

The Page Curve Is Not Enough: Universality, Ensemble Dependence, and Microscopic Ambiguity in Black Hole Information Recovery

Bernd Johannes Wuebben

New York, NY, USA

E-mail: wuebben@gmail.com

ABSTRACT: A Page curve constrains radiation entanglement, but its implications for microscopic information recovery depend on which state, ensemble, and channel data are specified. We examine this dependence through explicit finite-dimensional constructions and controlled random-state limits. Two density-matrix ensembles can have the same averaged spectral measure, every averaged replica moment, and the same mean von Neumann entropy, while their entropy and purity fluctuations differ. Fixed Marchenko–Pastur spectra and normalized Wishart spectra give a related large-dimension comparison. At the channel level, an isometry and a replacement channel can reproduce the complete nested radiation trajectory for one input while having different correlations with an external reference. For normalized states, annealed and quenched Rényi entropies have the same von Neumann limit; the derivative of their finite-order difference measures the entropy variance. We derive normalized replica-covariance identities and illustrate their normalization with the standard Haar purity covariance, separating these quantities from entropy of an averaged state. Finally, Gaussian and quartic fluctuations with equal variance give different entropy-transition profiles, illustrating the additional role of joint branch fluctuations and reservoir constraints. These calculations delimit what entropy and spectral agreement establish and specify further comparisons needed in gravitational models.

KEYWORDS: Black hole information paradox, Page curve, replica wormholes, entanglement entropy, universality, Marchenko–Pastur law, free probability, annealed-versus-quenched entropy

Contents

1	Introduction	1
1.1	Explicit comparisons	1
1.2	Relation to established results	2
1.3	Organization and conventions	2
2	What the Page curve establishes	2
2.1	Entropy of a state and recovery of an input	3
2.2	An exact rising and falling entropy curve	4
2.3	An external reference distinguishes the channels	4
2.4	Spectral information and preparation averages	5
3	Identical averaged spectra and distinct fluctuations	6
4	The normalized random-state benchmark	7
4.1	Exact normalization and finite-dimensional moments	7
4.2	The limiting density and entropy	8
4.3	Which ingredients are universal?	9
4.4	What a matching density establishes	9
5	Thermodynamic fluctuations and the observable being averaged	9
5.1	Averaging two competing branches	9
5.2	Equal widths and different crossover shapes	10
5.3	Regular phases, finite reservoirs, and criticality	11
5.4	Averaging sector entropies and combining spectra	12
6	Design matching and shared spectral limits	13
6.1	Finite designs determine finitely many moments	13
6.2	From moments to entropy: the additional assumptions	14
6.3	A fixed-spectrum ensemble with the same limiting density	14
6.4	Information absent even from the complete spectrum	15
6.5	A conditional organization of observables	16
7	Rényi averages and entropy fluctuations	17
7.1	The two Rényi averages	17
7.2	Averaging a spectrum and averaging a state	18
7.3	An exactly evaluated two-state ensemble	19
7.4	Normalized replica covariances	20
7.5	Concentration and its limits	20
7.6	Fluctuations within and between preparations	21
7.7	Application to gravitational ensembles	21
8	Comparisons between gravitational descriptions	22
8.1	What an entropy calculation establishes	22
8.2	The objects supplied by different approaches	23
8.3	A comparison table with explicit status	23
8.4	Precision, normalization, and accessible observables	24

9	From equilibrium states to evaporation dynamics	25
9.1	Consistency of successive radiation states	25
9.2	Fixed total energy does not fix the black-hole energy	26
9.3	A fluctuation width is not an evaporation timescale	26
9.4	What a dynamical comparison should specify	27
10	Further calculations	27
10.1	Normalized two-trace fluctuations in a common preparation	27
10.2	Generalized-entropy response near a critical point	28
10.3	Spectral comparisons with matched conservation laws	28
10.4	Reference correlations and recovery	28
10.5	Reproducibility of the worked comparisons	28
11	Conclusion	29
A	Review of Free Probability and Non-Crossing Partitions	29
A.1	Non-crossing partitions	29
A.2	Catalan numbers	30
A.3	Narayana numbers	30
A.4	The Möbius function on the non-crossing partition lattice	31
A.5	Free cumulants and the moment–cumulant relation	31
A.6	The R -transform	31
A.7	The free Poisson (Marchenko–Pastur) distribution	32
A.8	Why this matters for replica wormholes	32
B	Normalized moments and the spectral comparison	33
B.1	The symmetric-projector formula	33
B.2	Planar terms and finite-dimensional corrections	34
B.3	Evaluation of the Marchenko–Pastur entropy integral	34
B.4	Quantile normalization and convergence	35
B.5	Reproducibility of the comparison	35
C	Fluctuation profiles and finite-reservoir limits	35
C.1	An integrated-tail representation	36
C.2	Quartic moments and standardization	36
C.3	Reservoir assumptions and response of two branches	36
D	Normalized Haar replica covariances	37
D.1	The Haar moment identity	37
D.2	The fourth replica moment	38
D.3	An exact two-qubit Rényi gap	39
D.4	Why normalization affects the leading covariance	39
D.5	From moment covariances to entropy variance	39

1 Introduction

The entropy of Hawking radiation provides a basic test of a proposed description of black-hole evaporation. Hawking’s semiclassical calculation led to a loss of purity [11]. For a pure state evolving unitarily into an accessible final radiation system, the radiation entropy must eventually return to zero. Page’s random-state calculation provides a quantitative benchmark for the intermediate entanglement [22]. Island and replica-wormhole calculations have reproduced the expected transition in concrete gravitational models and connected it to entanglement-wedge reconstruction [1, 24, 25]. The resulting entropy calculations place strong constraints on a semiclassical description.

The inverse question is different: what properties of the underlying quantum state, ensemble, or channel are fixed by an entropy curve? A scalar entropy discards eigenvectors and most eigenvalue information. An averaged spectral density retains more information, but discards correlations between eigenvalues in a realization. A trajectory for one input state does not specify the response to an unknown input or to an input entangled with a reference. These distinctions matter when a successful entropy calculation is used to compare microscopic mechanisms of information recovery.

We study these questions through finite-dimensional constructions and controlled random-state limits. The purpose is to determine which inferences require additional data. We do not assume that the principal approaches to the information problem form mutually exclusive theories, or that each has a unique radiation spectrum. A comparison between two gravity models must specify their states, observables, and constraints before the mathematical examples below can be applied.

1.1 Explicit comparisons

The first comparison concerns channels. Given one input and an isometric evolution, a replacement channel can reproduce the complete radiation state at every nested release step for that input. The replacement channel discards unknown input information. A reference initially maximally entangled with a D -dimensional input distinguishes the final channels: its mutual information with the radiation is $2\log D$ for the isometry and zero for the replacement. An explicit sequence of released Bell pairs gives a rising and falling entropy curve in both cases. This is an elementary channel counterexample, with no claim that a replacement channel satisfies a gravitational locality or energy constraint. It states what radiation data from a single preparation, even when known exactly, can establish on their own.

The second comparison concerns ensembles. We construct two four-dimensional density-matrix ensembles with the same averaged spectral measure. Consequently they have the same averaged replica moments at every positive order and the same mean von Neumann entropy. One has fluctuating entropy and purity, while the other has a fixed spectrum and hence no spectral fluctuations. The agreement holds at finite dimension. A separate comparison between fixed Marchenko–Pastur quantiles and normalized Wishart matrices shows the same distinction in the usual large-dimension Page regime. Their limiting density and entropy correction agree, while their purity variances differ. These examples isolate the information contained in one-trace moments from the information in connected two-trace moments.

