\{m_i\}} \left(\frac{d_A}{d_B}\right)^{\sum_i m_i-1} \prod_{m_i \neq 0} \, _2F_1\left(\frac{i(\alpha-1)}{\alpha}, 1+i-\frac{i}{\alpha}; 2; \frac{d_A}{d_B}\right)^{m_i}. \tag{89}$$

This formula still presents a daunting task to evaluate in terms of elementary functions for generic α, though it provides a tractable, controlled expansion in d_A/d_B. This is because, for small d_A/d_B, the hypergeometric function is close to one. We then must consider the smallest values of $\sum_i m_i$. First, taking only the leading term with $\sum_i m_i = 1$ (cyclic permutation), we obtain

$$\overline{\mathrm{Tr}[(\sigma_A^{(1-\alpha)/2\alpha} \rho_A \sigma_A^{(1-\alpha)/2\alpha})^\alpha]} = \, _2F_1\left(\alpha-1, \alpha; 2; \frac{d_A}{d_B}\right) + O\left(\frac{d_A}{d_B}\right). \tag{90}$$

This is not terribly useful because, as explained above, to this order, the right-hand side is exactly one, which would lead to the SRRE being identically zero. To find a nontrivial result, we need the next term where $\sum_i m_i = 2$, which can be achieved in many ways.

$$\overline{\mathrm{Tr}[(\sigma_A^{(1-\alpha)/2\alpha} \rho_A \sigma_A^{(1-\alpha)/2\alpha})^\alpha]} = \frac{d_A}{d_B} \sum_{j=1}^{\lfloor \alpha/2 \rfloor} K_{\{m_j, m_{\alpha-j}=1\}} \, _2F_1\left(\frac{j(\alpha-1)}{\alpha}, 1+j-\frac{j}{\alpha}; 2; \frac{d_A}{d_B}\right)$$

$$\times \, _2F_1\left(\frac{(\alpha-j)(\alpha-1)}{\alpha}, 1+(\alpha-j)-\frac{\alpha-j}{\alpha}; 2; \frac{d_A}{d_B}\right)$$

$$+ \, _2F_1\left(\alpha-1, \alpha; 2; \frac{d_A}{d_B}\right) + O\left(\frac{d_A}{d_B}\right)^2, \tag{91}$$

where the floor function in the sum ensures that we do not double count. The Kreweras number is

$$K_{\{m_j, m_{\alpha-j}=1\}} = \begin{cases} \alpha, & j \neq \dfrac{\alpha}{2}, \\ \dfrac{\alpha}{2}, & j = \dfrac{\alpha}{2}; \end{cases} \qquad (92)$$

$j = \alpha/2$ will only occur when α is even. The exact forms of the hypergeometric functions in the sum are not important at this order because, for small d_A/d_B, they are all close to one. Therefore, only the Kreweras number is important. We can easily compute the sum at this order for any integer α and find that this parity effect disappears,

$$\overline{\text{Tr}\big[(\sigma_A^{(1-\alpha)/2\alpha}\rho_A\sigma_A^{(1-\alpha)/2\alpha})^\alpha\big]}$$
$$= 1 + \alpha(\alpha - 1)\frac{d_A}{d_B} + O\left(\frac{d_A}{d_B}\right)^2, \qquad (93)$$

leading to a SRRE of

$$\overline{\tilde{D}_\alpha(\rho_A||\sigma_A)} = \alpha\frac{d_A}{d_B} + O\left(\frac{d_A}{d_B}\right)^2. \qquad (94)$$

Note that this agrees with the previously derived von Neumann relative entropy in the relevant $\alpha \to 1$ limit.

Moreover, it obeys the data processing inequality for all positive α if we take the quantum channel to be the partial trace.

The Uhlmann fidelity is found [68] by setting $\alpha = 1/2$:

$$\overline{F(\rho_A||\sigma_A)} = 1 - \frac{d_A}{2d_B} + O\left(\frac{d_A}{d_B}\right)^2. \qquad (95)$$

At this order, the Uhlmann fidelity is identical to Holevo's just-as-good fidelity, Eq. (76). The value of the fidelity exactly at $d_A = d_B$ was found in Ref. [66] to be 9/16 and additional results may be found in Ref. [69].

We can also evaluate the SRRE exactly for any integer moment using Eq. (89). Here, we work out the least tedious case of $\alpha = 2$ that is also known as the *collision relative entropy* [70]. In this case, Eq. (91) is actually exact and does not contain $O(d_A/d_B)^2$ corrections. We only sum over $j = 1$, so

$$\overline{\text{Tr}[(\sigma_A^{-1/4}\rho_A\sigma_A^{-1/4})^2]} = \frac{d_A}{d_B}\,_2F_1\left(\frac{1}{2}, \frac{3}{2}; 2; \frac{d_A}{d_B}\right)^2$$
$$+ \frac{1}{1 - d_A/d_B}, \qquad (96)$$

leading to a SRRE of

$$\overline{\tilde{D}_2(\rho_A||\sigma_A)} = \begin{cases} \log\left[\dfrac{d_A}{d_B}\,_2F_1\left(\dfrac{1}{2}, \dfrac{3}{2}; 2; \dfrac{d_A}{d_B}\right)^2 + \dfrac{1}{1 - d_A/d_B}\right], & d_A < d_B, \\ \infty, & d_A > d_B, \end{cases} \qquad (97)$$

where we have set the SRRE to infinity when $d_A > d_B$ because the von Neumann relative entropy is infinite in this regime and the SRREs are monotonically increasing with α. It is straightforward to evaluate the higher integer SRREs if desired.

It is equally important to investigate the opposite regime where d_A/d_B is large. In this case, the inner sum in Eq. (85) gives

$$\overline{\text{Tr}[(\sigma_A^m\rho_A\sigma_A^m)^\alpha]} = \frac{\sum_{\tau \in \text{NC}_\alpha} d_B^{C(\tau)} \prod_{\zeta \in \text{cyc}(\eta^{-1}\circ\tau)} d_A^{2|\zeta|m} d_B\,_2F_1(-2|\zeta|m, 1 - 2|\zeta|m; 2; d_B/d_A)}{(d_A d_B)^{\alpha(2m+1)}}$$

$$= \frac{\sum_{\tau \in \text{NC}_\alpha} (d_B/d_A)^{C(\tau)+C(\eta^{-1}\circ\tau)} d_A^{\alpha(2m+1)+1} \prod_{\zeta \in \text{cyc}(\eta^{-1}\circ\tau)}\,_2F_1(-2|\zeta|m, 1 - 2|\zeta|m; 2; d_B/d_A)}{(d_A d_B)^{\alpha(2m+1)}}. \qquad (98)$$

Now that we have done the sums over the m permutations, we can safely take $m \to (1 - \alpha)/2\alpha$ and rewrite the sum in terms of Kreweras numbers as

$$\overline{\text{Tr}\big[(\sigma_A^{(1-\alpha)/2\alpha}\rho_A\sigma_A^{(1-\alpha)/2\alpha})^\alpha\big]} = \left(\frac{d_B}{d_A}\right)^\alpha \sum_{m_1,\ldots,m_\alpha=0}^{\alpha} K_{\{m_i\}} \prod_{m_i \neq 0}\,_2F_1\left(\frac{i(\alpha-1)}{\alpha}, 1 + \frac{i(\alpha-1)}{\alpha}; 2; \frac{d_B}{d_A}\right)^{m_i}. \qquad (99)$$

This is an exact formula, but is difficult to evaluate away from limits. For large d_A/d_B, all of the hypergeometric functions are close to one so all that matters is the total number of noncrossing permutations, which is given by the Catalan number

$$C_\alpha := \frac{1}{\alpha + 1}\binom{2\alpha}{\alpha}. \tag{100}$$

Therefore, at leading order, we have

$$\overline{\mathrm{Tr}\big[(\sigma_A^{(1-\alpha)/2\alpha}\rho_A\sigma_A^{(1-\alpha)/2\alpha})^\alpha\big]} = C_\alpha\left(\frac{d_B}{d_A}\right)^\alpha + O\left(\frac{d_B}{d_A}\right)^{\alpha+1}. \tag{101}$$

Unlike for small d_A/d_B, there are no additional terms at leading order. The SRRE is thus

$$\overline{\tilde{D}_\alpha(\rho_A||\sigma_A)} = \frac{\alpha}{\alpha - 1}\log\left[\frac{d_B}{d_A}\right] + \frac{1}{\alpha - 1}\log[C_\alpha]$$
$$+ O\left(\frac{d_B}{d_A}\right). \tag{102}$$

It is important to note that this is only well defined for $\alpha < 1$. This is to be expected because of rank deficiency. In the well-defined regime, the SRRE is monotonic in α and manifestly obeys the data processing inequality.

We can evaluate the asymptotic expression at $\alpha = 1/2$ to find the Uhlmann fidelity

$$\overline{F(\rho_A||\sigma_A)} = \frac{64 d_B}{9\pi^2 d_A} + O\left(\frac{d_B}{d_A}\right)^{3/2}. \tag{103}$$

The prefactor comes from the Catalan number that is nonintegral for noninteger α. We see that the fidelity is inversely proportional to d_A/d_B, decaying to zero when subsystem A occupies most of the Hilbert space. The full spectrum of $\sigma_A^{(1-\alpha)/2\alpha}\rho_A\sigma_A^{(1-\alpha)/2\alpha}$, and hence the fidelity, may be evaluated using techniques of free probability theory. This is completed in Appendix A. The answer is the free multiplicative convolution of two Marchenko-Pastur distributions.

D. Trace distance

The final distinguishability measure we discuss is the trace distance. This is the ideal measure when discussing one-shot state discrimination (24). More general than Eq. (17), we can define an α-norm version of the trace distance as

$$T_\alpha(\rho_A||\sigma_A) := \frac{1}{2^{1/\alpha}}|\rho_A - \sigma_A|_\alpha, \tag{104}$$

where the α-norm of an operator, A, is defined as

$$|A|_\alpha := \big(\mathrm{Tr}\big[\sqrt{A^\dagger A}^\alpha\big]\big)^{1/\alpha}. \tag{105}$$

For even α and Hermitian A, we can dispose of the square root:

$$|A|_{\tilde{\alpha}} = (\mathrm{Tr}[A^{2\tilde{\alpha}}])^{1/2\tilde{\alpha}}, \quad 2\tilde{\alpha} := \alpha. \tag{106}$$

The trace norm is then the $\tilde{\alpha} \to 1/2$ limit of this expression.

The replica trick, while only requiring a single replica parameter, is quite difficult as we must compute all even powers of $\rho_A - \sigma_A$, which involves arbitrary mixing of ρ_A and σ_A [71]:

$$\mathrm{Tr}[(\rho_A - \sigma_A)^\alpha] = \sum_{\mathcal{S}\in\mathcal{P}(\{1,2,...,\alpha\})}(-1)^{|\mathcal{S}|}\mathrm{Tr}(\tau_{\mathcal{S}_1}\cdots\tau_{\mathcal{S}_\alpha}). \tag{107}$$

Here the sum runs over the power set of $\{1, 2, \ldots, \alpha\}$ and $|\cdot|$ is the cardinality of the subset. We have $\tau_{\mathcal{S}_i} = \rho_A$ if $i \in \mathcal{S}$ and $\tau_{\mathcal{S}_i} = \sigma_A$ if $i \notin \mathcal{S}$. Each term in the sum can be expressed as an appropriate summation over the symmetric group, though this is far from straightforward.

Consider the small d_A/d_B limit. In this case, the terms that maximize $C(\tau)$ will dominate the sum. This is when τ is the identity. This permutation is always present, regardless of $\mathcal{S}$ and universally contributes as $d_A^{1-\alpha}$, which in the $\alpha \to 1$ limit contributes at $O(1)$. To see if this contributes to the overall sum, we need to understand the cardinalities. The number of subsets with cardinality k is given by the binomial coefficient, so the identity contributes as

$$\overline{\mathrm{Tr}[(\rho_A - \sigma_A)^\alpha]} \supset \sum_{k=0}^{\alpha}\binom{\alpha}{k}(-1)^k d_A^{1-\alpha} = 0. \tag{108}$$

To get a nontrivial answer, we must therefore move beyond the identity permutation. This is to be expected because the trace distance will be small for small d_A/d_B and should not be $O(1)$. The next leading term is when $C(\tau) = \alpha - 1$, which corresponds to the identity on all sites except for two that are swapped; this is always noncrossing. This contributes universally as $d_A^{2-\alpha}d_B^{-1}$, which will lead to the $O(d_A/d_B)$ contribution. The combinatorics are slightly more complicated. If $|\mathcal{S}| = k$ then there are $\binom{k}{2} + \binom{\alpha-k}{2}$ ways to have a single pairing because we can only choose

pairings within the block of k ρ_A's or $n - k$ σ_A's. Therefore, the contribution at this order is

$$
\overline{\mathrm{Tr}[(\rho_A - \sigma_A)^\alpha]}
$$
$$
\supset \sum_{k=0}^{\alpha} \binom{\alpha}{k}\left[\binom{k}{2} + \binom{\alpha - k}{2}\right](-1)^k d_A^{2-\alpha} d_B^{-1}
$$
$$
= 2\delta_{\alpha,2} d_A^{2-\alpha} d_B^{-1}. \tag{109}
$$

This is not an analytic function, so the $\tilde{\alpha} \to 1/2$ limit is quite ambiguous. We are free to work to higher orders, though we will argue that this will not help our cause.

In the small d_B/d_A regime, the expansion is more involved. The leading terms come from maximizing $C(\eta^{-1} \circ \tau)$. We can only have $C(\eta^{-1} \circ \tau) = \alpha$ in the case that $\mathcal{S} = \{1, 2, \ldots, \alpha\}$ or is empty because otherwise $\tau = \eta$ will not be an allowed permutation. At this order, we therefore have

$$
\overline{\mathrm{Tr}[(\rho_A - \sigma_A)^\alpha]} = \begin{cases} 2d_B^{1-\alpha}, & \alpha \in 2\mathbb{Z}, \\ 0 & \alpha \in 2\mathbb{Z} + 1. \end{cases} \tag{110}
$$

The parity effect arises from the exponent of the sign in the sum. Analytically continuing the even integers to one, we find that

$$
\overline{T(\rho_A || \sigma_A)} = 1 + O\left(\frac{d_B}{d_A}\right), \tag{111}
$$

meaning that the states are nearly maximally distant. To understand how the trace distance approaches one, we need to work at the next order. The noncrossing permutations that give $C(\eta^{-1} \circ \tau) = \alpha - 1$ are those that are of the form $\eta_{\alpha_1} \times \eta_{\alpha_2}$. This means that the ρ_A and the σ_A must be in disjoint blocks, i.e., $\mathcal{S}$ is a set only containing consecutive integers. There are $(\alpha - 1)$ ways to partition α into nonzero integers α_1 and α_2 if we define the tuple (α_1, α_2) to be distinct from (α_2, α_1). There is an additional factor of α coming from the rotations of $\mathcal{S}$ to $\mathcal{S} + 1$, leading to

$$
\overline{\mathrm{Tr}[(\rho_A - \sigma_A)^\alpha]} \supset \sum_{k=1}^{\alpha-1} \alpha(-1)^k d_A^{-1} d_B^{2-\alpha}
$$
$$
= \begin{cases} -\alpha d_A^{-1} d_B^{2-\alpha}, & \alpha \in 2\mathbb{Z}, \\ 0 & \alpha \in 2\mathbb{Z} + 1. \end{cases} \tag{112}
$$

Finally, when $\mathcal{S} = \{1, 2, \ldots, \alpha\}$ or $\mathcal{S} = \emptyset$, we again can partition the elements into α_1 and α_2 size blocks, but this time (α_1, α_2) and (α_2, α_1) are indistinguishable. Therefore, for even α, there are $\alpha/2$ possibilities while for odd α, there are only $(\alpha - 1)/2$ possibilities plus the rotation factors [72], leading to

$$
\overline{\mathrm{Tr}[(\rho_A - \sigma_A)^\alpha]} \supset \begin{cases} \alpha(\alpha - 1) d_A^{-1} d_B^{2-\alpha}, & \alpha \in 2\mathbb{Z}, \\ 0 & \alpha \in 2\mathbb{Z} + 1, \end{cases} \tag{113}
$$

where the odd terms are trivial because the $\mathcal{S} = \{1, 2, \ldots, \alpha\}$ and $\mathcal{S} = \emptyset$ terms exactly cancel in the sum due to the power of the sign. Taking the $\alpha \to 1$ limit of even α, we find the trace norm at this order to be

$$
\overline{T(\rho_A || \sigma_A)} = 1 - \frac{d_B}{2d_A} + O\left(\frac{d_B}{d_A}\right)^2. \tag{114}
$$

The trace distance may be evaluated away from limits using free probability techniques, as we review in Appendix A [73], giving

$$
\overline{T(\rho_A || \sigma_A)} = \begin{cases} \dfrac{\sqrt{d_A(2d_B - d_A)}(d_A + d_B) + d_B(4d_A - 2d_B)\sin^{-1}(\sqrt{d_A/2d_B})}{2\pi d_A d_B}, & d_A < 2d_B, \\ 1 - \dfrac{d_B}{2d_A}, & d_A > 2d_B. \end{cases} \tag{115}
$$

Interestingly, our asymptotic formula was exact. The trace distance at $d_A = d_B$ was found in Ref. [66] to be $1/2 + 1/\pi$.