The third comparison concerns averaging and thermodynamic constraints. For normalized finite-dimensional density matrices, the annealed and quenched Rényi entropies have the same von Neumann limit. Their finite-order difference is nonnegative when defined as quenched minus annealed, and its derivative at order one is half the ensemble variance of the entropy. We derive this relation and evaluate examples with the normalization kept inside the ensemble average. We also compare Gaussian and quartic entropy differences of equal variance: their averaged-minimum crossover profiles differ. The full profile can identify the scalar fluctuation law under a fixed-law assumption, but a marginal

heat capacity cannot determine the joint entropy-branch fluctuations or the associated evaporation dynamics.

1.2 Relation to established results

Most of the mathematical ingredients are established. Haar-state entanglement and its fluctuations provide the finite-dimensional benchmarks [4, 10, 15, 22]. Unitary designs specify which polynomial averages match Haar measure; random-circuit results make explicit the dependence on the required order [6]. We use a finite-moment statement with its error dependence, rather than infer an entire entropy trajectory from a fixed design order. The Marchenko–Pastur law and its non-crossing moment expansion are standard random-matrix and free-probability results [21].

On the gravitational side, ref. [25] supplies a concrete replica model and also studies reconstruction through the Petz map. Thus the scope of that work already extends beyond its entropy curve. The ensemble interpretation of gravitational path integrals has its own state and factorization questions [17, 26]. Bousso and Miyaji explicitly calculate fluctuations of the radiation entropy in a gravitational model [5]. Our normalized covariance formulas and elementary examples clarify which quantities such calculations compare; the use of entropy fluctuations as a probe is not new here.

Energy constraints and fluctuations near competing entropy branches have likewise been studied in chaotic eigenstates and fixed-area constructions [18, 20]. Paper I derives a broad-distribution correction between an exact planar spectrum and a level-resolved diagonal approximation in a specified canonical random-state model [32]. Paper II studies the regular, charged-critical, and finite-reservoir fluctuation laws in a thermodynamic density model [33]. We restate the inputs needed here and distinguish these two calculations from ensemble variances of normalized radiation entropies.

The contribution of the present paper is a constructive comparison of these different inverse questions. The finite examples make their logical separation explicit, while the random-state and thermodynamic calculations provide quantitative tests within specified models. The results do not select a microscopic theory of gravity. They specify additional comparisons needed before matching entropy curves can be used to draw that conclusion.

1.3 Organization and conventions

Section 2 treats channels and reference correlations. Section 3 gives the exact finite-dimensional spectral comparison. Section 4 sets out the random-state moment calculation and the scope of design matching. Section 5 treats fluctuating entropy branches and conserved energy. Design bounds and the comparison at a common spectral limit are developed in section 6, and section 7 derives the averaging and covariance relations. Sections 8 and 9 discuss gravitational comparisons and time-dependent interpretation. The remaining research questions are collected in section 10. The appendices give the combinatorial, spectral, and probability calculations used in the text.

All logarithms are natural. Density matrices have unit trace in each realization unless an unnormalized operator is explicitly introduced. Averages refer to the ensemble named in the relevant calculation. A finite-dimensional identity, a large-dimension limit, and a semiclassical gravity approximation are stated separately. Throughout, $\gamma = d_R/d_B$ is the radiation-to-black-hole dimension ratio.

2 What the Page curve establishes

The island calculations recover a turnover of the fine-grained radiation entropy in controlled semiclassical settings [1, 2, 24, 25]. This is an important constraint on the gravitational description. To state what follows from the entropy curve alone, however, one must distinguish the state produced by a

specified preparation from the channel that acts on unknown quantum information. The distinction is standard in quantum information theory; the examples below make its consequences for a Page curve explicit. They are finite-dimensional constructions, without a claim to model local or energy-conserving gravitational dynamics.

2.1 Entropy of a state and recovery of an input

Let A be an input system of dimension $D \geq 2$. The emitted radiation after k steps is $R_k = b_1 \otimes \cdots \otimes b_k$, and B_k denotes the remaining system. We use a finite-dimensional tensor factorization; the identification of the corresponding radiation algebra in gravity is an additional part of any gravitational application. For a pure state on $B_k R_k$, the radiation entropy satisfies

$$S(R_k) = -\text{tr}(\rho_{R_k} \ln \rho_{R_k}), \quad S(R_k) = S(B_k) \leq \min\{\ln d_{R_k}, \ln d_{B_k}\}. \quad (2.1)$$

Approach to this bound is a typicality statement, as in the random-state calculation of Page [22]. Unitarity alone does not require an initial rise, a unique maximum, or a particular turnover profile. A unitary evolution may, for example, emit a product state for a specially chosen input. A Page-like curve combines the endpoint purity expected under complete unitary evaporation with assumptions about intermediate entanglement.

Write $\mathcal{N}_k$ for the completely positive trace-preserving channel from A to R_k . For a specified input $\psi_0 = |\psi_0\rangle\langle\psi_0|$, the entropy data are the numbers $S[\mathcal{N}_k(\psi_0)]$. Information recovery asks whether the action of $\mathcal{N}_k$ on an input space can be reversed, including when the input is entangled with a reference. A single-input entropy curve is therefore a restriction of the data relevant to recovery. The following elementary construction shows that even complete knowledge of the radiation density matrices for that input does not remove this restriction.

Proposition 2.1 (Identical radiation histories for one input). *Let $R_N = b_1 \otimes \cdots \otimes b_N$ have dimension at least D , and let $|\Omega\rangle \in R_N$ and $|\psi_0\rangle \in A$ be normalized. There exist an isometric channel $\mathcal{V}$ and a constant-output channel $\mathcal{C}$ from A to R_N such that their reductions to every R_k coincide on ψ_0 . The channel $\mathcal{V}$ has a left inverse on the input space, whereas $\mathcal{C}$ transmits no information about the input.*

Proof. Extend $|\psi_0\rangle$ to an orthonormal basis of A , and extend $|\Omega\rangle$ to an orthonormal set of D vectors in R_N . Mapping the first set to the second defines an isometry V with $V|\psi_0\rangle = |\Omega\rangle$. Define

$$\mathcal{V}(X) = V X V^\dagger, \quad \mathcal{C}(X) = \text{tr}(X) |\Omega\rangle\langle\Omega|. \quad (2.2)$$

Both maps are channels. For every k ,

$$\mathcal{V}_k(\psi_0) = \mathcal{C}_k(\psi_0) = \text{tr}_{b_{k+1} \cdots b_N} |\Omega\rangle\langle\Omega|, \quad (2.3)$$

where the subscript denotes restriction of the output to R_k . The map $V^\dagger$ reverses $\mathcal{V}$ on its image and extends to a trace-preserving recovery channel on all of R_N . The output of $\mathcal{C}$ is independent of its input, so no recovery channel can reverse it on a space of dimension greater than one. $\square$

These are consistent histories of sequentially released subsystems: prepare the full output once and emit the factors $b_1, \dots, b_N$ in order. Equation (2.3) equates every radiation observable for the chosen input, including all Rényi entropies and all matrix elements in a specified basis. It does not assert equality for other inputs. In particular, it establishes a limitation of single-preparation data rather than an obstruction to quantum channel tomography.

2.2 An exact rising and falling entropy curve

An explicit choice of the output illustrates the proposition. Take $N = 2q$ qubits and pair qubit j with qubit $q + j$:

$$|\Omega\rangle = \bigotimes_{j=1}^q \frac{|0\rangle_j |0\rangle_{q+j} + |1\rangle_j |1\rangle_{q+j}}{\sqrt{2}}. \quad (2.4)$$

For $k \leq q$, every emitted qubit is one half of a Bell pair whose other half remains inside, so ρ_{R_k} is maximally mixed. For $k > q$, each additional emission completes one of those pairs. The number of pairs crossing the radiation cut is therefore $\min(k, 2q - k)$, giving exactly

$$S(R_k) = \min(k, 2q - k) \ln 2, \quad 0 \leq k \leq 2q. \quad (2.5)$$

For six qubits, the entropies in units of $\ln 2$ are 0, 1, 2, 3, 2, 1, 0. The final radiation state is pure. The proposition produces this same trajectory under an information-preserving channel and an input-erasing channel. The example is deliberately simple: its linear branches and flat nonzero entanglement spectrum are not the finite-size Haar profile. Any desired pure multipartite output, including a Haar-typical one, can instead be used in the proposition.

The constant-output channel is often called a *replacer channel*. It has a globally isometric realization

$$W|\psi\rangle_A = |\Omega\rangle_{R_N} \otimes |\psi\rangle_E, \quad \mathcal{C}(X) = \text{tr}_E(WXW^\dagger), \quad (2.6)$$

where E is an unobserved system of dimension D . The input has moved to E ; it has not been destroyed by the enlarged evolution. Thus the example does not contradict unitarity of a closed system. It shows that purity of the observed radiation does not certify that the radiation contains the input. If one independently assumes that the complete evolution is an isometry and that no complementary output remains, recovery follows from those assumptions. They are not implied by the curve. Equation (2.6) also identifies why this construction is not a model of complete black-hole evaporation with a proven absence of remnants or hidden sectors.

2.3 An external reference distinguishes the channels

The relevant distinction can be measured by entangling the input with a reference, as in quantum communication and the Hayden–Preskill recovery problem [13, 27]. Let

$$|\Phi_D\rangle_{QA} = \frac{1}{\sqrt{D}} \sum_{i=1}^D |i\rangle_Q |i\rangle_A, \quad |\Psi_V\rangle = (I_Q \otimes V)|\Phi_D\rangle. \quad (2.7)$$

The two final joint states are

$$\rho_{QR_N}^V = |\Psi_V\rangle\langle\Psi_V|, \quad \rho_{QR_N}^C = \frac{I_Q}{D} \otimes |\Omega\rangle\langle\Omega|. \quad (2.8)$$

Consequently their mutual information, defined by $I(Q : R) = S(Q) + S(R) - S(QR)$, is

$$I(Q : R_N)_V = 2 \ln D, \quad I(Q : R_N)_C = 0. \quad (2.9)$$

This is a different input preparation from the pure-state experiment in equation (2.3). For the reference-entangled input, an information-preserving channel produces mixed radiation entangled with Q ; requiring its radiation entropy to vanish would be the wrong recovery criterion.

One can also compare with the ideal encoded state through

$$F_V(\mathcal{N}) = \langle \Psi_V | (\text{id}_Q \otimes \mathcal{N})(|\Phi_D\rangle\langle\Phi_D|) | \Psi_V \rangle. \quad (2.10)$$

For the two channels,

$$F_V(\mathcal{V}) = 1, \quad F_V(\mathcal{C}) = \frac{1}{D^2}. \quad (2.11)$$

Indeed, substituting the second state in equation (2.8) gives $D^{-2} \sum_i |\langle \Omega | V | i \rangle|^2 = D^{-2}$, because $|\Omega\rangle$ lies in the image of V . Applying a decoder that inverts V on this image gives the usual entanglement fidelity for the maximally mixed input [27]. Every decoder applied to a replacer still produces a constant-output channel, whose entanglement fidelity for that input is D^{-2} . These comparisons distinguish transmission of quantum information from purification of a specified output state without requiring a reconstruction of every matrix element of the channel.

Full reference-assisted state tomography supplies a stronger positive result. For the maximally entangled input above, define the normalized Choi state of a channel $\mathcal{N} : A \rightarrow R$ by

$$J_{\mathcal{N}} = (\text{id}_Q \otimes \mathcal{N})(|\Phi_D\rangle\langle\Phi_D|) = \frac{1}{D} \sum_{i,j} |i\rangle\langle j|_Q \otimes \mathcal{N}(|i\rangle\langle j|). \quad (2.12)$$

The entire state $J_{\mathcal{N}}$ determines the channel on every input:

$$\mathcal{N}(\omega) = D \text{Tr}_Q[(\omega^T \otimes I_R) J_{\mathcal{N}}], \quad (2.13)$$

where the transpose is taken in the basis defining $|\Phi_D\rangle$. This is the standard channel-state correspondence [7]; the formula follows by contracting the displayed matrix units. Thus one suitably entangled preparation with complete joint-state information can determine a channel. Entropy or selected reference correlators retain only part of that information. The single-input counterexample concerns a fixed unentangled input, and does not apply to tomography of $J_{\mathcal{N}}$.

2.4 Spectral information and preparation averages

Even at a fixed time, the complete entanglement spectrum does not determine the density matrix relative to specified physical observables. For example, let

$$\rho = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix}, \quad H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \rho' = H \rho H^\dagger, \quad (2.14)$$

with $0 < p < 1$ and $p \neq 1/2$. These states have identical Rényi entropies at every order. Holding the observable $Z = \text{diag}(1, -1)$ fixed, however,

$$\text{tr}(\rho Z) = 2p - 1, \quad \text{tr}(\rho' Z) = 0. \quad (2.15)$$

The comparison concerns different states in the same physical basis, not a simultaneous relabeling of states and observables. Spectral statistics can distinguish some state ensembles, but they do not specify the eigenvectors, an encoding map, or the dynamics that produced the states. Proposition 2.1 gives the stronger statement that even full state information for a fixed unentangled input cannot determine the channel on other inputs.

A separate distinction concerns averaging. In matrix-integral descriptions of JT gravity, the gravitational observables have an ensemble interpretation [17, 26]. Other averages can refer to states or circuits within a fixed theory. These operations should be specified independently of whether the

underlying Hamiltonian is fixed. For normalized density matrices in a common finite-dimensional space, differentiation at replica number one gives exactly

$$-\partial_n \ln \mathbb{E} \operatorname{tr} \rho^n|_{n=1} = \mathbb{E}[-\operatorname{tr}(\rho \ln \rho)]. \quad (2.16)$$

The equality follows from $\operatorname{tr} \rho = 1$ and differentiation under the average; $x \ln x$ is bounded on $[0, 1]$. The entropy of the averaged state, $S(\mathbb{E}\rho)$, is a different observable. Approximating replica moments or taking a dimension limit before the derivative can require additional control, but it does not change this exact finite-dimensional identity. Section 7 treats the finite-order Rényi distinction and entropy fluctuations.

The scope of the argument is consequently precise. A Page curve tests a property of radiation states in a specified preparation. Additional spectral information tests further properties of those states and of their distribution. Recovery tests how the radiation depends on an input, including its entanglement with a reference. The examples in this section separate these questions; they do not assert that two fully specified gravitational theories agree on every such test.

3 Identical averaged spectra and distinct fluctuations

An averaged spectral measure contains the expectation of every linear spectral statistic, including the von Neumann entropy. It need not determine the covariance of two such statistics. The following four-dimensional example demonstrates this distinction exactly, without a limiting argument. It is an explicit application of the standard difference between one-point and two-point probability data.

Set $c = 1/4$ and choose $0 \leq u < v < c$. Define three normalized spectra,

$$\begin{aligned} L_u &= (c + u, c - u, c + u, c - u), \\ L_v &= (c + v, c - v, c + v, c - v), \\ L_{uv} &= (c + u, c - u, c + v, c - v). \end{aligned} \quad (3.1)$$

Ensemble $\mathcal{A}$ chooses L_u or L_v with equal probabilities; ensemble $\mathcal{B}$ always uses L_{uv} . In each case conjugate the diagonal state by an independent Haar unitary in $U(4)$. All states are full rank and can be purified with an environment of dimension at least four. Each ensemble has averaged density matrix $I_4/4$.