Without free probability, we can still use bounds from Sec. II to place strong constraints on the trace distance. In particular, this is helpful for the small d_A/d_B where we were unable to find an analytic answer. First, we use Pinsker's inequality (20), with relative entropy (67) as an upper bound for $d_A < d_B$, to obtain

$$
\overline{T(\rho_A || \sigma_A)} \leq \sqrt{2 + \frac{d_A}{d_B} + 2\left(\frac{d_B}{d_A} - 1\right)\log\left(1 - \frac{d_A}{d_B}\right)}. \tag{116}
$$

This is only a useful bound when the right-hand side is less than one. Recall that, when $d_A = d_B$, the right-hand side will be $\sqrt{3} > 1$. For small d_A/d_B, we have

$$\overline{T(\rho_A||\sigma_A)} \leq \sqrt{\frac{2d_A}{d_B}} + O\left(\frac{d_A}{d_B}\right)^{3/2}. \qquad (117)$$

This means that the trace distance is very small, though without a lower bound, we cannot yet say that we could not find the leading order trace distance from the above expansion. We can determine a lower bound using Holevo's just-as-good fidelity that is, in general, stronger than the lower bound from the Uhlmann fidelity:

$$1 - {}_2F_1\left(\frac{1}{2}, -\frac{1}{2}; 2; \frac{d_A}{d_B}\right)^2$$

$$\leq \overline{T(\rho_A||\sigma_A)} \leq \sqrt{1 - {}_2F_1\left(\frac{1}{2}, -\frac{1}{2}; 2; \frac{d_A}{d_B}\right)^4}. \qquad (118)$$

Here we have also included the upper bound that is, in general, weaker than the upper bound from the Uhlmann fidelity. For small d_A/d_B, this gives

$$\frac{d_A}{4d_B} + O\left(\frac{d_A}{d_B}\right)^2 \leq \overline{T(\rho_A||\sigma_A)} \leq \sqrt{\frac{d_A}{2d_B}} + O\left(\frac{d_A}{d_B}\right)^{3/2}. \qquad (119)$$

This is a stronger upper bound than from Pinsker's inequality, though the scaling is still not nailed down, only constrained to between linear and square root with d_A/d_B. Because the scaling is at most linear, we cannot hope to find the leading order behavior from the expansion at the beginning of this subsection because the linear term was zero if we are to trust the "continuation." The Uhlmann fidelity for small d_A/d_B does not strengthen the upper bound at leading order.

We also want to characterize the $d_A > d_B$ regime. While Pinsker's inequality does not help here because the relative entropy is infinite, Holevo's just-as-good fidelity sets the

nontrivial upper and lower bounds

$$1 - \frac{d_B}{d_A} {}_2F_1\left(\frac{1}{2}, -\frac{1}{2}; 2; \frac{d_B}{d_A}\right)^2$$

$$\leq \overline{T(\rho_A||\sigma_A)} \leq \sqrt{1 - \frac{d_B^2}{d_A^2} {}_2F_1\left(\frac{1}{2}, -\frac{1}{2}; 2; \frac{d_B}{d_A}\right)^4}. \qquad (120)$$

For small d_B/d_A, this is

$$1 - \frac{d_B}{d_A} + O\left(\frac{d_B}{d_A}\right)^2 \leq \overline{T(\rho_A||\sigma_A)} \leq 1 - \frac{d_B^2}{2d_A^2} + O\left(\frac{d_B}{d_A}\right)^3, \qquad (121)$$

meaning that the states are almost as far away from each other as possible, approaching one exponentially in $N_A - N_B$. The upper bound can be improved by the Uhlmann fidelity at leading order such that the scaling behavior is completely fixed:

$$1 - \frac{d_B}{d_A} + O\left(\frac{d_B}{d_A}\right)^2$$

$$\leq \overline{T(\rho_A||\sigma_A)} \leq 1 - \frac{32d_B}{9\pi^2 d_A} + O\left(\frac{d_B}{d_A}\right)^2. \qquad (122)$$

These bounds are consistent with analytic expressions (114) and (115).

E. Small-N numerics

All of our computations thus far have been in the limit where both d_A and d_B are large. It is important to ask whether these asymptotic results are accurate when d_A and d_B are finite. One motivation is if these predictions can be observed in experiments and noisy intermediate-scale quantum technology [74]. Of course, the Hilbert space dimensions are exponentially large in the number of qubits, so there is hope that our results are predictive for small-scale experiments. In this section, we numerically compute
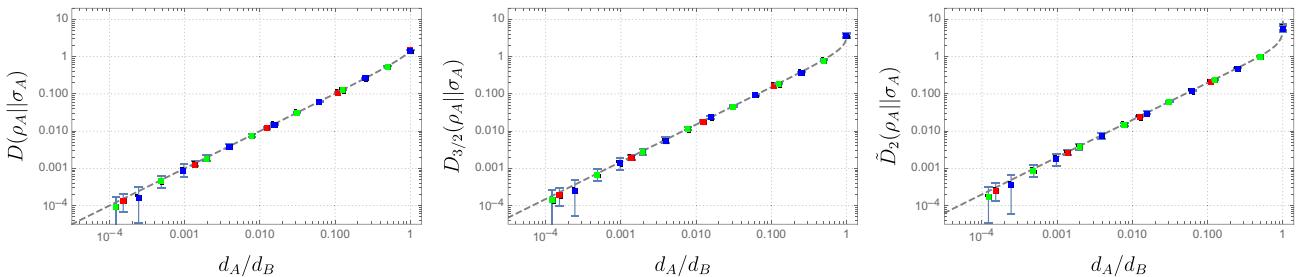
FIG. 1. Three representative examples of relative entropies are shown, all of which with $\alpha \geq 1$, so they are infinite for $d_A/d_B > 1$. Left: von Neumann relative entropy with the dashed line given by Eq. (67). Center: PRRE with $\alpha = 3/2$ with the dashed line given by Eq. (74). Right: SRRE with $\alpha = 2$ with the dashed line given by Eq. (97). The data are given for total Hilbert space dimensions of 2^{14} (blue), 2^{15} (green), and 3^{10} (red). The error bars represent the statistical fluctuations in the 10^3 disorder realizations that decay for large Hilbert spaces.
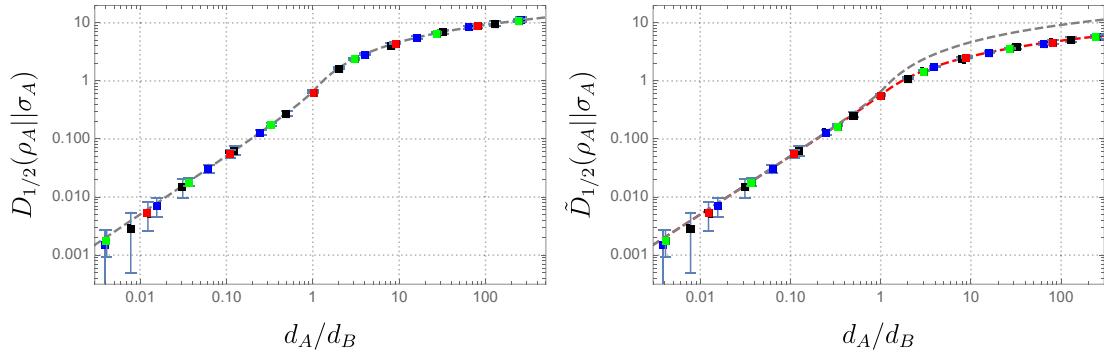
FIG. 2.   Representative examples of relative entropies are shown for $\alpha < 1$ such that there are no divergences. Left: PRRE for $\alpha = 1/2$ (related to Holevo's just-as-good fidelity) with the dashed line given by Eq. (74). Right: SRRE for $\alpha = 1/2$ (related to the Uhlmann fidelity) with the dashed red line given by the answer from spectrum (A15). The gray line is the upper bound from the PRRE. The data are given for total Hilbert space dimensions 2^9 (black), 2^{10} (blue), 3^6 (red), and 3^7 (green). The error bars represent the statistical fluctuations in the 10^3 disorder realizations that decay for large Hilbert spaces.

the various distance measures and compare to the asymptotic formulas. This serves as a further consistency check of our results, which we find to be extraordinarily accurate.

In Fig. 1, we plot the von Neumann relative entropy, $D_{3/2}$, and $\tilde{D}_2$. All of these quantities are infinite for $d_A > d_B$ due to the rank deficiencies in the reduced density matrix. For this reason, we are able to sample very large Hilbert space dimensions because the bottleneck on classical computers is d_A and not the total system size. We find very accurate agreement between the exact large-N predictions and the small-N numerics. The fluctuations in the entropies are noticeably larger for small d_A because of the subleading corrections that we have thus far ignored.

In Fig. 2, we investigate the other regime by plotting $D_{1/2}$ and $\tilde{D}_{1/2}$. These quantities are related to Holevo's just-as-good and Uhlmann fidelities, respectively, and are therefore well defined in the $d_A > d_B$ regime. This limits the Hilbert space sizes we can probe, though we still find very accurate agreement with the large-N analysis.

Finally, in Fig. 3, we plot the trace distance, examining both the small d_A/d_B and large d_A/d_B regimes. The large-N expressions precisely agree with numerics and are bounded within the fidelities.

IV. DISTINGUISHING BLACK HOLES

While studying random states is interesting in its own right, the physical implications of our results becomes significantly richer when we apply them to gravitational systems. We explain how the connection between random states and gravitational systems is more than an analogy and in some ways, quantitatively identical.

A. Fixed-area states in holography

In quantum field theory, we compute the moments of reduced density matrices by evaluating the partition function on certain replica manifolds [75,76]. These are glued according to the relevant trace structure. If the quantum field theory is holographic, we may map the calculation
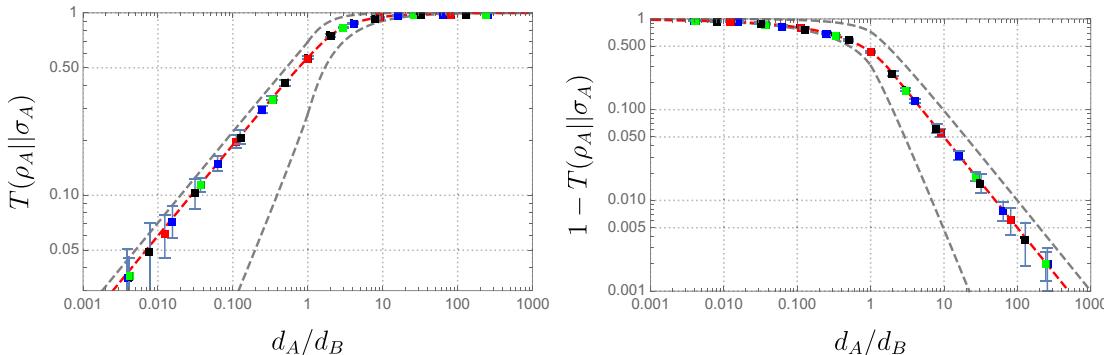
FIG. 3.   Left: the trace distance is shown with the dashed red line given by Eq. (115) and gray lines given by the bounds from fidelities (118) and (120). Right: one minus the trace distance is shown to display the approach to one. The data are given for total Hilbert space dimensions 2^9 (black), 2^{10} (blue), 3^6 (red), and 3^7 (green). The error bars represent the statistical fluctuations in the 10^3 disorder realizations that decay for large Hilbert spaces.

to an evaluation of the gravitational path integral with boundary conditions prescribed by this trace structure. In the gravitational path integral, we are instructed to sum over all geometries with the given boundary conditions. In the derivation of the Ryu-Takayanagi formula [77], only replica symmetric geometries were considered. In contrast, we find that replica symmetry breaking saddles are important for the evaluation of relative entropies.

In general, it is very difficult to evaluate the gravitational path integral for multiple replicas. This is because the nontrivial coupling between the replicas leads to backreaction, changing the bulk geometry [78]. A great simplification in the gravitational path integral can be made if we focus on "fixed area states" [9,10]. These are states where the area of one or more surfaces is fixed and not integrated over. In general, the different replicas will not backreact among themselves, so we are left with copies of the original bulk geometry except for potential conical singularities appearing at the locations of the fixed surfaces.

As an example, consider the Rényi entropies of a region on the boundary of a pure state black hole background [79]. There exist two extremal surfaces that are candidate Ryu-Takayanagi surfaces, γ_1 and γ_2, each wrapping the black hole horizon in topologically distinct manners. Denote the areas of these two surfaces A_1 and A_2, respectively. The moments of the reduced density matrix are

$$\text{Tr}[\rho_A^\alpha] = \frac{\mathcal{Z}(\rho_A^\alpha)}{\mathcal{Z}(\rho_A)^\alpha}, \tag{123}$$

where the numerator is the gravitational path integral on the replicated geometry and the denominator is the path integral on a single copy, necessary for normalization. Because the geometry is identical in both geometries away from the conical singularities, the numerator and denominator will almost completely cancel. The nontrivial terms come from the actions of the conical singularities that are determined by their opening angles

$$\text{Tr}[\rho_A^\alpha] = \sum_{\tau \in S_\alpha} e^{[C(\eta^{-1} \circ \tau) - \alpha]A_1/4G_N + [C(\tau) - \alpha]A_2/4G_N}$$
$$\times \text{Tr}[\rho_b^{n_1}]\text{Tr}[\rho_b^{n_2}] \cdots \text{Tr}[\rho_b^{n_{C(\tau)}}], \tag{124}$$

where G_N is Newton's constant, the n_k are the lengths of the cycles in τ, and ρ_b is the bulk state labeling the black hole microstate. These account for the bulk entropy term in the FLM formula [80]. The sum over the permutation group arises from all of the ways the replicas may be glued together in the codimension-one region bounded by the two fixed surfaces (see Fig. 4). We have chosen the bulk state to be pure such that all of the bulk traces are one; hence,

$$\text{Tr}[\rho_A^\alpha] = \sum_{\tau \in S_\alpha} e^{[C(\eta^{-1} \circ \tau) - \alpha]A_1/4G_N + [C(\tau) - \alpha]A_2/4G_N}. \tag{125}$$
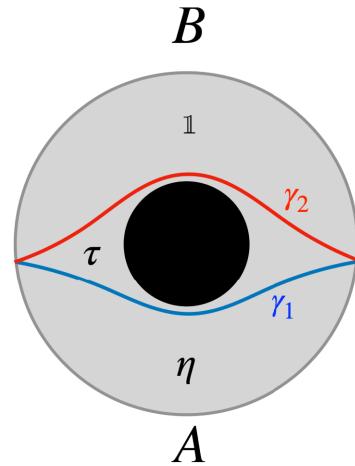
FIG. 4. An AdS black hole is shown with the asymptotic boundary partitioned into two regions A and B. The two candidate Ryu-Takayanagi surfaces are shown in blue and red, respectively. In the gravitational replica trick, the region bounded by γ_1 and A is glued cyclically (η) among the replicas. The region bounded by γ_2 and B is glued according to the identity ($\mathbb{1}$). The region between γ_1 and γ_2 is not fixed by the asymptotic boundary conditions and may be glued among the replicas according to arbitrary S_α permutations (τ).

This sum should now look familiar as it is identical to the sum needed for the Rényi entropies of Haar random states, Eq. (56), once identifying $d_A \leftrightarrow e^{A_1/4G_N}$ and $d_B \leftrightarrow e^{A_2/4G_N}$. In this way, entropies in fixed-area states in holography are identical to entropies in Haar random states [81].

This connection becomes even richer when we consider more than one gravitational state to compute the relative entropies. Consider the following moments needed for the von Neumann relative entropy:

$$\text{Tr}[\rho_A \sigma_A^{\alpha-1}] = \frac{\mathcal{Z}(\rho_A \sigma_A^{\alpha-1})}{\mathcal{Z}(\rho_A)\mathcal{Z}(\sigma_A)^{\alpha-1}}. \tag{126}$$

Both states have fixed areas and the same semiclassical geometry, but come from different black hole microstates, ρ_b and σ_b. In the language of Refs. [82,83], they are orthogonal states in the same code subspace. Just as before, the gravitational path integral instructs us to sum over all topologies, meaning that the region between the two fixed-area surfaces can be glued according to any S_α permutation.

$$\text{Tr}[\rho_A \sigma_A^{\alpha-1}] = \sum_{\tau \in S_\alpha} e^{[C(\eta^{-1} \circ \tau) - \alpha]A_1/4G_N + [C(\tau) - \alpha]A_2/4G_N}$$
$$\times \text{Tr}[\rho_b \sigma_b^{n_1 - 1}]\text{Tr}[\sigma_b^{n_2}] \cdots \text{Tr}[\sigma_b^{n_{C(\tau)}}]. \tag{127}$$

This is generally different from the calculation in Haar random states that only contained a sum over a subgroup $\mathbb{1} \times S_{\alpha-1}$, Eq. (62). However, because ρ_b and σ_b are orthogonal, $\text{Tr}[\rho_b \sigma_b^{n_1-1}]$ is only nonzero if $n_1 = 1$. This reduces the sum to

$$\text{Tr}[\rho_A \sigma_A^{\alpha-1}] = \sum_{\tau \in \mathbb{1} \times S_{\alpha-1}} e^{[C(\eta^{-1} \circ \tau) - \alpha] A_1 / 4 G_N + [C(\tau) - \alpha] A_2 / 4 G_N},$$

(128)

which is identical to (62) under the same identification. A nearly identical argument holds for the PRRE, SRRE, and trace distance. Therefore we conclude that not only do fixed-area states have the same entropies as Haar random states, but they also have identical Hilbert space geometries.

These results have interesting implications for the distinguishability of black hole microstates. Namely, the asymptotic observer with arbitrarily small information about the state (small A) is able to distinguish between any black hole microstates. This is surprising because we usually consider all black holes to look the same from outside the horizon to any observer, especially local observers. The catch is that the microstates are only distinguishable nonperturbatively in Newton's constant, $O(e^{-1/G_N})$. This is because all distinguishability measures are linear in d_A/d_B for small region A that translates to proportional to $e^{(A_1-A_2)/4G_N}$. This means that while distinguishability is in principle possible, the error rates in state discrimination will be very high unless the observer has an exponentially large number of copies of the system. The distinguishability is nonperturbatively small up until region A is roughly one qubit less than half the boundary system, at which point it becomes $O(1)$. When the observer has access to more than half of the boundary, the black hole microstates become completely distinguishable up to nonperturbatively small corrections.

We also note that these results represent nonperturbative corrections to the Jafferis-Lewkowycz-Maldacena-Suh (JLMS) formula [84] that asserts that the boundary relative entropy equals the bulk relative entropy within the entanglement wedge. We have considered bulk states that are pure, orthogonal, and localized between the two extremal surfaces; the bulk states are identical outside of the black hole [85]. When A is sufficiently small, the black hole is not within its entanglement wedge so the bulk states are identical, i.e., the bulk relative entropy is zero. Therefore, the JLMS formula asserts that boundary relative entropy is zero. We have shown that there are nonperturbative corrections to this statement.

B. The PSSY model and replica wormholes

In a landmark achievement, the Page curve [7] for an evaporating black hole was computed for the first time in two independent papers [86,87]. The key mechanism that "fixed" Hawking's calculation was the inclusion of certain wormhole saddles in the gravitational path integral, referred to as "replica wormholes" [11,88]. Using the toy model of black hole evaporation presented in Ref. [11] (PSSY), we now show the role of replica wormholes in calculations of relative entropy. This elucidates how the assumptions of Hawking fail. We call this a violation of the no-hair theorem, which is a nonperturbative effect and therefore not present in Hawking's calculation.

The PSSY model consists of two-dimension Jackiw-Teitelboim gravity decorated with end of the world (EOW) branes with k flavors. The Euclidean action is given by

$$I = -\frac{S_0}{2\pi}\left[\frac{1}{2}\int_{\mathcal{M}} \sqrt{g}R + \int_{\partial\mathcal{M}} \sqrt{h}K\right]$$
$$-\left[\frac{1}{2}\int_{\mathcal{M}} \sqrt{g}\phi(R+2) + \int_{\partial\mathcal{M}} \sqrt{h}\phi K\right]$$
$$+\mu \int_{\text{brane}} ds,$$

(129)

where S_0 is the large ground state entropy, g (h) is the bulk (asymptotic boundary) metric with curvature R (K), ϕ is the dilation, and μ is the tension of the EOW brane.

The EOW brane has $k \gg 1$ internal microstates. The global states on the black hole and radiation that we consider are of a maximally entangled form

$$|\Psi\rangle = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} |\psi_i\rangle_B |i\rangle_R,$$

(130)

where $|i\rangle_R$ represents an orthonormal basis of the states of the radiation. Consider a second microstate

$$|\Psi'\rangle = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} |\psi_i\rangle_B |i+1\rangle_R,$$

(131)

where $i + k \sim i$ is implied. The definitions of these states are not microscopic in the sense that the $|\psi_i\rangle_B$ are defined by a gravitational path integral and are not exactly orthogonal. As for the fixed-area state calculation, they may also be thought of as being orthogonal in the code subspace. Because of the nonorthogonality, the reduced states of Eqs. (130) and (131) on the radiation are not *a priori* identical even though the states appear to be related only by a local unitary transformation on B [89].

The overlap between these two states is

$$\langle\Psi'|\Psi\rangle = \frac{1}{k}\sum_{i,l=1}^{k}\langle\psi_l|\psi_i\rangle_B\langle l|i+1\rangle_R = \frac{1}{k}\sum_{i=1}^{k}\langle\psi_{i+1}|\psi_i\rangle_B.$$

(132)

The overlap on the right-hand side is given by a gravity amplitude

$$\langle \psi_i | \psi_j \rangle_B = \overset{i}{\circ}\!\!\cdots\!\!\underset{\longleftarrow}{}\!\!\cdots\!\!\overset{j}{\circ}. \tag{133}$$

Because $i \neq i+1$, connecting the brane in the gravity diagram is incompatible and the amplitude is zero. This means that $|\Psi\rangle$ and $|\Psi'\rangle$ are roughly orthogonal, but there are important caveats to this statement because the overlap should be thought of as an ensemble averaged statement. In particular, $|\langle \psi_i | \psi_{i+1} \rangle_B|^2$ is nonzero. This is completely analogous to the Haar random story where, on average, two independently chosen vectors will have zero overlap, but the variance is nonzero. The analog of Eq. (133) for random matrices is

$$\langle \psi_i | \psi_{j \neq i} \rangle_B = \boxed{\;\vdots\;\text{----}\;\vdots\;}. \tag{134}$$

Because, when ensemble averaging, we cannot contract black and red indices, the average, $\overline{\langle \psi_i | \psi_{j \neq i} \rangle_B}$, equals zero. In complete analogy, the ensemble average of

$$|\langle \psi_i | \psi_{j \neq i} \rangle_B|^2 = \boxed{\;\vdots\;\text{----}\;\vdots\;}\;\boxed{\;\vdots\;\text{----}\;\vdots\;} \tag{135}$$

is nonzero (though very small) because we may now contract red with red and black with black.

We are interested in the relative entropy of the radiation for two different microstates. Hawking and even the island formula papers assumed that the radiation is seen as purely thermal [90] before the Page time in accordance with the no-hair theorem, i.e., all black holes of the same mass, charge, and angular momentum look the same from the outside. After the Page time, while the island formula papers did not assume the radiation to be purely thermal, there was no difference between the calculations for different microstates of the black hole. From one perspective, this is great because unitarity can be realized without

knowing the microscopic theory. On the other hand, it is disappointing because it bypasses the question of why all initial states appear to lead to the same final state. We resolve this part of the information problem within the PSSY model and believe analogous results should hold in more realistic models of black hole evaporation.

The reduced density matrix on the radiation for the first state is given by

$$\rho_R := \text{Tr}_B |\Psi\rangle\langle\Psi| = \frac{1}{k} \sum_{i,j=1}^{k} \langle \psi_j | \psi_i \rangle_B |i\rangle\langle j|_R \tag{136}$$

and similarly for the second state is given by

$$\rho_R' := \text{Tr}_B |\Psi'\rangle\langle\Psi'| = \frac{1}{k} \sum_{i,j=1}^{k} \langle \psi_j | \psi_i \rangle_B |i+1\rangle\langle j+1|_R. \tag{137}$$

From here on out, we drop the subscripts labeling the Hilbert spaces as it should be clear.

We now compute the PRRE between two states of the radiation using the replica trick. We have

$$\text{Tr}[\rho_R^\alpha \rho_R'^m]$$

$$= \frac{1}{k^{\alpha+m}} \sum_{i_1,\dots,i_{\alpha+m},j_1,\dots,j_{\alpha+m}=1}^{k} \langle \psi_{j_1} | \psi_{i_1} \rangle \dots \langle \psi_{j_{\alpha+m}} | \psi_{i_{\alpha+m}} \rangle$$

$$\times \delta_{j_1 i_2} \delta_{j_2 i_3} \cdots \delta_{j_{\alpha-1} i_\alpha} \delta_{j_\alpha i_{\alpha+1}} \delta_{j_{\alpha+1}, i_{\alpha+2}} \cdots$$

$$\times \delta_{j_{\alpha+m-1}, i_{\alpha+m}} \delta_{j_{\alpha+m+1}, i_1}$$

$$= \frac{1}{k^{\alpha+m}} \sum_{i_1,\dots,i_{\alpha+m}=1}^{k} \langle \psi_{i_2} | \psi_{i_1} \rangle \langle \psi_{i_3} | \psi_{i_2} \rangle \cdots$$

$$\times \langle \psi_{i_\alpha} | \psi_{i_{\alpha-1}} \rangle \langle \psi_{i_{\alpha+1}+1} | \psi_{i_\alpha} \rangle$$

$$\times \langle \psi_{i_{\alpha+2}} | \psi_{i_{\alpha+1}} \rangle \langle \psi_{i_{\alpha+3}} | \psi_{i_{\alpha+2}} \rangle \cdots$$

$$\times \langle \psi_{i_{\alpha+m}} | \psi_{i_{\alpha+m-1}} \rangle \langle \psi_{i_1-1} | \psi_{i_{\alpha+m}} \rangle. \tag{138}$$
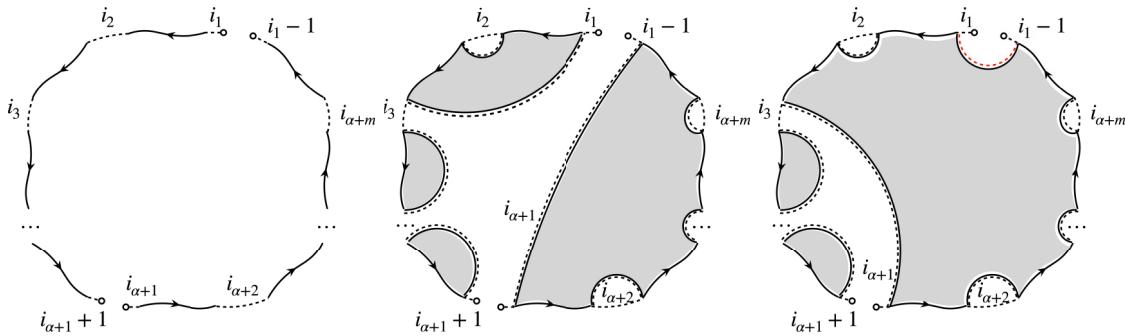

FIG. 5. Left: the boundary conditions for the path integral in Eq. (138). Center: an example of a legal way of filling in the geometry. This geometry is also planar, so it contributes at leading order. Right: an example of an illegal way of filling in the geometry because the EOW brane with label i_1 cannot be connected to the EOW brane with label $i_1 - 1$. It is not hard to convince oneself that every diagram that connects the left n and right m boundaries will always lead to an inconsistency as in the right diagram.

This is a more complicated but still tractable gravitational path integral. As shown in Fig. 5, the sum is over only the noncrossing permutations in the $S_\alpha \times S_m$ subgroup of $S_{\alpha+m}$ due to the boundary conditions on the EOW brane i_1 and $i_1 - 1$ being incompatible as well as $i_{\alpha+1}$ and $i_{\alpha+1} + 1$ being incompatible. There are crossing permutations in $S_{\alpha+m}/S_\alpha \times S_m$ that are compatible with the boundary conditions, but these are subleading. The only other difference in diagrams from the random matrix theory calculation is that the geometries are allowed to have additional handles. This, however, is unimportant because each handle will contribute a factor of e^{-2S_0} and we have assumed the ground state entropy to be large.

To compute the PRRE, we must evaluate the grav-

itational path integral on these replica geometries. For simplicity, we consider the case where the black hole is in the microcanonical ensemble, i.e., instead of fixing the lengths of the boundary, we fix the energy, E. The path integral of an n-boundary wormhole is [11]

$$\mathbf{Z}_n = e^{\mathbf{S}} y\left(\sqrt{2E}\right)^n, \tag{139}$$

where $\mathbf{S}$ is the microcanonical entropy at energy E. Because of the simply power of n, after normalizing the density matrix, the function y will drop out of the final answer. All calculations are then identical to random matrix theory with the identification of $k \leftrightarrow d_A$ and $e^{\mathbf{S}} \leftrightarrow d_B$ [91]. We choose to only write the PRRE

$$D_\alpha(\rho_R || \rho_R') = \frac{1}{\alpha - 1} \begin{cases} \log\left[{}_2F_1\left(1 - \alpha, -\alpha; 2; \frac{k}{e^{\mathbf{S}}}\right) {}_2F_1\left(\alpha - 1, \alpha; 2; \frac{k}{e^{\mathbf{S}}}\right) \right], & k < e^{\mathbf{S}}, \\ \log\left[\frac{e^{\mathbf{S}}}{k} {}_2F_1\left(1 - \alpha, -\alpha; 2; \frac{e^{\mathbf{S}}}{k}\right) {}_2F_1\left(\alpha - 1, \alpha; 2; \frac{e^{\mathbf{S}}}{k}\right) \right], & k > e^{\mathbf{S}}. \end{cases} \tag{140}$$

The quantum Chernoff bound asserts that

$$\lim_{n \to \infty} -\frac{\log\{\min_A[\alpha_n(A) + \beta_n(A)]\}}{n} = \max_{\alpha \in (0,1)} \begin{cases} -\log\left[{}_2F_1\left(1 - \alpha, -\alpha; 2; \frac{k}{e^{\mathbf{S}}}\right) {}_2F_1\left(\alpha - 1, \alpha; 2; \frac{k}{e^{\mathbf{S}}}\right) \right], & k < e^{\mathbf{S}}, \\ -\log\left[\frac{e^{\mathbf{S}} {}_2F_1(1 - \alpha, -\alpha; 2; e^{\mathbf{S}}/k) {}_2F_1(\alpha - 1, \alpha; 2; e^{\mathbf{S}}/k)}{k} \right], & k > e^{\mathbf{S}}, \end{cases}$$

$$= \begin{cases} -2\log\left[{}_2F_1\left(\frac{1}{2}, -\frac{1}{2}; 2; \frac{k}{e^{\mathbf{S}}}\right) \right], & k < e^{\mathbf{S}}, \\ -\log\left[\frac{e^{\mathbf{S}}}{k} {}_2F_1\left(\frac{1}{2}, -\frac{1}{2}; 2; \frac{e^{\mathbf{S}}}{k}\right)^2 \right], & k > e^{\mathbf{S}}, \end{cases} \tag{141}$$

where we found the maximum to be at $\alpha = 1/2$, i.e., Holevo's just-as-good fidelity

$$\min_A[\alpha_n(A) + \beta_n(A)] \leq F_H(\rho_R || \rho_R')^{n/2}. \tag{142}$$

While Eq. (142) saturates at large n, the right-hand side holds as an upper bound for all integer n. Holevo's just-as-good fidelity, plotted in Fig. 6, is

$$F_H(\rho_R || \rho_R') = \begin{cases} {}_2F_1\left(-\frac{1}{2}, \frac{1}{2}; 2; \frac{k}{e^{\mathbf{S}}}\right)^4, & k < e^{\mathbf{S}}, \\ \frac{e^{2\mathbf{S}} {}_2F_1[-1/2, 1/2; 2; e^{\mathbf{S}}/k]^4}{k^2}, & k > e^{\mathbf{S}}. \end{cases} \tag{143}$$

When observing a black hole from the outside, our task is not as simple as distinguishing *two* states. Rather, we

need to distinguish between all $e^{\mathbf{S}} \gg 2$ states of the black hole. On the face of it, this seems like an insurmountable task. However, using the multiple quantum Chernoff bound (36) and the normal distribution for relative entropies leading to Eq. (80), we determine that our asymptotic error in the multistate discrimination is identical, at leading order, to that of the two-state discrimination.

This has important implications on the nature of black hole evaporation that have not been addressed in the calculations of the entropy. The following island formula (or quantum Ryu-Takayanagi formula) was the main tool in recent calculations of entropy of Hawking radiation [92,93]:

$$S_{vN}(\rho_R) = \min_\chi \left[\frac{A_\chi}{4G_N} + S_{\text{semi-cl}}(\Sigma_\chi) \right]. \tag{144}$$
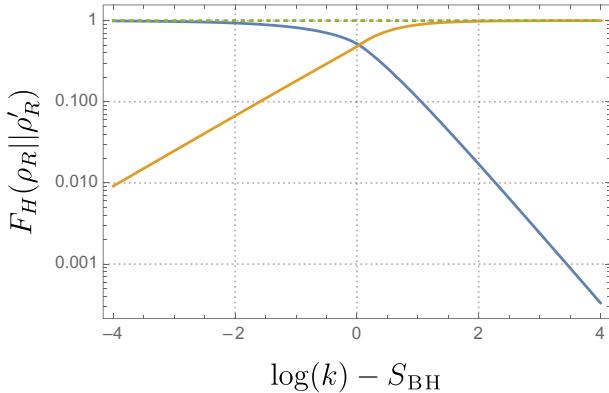
FIG. 6. Holevo's just-as-good fidelity (blue) and one minus the fidelity (orange) are shown for the PSSY model following Eq. (143). Before the Page time ($\log[k] = S_{\mathrm{BH}}$), the fidelity is exponentially close to one. After the Page time, the fidelity exponentially decays to zero. BH, Black Hole.

Here A_χ is the area of the codimension-two quantum extremal surface χ and $S_{\mathrm{semi\text{-}cl}}(\Sigma_\chi)$ is the von Neumann entropy of the bulk quantum fields in the codimension-one region, Σ_χ, bounded by χ in the bulk, i.e., the entanglement wedge of the radiation, R. While this formula accurately computes the von Neumann entropy of the radiation, restoring *consistency* with unitarity, it leaves more to be desired. In particular, the bulk entropy term is completely semiclassical and given by a quantum field theory calculation in curved space. The calculation is agnostic to the details of the black hole microstate. One of the central pieces of the apparent paradox was that Hawking radiation always looked the same on the outside regardless of the dynamics in the black hole interior, the phenomenon of *no hair*. Equation (143) instead tells us that there is *detectable* hair in the radiation even before the Page time. In fact, information about the particular microstate is present even in the first Hawking quantum. Because the fidelity of the radiation for any two different microstates is strictly less than one, we can always tell the difference between Hawking radiation coming from black holes that are in different microstates, even if they have the same macroscopic parameters mass, charge, and angular momentum.

The caveat is that the difference between states of the radiation coming from distinct black hole microstates is exponentially small in the black hole entropy, i.e., the deviation of the fidelity from one before the Page time is $O(e^{-\mathbf{S}})$. This means that while in principle possible, any reasonable observer will be hard pressed to observe this difference. If we want the probability of error in distinguishing the black hole microstates to be less than ϵ, we need an $O(e^{\mathbf{S}} \log \epsilon)$ number of copies of the state of the radiation, more precisely,

$$n = \frac{\log(\epsilon)}{2 \log({}_2F_1[-1/2, 1/2; 2; k/e^{\mathbf{S}}])}. \quad (145)$$

After the Page time, there is a different caveat. The fidelity is exponentially close to zero, so the states are essentially fully distinguishable. Precisely, with just one copy of the state, the error probability is bounded above as

$$\min_A [\alpha_n(A) + \beta_n(A)]$$
$$\leq \frac{e^{\mathbf{S}} \, {}_2F_1[-1/2, 1/2; 2; e^{\mathbf{S}}/k]^2}{k} = O(k^{-1}). \quad (146)$$

The issue is that the amount of radiation needed to perform this discrimination is of the order of the size of the black hole. This means that the observer will have to perform a very complex computation, which again is not so feasible in practice.