Proposition 3.1 (Exact matching of averaged spectral data). *The ensembles $\mathcal{A}$ and $\mathcal{B}$ have identical averaged spectral measures. In particular, for every function g defined on the four eigenvalue values,*

$$\mathbb{E}_{\mathcal{A}} \operatorname{Tr} g(\rho) = \mathbb{E}_{\mathcal{B}} \operatorname{Tr} g(\rho). \quad (3.2)$$

They therefore have identical $\mathbb{E} \operatorname{Tr} \rho^n$ for all real $n > 0$ and identical mean von Neumann entropy. Nonetheless,

$$\operatorname{Var}_{\mathcal{A}}(\operatorname{Tr} \rho^2) = 4(u^2 - v^2)^2 > 0, \quad \operatorname{Var}_{\mathcal{B}}(\operatorname{Tr} \rho^2) = 0, \quad (3.3)$$

$$\operatorname{Var}_{\mathcal{A}} S(\rho) = [f(u) - f(v)]^2 > 0, \quad \operatorname{Var}_{\mathcal{B}} S(\rho) = 0, \quad (3.4)$$

where $f(t) = -(c + t) \log(c + t) - (c - t) \log(c - t)$.

Proof. Each of $c + u, c - u, c + v, c - v$ occurs once in the average of the counting measures of L_u and L_v , and once in the counting measure of L_{uv} . This proves Eq. (3.2). The purity in ensemble $\mathcal{A}$ takes the values $1/4 + 4u^2$ and $1/4 + 4v^2$, whereas it is constant in $\mathcal{B}$. The variance of two equally probable

numbers is one quarter of their squared difference, giving Eq. (3.3). The corresponding entropies are $2f(u)$ and $2f(v)$ in $\mathcal{A}$ and $f(u) + f(v)$ in $\mathcal{B}$. Finally, $f'(t) = \log[(c-t)/(c+t)] < 0$ for $0 < t < c$, so the entropy variance in $\mathcal{A}$ is strictly positive. $\square$

For example, $u = 0$ and $v = 1/8$ give the common mean purity $9/32$ and common mean entropy

$$\mathbb{E}S = \frac{5}{2} \log 2 - \frac{3}{8} \log 3, \quad \text{Var}_{\mathcal{A}}(\text{Tr}\rho^2) = \frac{1}{1024}, \quad \text{Var}_{\mathcal{A}}S = \frac{1}{4} \left(\frac{3}{4} \log 3 - \log 2 \right)^2. \quad (3.5)$$

Both variances vanish in $\mathcal{B}$. Since the averaged density matrix and mean entropy also agree, the quantity $S(\mathbb{E}\rho) - \mathbb{E}S(\rho)$ is identical in these two ensembles. That concavity gap therefore does not determine the entropy variance either.

This comparison is stronger than matching finitely many moments or a limiting density: it matches all averaged one-point spectral data at finite dimension. An exact mean von Neumann entropy and every averaged replica trace still leave fluctuation data undetermined. Products of traces, rather than additional single-trace orders alone, detect the difference. In particular, the annealed Rényi entropies agree in these ensembles, while their averaged Rényi entropies can differ because the logarithm is applied before averaging. The distinction and its relation to Eq. (3.4) are made precise in Section 7.

4 The normalized random-state benchmark

A useful comparison between microscopic models requires a precisely specified reference ensemble. We use the reduced state of a complex Haar-random pure state. Its finite-dimensional moments, limiting Marchenko–Pastur law, and entropy are standard results [14, 15, 21, 22, 28]. They also describe the equal-weight, fixed-dimension limit of the replica model studied by Penington et al. [25]. The role of this benchmark is to identify which spectral data coincide under specified assumptions. It does not assert that every gravitational replica calculation belongs to the same random-matrix ensemble.

4.1 Exact normalization and finite-dimensional moments

Let $\mathcal{H}_R = \mathbb{C}^m$ and $\mathcal{H}_B = \mathbb{C}^b$, with $m = d_R$, $b = d_B$, and $\gamma = m/b$. For a normalized Haar-random vector $|\psi\rangle \in \mathcal{H}_R \otimes \mathcal{H}_B$, set $\rho_R = \text{Tr}_B |\psi\rangle\langle\psi|$. Equivalently, if G is an $m \times b$ matrix of independent standard complex Gaussian entries, then

$$\rho_R = \frac{GG^\dagger}{\text{Tr}(GG^\dagger)}. \quad (4.1)$$

Normalization is imposed for each realization. Replacing the denominator by its expectation gives a different finite-dimensional random matrix. Although this replacement preserves the limiting density, it need not preserve the leading coefficient of a connected fluctuation.

Write $c(\pi)$ for the number of cycles of a permutation, let $\tau_k = (1\,2\cdots k)$, and define the rising factorial $(D)_k = D(D+1)\cdots(D+k-1)$ with $D = mb$. The exact moment is

$$\mathbb{E}_H \text{Tr}\rho_R^k = \frac{1}{(mb)_k} \sum_{\pi \in S_k} b^{c(\pi)} m^{c(\tau_k \pi)}. \quad (4.2)$$

Appendix B derives this expression from the symmetric projector and gives the corresponding formula for products of traces. In particular,

$$\mathbb{E}_H \text{Tr}\rho_R = 1, \quad \mathbb{E}_H \text{Tr}\rho_R^2 = \frac{m+b}{mb+1}. \quad (4.3)$$

The second identity is Lubkin's mean-purity formula [15]. Both identities hold at arbitrary positive integer dimensions, including $m = 1$ or $b = 1$. They provide simple normalization checks before any large-dimension expansion is made.

For fixed k and $m, b \rightarrow \infty$ with $m/b \rightarrow \gamma \in (0, \infty)$, Eq. (4.2) becomes

$$m^{k-1} \mathbb{E}_H \text{Tr} \rho_R^k \rightarrow M_k(\gamma) = \sum_{j=1}^k \frac{1}{k} \binom{k}{j} \binom{k}{j-1} \gamma^{k-j}. \quad (4.4)$$

The power multiplying the physical moment is m^{k-1} , because the spectral measure below describes $m\lambda$, where λ is an eigenvalue of ρ_R . Thus $M_2 = 1 + \gamma$ corresponds to the physical purity $m^{-1} + b^{-1}$ at leading order. The finite-dimensional formula and its limiting form should be distinguished throughout.

The leading terms in the permutation sum satisfy $c(\pi) + c(\tau_k \pi) = k + 1$ and are indexed by non-crossing partitions. Their Narayana multiplicities give Eq. (4.4). Keeping only the identity and the inverse full cycle, for $k \geq 2$, would instead give

$$(\mathbb{E} \text{Tr} \rho_R^k)_{\text{two terms}} \sim m^{1-k} + b^{1-k} = m^{1-k}(1 + \gamma^{k-1}). \quad (4.5)$$

At $\gamma = 1$, the full limiting moment is the Catalan number $C_k = (k+1)^{-1} \binom{2k}{k}$. The ratio of the complete leading sum to the two retained terms is $C_k/2$, equal to 1, 5/2, 7 for $k = 2, 3, 4$. Consequently, the two terms reproduce purity at leading order but miss the higher-moment transition even within this benchmark. These are standard non-crossing-partition facts concerning the stated random-state benchmark.

4.2 The limiting density and entropy

For eigenvalues λ_i of ρ_R , define the empirical probability measure $\nu_m = m^{-1} \sum_{i=1}^m \delta_{m\lambda_i}$. Its limiting Marchenko–Pastur measure is

$$d\mu_\gamma(x) = \frac{\sqrt{(x_+ - x)(x - x_-)}}{2\pi\gamma x} \mathbf{1}_{[x_-, x_+]}(x) dx + (1 - \gamma^{-1})_+ \delta_0(dx), \quad x_\pm = (1 \pm \sqrt{\gamma})^2. \quad (4.6)$$

It has total mass one and first moment one. The atom for $\gamma > 1$ accounts for the $m - b$ zero eigenvalues of a rank- b radiation state. The density is a law for the *scaled* eigenvalues x , rather than for the eigenvalues λ themselves.