Now, consider what we would have concluded if we did not include replica wormholes in the gravitational path integral. This is the analog of Hawking's calculation of the state of the radiation that led to information loss. Removing replica wormholes corresponds to only including the identity permutation in the sum. This means that

$$\mathrm{Tr}[\rho_R^\alpha \rho_R'^m] = k^{1-\alpha-m}, \quad (147)$$

leading to all PRREs being identically zero, regardless of how much radiation is collected. This is consistent with the initial paradox where the radiation was thought to be in the same state regardless of the black hole microstate. In fact, this was clear from Eqs. (136) and (137) because, if the states of the black hole are orthogonal, the reduced density matrices on R would be identical.

Finally, we note that the computation of the relative entropy between two states in the PSSY model was recently studied as a way to detect the violation of global symmetries in theories of quantum gravity [94]. The simpler quantity $\mathrm{Tr}[\rho_A \sigma_A]$ was evaluated as a proxy with the full relative entropy computation left as an open question. Equation (140) is the (generalized) solution to this question. While an $O(1)$ answer was anticipated for the relative entropy after the Page time in Ref. [94], we conclude that the relative entropy is indeed infinite. It is only $O(1)$ slightly prior to the Page time and exponentially small but finite at earlier times.

V. TENSOR NETWORKS

Tensor networks represent a generalization of the states we have considered, adding in the ingredient of locality. As such, tensor networks have been particularly useful as toy models of holographic duality [56,95]. They are also independently interesting as presenting new classes of ensembles of random states with novel spectral properties [55]. In this section, we generalize the computations of Sec. III to generic random tensor networks, finding qualitatively new phenomena. A specific application of these

results is for the random tensor networks used for modeling holography. We clarify which random states faithfully represent holographic states and which do not.

We begin with the warm-up example of a random tensor network with two tensors, T_1 and T_2, contracted together as

$$\text{—}\boxed{T_1}\text{—}\boxed{T_2}\text{—}. \tag{148}$$

This is the simplest generalization of the single-tensor network, i.e., Haar random state

$$\text{—}\boxed{T}\text{—}. \tag{149}$$

The two-tensor network has one additional degree of freedom, the dimension of the internal bond, d_b. For T_1 and T_2 independently Gaussian, it is straightforward to generalize the diagrammatic approach. The state is now

$$|\Psi\rangle := \Big| \; \cdots\!\!\rightarrow\!\!\cdots \; \vdots \;, \tag{150}$$

where the dotted line is for d_b and is always contracted. The arrow indicates that the dotted lines must be connected in a way that has all arrows with the same orientation. The reduced density matrix is

$$\rho_A := \Big| \; \cdots\!\!\rightarrow\!\!\cdots \;\; \vdots\text{----}\vdots \;\; \cdots\!\!\rightarrow\!\!\cdots \; \Big|. \tag{151}$$

Note the directions of the arrows. We can see that the normalization associated with each density matrix is $(d_A d_B d_b)^{-1}$. When taking the average of the moments, we now have a double sum over the permutation group, corresponding to the two random tensors. For example, the purity moments will be

$$\overline{\text{Tr}[\rho_A^\alpha]} = \frac{1}{(d_A d_B d_b)^\alpha} \sum_{\tau_1,\tau_2 \in S_\alpha} d_A^{C(\eta^{-1}\circ\tau_1)} d_B^{C(\tau_2)} d_b^{C(\tau_1^{-1}\circ\tau_2)}. \tag{152}$$

To solve this equation at leading order, we need to maximize the exponents. That is, we must find the set of permutations, $\{\tau_1, \tau_2\}$, that maximize $C(\eta^{-1}\circ\tau_1) + C(\tau_2) + C(\tau_1^{-1}\circ\tau_2)$. This is already a significantly harder problem than the single-tensor network where the answer is that τ must be a noncrossing permutation. Interestingly, this maximization may be rephrased as a classical network flow problem [55]. We attach a "source" to B and a "sink" to A and determine the maximal flow, $w_{\text{max-flow}}$, of the network where each edge has a weight corresponding to the logarithm of the Hilbert space dimension. We apply the Ford-Fulkerson method in which, one at a time, we take a path from the source to the sink through the tensors, subtracting the weight of the edges by one as we go along the

path [96]. Each one of these paths is called an *augmenting path*. We repeat this process until there are no more paths from the source to the sink such that we are left with a *residual network*. The rules for each permutation are as follows.

1. All the τ_i are noncrossing.
2. The τ_i are nondecreasing along each augmenting path in the network, i.e., each permutation is contained within all permutations further along the path.
3. All the τ_i in the connected component of the source in the residual network are set to the identity.
4. All the τ_i in the connected component of the sink in the residual network are set to η.
5. All the τ_i in the same connected component are identical.

At leading order, the moments will then be

$$\overline{\text{Tr}[\rho_A^\alpha]} = N^{-(\alpha-1)w_{\text{max-flow}}}$$
$$\times \sum_{\{\tau_1,\tau_2\}} \tilde{d}_A^{C(\eta^{-1}\circ\tau_1)-\alpha} \tilde{d}_B^{C(\tau_2)-\alpha} \tilde{d}_b^{C(\tau_1^{-1}\circ\tau_2)-\alpha}, \tag{153}$$

where $\{\tau_1, \tau_2\}$ is the set of permutations obeying the constraints and the dimensions with tildes are $O(1)$ due to multiples of N, a large parameter, being pulled out. For example, if $d_A = O(N^2)$ then $\tilde{d}_A := d_A N^{-2}$. In the special case that all the $\tilde{d}_i$ are one, we have

$$\overline{\text{Tr}[\rho_A^\alpha]} = F_\alpha N^{-(\alpha-1)w_{\text{max-flow}}}, \tag{154}$$

where F_α represents the number of paths satisfying the constraints.

For example, consider the case where $d_A = d_B = d_b = N$. There will be a single augmenting path ($w_{\text{max-flow}} = 1$)

$$\boxed{\text{source}} \Longrightarrow \boxed{T_2} \Longrightarrow \boxed{T_1} \Longrightarrow \boxed{\text{sink}} \tag{155}$$

such that the resulting network will consist of disconnected tensors with the constraint that $\tau_1 \leq \tau_2 \in \text{NC}_\alpha$:

$$\boxed{\text{source}} \qquad \boxed{T_2} \qquad \boxed{T_1} \qquad \boxed{\text{sink}}. \tag{156}$$

The number of such permutations is given by the second Fuss-Catalan number

$$\text{FC}_\alpha^{(2)} := \frac{1}{2\alpha+1}\binom{3\alpha}{\alpha}, \tag{157}$$

so the moments will be given by

$$\overline{\text{Tr}[\rho_A^\alpha]} = \frac{1}{2\alpha+1}\binom{3\alpha}{\alpha}N^{1-\alpha}. \tag{158}$$

The associated von Neumann entropy is

$$S_{vN}(A) = \log[N] - \frac{5}{6}. \tag{159}$$

This generalizes Page's formula.

Had we instead taken, for example, $\sqrt{d_A} = d_B = d_b = N$, the same augmenting path would have led to the residual network

$$\boxed{\text{source}} \qquad \boxed{T_2} \qquad \boxed{T_1}\!\!-\!\!\!-\!\!\!-\!\!\boxed{\text{sink}}. \tag{160}$$

Because T_1 is still connected to the sink, τ_1 will be set to η while τ_2 can be any noncrossing permutation of which there are a Catalan number's worth, leading to

$$S_{vN}(A) = \log[N] - \frac{1}{2}. \tag{161}$$

More generally, we can have a tensor network with n tensors, $\{T_1, T_2, \ldots, T_n\}$. A set of indices of these tensors will be contracted. We refer to the dimensions of these indices by the tensors they connect, e.g., d_{ij}. There is also a set of uncontracted indices that correspond to systems A and B. We refer to the dimensions of these indices as d_{Ai} and d_{Bi} that label the subsystem they belong to and the tensors that they are indices of. The purity moments can then be expressed as a sum over n permutation elements:

$$\overline{\text{Tr}[\rho_A^\alpha]} = \frac{1}{(\prod_i d_{Ai} d_{Bi} \prod_j d_{ij})^\alpha} \sum_{\tau_1, \ldots, \tau_n \in S_\alpha} \prod_{i=1}^n d_{Ai}^{C(\eta^{-1} \circ \tau_i)} d_{Bi}^{C(\tau_i)}$$
$$\times \prod_{j=1}^n d_{ij}^{C(\tau_i^{-1} \circ \tau_j)}. \tag{162}$$

Here, we must maximize the more complicated exponent that can also be formulated as a network flow problem. Equation (153) is generalized to

$$\overline{\text{Tr}[\rho_A^\alpha]} = N^{-(\alpha-1)w_{\text{max-flow}}} \sum_{\{\tau_i\}} \prod_i^n \tilde{d}_{Ai}^{C(\eta^{-1} \circ \tau_i) - \alpha} \tilde{d}_{Bi}^{C(\tau_i) - \alpha}$$
$$\times \prod_{j=1}^n \tilde{d}_{ij}^{C(\tau_i^{-1} \circ \tau_j) - \alpha}, \tag{163}$$

where $\{\tau_i\}$ is the set of permutations obeying the updated rules. Equation (154) still applies, though the combinatorics may become significantly more difficult. If we are not concerned with the $O(1)$ constant, we only need to determine the maximal flow. By the max-flow min-cut theorem, the maximal flow from the source to the sink will always be equal to the minimal cut, γ_A, in the network needed to separate the source and sink into disconnected

components [96,97]. The von Neumann entropy is then

$$S_{vN}(A) = \gamma_A \log[N] + O(1). \tag{164}$$

We can now generalize this, as before, to relative entropy. We explicitly compute the PRRE. This only changes the permutation allowed in the sum

$$\overline{\text{Tr}[\rho_A^\alpha \sigma_A^m]} = \frac{1}{(\prod_i d_{Ai} d_{Bi} \prod_j d_{ij})^{\alpha+m}} \sum_{\tau_1, \ldots, \tau_n \in S_\alpha \times S_m}$$
$$\times \prod_{i=1}^n d_{Ai}^{C(\eta^{-1} \circ \tau_i)} d_{Bi}^{C(\tau_i)} \prod_{j=1}^n d_{ij}^{C(\tau_i^{-1} \circ \tau_j)}. \tag{165}$$

The key difference between this replica trick and that for Rényi entropies is that η is not an allowed permutation in the sum. This effects all of the $C(\eta^{-1} \circ \tau_i)$ terms because they are maximized not by $\alpha + m$, but $\alpha + m - 1$, which occurs with $\tau_i = \eta_\alpha \times \eta_m \in S_\alpha \times S_m$. This changes rule 1 of the Ford-Fulkerson algorithm to "All the τ_i are in $\text{NC}_\alpha \times \text{NC}_m$" and rule 4 to "All the τ_i in the connected component of the sink in the residual network are set to $\eta_\alpha \times \eta_m$." The moments are then

$$\overline{\text{Tr}[\rho_A^\alpha \sigma_A^m]} = N^{-(\alpha+m-1)\gamma_A} N^{-(E_A - \gamma_A)}$$
$$\times \sum_{\{\tau_i\}} \prod_{i=1}^n \tilde{d}_{Ai}^{C(\eta^{-1} \circ \tau_i) - \alpha - m} \tilde{d}_{Bi}^{C(\tau_i) - \alpha - m}$$
$$\times \prod_{j=1}^n \tilde{d}_{ij}^{C(\tau_i^{-1} \circ \tau_j) - \alpha - m}, \tag{166}$$

where E_A is the weight of the external A edges before applying the Ford-Fulkerson algorithm, which, in the single-tensor case simply equaled the maximum flow. In the special case where all the $\tilde{d}_i$ are one,

$$\overline{\text{Tr}[\rho_A^\alpha \sigma_A^m]} = F_{\alpha,m} N^{-(\alpha+m-1)\gamma_A} N^{\gamma_A - E_A}. \tag{167}$$

First, consider the two-tensor network when all dimensions equal N. There is a single augmenting path and $E_A = \gamma_A$. However, due to the restriction to $S_\alpha \times S_m$, τ_1 and τ_2 are restricted to noncrossing within the subgroup, such that

$$F_{\alpha,m} = \text{FC}_\alpha^{(2)} \text{FC}_m^{(2)}, \tag{168}$$

which is much smaller than $\text{FC}_{\alpha+m}$. The von Neumann relative entropy is then given by

$$D(\rho_A || \sigma_A) = \frac{17}{6}, \tag{169}$$

which should be compared with 3/2 that was found for the single-tensor network. Apparently, adding a random tensor makes the state more distinguishable. Generalizing this

conclusion, if the tensor network is a string of n tensors

$$\underline{\quad}\boxed{T_1}\!-\!\boxed{T_2}\!-\!\boxed{T_3}\cdots\!\cdots\!\boxed{T_n}\underline{\quad}, \qquad (170)$$

the combinatorial factor is given by a product of the nth Fuss-Catalan number

$$
\begin{aligned}
F_{\alpha,m} &= \mathrm{FC}_\alpha^{(n)}\mathrm{FC}_m^{(n)} \\
&= \frac{1}{(n\alpha+1)(nm+1)}\binom{(n+1)\alpha}{\alpha}\binom{(n+1)m}{m}.
\end{aligned}
\qquad (171)
$$

The relative entropy is then

$$D(\rho_A||\sigma_A) = H_{n+1} + n - 1, \qquad (172)$$

where $H_n := \sum_{k=1}^{n} 1/k$ is the harmonic number. This function is monotonically increasing in n.

If we are not concerned with the $O(1)$ contribution, the PRRE will generally be given by

$$D_\alpha(\rho_A||\sigma_A) = \frac{\gamma_A - E_A}{\alpha-1}\log[N] + O(1). \qquad (173)$$

This implies that the quantum relative entropy is always divergent if the A edges do not coincide with the minimal cut. We will come back to this point shortly. Similarly, note that Holevo's just-as-good fidelity is exponentially small in this case,

$$F_H(\rho_A||\sigma_A) = N^{-2(E_A-\gamma_A)}. \qquad (174)$$

This means that, for many tensor networks, two independent states will be easily distinguishable, no matter the relative size of A and B. Note that, when $E_A = \gamma_A$, the $O(1)$ and subleading terms are very interesting and are the main topic of this paper.

Recall that holographic random tensor networks are tensor networks composed of random tensors that are arranged geometrically as discretized hyperbolic space (see Fig. 7) [56]. Because of the negative curvature of this space, the minimal surfaces for boundary regions always lie in the bulk. This means that we always have $E_A > \gamma_A$, so independent states will always be completely distinguishable. This seems to be in tension with the holographic results of Sec. IV. As it seems, single-tensor networks, which have no built in locality, exactly match holographic states while the tensor networks that naively look like anti–de Sitter space do not share any information theoretic properties with holography except for the entropy.

At face value, the above conclusions are a bit unsettling. Fortunately, this can be remedied by more carefully stating how a tensor network should model holographic states. Tensor network models represent the holographic
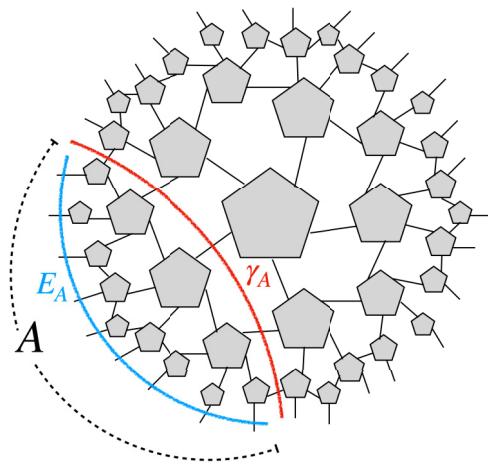
FIG. 7. A discretization of hyperbolic space is shown as a tensor network. For boundary subregion A, the minimal cut through the network, γ_A, always dips into the bulk and is smaller than the boundary cut E_A.

map as a quantum error correcting code where the bulk degrees of freedom play the role of "logical qubits" that are protected by being embedded in the larger boundary Hilbert space. The logical qubits live in a code subspace. In the random tensor networks we have been considering, the code subspace (the ensemble of states we are sampling from) is identical to the Hilbert space on the boundary. This equality between bulk effective field theory and boundary Hilbert space dimensions only occurs in AdS/CFT when one has a large black hole whose horizon approaches the asymptotic boundary of the space. This is the reason for the requirement that $E_A = \gamma_A$; all minimal surfaces in the large black hole geometry hug the asymptotic boundary. In order to model other holographic states using tensor networks, we must make the code subspace significantly smaller than the total Hilbert space. Additionally, the bulk density matrices should not be orthogonal. For example, when considering perturbations about vacuum AdS, the total Hilbert space dimension is $O(e^{-1/G_N})$ while the code subspace is $O(e^{G_N^0})$. In practice, this means that, for the two states ρ and σ, we must take the random tensors to be correlated with each other, i.e., the measure for each random tensor only has support on a proper subset of the Hilbert space.

Another important class of random tensor networks is random unitary circuits. In these tensor networks, all tensors are random unitary operators drawn from the Haar measure. Such networks have been the focus of intense study because they present an exactly solvable minimal model of chaotic many-body dynamics, only preserving locality and unitarity. Using the replica trick and Weingarten calculus, many measures of entanglement and operator growth have been computed using geometric quantities in these circuits. Analogously, the distinguishibility

measures discussed in this paper will be computable. For the dynamics of evolving a state from a product state with a random unitary circuit, twice the time (depth of the circuit) plays the role of γ_A when it is smaller than the length of the region A, l_A, which plays the role of E_A. We will therefore find that states are easily distinguishable for times $t < l_A/2$ and very hard to distinguish afterwards. This describes the process of thermalization where different initial states become indistinguishable at late times. We note that this is the exact time at which the entanglement entropy saturates after a quench [98,99]. The details of this calculation, including the precise approach to equilibrium, are left to future work.

VI. SUBSYSTEM EIGENSTATE THERMALIZATION

The ETH was a major development in understanding the emergence of thermal physics from isolated quantum many-body systems in pure states [100–102]. The statement of eigenstate thermalization is that, given two energy eigenstates, $|E_i\rangle$ and $|E_j\rangle$, and a "simple" few-body operator $\mathcal{O}$, the expectation value varies smoothly with the macroscopic, thermodynamic quantities such as energy

$$\langle E_i|\mathcal{O}|E_j\rangle = f_{\mathcal{O}}(E)\delta_{ij} + e^{-S(E)/2}R_{ij}, \quad E := \frac{E_i + E_j}{2}, \tag{175}$$

where $f_{\mathcal{O}}(E)$ is a smooth function of the energy, $S(E)$ is the thermodynamic entropy, and R_{ij} is an $O(1)$ pseudo-random matrix. The ETH is expected to hold for generic nonintegrable systems and violated in integrable systems. In words, it states that expectation values of simple observables appear thermal, up to exponentially small corrections.

Note that the standard, *local* ETH is a statement only about local or few-body operators. A significant strengthening of the ETH can be made by asserting that the entire reduced density matrix supported on a finite spatial region appears thermal. More precisely, the subsystem eigenstate thermalization hypothesis states that the reduced density matrices of eigenstates, $\rho_A(\psi)$, are exponentially close in trace distance to a universal thermal density matrix, $\rho_{\text{univ}}(E)$, that only depends on the total energy [15,103]

$$|\rho_A(\psi) - \rho_{\text{univ}}(E)|_1 = O(e^{-S(E)/2}). \tag{176}$$

In addition, "off-diagonal" matrices are exponentially suppressed:

$$|\text{Tr}_B[|E_i\rangle\langle E_j|]|_1 = \delta_{ij} + O(e^{-S(E)/2}). \tag{177}$$

These conditions imply the local ETH for *all* operators in region A and are significantly stronger [104].

It is important to understand which systems obey the subsystem ETH. Of course, for all systems, when A is the entire system, the subsystem ETH completely fails because the distance between a pure state and a thermal state is $O(1)$. It is then nontrivial to determine at which point the subsystem ETH breaks down and thermal physics no longer applies. In the following, we show that holographic CFTs obey the subsystem ETH whenever A is smaller than half the total system size. More generally, we find that generic chaotic Hamiltonian systems, whose eigenstate ansatzes were put forward in Refs. [12–14], obey the subsystem ETH.

A. Generic chaotic Hamiltonians

We use the following ansatz for the tensor product decomposition of energy eigenstates of energy E for generic chaotic quantum many-body systems [12]:

$$|E\rangle = \mathcal{N}^{-1/2} \sum_{E-\Delta < E_i+E_J < E+\Delta} c_{iJ}|E_i\rangle_A|E_J\rangle_B. \tag{178}$$

Here $\Delta \ll E$, $\mathcal{N}$ is the normalization, $|E_i\rangle_A$ and $|E_J\rangle_B$ are subsystem energy eigenstates [105], and the coefficients are complex Gaussian random variables

$$\overline{c_{iJ}c^*_{i'J'}} = \delta_{ii'}\delta_{JJ'}, \tag{179}$$

where, with the proper normalization, the variance is set to one. The reduced density matrix is

$$\rho_A = \mathcal{N}^{-1} \sum_{E_i-2\Delta < E_j < E_i+2\Delta}$$
$$\times \sum_{E-E_i-\Delta < E_J < E-E_i+\Delta} c_{iJ}c^*_{jJ}|E_i\rangle\langle E_j|_A. \tag{180}$$

Using this ansatz for the reduced density matrix, we perform the replica trick for the PRRE. In analogy with Refs. [12–14,106] where the Rényi and von Neumann entropies were evaluated for this ansatz, we find, in analogy with Sec. B, that in the thermodynamic limit, after ensemble averaging,

$$\overline{\text{Tr}[\rho_A^\alpha \sigma_A^{1-\alpha}]} = \mathcal{N}^{-1} \int d\mathcal{E} e^{S_A(\mathcal{E})+S_B(E-\mathcal{E})} G_\alpha(\mathcal{E}), \tag{181}$$

where

$$G_\alpha(\mathcal{E}) := \begin{cases} {}_2F_1(1-\alpha, -\alpha; 2; e^{S_A(\mathcal{E})-S_B(E-\mathcal{E})}) \, {}_2F_1(\alpha-1, \alpha; 2; e^{S_A(\mathcal{E})-S_B(E-\mathcal{E})}), & S_A(\mathcal{E}) < S_B(E-\mathcal{E}), \\ e^{S_B(E-\mathcal{E})-S_A(\mathcal{E})} \, {}_2F_1(1-\alpha, -\alpha; 2; e^{S_B(E-\mathcal{E})-S_A(\mathcal{E})}) \, {}_2F_1(\alpha-1, \alpha; 2; e^{S_B(E-\mathcal{E})-S_A(\mathcal{E})}), & S_A(\mathcal{E}) > S_B(E-\mathcal{E}), \end{cases}$$
$$(182)$$

and

$$\mathcal{N} = \int d\mathcal{E}\, e^{S_A(\mathcal{E})+S_B(E-\mathcal{E})}. \qquad (183)$$

We use the following ansatz for the thermodynamic entropies:

$$S_A(\mathcal{E}) = fVs\left(\frac{\mathcal{E}}{Vf}\right),$$

$$S_B(E-\mathcal{E}) = (1-f)Vs\left(\frac{E-\mathcal{E}}{V(1-f)}\right). \qquad (184)$$

Here $s(u)$ is the entropy density, V is the volume of the total system, and f is the fractional volume of subsystem A. The saddle point equation for the main integral is

$$s'\left(\frac{\mathcal{E}_1}{Vf}\right) = s'\left(\frac{E-\mathcal{E}_1}{V(1-f)}\right) - \frac{\partial_\mathcal{E} G_\alpha(\mathcal{E}_1)}{G_\alpha(\mathcal{E}_1)}, \qquad (185)$$

and, for the normalization,

$$s'\left(\frac{\mathcal{E}_2}{Vf}\right) = s'\left(\frac{E-\mathcal{E}_2}{V(1-f)}\right). \qquad (186)$$

We can now evaluate the PRRE in various regimes. The saddle point equation for the normalization is simple to solve because $s'(u)$ is single valued; thus,

$$\mathcal{E}_2 = fE. \qquad (187)$$

This is not so surprising as it implies a constant energy density. The normalization is then evaluated to

$$\mathcal{N} = \sqrt{\frac{2\pi Vf(1-f)}{s''(E/V)}} e^{Vs(E/V)}. \qquad (188)$$

We now specify to $f < 1/2$ where we claim that $\mathcal{E}_1 < E/2$. In this regime, we can expand the hypergeometric function

$$G_\alpha(\mathcal{E}) \simeq 1 + \alpha(\alpha-1)e^{S_A(\mathcal{E})-S_B(E-\mathcal{E})}. \qquad (189)$$

In this approximation,

$$\frac{\partial_\mathcal{E} G_\alpha(\mathcal{E}_1)}{G_\alpha(\mathcal{E}_1)} = \frac{(\alpha-1)\alpha e^{fVs(\mathcal{E}_1/fV)}\{s'[(E-\mathcal{E}_1)/(1-f)V] + s'(\mathcal{E}_1/fV)\}}{e^{(1-f)Vs[(E-\mathcal{E}_1)/(1-f)V]} + (\alpha-1)\alpha e^{fVs(\mathcal{E}_1/fV)}}. \qquad (190)$$

This term is exponentially small for $f < 1/2$. Therefore, the saddle point equation for $\mathcal{E}_1$ can be treated as an expansion around $\mathcal{E}_2$,

$$\mathcal{E}_1 = \mathcal{E}_2(1+\delta), \quad \delta \ll 1. \qquad (191)$$

To leading order, the saddle point equation is

$$\frac{Es''(E/V)}{V}\delta = -\frac{Es''(E/V)f}{V(1-f)}\delta - \frac{2(\alpha-1)\alpha e^{(2f-1)Vs(E/V)}s'(E/V)}{(\alpha-1)\alpha e^{(2f-1)Vs(E/V)} + 1}$$

$$\simeq -\frac{Es''(E/V)f}{V(1-f)}\delta - 2(\alpha-1)\alpha e^{(2f-1)Vs(E/V)}s'\left(\frac{E}{V}\right). \qquad (192)$$

Solving for δ, we find self-consistency with the claim that $\mathcal{E}_1$ is very close to $\mathcal{E}_2$:

$$\delta \simeq \frac{2(\alpha-1)\alpha(f-1)Ve^{(2f-1)Vs(E/V)}s'(E/V)}{Es''(E/V)}. \qquad (193)$$

The saddle point solution is then

$$\int d\mathcal{E} e^{S_A(\mathcal{E})+S_B(E-\mathcal{E})} G_\alpha(\mathcal{E})$$

$$\simeq \sqrt{\frac{2\pi V f (1-f)}{s''(E/V)}} e^{V s(E/V)+\alpha(\alpha-1)e^{(2f-1)V s(E/V)}+O(\delta^2)}. \tag{194}$$

Therefore, the moments are

$$\overline{\text{Tr}[\rho_A^\alpha \sigma_A^{1-\alpha}]} \simeq e^{\alpha(\alpha-1)e^{(2f-1)V s(E/V)}}, \tag{195}$$

so the PRRE is exponentially suppressed in the entropy for all values of α,

$$\overline{D_\alpha(\rho_A||\sigma_A)} \simeq \alpha e^{(2f-1)V s(E/V)}. \tag{196}$$

This places the following strict bounds on the trace distance:

$$\frac{e^{(2f-1)V s(E/V)}}{4} \leq \overline{T(\rho_A||\sigma_A)} \leq \frac{e^{(f-1/2)V s(E/V)}}{\sqrt{2}}. \tag{197}$$

This provides further evidence for the refined subsystem eigenstate thermalization hypothesis in Ref. [15] that postulated the scaling of the trace distance on the size of subregion A because the upper bound precisely matches that scaling when replacing N_A, the number of qubits in region A, with the subsystem thermodynamic entropy.

Note also that these results generalize our result for random matrix theory because at infinite temperature, we can identify

$$fV s\left(\frac{E}{V}\right) = \log d_A, \quad (1-f) V s\left(\frac{E}{V}\right) = \log d_B. \tag{198}$$

Next, consider $f > 1/2$. Unfortunately, the maximum of the integral occurs right near the transition $S_A(\mathcal{E}_1) = S_B(E - \mathcal{E}_1)$. Because of this, we cannot simply make a saddle point approximation. However, it is straightforward to argue that the PRRE will be large for $\alpha < 1$ and infinite for $\alpha \geq 1$. Note that the integrands of the numerator and denominator of $\text{Tr}[\rho_A^\alpha \sigma_A^{1-\alpha}]$ are exponentially close for $S_A(\mathcal{E}_1) < S_B(E - \mathcal{E}_1)$, while the numerator is exponentially suppressed in relation to the denominator for $S_A(\mathcal{E}_1) > S_B(E - \mathcal{E}_1)$. Because the saddle point for the denominator occurs when $S_A(\mathcal{E}_1) > S_B(E - \mathcal{E}_1)$, $\text{Tr}[\rho_A^\alpha \sigma_A^{1-\alpha}]$ will be exponentially small, i.e., $\log(\text{Tr}[\rho_A^\alpha \sigma_A^{1-\alpha}])$ will be negative and of the order of the entropy. Because of the factor of $(\alpha - 1)^{-1}$ in the PRRE, this implies that the PRRE for $\alpha < 1$ is of the order of the entropy and ill defined (infinite) for $\alpha \geq 1$, in analogy with the random matrix theory result, Eq. (74).

We have found that whenever A is less than half the total system size, the PRRE and therefore the subsystem trace distance between any two eigenstates of the same energy is exponentially suppressed in the entropy. The trace distance is a metric on the space of density matrices, so these eigenstates lie within a ball with radius $O(e^{-S(E)})$. The universal density matrix then must also lie within this ball such that Eq. (176) is satisfied.

For Eq. (177), we need to perform an additional computation. The off-diagonal $(i \neq j)$ matrix for two random states is represented as

$$\text{Tr}_B\left[|\psi_i\rangle \langle\psi_j|\right] := |\;\overset{\ulcorner----\urcorner}{}\;|. \tag{199}$$

To compute the trace norm, we need the integer powers of

$$\text{Tr}_B\left[|\psi_i\rangle \langle\psi_j|\right] \text{Tr}_B\left[|\psi_j\rangle \langle\psi_i|\right] = |\;\overset{\ulcorner----\urcorner}{\underset{\llcorner\_\lrcorner}{}}\;\overset{\ulcorner----\urcorner}{}\;| \tag{200}$$

The moments are given by a new sum over permutations

$$\overline{\text{Tr}\{(\text{Tr}_B[|\psi_i\rangle\langle\psi_j|]\text{Tr}_B[|\psi_j\rangle\langle\psi_i|])^\alpha\}}$$

$$= \frac{1}{(d_A d_B)^{2\alpha}} \sum_{\tau \in S_{\text{odd}}} d_A^{C(\eta^{-1}\circ\tau)} d_B^{C(\tau)}. \tag{201}$$

Here, S_{odd} represents the set of permutations where, within each cycle, the difference between consecutive numbers is always odd. For example, the cycle $(1, 2, 5, 6)$ is allowed, but $(2, 4, 5)$ is not. Crucially, the identity permutation is not an allowed permutation. If we want to maximize the number of dashed loops for small d_A/d_B, τ must be composed of α noncrossing two cycles. The degeneracy is given by the Catalan number so that

$$\overline{\text{Tr}\{(\text{Tr}_B[|\psi_i\rangle\langle\psi_j|]\text{Tr}_B[|\psi_j\rangle\langle\psi_i|])^\alpha\}}$$

$$= C_\alpha d_A^{1-\alpha} d_B^{-\alpha} + O(d_B^{-\alpha-1}). \tag{202}$$

The trace norm is the $\alpha \to 1/2$ limit, such that

$$\overline{|\text{Tr}_B[|\psi_i\rangle\langle\psi_j|]|_1} = \frac{8}{3\pi}\sqrt{\frac{d_A}{d_B}} + O\left(\frac{d_A}{d_B}\right)^{3/2}. \tag{203}$$

When d_A/d_B is large, the cyclic permutation is an allowed permutation and will dominate, leading to

$$\overline{\text{Tr}\{(\text{Tr}_B[|\psi_i\rangle\langle\psi_j|]\text{Tr}_B[|\psi_j\rangle\langle\psi_i|])^\alpha\}} = d_B^{1-2\alpha} + O(d_A^{-1}). \tag{204}$$

Taking the $\alpha \to 1/2$ limit tells us that $|\text{Tr}_B[|\psi_i\rangle\langle\psi_j|]|_1$ is exponentially close to one.

For finite energy eigenstates, this translates to

$$\overline{\mathrm{Tr}\{(\mathrm{Tr}_B[|\psi_i\rangle\langle\psi_j|]\mathrm{Tr}_B[|\psi_j\rangle\langle\psi_i|])^\alpha\}}$$
$$= \mathcal{N}^{-2\alpha} \int d\mathcal{E} e^{S_A(\mathcal{E})+S_B(E-\mathcal{E})} G_\alpha(\mathcal{E}), \qquad (205)$$

where $G_\alpha(\mathcal{E})$ is now given by

$$G_\alpha(\mathcal{E}) = \begin{cases} C_\alpha e^{\alpha S_A(\mathcal{E})+(\alpha-1)S_B(E-\mathcal{E})}, & e^{S_A(\mathcal{E})-S_B(E-\mathcal{E})} \ll 1, \\ e^{(2\alpha-1)S_A(\mathcal{E})}, & e^{S_A(\mathcal{E})-S_B(E-\mathcal{E})} \gg 1. \end{cases}$$
$$(206)$$

At $\alpha = 1/2$, for sufficiently small f, the saddle point will fall in the $e^{S_A(\mathcal{E})-S_B(E-\mathcal{E})} \ll 1$ regime, leading to a saddle point equation for the numerator of

$$3s'\left(\frac{\mathcal{E}_1}{Vf}\right) = s'\left(\frac{E-\mathcal{E}_1}{V(1-f)}\right). \qquad (207)$$

Because of the factor of 3, $\mathcal{E}_1$ will be larger than $\mathcal{E}_2$. We then find that

$$\overline{|\mathrm{Tr}_B[|\psi_i\rangle\langle\psi_j|]|_1} \simeq \frac{8e^{(3/2)S_A(\mathcal{E}_1)+(1/2)S_B(E-\mathcal{E}_1)}}{3\pi e^{S_A(\mathcal{E}_2)+S_B(E-\mathcal{E}_2)}}. \qquad (208)$$

This is exponentially small because $e^{S_A(\mathcal{E}_1)+S_B(E-\mathcal{E}_1)} \ll e^{S_A(\mathcal{E}_2)+S_B(E-\mathcal{E}_2)}$ and $S_A(\mathcal{E}_1) \ll S_B(E-\mathcal{E}_1)$. For sufficiently large f, the saddle point will fall in the other regime such that the numerator and denominator are identical at $\alpha = 1/2$, leading to $\overline{|\mathrm{Tr}_B[|\psi_i\rangle\langle\psi_j|]|_1} \simeq 1$ at leading order.

B. Holographic states

We could now simply posit that because holographic systems are believed to be chaotic, their eigenstates will also have a spatial decomposition according to Eq. (178) and, thus, will obey the subsystem ETH for $f < 1/2$. However, this line of reasoning is somewhat unsatisfying because it is not constructive. Instead, we implement a gravitational calculation, using the fixed-area state analysis from Sec. IV, to evaluate the PRRE in normal states without any areas fixed. Our strategy follows Ref. [106] in manipulating the gravitational path integral into a form identical to Eq. (181), deriving the validity of using

Eq. (178) for holographic eigenstates when computing the PRRE. Because of the similarities with Ref. [106], we keep the derivation brief, referring the interested reader to the original literature.

To begin, we make the assumption that black hole microstates can be represented as a random superposition of energy eigenstates in a microcanonical energy window $I_E = [E - \Delta, E + \Delta]$, i.e.,

$$|E, \hat{c}\rangle \propto \sum_{E_i \in I_E} c^i e^{-\beta E_i/2} |E_i\rangle, \qquad (209)$$

where β is an effective temperature. These states are believed to be holographically dual to black hole geometries with end-of-world branes specifying the microstate lying behind the horizon [99,107–110], similar to the PSSY model.