The entropy can be expressed without an analytic continuation of an integer-index polynomial:

$$S(\rho_R) = \log m - \int x \log x d\nu_m(x), \quad 0 \log 0 := 0. \quad (4.7)$$

For $0 < \gamma \leq 1$, the integral against μ_γ is $\gamma/2$. The exact mean entropy, conjectured by Page and subsequently proved [22, 28], is

$$\mathbb{E}_H S(\rho_R) = H_{mb} - H_b - \frac{m-1}{2b} = \log m - \frac{\gamma}{2} + o(1), \quad m \leq b, \quad (4.8)$$

where $H_j = \sum_{r=1}^j r^{-1}$. For $m > b$, exchange m and b , since the two reduced states of a pure bipartite state have the same nonzero spectrum. In terms of the dimension detuning $\delta = \log b - \log m$, the leading expression on either side is

$$\mathbb{E}_H S(\rho_R) = \min(\log m, \log b) - \frac{1}{2} e^{-|\delta|} + o(1). \quad (4.9)$$

Here δ parametrizes a family of bipartite ensembles. A physical time-dependent Page curve additionally requires a rule for the evolution of the subsystem dimensions and states; Eq. (4.9) alone does not provide that rule.

4.3 Which ingredients are universal?

The free moment–cumulant relation is

$$M_k = \sum_{\pi \in NC(k)} \prod_{V \in \pi} r_{|V|}. \quad (4.10)$$

For the measure in Eq. (4.6), $r_j = \gamma^{j-1}$ and hence $R(z) = \sum_{j \geq 1} r_j z^{j-1} = (1 - \gamma z)^{-1}$ [21, 29]. The two inputs are the non-crossing combinatorics and the particular cumulant weights. Planarity by itself fixes neither those weights nor freeness of an arbitrary family of operators. For example, a semicircular variable has only its second free cumulant nonzero and has a different planar moment sequence. Energy-dependent weights and constraints in a radiation model likewise require their own spectral calculation.

The relevant lesson of Paper I [32] is therefore conditional: in the equal-weight fixed-dimension limit, the complete planar calculation has the same leading spectral data as this normalized random-state benchmark. The thermally weighted planar spectrum treated there requires further input; it cannot in general be replaced by a Gaussian average of the fixed-dimension entropy at order one. Section 5 treats the thermodynamic assumptions separately. Neither a genus expansion nor a matching mean heat capacity alone identifies the complete radiation ensemble.

4.4 What a matching density establishes

A limiting density determines its integrable one-point spectral statistics. It does not determine the joint distribution of eigenvalues, their finite-dimensional covariance, the eigenvectors, or correlations with an external reference. In particular, two ensembles may have the same limiting moments and entropy while assigning different probabilities to deviations from those values. Section 6 constructs such a pair explicitly, using the density in Eq. (4.6). This makes the limitation of the Page benchmark a statement about specified mathematical objects, without assuming that a named microscopic proposal realizes either ensemble.

5 Thermodynamic fluctuations and the observable being averaged

A Page transition involves a comparison of entropies. The distribution of that comparison depends on both the preparation and the observable. A heat capacity supplies useful information about energy fluctuations in a specified equilibrium ensemble, but does not specify an arbitrary entanglement-entropy crossover. This section separates these ingredients and explains the role of the results of Papers I and II [32, 33]. All entropies are dimensionless, with Boltzmann’s constant set to one.

5.1 Averaging two competing branches

Let F_1 and F_2 be integrable random entropy branches. Write $F_i = \overline{F}_i + \eta_i$, where $\mathbb{E}\eta_i = 0$, and define

$$\Delta = \overline{F}_1 - \overline{F}_2, \quad \xi = \eta_1 - \eta_2. \quad (5.1)$$

The exact algebraic identity

$$\mathbb{E} \min(F_1, F_2) = \overline{F}_2 + \mathbb{E} \min(\Delta + \xi, 0) \quad (5.2)$$

does not require independence between the branches. In a model in which an entropy is approximated by the minimum of two branches, it specifies the correction caused by fluctuations. The identification

of those branches with generalized entropies in gravity requires a separate state and replica calculation; it is not part of the algebraic identity. Fixed-area and chaotic-eigenstate treatments provide concrete settings for such calculations [18, 20].

The deficit relative to the minimum of the means is

$$B_\xi(\Delta) = \min(\Delta, 0) - \mathbb{E} \min(\Delta + \xi, 0) = \frac{\mathbb{E}|\Delta + \xi| - |\Delta|}{2} \geq 0. \quad (5.3)$$

The last inequality follows from convexity of the absolute value. At the crossing, $B_\xi(0) = \mathbb{E}|\xi|/2$. In particular, the width measured by $\sqrt{\text{Var}\xi}$ and the central deficit measure different features of the law. For a Gaussian difference of standard deviation σ ,

$$B_G(\Delta) = \sigma\phi(\Delta/\sigma) - |\Delta|\Phi(-|\Delta|/\sigma), \quad B_G(0) = \frac{\sigma}{\sqrt{2\pi}}, \quad (5.4)$$

where ϕ and Φ are the standard normal density and distribution function. The zero-width limit is taken separately and gives $B = 0$.

There is also a positive identification result within this model. Hold the law of ξ fixed while varying Δ and put $h(\Delta) = \mathbb{E} \min(\Delta + \xi, 0)$. If ξ has density f_ξ , then

$$h'(\Delta) = \mathbb{P}(\xi < -\Delta), \quad h''(\Delta) = -f_\xi(-\Delta). \quad (5.5)$$

The first equality follows by differentiating a Lipschitz integrand; the second follows wherever the density is continuous. With atoms, one-sided derivatives give the corresponding distribution function. Thus an exact crossover as a function of a calibrated entropy difference can determine this one-dimensional fluctuation law. This observation also specifies the limits of the inference: it identifies neither the joint law of F_1, F_2 nor the channel that prepares them. If the law changes with Δ , its variation contributes to the derivative, and (5.5) is no longer an inversion formula for a single distribution.

5.2 Equal widths and different crossover shapes

Consider a standard normal variable Y_G and the unit-variance quartic variable $Y_Q = X/\sqrt{v}$, where

$$p_X(x) = \frac{2}{\Gamma(1/4)} e^{-x^4}, \quad v = \frac{\Gamma(3/4)}{\Gamma(1/4)}. \quad (5.6)$$

Both variables have mean zero and variance one. Their central deficits are nevertheless different:

$$B_{Y_G}(0) = \frac{1}{\sqrt{2\pi}} \simeq 0.398942, \quad B_{Y_Q}(0) = 2^{-5/4} \simeq 0.420448. \quad (5.7)$$

The quartic value follows from the gamma-function moments in appendix C. Their curvature profiles also differ, because they reproduce different densities in (5.5). Figure 1 compares the two complete functions without fitting either shape.

Even the variance of a branch difference requires joint information:

$$\text{Var}\xi = \text{Var}\eta_1 + \text{Var}\eta_2 - 2\text{Cov}(\eta_1, \eta_2). \quad (5.8)$$

For two centered Gaussian branches, each with variance s^2 and correlation c , this gives $\text{Var}\xi = 2s^2(1-c)$. With the same marginal branch laws, the crossover width therefore ranges from zero at $c = 1$ to $2s$ at $c = -1$. These endpoints describe degenerate but well-defined joint Gaussian laws. A common fluctuating contribution cancels from the difference, whereas an anticorrelated contribution is amplified.