The corresponding density matrix is represented as the path integral on a strip of width β with boundary conditions determined by $\hat{c}$. We consider two microstates in the same energy window, corresponding to two independent sets of Gaussian random variables $\hat{c}$ and $\hat{d}$. As argued in Ref. [106], after disorder averaging, the random variables match up the boundary conditions of the strips according to the same Wick contractions previously discussed for Haar random states. Therefore, the path integral is given by a sum over all allowed Wick contractions

$$\mathrm{Tr}[\rho_A^\alpha \sigma_A^m] = \frac{\mathcal{Z}_{\alpha,m}}{\mathcal{Z}_{1,0}^\alpha \mathcal{Z}_{0,1}^m}, \quad \mathcal{Z}_{\alpha,m} \simeq \sum_{\mathcal{M}_i} \int \mathcal{D}\phi e^{-I_E(\mathcal{M}_i,\phi)},$$
$$(210)$$

where the $\mathcal{M}_i$ are the replica manifolds and ϕ are all quantum fields. Using the holographic dictionary, this is a sum over bulk geometries with asymptotic boundary conditions $\mathcal{M}_i$. In these bulk geometries, the EOW branes have disappeared due to the disorder averaging. This is an alternative way to see the reduction of bulk saddles from $S_{\alpha+m}$ to $S_\alpha \times S_m$. In general, solving the bulk equations of motion to evaluate the path integrals on shell is very difficult. We use the "double-defect" construction of Ref. [106] to separate the action into a bulk contribution $I_{\mathrm{bulk}}(g, \phi, E)$ and actions, $I_{\mathrm{brane}}(\Sigma_1)$ and $I_{\mathrm{brane}}(\Sigma_2)$, for cosmic branes, Σ_1 and Σ_2, that are located at the two extremal surfaces

$$\mathcal{Z}_{\alpha,m} = \sum_{\mathcal{M}_i} \int \mathcal{D}\phi \mathcal{D}g \mathcal{D}\Sigma_1 \mathcal{D}\Sigma_2 e^{-(\alpha+m)I_{\mathrm{bulk}}(g,\phi,E)-[(\alpha+m-k_i)/4G_N]I_{\mathrm{brane}}(\Sigma_1)-[(k_i-1)/4G_N]I_{\mathrm{brane}}(\Sigma_2)}, \qquad (211)$$

where k_i plays the role of the number of cycles in the Wick contraction corresponding to $\mathcal{M}_i$. The sum over $\mathcal{M}_i$ may be done prior to the path integral and we also take $m \to 1 - \alpha$ to arrive at

$$\mathcal{Z}_{\alpha,m} = \int \mathcal{D}\phi \mathcal{D}g \mathcal{D}\Sigma_1 \mathcal{D}\Sigma_2 e^{-I_{\mathrm{bulk}}(g,\phi,E)} G_\alpha(\Sigma_1, \Sigma_2), \qquad (212)$$

where

$$G_\alpha(\Sigma_1, \Sigma_2) = \begin{cases} {}_2F_1(1-\alpha, -\alpha; 2; e^{\Delta I_{\text{brane}}/4G_N}) \, {}_2F_1(\alpha - 1, \alpha; 2; e^{\Delta I_{\text{brane}}/4G_N}), & \Delta I_{\text{brane}} < 0, \\ e^{-\Delta I_{\text{brane}}/4G_N} \, {}_2F_1(1-\alpha, -\alpha; 2; e^{-\Delta I_{\text{brane}}/4G_N}) \, {}_2F_1(\alpha - 1, \alpha; 2; e^{-\Delta I_{\text{brane}}/4G_N}), & \Delta I_{\text{brane}} > 0, \end{cases} \tag{213}$$

and $\Delta I_{\text{brane}} := I_{\text{brane}}(\Sigma_1) - I_{\text{brane}}(\Sigma_2)$. Next, we make use of the fixed-area basis by postponing the integrals over the areas of the branes until the very end, such that the integral is rewritten as

$$\mathcal{Z}_{\alpha, 1-\alpha} = \int dA_1 dA_2 P(A_1, A_2) G_\alpha(A_1, A_2), \tag{214}$$

where $G_\alpha(A_1, A_2)$ is identical to $G_\alpha(\Sigma_1, \Sigma_2)$ with $I_{\text{brane}}(\Sigma_1) \leftrightarrow A_1$ and $I_{\text{brane}}(\Sigma_2) \leftrightarrow A_2$, and $P(A_1, A_2)$, the (unnormalized) probability of being in the state with areas A_1 and A_2, is given by

$$P(A_1, A_2) := \int \mathcal{D}g \mathcal{D}\phi \Big|_{A_{\Sigma_1}=A_1, A_{\Sigma_2}=A_2} e^{-I_{\text{bulk}}(g, \phi, E)}. \tag{215}$$

For high-energy eigenstates, $P(A_1, A_2)$ will localize to a trajectory where A_2 is a function of A_1 [106] such that

$$P(A_1, A_2) \simeq \delta_{A_2, A_2(A_1)} \int \mathcal{D}g \mathcal{D}\phi \Big|_{A_{\Sigma_1}=A_1} e^{-I_{\text{bulk}}(g, \phi, E)}. \tag{216}$$

Finally, to compare with Eq. (181), we want to change the integration variable from A_1 to the energy density in region A, $\mathcal{E}$. Using the Bekenstein-Hawking formula [2], we write the areas in terms of entropy densities

$$\frac{A_1}{4G_N} = fVs\left(\frac{\mathcal{E}(A_1, A_2)}{fV}\right) + A_\infty,$$

$$\frac{A_2}{4G_N} = (1-f)Vs\left(E - \frac{\mathcal{E}(A_1, A_2)}{(1-f)V}\right) + A_\infty, \tag{217}$$

where A_∞ is the divergent piece of the Ryu-Takayanagi surface that approximately cancels in all expressions because we are in the high-energy limit where the surfaces are approximately purely radial until they reach the horizon and subsequently tightly wrap the horizon. Here $\mathcal{E}(A_1, A_2)$ is the ADM energy that is a function of the horizon area $A_{\text{BH}} \simeq A_1 + A_2 - 2A_\infty$. In a saddle point approximation, the probability then becomes [106]

$$P(\mathcal{E}) \simeq e^{S_A(\mathcal{E}) + S_B(E-\mathcal{E})}. \tag{218}$$

In total, we find that

$$\frac{\mathcal{Z}_{\alpha, 1-\alpha}}{\mathcal{Z}_{1,0}^\alpha \mathcal{Z}_{0,1}^{1-\alpha}} \simeq \frac{\int d\mathcal{E} \, e^{S_A(\mathcal{E}) + S_B(E-\mathcal{E})} G_\alpha(\mathcal{E})}{\int d\mathcal{E} \, e^{S_A(\mathcal{E}) + S_B(E-\mathcal{E})}}. \tag{219}$$

It should now be evident that this formula is identical to Eq. (181), so we conclude that the PRRE for holographic

theories is exponentially small in the entropy when $f < 1/2$ and subsystem eigenstate thermalization will hold.

VII. DISCUSSION

We take the opportunity to now comment on future directions that are out of the scope of this work but deserve attention.

On the formal end, one may be worried that the basis of our calculations have assumed finite Hilbert space dimensions and tensor factorization of the Hilbert space into $\mathcal{H}_A \otimes \mathcal{H}_B$. This is not true in quantum field theory and is the reason why reduced density matrices and von Neumann entropies are not well defined. However, the relative entropy and related quantities are well-defined quantities in the continuum using modular theory (see, for example, Refs. [111,112]). For this reason, we expect that our calculations that assumed tensor factorization are accurate even though we ignored this subtlety. However, this expectation is not guaranteed. In particular, the rank deficiency in the reduced density matrices that led to infinite relative entropy has no analog in the continuum because subregions are described by type-III von Neumann algebras that are roughly "full rank."

More practically, we were only able to compute the relative entropies of the simple PSSY toy model of black hole evaporation. We expect that this model captures the essential features of the distinguishability of evaporating black holes, like the wormhole contributions that restore unitarity. It is an important future direction to complete analogous calculations in more realistic models of black hole evaporation. Though a complete calculation, including all saddles like in Jackiw-Teitelboim gravity, is most likely out of reach, one may be able to identify the saddles that lead to $O(e^{-1/G_N})$ fidelity prior to the Page time and $O(1)$ fidelity after the Page time.

For nonevaporating black holes, our calculation was limited to large black holes away from phase transitions. It is clearly of interest to generalize these holographic calculations to smaller, normal black holes. Moreover, there may be interesting crossover behavior in relative entropies near the phase transitions. In the Page states, the transitions were $O(1)$ [e.g., for relative entropy, the $D(\rho_A || \sigma_A) = 3/2$ at the transition]. However, for finite energy states, there may be enhanced corrections analogous to the $O(\sqrt{V})$ corrections in the von Neumann entropy [13,106,113,114]. These should be visible in chaotic, nonholographic systems as well. Preliminary numerical results were given in

Ref. [3], though a more systematic study of both integrable and chaotic Hamiltonians is warranted.

Finally, it would be interesting to study random tensor networks that act as quantum error correcting codes from "bulk" Hilbert spaces to "boundary" Hilbert spaces. These quantum error correcting codes are more closely related to the semiclassical holographic states than the tensor network states studied in Sec. V. Using these tensor networks, various interesting corrections to the JLMS formula [84] may be explored. We expect this to have important implications for approximate quantum error correction and entanglement wedge reconstruction in AdS/CFT. We furthermore hope that this line of study may impact the development of approximate quantum error correcting codes that may be useful for quantum computation.

APPENDIX A: DISTINGUISHIBILITY FROM FREE PROBABILITY THEORY

We have tried to avoid the use of free probability theory throughout this paper because while it is immensely powerful, it is very technical and nonintuitive for the wide audience that this paper is intended for. Moreover, it lacks the clear connections to gravitational and general chaotic systems that the replica trick has. Alas, for exact calculations of the trace distance and Uhlmann fidelity, we were unable to push our diagrammatic techniques far enough, so we address distinguishability measures using free probability in this appendix.

We begin a minimal introduction of the tools we use from applications of free probability theory as applied to random matrices. The interested reader is encouraged to consult Refs. [115,116]. First, we define a unital linear form, τ, acting on a vector space. For $N \times N$ matrices, X, this can be taken to be the normalized trace $\tau(X) := (1/N)\text{Tr}[X]$. Random variables X and Y are *free* if they satisfy

$$\tau[P_1(X)Q_1(Y)\cdots P_k(X)Q_k(Y)] = 0 \quad (A1)$$

for any set of polynomials $\{P, Q\}$ when $\tau[P_i(X)] = \tau[Q_i(Y)] = 0$. This, in turn, implies a factorization property

$$\tau[P(X)Q(Y)] = \tau[P(X)]\tau[Q(Y)], \quad (A2)$$

which can be seen by taking $P_1(X) = P(X) - \tau[P(X)]$ and $Q_1(X) = Q(X) - \tau[Q(X)]$. There are several useful ways to package the moments of a free random variable, $m_n := \tau(X^n)$. These enable the evaluation of empirical spectral measures for random matrices. For a free random variable X, we define the moment function as a formal power series

$$M_X(z) := \sum_{m=0}^{\infty} \tau(X^n)z^m, \quad z \in \mathbb{C}. \quad (A3)$$

This is related to Green's function or the *Cauchy transform* as

$$G_X(z) := \frac{1 + M_X(z)}{z}. \quad (A4)$$

From the Cauchy transform, one may extract the spectral measure using a Stieltjes transformation,

$$\mu_X(\lambda) = -\frac{1}{\pi} \lim_{\epsilon \to 0} \text{Im}[G_X(\lambda + i\epsilon)]. \quad (A5)$$

Sometimes, we know that random variables X and Y have spectral measures ρ_X and ρ_Y, but want to know the spectral measure of their sum, $X + Y$, or product, XY. The spectral measure of their sum is defined as the free convolution $\mu_{X+Y} := \mu_X \boxplus \mu_Y$, while the spectral measure of their product is defined as the free multiplicative product $\mu_{XY} := \mu_X \boxtimes \mu_Y$. In order to obtain the free convolution, it is convenient to introduce the *R transform*

$$R_X[G_X(z)] + \frac{1}{G_X(z)} := z. \quad (A6)$$

The R transform of a sum of free random variables is given by the sum of their individual R transforms,

$$R_{X+Y}(z) = R_X(z) + R_Y(z). \quad (A7)$$

For the free multiplicative product, it is convenient to introduce the *S transform*

$$zM_X(z)S_X[M_X(z)] := 1 + M_X(z). \quad (A8)$$

The S transform of a product of free random variables is given by the product of their individual S transforms,

$$S_{XY}(z) = S_X(z)S_Y(z). \quad (A9)$$

The key to the usefulness of free probability for us is that Wishart random matrices are free random variables asymptotically, as the dimensions become large. Their empirical spectral measure is given by the Marchenko-Pastur

distribution

$$\mu_{\text{MP}}^{(c)}(x) = \max\left[1 - \frac{1}{c}, 0\right]\delta(x)$$
$$+ \frac{\sqrt{[x - c(1 - c^{-1/2})^2][c(1 + c^{-1/2})^2 - x]}}{2\pi c x},$$
(A10)

where $c := d_A/d_B$ is the rectangular parameter and $x := d_A\lambda$.

1. Petz Rényi relative entropy

For the PRRE, we need to evaluate $\overline{\text{Tr}[\rho_A^\alpha \sigma_A^{1-\alpha}]}$. Because ρ_A and σ_A are asymptotically free random variables with Marchenko-Pastur distributions, $\mu_{\text{MP}}^{(c)}$, this factorizes according to Eq. (A2) as

$$\overline{\text{Tr}[\rho_A^\alpha \sigma_A^{1-\alpha}]} = \frac{\overline{\text{Tr}[\rho_A^\alpha]}\,\overline{\text{Tr}[\sigma_A^{1-\alpha}]}}{d_A}$$
$$= \int dx\,\mu_{\text{MP}}^{(c)}(x)x^\alpha \int dx\,\mu_{\text{MP}}^{(c)}(x)x^{1-\alpha}, \quad \text{(A11)}$$

These integrals may be evaluated to reproduce Eq. (73), the result from the replica trick.

2. Sandwiched Rényi relative entropy

For the SRRE, we need to evaluate the averaged moments of $\sigma_A^{(1-\alpha)/2\alpha}\rho_A\sigma_A^{(1-\alpha)/2\alpha}$. Using the replica trick, we succeeded for integer α, but were unable to analytically continue to real valued α in order to evaluate the Uhlmann fidelity at $\alpha = 1/2$. Here, we evaluate the spectrum of $\sigma_A^{1/2}\rho_A\sigma_A^{1/2}$ (equivalently $\rho_A\sigma_A$) to accomplish this goal. This is the free multiplicative convolution of Marchenko-Pastur laws $\mu_{\text{MP}}^{(c)} \boxtimes \mu_{\text{MP}}^{(c)}$ and has been called a generalized Fuss-Catalan distribution [117]. The Fuss-Catalan distributions themselves were first derived in Ref. [118]. The S transform for the Marchenko-Pastur distribution is given by

$$S_{\rho_A}(z) = \frac{1}{1 + cM_{\rho_A}(z)}.$$
(A12)

Therefore, the S transform for $\rho_A\sigma_A$ is

$$S_{\rho_A\sigma_A}(z) = \frac{1}{[1 + cM_{\rho_A}(z)]^2}.$$
(A13)

Substituting this into Eq. (A8), we find that

$$zM_{\rho_A\sigma_A}(z) = [1 + M_{\rho_A\sigma_A}(z)][1 + cM_{\rho_A\sigma_A}(z)]^2. \quad \text{(A14)}$$

Taking the correct root of this cubic equation and taking the Stieltjes transformation, we find the spectral measure

$$\mu_{\text{MP}}^{(c)} \boxtimes \mu_{\text{MP}}^{(c)}(x) = \max\left[0, \frac{c-1}{c}\right]\delta(x) - \frac{1}{\pi}\begin{cases} \dfrac{1}{\sqrt{3} \times 2^{2/3}x}\dfrac{3x + (c-1)^2}{(\sqrt{P} + A)^{1/3}} - \dfrac{\sqrt{3} \times 2^{2/3}}{12c^2x}(\sqrt{P} + A)^{1/3}, & (x_1, x_2), \\ 0, & \text{otherwise,} \end{cases}$$
(A15)

where we have defined

$$A = c^3[9(c+2)x + 2(c-1)^3],$$
$$P = 27c^6 x\{[8 - (c-20)c]x + 4(c-1)^3 - 4x^2\},$$
(A16)

and endpoints

$$x_1 = \max\left[0, \frac{8 + 20c - c^2 - \sqrt{c}(8+c)^{3/2}}{8}\right],$$
$$x_2 = \frac{8 + 20c - c^2 + \sqrt{c}(8+c)^{3/2}}{8}.$$
(A17)

The Uhlmann fidelity is then given by the integral

$$F(\rho_A||\sigma_A) = \left(\int_{x_1}^{x_2} dx\,\mu_{\text{MP}}^{(c)} \boxtimes \mu_{\text{MP}}^{(c)}(x)\sqrt{x}\right)^2.$$
(A18)

We compare the derived distribution with numerics in Fig. 8.