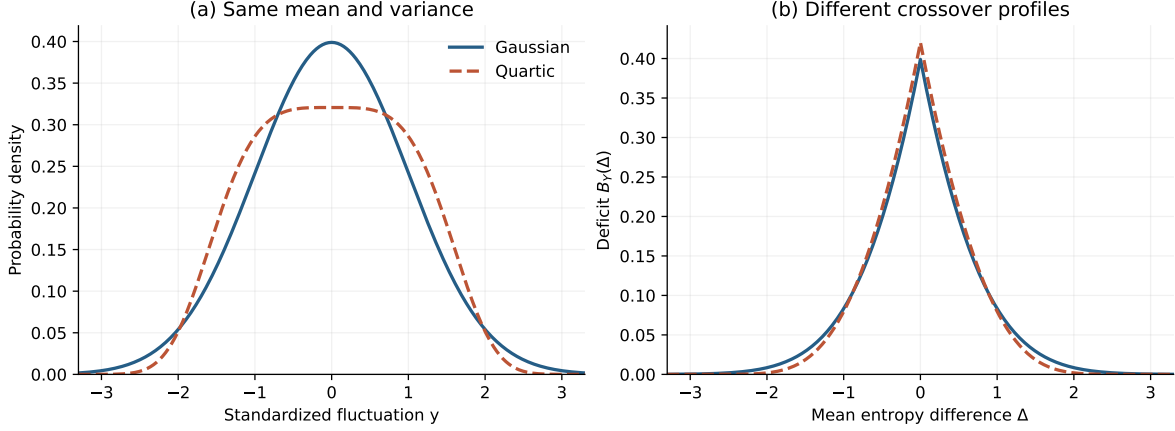


Figure 1. Equal variance does not determine the averaged-minimum crossover. Left: standard normal and standardized quartic densities, each with mean zero and variance one. Right: their deficits $B_Y(\Delta)$ from (5.3), evaluated by deterministic quadrature. The central values are given in (5.7). These are probability models for a fluctuating entropy difference; no evaporation trajectory is assumed.

5.3 Regular phases, finite reservoirs, and criticality

For a temperature-independent density of states, the exact canonical identity is $\text{Var}E = T^2 C_{\text{can}}$, where $C_{\text{can}} = \partial_T \mathbb{E}E$. At a stable, sharply concentrated saddle, C_{can} agrees to leading order with the saddle heat capacity $C_B > 0$. Linearizing a thermodynamic entropy using $dS_B/dE = 1/T$ then gives $\text{Var}S_B \simeq C_B$ when higher-order fluctuations are negligible. This concerns the fluctuating thermodynamic function $S_B(E)$. The capacity of entanglement of a normalized state is instead

$$C_{\text{ent}}(\rho) = \text{Tr}\rho(\log \rho)^2 - (\text{Tr}\rho \log \rho)^2. \quad (5.9)$$

For a Gibbs state it equals $\beta^2 \text{Var}E$; for a general radiation state it need not equal a black-hole heat capacity. It also differs from the ensemble variance of $S(\rho)$ studied in section 7.

Fixing the energy of a composite system permits exchange between its subsystems. In a density-of-states approximation,

$$p(E_B) \propto \exp\{S_B(E_B) + S_R(E_{\text{tot}} - E_B)\}. \quad (5.10)$$

At a common-temperature stationary point, the Gaussian variance is $T^2 C_{\text{eff}}$, with

$$C_{\text{eff}} = (C_B^{-1} + C_R^{-1})^{-1}, \quad C_B^{-1} + C_R^{-1} > 0. \quad (5.11)$$

The inequality is the local stability condition. The linear entropy variations are $\delta S_B = \delta E_B/T$ and $\delta S_R = -\delta E_B/T$, so that $\text{Var}(S_B - S_R) \simeq 4C_{\text{eff}}$. These relations underlie enhanced entanglement corrections in energy-constrained chaotic systems [20]. For a negative C_B and positive C_R , local stability is possible if $C_R < |C_B|$. Global phase dominance and an actual physical confinement mechanism are additional questions. Fixing E_B itself removes this classical exchange fluctuation; fixing E_{tot} does not.

Paper II gives a non-Gaussian example in its specified RN–AdS thermodynamic density model [33]. In $d+1 \geq 4$ bulk dimensions at the charged critical point, set $N = S_c$ and

$$a_d = \frac{(d-1)(2d-3)}{24}, \quad L_N = (d-1)a_d^{-1/4}N^{3/4}. \quad (5.12)$$

The centered black-hole entropy divided by L_N tends to the quartic law in (5.6). A regular finite reservoir changes this to

$$p_u(x) = Z_u^{-1} e^{-x^4 - ux^2}, \quad u = \lim_{N \rightarrow \infty} \frac{L_N^2}{2C_R}. \quad (5.13)$$

The result assumes the local reservoir expansion and the tail and moment control stated in appendix C. A reservoir of capacity much larger than $N^{3/2}$ retains the canonical quartic law. Capacities of order $N^{3/2}$ change its shape. A diverging but smaller regular reservoir produces a Gaussian law on its own narrower scale. These are different preparations of the same critical black-hole family.

To transfer such a law to generalized entropies $F_{1,N}, F_{2,N}$, one must also control their response to the critical mode. For example, if

$$\frac{(F_{1,N} - F_{2,N}) - \mathbb{E}(F_{1,N} - F_{2,N})}{L_N} = aX_N + r_N, \quad \mathbb{E}r_N^2 \rightarrow 0, \quad (5.14)$$

with X_N converging in distribution and second moment to p_u , then the limiting variance is $a^2 \text{Var}_{p_u} X$ and the averaged-minimum profile follows from the law of aX . This is a conditional transfer statement. If $a = 0$, this mode contributes no leading width on the L_N scale. A thermodynamic critical point alone does not determine a .

The charged critical point must also be distinguished from a first-order transition and a metastable spinodal. At the neutral AdS temperature minimum, the Gaussian saddle expansion fails as two black-hole branches meet. The Hawking–Page transition instead compares the free energies of thermal AdS and a black hole [12]. These events occur at different radii and support different fluctuation analyses.

5.4 Averaging sector entropies and combining spectra

There is a further distinction between a classical mixture of entropy values and the spectrum of the resulting quantum state. If radiation resolves mutually orthogonal sectors,

$$\rho_R = \bigoplus_{\alpha} p_{\alpha} \rho_{R,\alpha}, \quad S(\rho_R) = H(p) + \sum_{\alpha} p_{\alpha} S(\rho_{R,\alpha}), \quad (5.15)$$

where $H(p) = -\sum_{\alpha} p_{\alpha} \log p_{\alpha}$. The sector entropy is part of the answer. The identity follows by substituting the eigenvalues $p_{\alpha} \lambda_{\alpha j}$ and does not require random matrices. A fixed total energy can generate this structure because different radiation energies occupy orthogonal subspaces.

The canonical random-state model in Paper I has a different spectral structure: black-hole levels contribute to a common radiation space [25, 31, 32]. Let x denote a normalized black-hole level weight, sampled with that weight, and let m be the radiation dimension. Set $K = -\log x$. Its level-resolved diagonal approximation is

$$S_{\text{diag}} = \mathbb{E}_x \left[\min(\log m, K) - \frac{1}{2} e^{-|\log m - K|} \right]. \quad (5.16)$$

The planar spectrum combines these contributions before computing the entropy. In the broad-distribution limit established in Paper I,

$$S_{\text{planar}} - S_{\text{diag}} \rightarrow -q \log q, \quad q = \lim \mathbb{P}(K > \log m) \in (0, 1). \quad (5.17)$$

Here broad distribution means $\mathbb{P}(|K - \log m| \leq M) \rightarrow 0$ for every fixed M , in the stated planar model. At a symmetric crossing $q = 1/2$, the difference is $(\log 2)/2$. This imported result concerns

radiation that does not resolve the level sector. It is not an annealed–quenched difference, and cannot be transferred to the direct sum in (5.15). Nor is it automatically the complete order-one gravitational correction after bulk entropy and other saddle fluctuations are included.

These examples specify several observables that can differ even when coarse thermodynamic data agree. They also show why comparisons should state the preparation, the branch difference, and the operation used to combine sectors before assigning a crossover to a microscopic theory.

6 Design matching and shared spectral limits

The distinction between a spectral limit and an ensemble of states can be made explicit. We first state exactly what a finite design condition implies. The exact comparison in section 3 already separates averaged spectral data from fluctuations. We now place that distinction within a common limiting Marchenko–Pastur density and order-one entropy correction. This comparison concerns normalized bipartite quantum states. It supplies a counterexample to inference from the limiting density, rather than a classification of candidate quantum-gravity theories.

6.1 Finite designs determine finitely many moments

Let $\mathcal{E}$ be a probability distribution on normalized pure states of $\mathbb{C}^m \otimes \mathbb{C}^b$. Its k -copy moment operator is

$$\Omega_{\mathcal{E}}^{(k)} = \mathbb{E}_{\mathcal{E}}(|\psi\rangle\langle\psi|)^{\otimes k}. \quad (6.1)$$

An exact projective state t -design has $\Omega_{\mathcal{E}}^{(k)} = \Omega_{\text{H}}^{(k)}$ for $k \leq t$. An exact unitary t -design acting on any fixed pure input produces such a state design. For an approximate comparison, we specify the norm directly: throughout this subsection, the error at order k is

$$\epsilon_k = \|\Omega_{\mathcal{E}}^{(k)} - \Omega_{\text{H}}^{(k)}\|_1. \quad (6.2)$$

This definition avoids identifying distinct conventions for approximate unitary designs. A circuit theorem stated in another norm must be converted to Eq. (6.2) before using the following bound.

Proposition 6.1 (Reduced-state moment comparison). *For each positive integer k , the ensemble in Eq. (6.1) satisfies*

$$|\mathbb{E}_{\mathcal{E}} \text{Tr} \rho_R^k - \mathbb{E}_{\text{H}} \text{Tr} \rho_R^k| \leq \epsilon_k. \quad (6.3)$$

Consequently the difference of the scaled moments is bounded by $m^{k-1}\epsilon_k$. More generally, the difference of expectations of $\prod_{a=1}^r \text{Tr} \rho_R^{k_a}$ is at most ϵ_K , where $K = \sum_a k_a$. An exact state t -design therefore reproduces these Haar expectations whenever $K \leq t$.

Proof. Let $T_{R,k}$ cyclically permute the k copies of $\mathcal{H}_R$ and act trivially on the B factors. The permutation trace identity gives

$$\text{Tr} \rho_R^k = \text{Tr}[(|\psi\rangle\langle\psi|)^{\otimes k} T_{R,k}]. \quad (6.4)$$

Since $T_{R,k}$ is unitary, its operator norm equals one. Trace-norm and operator-norm duality then give Eq. (6.3). For a product of traces, replace the single cycle by the permutation with disjoint cycles of lengths $k_1, \dots, k_r$; its norm is also one. $\square$

The proposition is an elementary consequence of moment matching, consistent with the general theory of unitary integration and designs [6, 8]. It isolates two limitations that matter in applications. First, a fixed design order controls only a fixed list of moments. It gives no equality of the full spectrum or of an entire time-dependent entropy process. Second, a small unscaled error can become appreciable after multiplication by m^{k-1} . To reproduce the k th Marchenko–Pastur moment in a large-dimension limit, a sufficient condition is $m^{k-1}\epsilon_k \rightarrow 0$, in addition to the Haar limit in Eq. (4.4). This requirement must be checked at the dimensions and orders being used.

Products of traces also expose a difference between an average and a fluctuation. Purity is determined by a two-copy moment, whereas its variance requires the four-copy moment. More explicitly, because $0 \leq \text{Tr} \rho_R^2 \leq 1$,

$$|\text{Var}_{\mathcal{E}}(\text{Tr} \rho_R^2) - \text{Var}_{\text{H}}(\text{Tr} \rho_R^2)| \leq \epsilon_4 + 2\epsilon_2. \quad (6.5)$$

To resolve the Haar variance of order m^{-4} at fixed aspect ratio, this bound is informative only when its right-hand side is smaller than that scale. A statement that an ensemble is approximately Haar without its order, norm, and error is therefore insufficient for a fluctuation comparison.

The condition is also distinct from a statement about scrambling time. The local-circuit results of Brandão, Harrow, and Horodecki [6] have explicit circuit assumptions and dependence on the design order. The spread of an initially local perturbation does not by itself establish arbitrary-order moment matching. Moreover, matching the output distribution for one chosen input does not determine a quantum channel on arbitrary inputs entangled with a reference. The operational distinction developed in Section 2 consequently remains even if a particularly accurate spectral design has been established.

6.2 From moments to entropy: the additional assumptions

The passage from moments to entropy is possible under explicit conditions; it is not a consequence of finitely many moment equalities. Suppose a sequence of empirical scaled spectral measures ν_m converges weakly in probability to μ_γ , and that

$$\sup_m \mathbb{E} \int x^2 d\nu_m(x) < \infty. \quad (6.6)$$

Then

$$S(\rho_R) - \log m \longrightarrow - \int x \log x d\mu_\gamma(x) \quad (6.7)$$

in probability and in mean. Indeed, $x \log x$ extends continuously to zero and its negative part is bounded. For $K > e$,

$$\mathbb{E} \int_{x>K} x \log x d\nu_m(x) \leq \frac{\log K}{K} \mathbb{E} \int x^2 d\nu_m(x). \quad (6.8)$$

Truncation at K , weak convergence, and then $K \rightarrow \infty$ prove the claim. This is a useful sufficient condition because the normalized Haar ensemble has $\mathbb{E} \int x^2 d\nu_m = m(m+b)/(mb+1)$, bounded at fixed aspect ratio. More generally, a moment argument for weak convergence requires arbitrarily high fixed orders as the dimension increases, and concentration or another control of the random measure. A finite set of matching averaged moments does not provide these assumptions.

6.3 A fixed-spectrum ensemble with the same limiting density

Fix $\gamma \in (0, 1]$ and choose integers $b_m \geq m$ with $m/b_m \rightarrow \gamma$. Let F_γ be the cumulative distribution function of the continuous probability measure in Eq. (4.6). Set

$$q_{i,m} = F_\gamma^{-1}\left(\frac{i-1/2}{m}\right), \quad \bar{q}_m = \frac{1}{m} \sum_{i=1}^m q_{i,m}, \quad \lambda_{i,m}^{\text{F}} = \frac{q_{i,m}}{m\bar{q}_m}. \quad (6.9)$$

All eigenvalues are nonnegative and sum to one. Define two ensembles on the same radiation Hilbert space:

$$\rho_R^F = U \text{diag}(\lambda_{1,m}^F, \dots, \lambda_{m,m}^F) U^\dagger, \quad \rho_R^W = \frac{GG^\dagger}{\text{Tr}(GG^\dagger)}. \quad (6.10)$$

Here U is Haar-distributed on $U(m)$ and G is an $m \times b_m$ complex Gaussian matrix. The superscripts distinguish a fixed spectrum with random eigenvectors from the normalized Wishart spectrum. Both ensembles are invariant under radiation-basis rotations and satisfy $\mathbb{E}\rho_R = I_m/m$. Both can be realized as reductions of normalized pure states on $\mathbb{C}^m \otimes \mathbb{C}^{b_m}$. For the first, an explicit purification is $\sum_i \sqrt{\lambda_{i,m}^F} U|i\rangle_R |i\rangle_B$. Thus the comparison does not rely on nonpositive matrices or on normalization only after averaging.

Proposition 6.2 (Equal limiting density, unequal purity fluctuations). *For the ensembles in Eq. (6.10), the empirical measures of $m\lambda$ converge to μ_γ , deterministically for F and in probability for W. Their entropy deficits both tend to $\gamma/2$:*

$$\log m - S(\rho_R^F) \rightarrow \frac{\gamma}{2}, \quad \log m - S(\rho_R^W) \rightarrow \frac{\gamma}{2} \quad \text{in probability and in mean.} \quad (6.11)$$

Nevertheless, for $m, b_m > 1$,

$$\text{Var}_F(\text{Tr}\rho_R^2) = 0, \quad (6.12)$$

$$\text{Var}_W(\text{Tr}\rho_R^2) = \frac{2(m^2 - 1)(b_m^2 - 1)}{(mb_m + 1)^2(mb_m + 2)(mb_m + 3)} > 0. \quad (6.13)$$

In particular, $m^4 \text{Var}_W(\text{Tr}\rho_R^2) \rightarrow 2\gamma^2$, whereas the corresponding fixed-spectrum quantity vanishes at every dimension.