3. Trace distance

The trace distance is defined using the trace norm of the difference of ρ_A and σ_A. This is tailor-made for a computation in free probability because of free convolution. The R transform for the Marchenko-Pastur distribution is

$$\mathcal{R}_{\rho_A}(z) = \frac{1}{1 - cz}.$$
(A19)
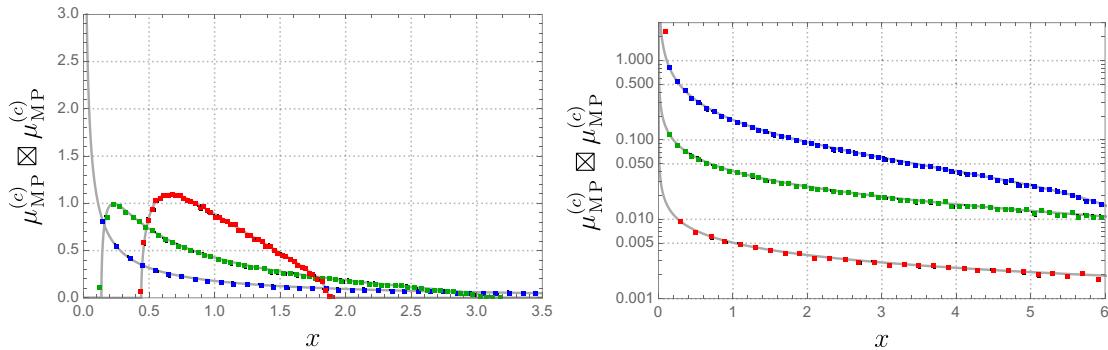
FIG. 8.   The probability density function for the generalized Fuss-Catalan distribution is shown (gray lines) with comparison to numerics. On the left, the blue, green, and red dots correspond to $c = 2^0, 2^{-2}, 2^{-4}$, respectively. On the right, the blue, green, and red dots correspond to $c = 2^0, 2^2, 2^4$, respectively. The total system size is 2^{16} and we disorder average over 10^3 realizations.

Because we are taking the difference, we must rescale the second R transform as

$$\mathcal{R}_{-\sigma_A}(z) = -\mathcal{R}_{\sigma_A}(-z) = -\frac{1}{1 + cz}. \qquad (A20)$$

Therefore, the R transform of the difference is

$$\mathcal{R}_{\rho_A - \sigma_A}(z) = \frac{2cz}{1 - c^2 z^2}. \qquad (A21)$$

$$\mu_{\mathrm{MP}}^{(c)} \boxplus \mathcal{D}_{-1}[\mu_{\mathrm{MP}}^{(c)}](x) = \begin{cases} \dfrac{\sqrt{(2-c)^2 + 3x^2}}{\sqrt{3}\pi c|x|} \sinh\left[\dfrac{\log[\eta(x) + \sqrt{\eta^2(x) - 1}]}{3}\right], & |x| \in (x_-, x_+), \\ \max\left[0, 1 - \dfrac{2}{c}\right]\delta(x), & \text{otherwise}, \end{cases} \qquad (A23)$$

where $\mathcal{D}_{-1}$ represents the rescaling, the function $\eta(x)$ is given by

$$\eta(x) := \frac{9(c+1)^2 x^2 + (2-c)^2}{[(2-c)^2 + 3x^2]^{3/2}}, \qquad (A24)$$

and the endpoints of the spectrum are

$$x_\pm = \max\left[0, \frac{1}{4}(\sqrt{4c+1} \pm 3)^{3/2}(\sqrt{4c+1} \mp 1)^{1/2}\right]. \qquad (A25)$$

One can then evaluate the trace distance from the integral

$$T(\rho_A || \sigma_A) = \frac{1}{2}\int dx\, \mu_{\mathrm{MP}}^{(c)} \boxplus \mathcal{D}_{-1}[\mu_{\mathrm{MP}}^{(c)}](x)|x|, \qquad (A26)$$

leading to Eq. (115).

Inverting the R transform, one finds the Cauchy transform

$$\frac{2cG(z)}{1 - c^2 G(z)^2} + \frac{1}{G(z)} = z. \qquad (A22)$$

Once again, one can take the correct root of the cubic equation and take the Stieltjes transformation to find the spectral measure [66,73]

APPENDIX B: COMMUTATION OF ENSEMBLE AVERAGE AND LOGARITHM

We have been using the replica trick throughout the paper to compute relative entropies. This has involved evaluating ensemble averages of traces of powers of density matrices and then taking a logarithm. In general, the ensemble average and logarithm do not commute. In this appendix, we show that the two operations approximately commute in the large-N limit. To properly take the average of a logarithm, we need the additional replica trick

$$\overline{\log(\mathrm{Tr}[\rho_A^\alpha \sigma_A^m])} = \lim_{q \to 0} \frac{\overline{(\mathrm{Tr}[\rho_A^\alpha \sigma_A^m])^q} - 1}{q}. \qquad (B1)$$

For illustration, we work with the PRRE though the argument is the same for all other quantities. In diagrams, the necessary moments are

$$\overline{\left(\mathrm{Tr}\left[\rho_A^\alpha \sigma_A^m\right]\right)^q} = \underbrace{} \quad \cdots \quad \underbrace{}, \tag{B2}$$

where there are q total blocks. As a sum over permutations, this is

$$\overline{(\mathrm{Tr}[\rho_A^\alpha \sigma_A^m])^q}$$

$$= \frac{1}{(d_A d_B)^{q(\alpha+m)}} \sum_{\tau \in (S_\alpha \times S_m)^{\times q}} d_A^{C((\eta^{-1})^{\times q} \circ \tau)} d_B^{C(\tau)}, \tag{B3}$$

where in cycle notation

$$(\eta^{-1})^{\times q} = \prod_{i=0}^{q-1}(1+q, 2+q, \ldots, \alpha+m+q). \tag{B4}$$

The leading terms come from noncrossing contractions within each block independently, leading to $C[(\eta^{-1})^{\times q} \circ \tau] + C(\tau) = q(\alpha+m+1)$. Contractions that connect the blocks will be subleading with $C[(\eta^{-1})^{\times q} \circ \tau] + C(\tau) \leq q(\alpha+m+1) - 2$. Therefore, the ensemble average factorizes at leading order as

$$\overline{(\mathrm{Tr}[\rho_A^\alpha \sigma_A^m])^q} \simeq \overline{(\mathrm{Tr}[\rho_A^\alpha \sigma_A^m])}^q. \tag{B5}$$

Using Eq. (B1), this implies that the ensemble average and logarithm approximately commute at large N.

APPENDIX C: EQUIVALENCE WITH HAAR UNITARY TENSOR NETWORKS

Frequently, random tensor networks are constructed by projected Haar unitary states. This is in fact equivalent to the Gaussian random networks we use. The reason is the following. In the Haar random construction, every vertex of degree k is a state of k qudits projected to a random state $U|0\rangle$, where U is a Haar random unitary and $|0\rangle$ is any state. This gives the state $\langle 0|^{\otimes k}U^\dagger|i_1\rangle \cdots |i_k\rangle$. Denoting the set of k indices by one index i, this exactly corresponds to the Gaussian tensor network with the identification $U_{i,0} \leftrightarrow X_i^*$. Every edge corresponds to a maximally entangled pair in the projected Haar random network, which is just the index contraction in the Gaussian network. The projected unitaries indeed have a Gaussian distribution, since (see, e.g., Ref. [119])

$$\langle X_{i_1}^* \cdots X_{i_n}^* X_{j_1} \cdots X_{j_n} \rangle$$

$$= \langle U_{i_1,0} \cdots U_{i_n,0} U_{0,j_1}^\dagger \cdots U_{0,j_n}^\dagger \rangle$$

$$= \sum_{\sigma,\tau \in S_n} \delta_{i_1 j_{\sigma(1)}} \cdots \delta_{i_n j_{\sigma(n)}} \mathrm{Wg}(n, \tau \circ \sigma^{-1})$$

$$\propto \sum_\sigma \delta_{i_1 j_{\sigma(1)}} \cdots \delta_{i_n j_{\sigma(n)}} \tag{C1}$$

and zero for a different number of X and X^*'s, just as for Gaussian variables.

APPENDIX D: INTERPOLATING BETWEEN QSD AND QHT

In the main text, we characterized the asymptotic error rates in distinguishing states in the totally symmetric (QSD) and totally asymmetric (QHT) cases with the quantum Chernoff distance and relative entropy, respectively. It is natural to ask if the there is a way to interpolate between the two. This was addressed in Ref. [120] where the type-II error $\beta(A)$ was optimized given the constraint that $\alpha(A) \leq \varepsilon^{1-s}\beta(A)^s$ for $s \geq 0$ and $\varepsilon > 0$. Note that this coincides with QHT for $s = 0$ and QSD for $s = 1$. This is referred to as s-hypothesis testing and the error rate was proven to be given by the s-quantum divergence defined as

$$\xi_s(\rho||\sigma) := \max_{0 \leq \alpha \leq 1} \frac{\log(\mathrm{Tr}[\rho^\alpha \sigma^{1-\alpha}])}{\alpha(1-s)-1}. \tag{D1}$$

It is instructive to examine the two familiar limits. When $s = 0$, the quantity being maximized is the PRRE. We know that the PRRE is monotonically increasing with α so $\xi_0(\rho||\sigma)$ is given by the relative entropy in accordance with our expectation. When $s = 1$, the right-hand side becomes the definition of the quantum Chernoff distance. Using Eq. (74), we can evaluate the s-quantum divergence for random states. We plot the value of α that maximizes the right-hand side as a function of s in Fig. 9. This function monotonically decreases from one at $s = 0$ to zero at $s = \infty$, passing through $\alpha = 1/2$ at $s = 1$.
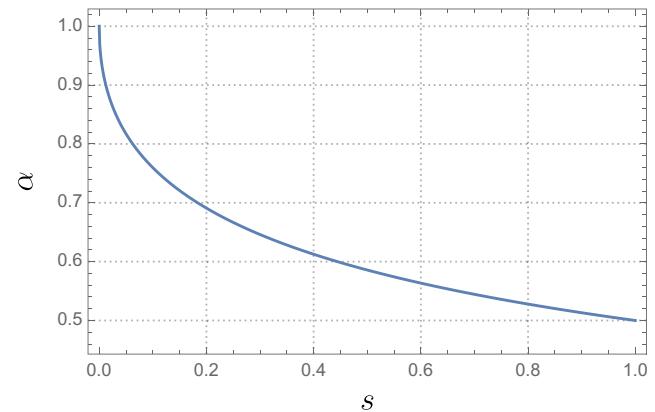

FIG. 9. The interpolation between QHT ($s = 0$) and QSD ($s = 1$) is shown for $d_A < d_B$.

APPENDIX E: ENTANGLEMENT PLATEAU

Ref. [3] was not the first attempt in the literature to use related information theoretic quantities to try to distinguish black hole microstates. A selection of previous works include Refs. [104,121–125]. In particular, Bao and Ooguri [121] used the Holevo information to distinguish black hole microstates. The Holevo information of A for an ensemble of density matrices $\{\rho_i\}$ is given by the average relative entropy from each microstate to the ensemble average

$$\chi(A) := \sum_i p_i D\left(\rho_{A,i} \left\| \sum_j p_j \rho_{A,j}\right.\right)$$
$$= S_{vN}\left(\sum_i p_i \rho_{A,i}\right) - \sum_i p_i S_{vN}(\rho_{A,i}). \quad (E1)$$

This is bounded above by the Shannon entropy, $-\sum_i p_i \log p_i$, which happens to be the black hole entropy, S_{BH}, in this case.

The authors computed each piece holographically to leading order in Newton's constant using the Ryu-Takayanagi formula [126,127] and determined that there are three phases of $\chi(A)$. The microstate entropies were computed from the two extremal surfaces of the previous section, so the transition occurred when A was at half the total system size. In contrast, the ensemble averaged state (Gibbs state) entropy does not transition to the second extremal surface at the same time. While this was not the perspective taken in the original paper, we attribute this to the black hole contributing to the bulk entropy in the FLM formula [80]. This transition occurs when A is much larger than its complement and discussed thoroughly in Ref. [128] where it is called the "entanglement plateaux." In summary, they found the Holevo information to be identically zero until the halfway point, then linearly increase

until a critical size when it saturates to its maximal value of the black hole entropy. The relevant extremal surfaces and behavior of the Holevo information are shown in Fig. 10.

Using fixed-area states, we can improve upon this analysis by computing nonperturbative corrections. As already reviewed in the previous section, all pure black hole microstates will have entropy

$$S_{vN}(\rho_{A,i}) = \frac{A_{1,2}}{4G_N} - \frac{e^{(A_{1,2} - A_{2,1})/4G_N}}{2}, \quad (E2)$$

where the first (second) subscript occurs when A_1 (A_2) is minimal.

For the entropy in the Gibbs state, we must account for the bulk entropy

$$\text{Tr}\left[\left(\sum_i p_i \rho_{A,i}\right)^n\right]$$
$$= \sum_{\tau \in S_n} \frac{e^{[C(\eta^{-1} \circ \tau)A_1 + C(\tau)A_2]/4G_N}}{e^{n(A_1 + A_2)/4G_N}} \text{Tr}[\rho_b^{n_1}] \cdots \text{Tr}[\rho_b^{n_{C(\tau)}}], \quad (E3)$$

where the n_i are the lengths of the cycles of τ. The bulk entropy terms manifestly go away when the black hole is in a pure state. For simplicity, let us take ρ_b to be the maximally mixed state with dimension $e^{S_{\text{BH}}}$. Finite temperature corrections are not immediately important to demonstrate our main conclusions but may be interesting. In this case, the sum simplifies to

$$\text{Tr}\left[\left(\sum_i p_i \rho_{A,i}\right)^n\right]$$
$$= \sum_{\tau \in S_n} \frac{e^{[C(\eta^{-1} \circ \tau)A_1/4G_N + C(\tau)(A_2/4G_N + S_{\text{BH}})]}}{e^{n[(A_1 + A_2)/4G_N + S_{\text{BH}}]}}. \quad (E4)$$


FIG. 10. Left: two competing extremal surfaces are shown for region A. In this case, there is no bulk entropy term because the black hole is in a pure state. Center: when the volume of A, V_A, becomes sufficiently large ($> V_{\text{crit}}$), the area of the red curve plus the black hole entropy is smaller than the area of the blue curve. The black hole entropy can either be viewed as a bulk entropy term or an additional area term. Right: the Holevo information up to nonperturbative corrections as a function of the volume of A.

This is just a renormalization of $A_2/4G_N$ to $A_2/4G_N + S_{\text{BH}}$. This makes sense because the alternative perspective would be to fix the area of the new extremal surface that includes the black hole horizon. Computing the sum as before, we find that

$$S_{vN}\left(\sum_i p_i \rho_{A,i}\right) = \begin{cases} \dfrac{A_1}{4G_N} - \dfrac{e^{(A_1-A_2)/4G_N - S_{\text{BH}}}}{2}, & \dfrac{A_1}{4G_N} < \dfrac{A_2}{4G_N} + S_{\text{BH}}, \\[3ex] \dfrac{A_2}{4G_N} + S_{\text{BH}} - \dfrac{e^{(A_2-A_1)/4G_N + S_{\text{BH}}}}{2}, & \dfrac{A_1}{4G_N} > \dfrac{A_2}{4G_N} + S_{\text{BH}}. \end{cases} \tag{E5}$$

In total, the Holevo information is

$$\chi(A) = \begin{cases} \dfrac{e^{(A_1-A_2)/4G_N}(1 - e^{-S_{\text{BH}}})}{2}, & A_1 < A_2, \\[3ex] \dfrac{A_1-A_2}{4G_N} + \dfrac{e^{(A_1-A_2)/4G_N - S_{\text{BH}}} - e^{(A_2-A_1)/4G_N}}{2}, & A_2 < A_1 < A_2 + 4G_N S_{\text{BH}}, \\[3ex] S_{\text{BH}} + \dfrac{e^{(A_2-A_1)/4G_N}(1 + e^{S_{\text{BH}}})}{2}, & A_1 > A_2 + 4G_N S_{\text{BH}}. \end{cases} \tag{E6}$$

We see that this is the same answer as in Ref. [121] up to nonperturbative corrections [129]. The most important improvements are the first regime where we show the Holevo information to be nonzero, which implies that microstates are distinguishable and the last regime where we show the Holevo information to be submaximal, which implies that microstates are not perfectly distinguishable.

[1] J. D. Bekenstein, Black holes and entropy, Phys. Rev. D **7**, 2333 (1973).

[2] S. W. Hawking, Particle creation by black holes, Commun. Math. Phys. **43**, 199 (1975).

[3] J. Kudler-Flam, Relative entropy of random states and black holes, Phys. Rev. L **126**, 171603 (2021).

[4] E. P. Wigner, Characteristic vectors of bordered matrices with infinite dimensions, Ann. Math. **62**, 548 (1955).

[5] B. Collins and I. Nechita, Random matrix techniques in quantum information theory, J. Math. Phys. **57**, 015215 (2016).

[6] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics, Adv. Phys. **65**, 239 (2016).

[7] D. N. Page, Information in black hole radiation, Phys. Rev. L **71**, 3743 (1993).

[8] P. Hayden and J. Preskill, Black holes as mirrors: Quantum information in random subsystems, J. High Energy Phys. **2007**, 120 (2007).

[9] X. Dong, D. Harlow, and D. Marolf, Flat entanglement spectra in fixed-area states of quantum gravity, J. High Energy Phys. **2019**, 240 (2019).

[10] C. Akers and P. Rath, Holographic Renyi entropy from quantum error correction, J. High Energy Phys. **2019**, 52 (2019).

[11] G. Penington, S. H. Shenker, D. Stanford, and Z. Yang, Replica wormholes and the black hole interior, ArXiv:1911.11977 (2019).

[12] J. M. Deutsch, Thermodynamic entropy of a many-body energy eigenstate, New J. Phys. **12**, 075021 (2010).

[13] C. Murthy and M. Srednicki, Structure of chaotic eigenstates and their entanglement entropy, Phys. Rev. E **100**, 022131 (2019).

[14] T.-C. Lu and T. Grover, Renyi entropy of chaotic eigenstates, ArXiv:1709.08784 (2017).

[15] A. Dymarsky, N. Lashkari, and H. Liu, Subsystem eigenstate thermalization hypothesis, Phys. Rev. E **97**, 012140 (2018).

[16] Positivity can actually be derived from monotonicity, though we choose to separate these conditions for added clarity.

[17] G. Lindblad, Completely positive maps and entropy inequalities, Commun. Math. Phys. **40**, 147 (1975).

[18] A. RENYI, On measures of entropy and information, Proc. Fourth Berkeley Symp. Math. Stat. Prob. **1**, 547 (1961).

[19] D. Petz, Quasi-entropies for finite quantum systems, Rep. Math. Phys. **23**, 57 (1986).

[20] E. H. Lieb, Convex trace functions and the Wigner-Yanase-Dyson conjecture, Adv. Math. (N. Y.) **11**, 267 (1973).

[21] A. Uhlmann, Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory, Commun. Math. Phys. **54**, 21 (1977).

[22] M. M. Wilde, Recoverability for Holevo's just-as-good fidelity, ArXiv:1801.02800 (2018).

[23] A. S. Kholevo, On quasiequivalence of locally normal states, Theor. Math. Phys. **13**, 1071 (1972).

[24] R. Jozsa, Fidelity for mixed quantum states, J. Mod. Opt. **41**, 2315 (1994).

[25] K. M. R. Audenaert and N. Datta, α-z-Rényi relative entropies, J. Math. Phys. **56**, 022202 (2015).

[26] 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**, 122203 (2013).

[27] M. M. Wilde, A. Winter, and D. Yang, Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy, Commun. Math. Phys. **331**, 593 (2014).

[28] M. Ohya and D. Petz, *Quantum Entropy and Its Use*, Theoretical and Mathematical Physics (Springer Berlin Heidelberg, 2004).

[29] C. A. Fuchs and J. van de Graaf, Cryptographic distinguishability measures for quantum mechanical states, arXiv:quant–ph/9712042 (1997).

[30] M. Hayashi, *Quantum Information An Introduction* (Springer, Berlin, Heidelberg, 2006).

[31] S. Khatri and M. M. Wilde, Principles of quantum communication theory: A modern approach, ArXiv:2011.04672 (2020).

[32] Recently, an interpolation QSD and QHT has been introduced [120]. We comment on this interpolation in Appendix D.

[33] C. W. Helstrom, Quantum detection and estimation theory, J. Stat. Phys. **1**, 231 (1969).

[34] C. W. Helstrom, *Quantum Detection and Estimation Theory*, Mathematics in Science and Engineering (Academic Press, New York, NY, 1976).

[35] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, L. Masanes, A. Acin, and F. Verstraete, Discriminating states: The quantum Chernoff bound, Phys. Rev. L **98**, 160501 (2007).

[36] M. Nussbaum and A. Szkoła, The Chernoff lower bound for symmetric quantum hypothesis testing, ArXiv:quant–ph/0607216 (2006).

[37] F. Hiai and D. Petz, The proper formula for relative entropy and its asymptotics in quantum probability, Commun. Math. Phys. **143**, 99 (1991).

[38] T. Ogawa and H. Nagaoka, Strong converse and Stein's lemma in quantum hypothesis testing (2005), p. 28.

[39] M. Hayashi, Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding, Phys. Rev. A **76**, 062301 (2007).

[40] H. Nagaoka, The converse part of the theorem for quantum Hoeffding bound, arXiv:quant–ph/0611289 (2006).

[41] M. Mosonyi and F. Hiai, On the quantum Renyi relative entropies and related capacity formulas, ArXiv:0912.1286 (2009).

[42] M. Mosonyi and T. Ogawa, Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies, Commun. Math. Phys. **334**, 1617 (2015).

[43] M. Nussbaum and A. Szkoła, Exponential error rates in multiple state discrimination on a quantum spin chain, J. Math. Phys. **51**, 072203 (2010).

[44] M. Nussbaum and A. Szkoła, in *Theory of Quantum Computation, Communication, and Cryptography*, edited by W. van Dam, V. M. Kendon, and S. Severini (Springer Berlin Heidelberg, Berlin, Heidelberg, 2011), p. 1.

[45] M. Nussbaum and A. Szkoła, An asymptotic error bound for testing multiple quantum hypotheses, ArXiv:1112.1529 (2011).

[46] K. M. R. Audenaert and M. Mosonyi, Upper bounds on the error probabilities and asymptotic error exponents in quantum multiple state discrimination, J. Math. Phys. **55**, 102201 (2014).

[47] K. Li, Discriminating quantum states: the multiple Chernoff distance, ArXiv:1508.06624 (2015).

[48] The quantum Sanov lemma provides the analogous asymmetric multiple state hypothesis testing result [130,131]. We also note intriguing new multiple state divergences obeying the data processing inequality whose operational meaning is not yet fully understood [132].

[49] H. Yuen, R. Kennedy, and M. Lax, Optimum testing of multiple hypotheses in quantum detection theory, IEEE Trans. Inf. Theory **21**, 125 (1975).

[50] K. Zyczkowski and H.-J. Sommers, Induced measures in the space of mixed quantum states, J. Phys. A Math. Gen. **34**, 7111 (2001).

[51] H.-J. Sommers and K. Zyczkowski, Statistical properties of random density matrices, J. Phys. A Math. Gen. **37**, 8457 (2004).

[52] K. Życzkowski, K. A. Penson, I. Nechita, and B. Collins, Generating random density matrices, J. Math. Phys. **52**, 062201 (2011).

[53] I. Nechita, Asymptotics of random density matrices, Ann. Henri Poincaré **8**, 1521 (2007).

[54] B. Collins and I. Nechita, Gaussianization and eigenvalue statistics for random quantum channels (III), ArXiv:0910.1768 (2009).

[55] B. Collins, I. Nechita, and K. Życzkowski, Random graph states, maximal flow and Fuss-Catalan distributions, J. Phys. A Math. Gen. **43**, 275303 (2010).

[56] P. Hayden, S. Nezami, X.-L. Qi, N. Thomas, M. Walter, and Z. Yang, Holographic duality from random tensor networks, J. High Energy Phys. **2016**, 9 (2016).

[57] E. Brézin and A. Zee, Universal relation between Green functions in random matrix theory, Nucl. Phys. B **453**, 531 (1995).

[58] J. Jurkiewicz, G. Łukaszewski, and M. A. Nowak, Diagrammatic approach to fluctuations in the Wishart ensemble, Acta Phys. Pol. B **39**, 799 (2008).

[59] H. Shapourian, S. Liu, J. Kudler-Flam, and A. Vishwanath, Entanglement negativity spectrum of random mixed states: A diagrammatic approach, ArXiv:2011.01277 (2020).

[60] N. Lashkari, Modular Hamiltonian for excited states in conformal field theory, Phys. Rev. L **117**, 041601 (2016).

[61] G. Kreweras, Sur les partitions non croisees d'un cycle, Discrete Math. **1**, 333 (1972).

[62] R. Simion, Noncrossing partitions, Discrete Math. **217**, 367 (2000).

[63] The two elements of the piecewise function are equivalent on the integers. The reason why we write it as a piecewise function is for ease of analytic continuation to noninteger values because the hypergeometric functions are entire when the argument is less than one.

[64] D. N. Page, Average entropy of a subsystem, Phys. Rev. L **71**, 1291 (1993).

[65] This can be generalized such that the auxiliary systems for σ and ρ are of different sizes d_{B_1} and d_{B_2}. In the diagrammatics, this corresponds to assigning different weights to

the black and red dashed lines. The resulting generalized sums are still tractable, though, we do not currently have use for these calculations in our applications to black holes because d_B corresponds to the size of the black hole, which is simple to measure by an outside observer, rendering σ_A and ρ_A easily distinguishable when $d_{B_1} \neq d_{B_2}$. Some exact results for this setup in the Wishart ensemble can be found in Refs. [133–135].

[66] Z. Puchała, Ł. Pawela, and K. Życzkowski, Distinguishability of generic quantum states, Phys. Rev. A **93**, 062112 (2016).

[67] This notation means that there are m_i cycles of length i.

[68] We note that an exact expression was recently found for the fidelity of two random density matrices, consistent with our large-N results [135]. Our result is complementary as the exact expression is very complicated and not tractable at large N.

[69] K. Życzkowski and H.-J. Sommers, Average fidelity between random quantum states, Phys. Rev. A **71**, 032313 (2005).

[70] R. Renner, PhD thesis, 2005.

[71] J. Zhang, P. Ruggiero, and P. Calabrese, Subsystem trace distance in quantum field theory, Phys. Rev. L **122**, 141602 (2019).

[72] The rotation factor for even α when $\alpha_1 = \alpha_2$ is only $\alpha/2$ because of indistinguishability.

[73] J. Mejía, C. Zapata, and A. Botero, The difference between two random mixed quantum states: Exact and asymptotic spectral analysis, ArXiv:1511.07278 (2015).

[74] J. Preskill, Quantum computing in the NISQ era and beyond, ArXiv:1801.00862 (2018).

[75] C. Holzhey, F. Larsen, and F. Wilczek, Geometric and renormalized entropy in conformal field theory, Nucl. Phys. B **424**, 443 (1994).

[76] P. Calabrese and J. Cardy, Entanglement entropy and quantum field theory, J. Stat. Mech.: Theory Exp. **2004**, 06002 (2004).

[77] A. Lewkowycz and J. Maldacena, Generalized gravitational entropy, J. High Energy Phys. **2013**, 90 (2013).

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

[79] On the CFT side of the duality, these states should be thought of as high-energy pure states.

[80] T. Faulkner, A. Lewkowycz, and J. Maldacena, Quantum corrections to holographic entanglement entropy, J. High Energy Phys. **2013**, 74 (2013).

[81] This was pointed out in Ref. [11].

[82] A. Almheiri, X. Dong, and D. Harlow, Bulk locality and quantum error correction in AdS/CFT, J. High Energy Phys. **2015**, 163 (2015).

[83] D. Harlow, The Ryu-Takayanagi formula from quantum error correction, Commun. Math. Phys. **354**, 865 (2017).

[84] D. L. Jafferis, A. Lewkowycz, J. Maldacena, and S. J. Suh, Relative entropy equals bulk relative entropy, J. High Energy Phys. **2016**, 4 (2016).

[85] Recently, the SRRE between a state with a single fixed-area surface and a state with two fixed-area surfaces was evaluated [136].

[86] G. Penington, Entanglement wedge reconstruction and the information paradox, J. High Energy Phys. **2020**, 2 (2020).

[87] A. Almheiri, N. Engelhardt, D. Marolf, and H. Maxfield, The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole, J. High Energy Phys. **2019**, 63 (2019).

[88] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, Replica wormholes and the entropy of Hawking radiation, J. High Energy Phys. **2020**, 13 (2020).

[89] We thank Daniel Harlow for comments on this point.

[90] Known graybody factors do not qualitatively change this statement in any meaningful way.

[91] For analogous reasons, the exact same analysis as in Sec. III can be made for the SRRE and trace distance, but we do not write these out explicitly to avoid repetition.

[92] N. Engelhardt and A. C. Wall, Quantum extremal surfaces: Holographic entanglement entropy beyond the classical regime, J. High Energy Phys. **2015**, 73 (2015).

[93] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, The entropy of Hawking radiation, ArXiv:2006.06872 (2020).

[94] Y. Chen and H. W. Lin, Signatures of global symmetry violation in relative entropies and replica wormholes, ArXiv:2011.06005 (2020).

[95] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill, Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence, J. High Energy Phys. **2015**, 149 (2015).

[96] L. R. Ford and D. R. Fulkerson, Maximal flow through a network, Can. J. Math. **8**, 399 (1956).

[97] P. Elias, A. Feinstein, and C. Shannon, A note on the maximum flow through a network, IRE Trans. Inf. Theory **2**, 117 (1956).

[98] P. Calabrese and J. Cardy, Evolution of entanglement entropy in one-dimensional systems, J. Stat. Mech.: Theory Exp. **2005**, 04010 (2005).

[99] T. Hartman and J. Maldacena, Time evolution of entanglement entropy from black hole interiors, J. High Energy Phys. **2013**, 14 (2013).

[100] J. M. Deutsch, Quantum statistical mechanics in a closed system, Phys. Rev. A **43**, 2046 (1991).

[101] M. Srednicki, Chaos and quantum thermalization, Phys. Rev. E **50**, 888 (1994).

[102] J. M. Deutsch, Eigenstate thermalization hypothesis, Rep. Prog. Phys. **81**, 082001 (2018).

[103] J. R. Garrison and T. Grover, Does a Single Eigenstate Encode the Full Hamiltonian?, Phys. Rev. X **8**, 021026 (2018).

[104] N. Lashkari, A. Dymarsky, and H. Liu, Eigenstate thermalization hypothesis in conformal field theory, ArXiv:1610.00302 (2016).

[105] In fact, it is not quite correct to consider these subsystem energy eigenstates due to the interaction terms in the Hamiltonian coupling A and B that lead to correlations near the boundary, as emphasized in Ref. [14]. It is more accurate to consider very similar states referred to as "many-body Berry (MBB) states" in Ref. [14]. These are constructed from perturbing an integrable Hamiltonian by an integrability breaking term. In MBB states, the subsystem eigenstates are not *energy* eigenstates, but local product states. The following calculation is unchanged.

[106] X. Dong and H. Wang, Enhanced corrections near holographic entanglement transitions: A chaotic case study, J. High Energy Phys. **2020**, 7 (2020).

[107] I. Kourkoulou and J. Maldacena, Pure states in the SYK model and nearly-AdS$_2$ gravity, ArXiv:1707.02325 (2017).

[108] A. Almheiri, A. Mousatov, and M. Shyani, Escaping the interiors of pure boundary-state black holes, ArXiv:1803.04434 (2018).

[109] S. Cooper, M. Rozali, B. Swingle, M. Van Raamsdonk, C. Waddell, and D. Wakeham, Black hole microstate cosmology, J. High Energy Phys. **2019**, 65 (2019).

[110] M. Miyaji, T. Takayanagi, and T. Ugajin, Spectrum of end of the world Branes in holographic BCFTs, ArXiv:2103.06893 (2021).

[111] S. Hollands and K. Sanders, Entanglement measures and their properties in quantum field theory, ArXiv:1702.04924 (2017).

[112] E. Witten, Notes on some entanglement properties of quantum field theory, ArXiv:1803.04993 (2018).

[113] L. Vidmar and M. Rigol, Entanglement entropy of eigenstates of quantum chaotic Hamiltonians, Phys. Rev. L **119**, 220603 (2017).

[114] D. Marolf, S. Wang, and Z. Wang, Probing phase transitions of holographic entanglement entropy with fixed area states, J. High Energy Phys. **2020**, 84 (2020).

[115] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability*, London Mathematical Society Lecture Note Series (Cambridge University Press, 2006).

[116] J. A. Mingo and R. Speicher, *Free Probability and Random Matrices* (Springer, New York, 2017).

[117] W. Młotkowski, M. A. Nowak, K. A. Penson, and K. Życzkowski, Spectral density of generalized Wishart matrices and free multiplicative convolution, Phys. Rev. E **92**, 012121 (2015).

[118] K. A. Penson and K. Życzkowski, Product of Ginibre matrices: Fuss-Catalan and Raney distributions, Phys. Rev. E **83**, 061118 (2011).

[119] B. Collins, Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral and free probability, arXiv:math–ph/0205010 (2002).

[120] R. Salzmann and N. Datta, Interpolating between symmetric and asymmetric hypothesis testing, ArXiv:2104.09553 (2021).

[121] N. Bao and H. Ooguri, Distinguishability of black hole microstates, Phys. Rev. D **96**, 066017 (2017).

[122] G. Sárosi and T. Ugajin, Relative entropy of excited states in conformal field theories of arbitrary dimensions, J. High Energy Phys. **2017**, 60 (2017).

[123] B. Michel and A. Puhm, Corrections in the relative entropy of black hole microstates, J. High Energy Phys. **2018**, 179 (2018).

[124] W.-Z. Guo, F.-L. Lin, and J. Zhang, Distinguishing black hole microstates using Holevo information, Phys. Rev. L **121**, 251603 (2018).

[125] I. Akal, Universality, intertwiners and black hole information, ArXiv:2010.12565 (2020).

[126] S. Ryu and T. Takayanagi, Holographic derivation of entanglement entropy from the anti de sitter space/conformal field theory correspondence, Phys. Rev. L **96**, 181602 (2006).

[127] S. Ryu and T. Takayanagi, Aspects of holographic entanglement entropy, J. High Energy Phys. **2006**, 045 (2006).

[128] V. E. Hubeny, H. Maxfield, M. Rangamani, and E. Tonni, Holographic entanglement plateaux, J. High Energy Phys. **2013**, 92 (2013).

[129] This result seems to be in tension with Ref. [124] where it was argued that there are $O(G_N^0)$ corrections to the Holevo information in AdS$_3$/CFT$_2$. It is not clear to us how to resolve this tension due to the very different flavors of assumptions that went into both calculations. This may be worth examining further.

[130] M. Hayashi, Optimal sequence of quantum measurements in the sense of Stein's lemma in quantum hypothesis testing, J. Phys. A Math. Gen. **35**, 10759 (2002).

[131] I. Bjelaković, J.-D. Deuschel, T. Krüger, R. Seiler, R. Siegmund-Schultze, and A. Szkoła, A quantum version of Sanov's theorem, Commun. Math. Phys. **260**, 659 (2005).

[132] K. Furuya, N. Lashkari, and S. Ouseph, Monotonic multistate quantum f-divergences, ArXiv:2103.09893 (2021).

[133] S. Kumar, Wishart and random density matrices: Analytical results for the mean-square Hilbert-Schmidt distance, Phys. Rev. A **102**, 012405 (2020).

[134] S. Kumar and S. Sai Charan, Spectral statistics for the difference of two Wishart matrices, J. Phys. A Math. Gen. **53**, 505202 (2020).

[135] A. Laha, A. Aggarwal, and S. Kumar, Random density matrices: Analytical results for mean root fidelity and mean square Bures distance, ArXiv:2105.02743 (2021).

[136] P. Hayden, O. Parrikar, and J. Sorce, The Markov gap for geometric reflected entropy, ArXiv:2107.00009 (2021).

[137] We note that both PRRE and SRRE can be described as specific cases of the α-z-relative entropies defined as [25]

$$D_{\alpha,z}(\rho||\sigma) := \frac{1}{\alpha-1}\log\{\mathrm{Tr}[(\sigma^{(1-\alpha)/2z}\rho^{\alpha/z}\sigma^{(1-\alpha)/2z})^z]\},$$

though we will not discuss this more general quantity. (SRRE) [26,27].