Proof. Midpoint quantiles form a Riemann approximation to the compactly supported measure μ_γ . Its first moment is one, so $\bar{q}_m \rightarrow 1$ and normalization in Eq. (6.9) does not change the limit. The normalized Wishart convergence is the standard Marchenko–Pastur limit. Equation (6.7) applies to Wishart; for the quantile ensemble, the scaled eigenvalues are uniformly bounded for all sufficiently large m , which supplies the same entropy convergence directly. The integral is $\gamma/2$, as evaluated in Appendix B. Every spectral statistic is constant under conjugation in ensemble F, proving Eq. (6.12). Equation (6.13) is the exact Haar purity variance [10]; its derivation from the normalized four-copy moment is reproduced in Appendix D. $\square$

The equalities here concern the limiting density and order-one entropy deficit. They do not assert equality of the two mean purities or mean entropies at finite dimension. Figure 2 shows this distinction: the entropy deficits approach their common limit while the appropriately rescaled purity variance remains different. The comparison establishes that the limiting density cannot determine fluctuation data. It makes no claim about a universal spacing law in every microscopic model, and it does not identify either ensemble with a complete model of black-hole evaporation.

6.4 Information absent even from the complete spectrum

At finite dimension, knowledge of every power trace determines the eigenvalues including multiplicities. It still leaves the eigenvectors undetermined. For example, if Λ is a nondegenerate diagonal state and V a unitary that mixes its eigenvectors, then Λ and $V\Lambda V^\dagger$ have identical entropies of every Rényi order. Their expectation of a specified radiation observable O need not agree:

$$\text{Tr}(O\Lambda) \neq \text{Tr}(OV\Lambda V^\dagger) \quad (6.14)$$

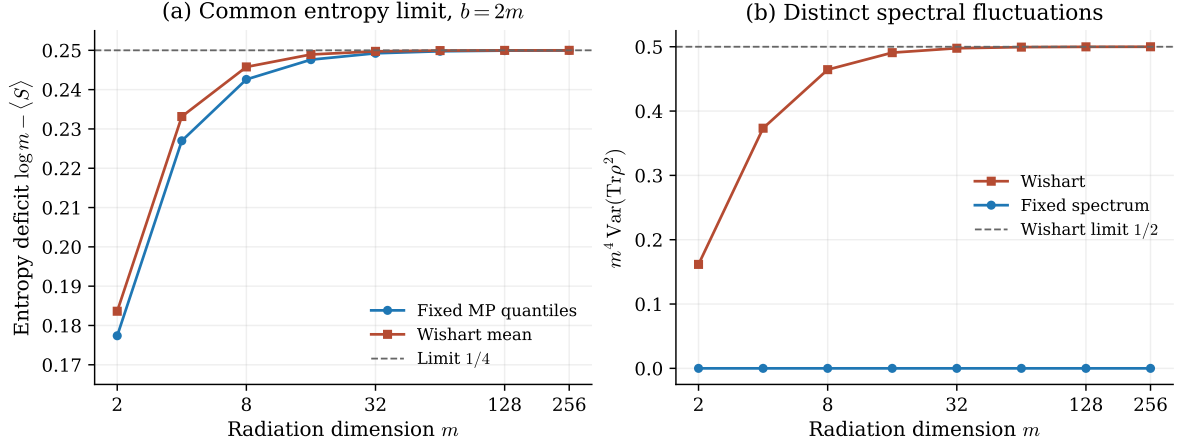


Figure 2. A common limiting entropy and different purity fluctuations. Left: the entropy deficit $\log m - S$ of the deterministic Marchenko–Pastur quantile spectrum and the exact mean deficit of the normalized Wishart ensemble, at $b = 2m$. Both tend to $1/4$. Right: the exact quantity $m^4 \text{Var}(\text{Tr} \rho^2)$ tends to $1/2$ for Wishart and is zero for the fixed spectrum. All curves are computed from the displayed quantiles and exact formulas; no fitted fluctuation profile is used.

for a suitable O . With $m = 2$, $\Lambda = \text{diag}(p, 1-p)$, $p \neq 1/2$, and V the Hadamard matrix, the expectation of the Pauli operator Z is $2p - 1$ in the first state and zero in the second. This exact example does not depend on a large-entropy approximation.

The choice of observable presupposes a physical basis or algebra of radiation measurements. Basis rotations are irrelevant when the task is only to classify spectra, but they become relevant once couplings, spatially localized operators, or a decoding map have been specified. Even an exact entanglement spectrum is therefore an incomplete specification of the radiation state. Specifying a state at one time, in turn, does not specify its dynamical map or its correlations with other systems. These are separate losses of information, and higher spectral resolution repairs only the first of them.

6.5 A conditional organization of observables

The preceding results suggest organizing comparisons by the data actually being matched. Table 1 lists the additional inputs needed at each stage. This is an organization of questions, rather than a theorem that all microscopic proposals must agree until some predetermined order in entropy. Two models can already disagree in their leading density if their constraints or statistical weights differ. Conversely, their fluctuation statistics can agree within a common universality class, so observing a variance does not automatically identify a unique mechanism.

Thermodynamic averaging introduces a further specification: the joint law of the fluctuating entropies or energies and their relation to the radiation state. Equality of dimensions or heat capacities does not supply that law. Section 5 describes this dependence, while Section 7 treats the distinction between averaged replica moments, spectral fluctuations, and the entropy of an averaged density matrix. Together with the explicit comparison above, these distinctions state the limitations of Page-curve matching without attributing uncomputed predictions to individual proposals.

Table 1. Information supplied by different comparisons. Each row requires its own assumptions; agreement in one row does not establish agreement in the next.

Data compared	What can be inferred	Information still required
Entropy at each time	Agreement of one scalar statistic	Spectrum, reference correlations, and channel action
Limiting scaled spectrum	Leading integrable spectral statistics	Finite-dimensional corrections and spectral covariance
Connected spectral moments	Fluctuations in a specified ensemble	Eigenvectors and realization of that ensemble
Selected radiation–reference correlators	The specified input-dependent correlations	Channel behavior beyond these correlators

7 Rényi averages and entropy fluctuations

An ensemble-averaged Page curve specifies a mean entropy. It does not specify how much the entropy varies between realizations, or how the radiation states themselves vary. These questions require different observables. We first distinguish the two customary orders of averaging Rényi entropies, then relate their difference to a normalized replica covariance. The identities are elementary at finite dimension, but their normalization and order of limits are essential when applying them to a gravitational replica calculation.

Throughout this section, ρ is a random density matrix on a fixed d -dimensional Hilbert space, normalized separately in every realization: $\rho \geq 0$ and $\text{tr } \rho = 1$ almost surely. The probability measure may describe random states, couplings, circuits, or another specified preparation. Its mathematical existence does not determine which of these physical interpretations is appropriate. Logarithms are natural, and $0 \log 0 = 0$.

7.1 The two Rényi averages

For real $n > 1$, define the moment $M_n(\rho) = \text{tr } \rho^n$ and the annealed and quenched Rényi entropies by

$$A_n = \frac{\log \mathbb{E} M_n}{1 - n}, \quad (7.1)$$

$$Q_n = \frac{\mathbb{E} \log M_n}{1 - n} = \mathbb{E} S_n(\rho). \quad (7.2)$$

The labels refer only to the position of the logarithm relative to the average. In particular, Q_n is still an ensemble average, rather than the entropy of a distinguished member. Concavity of the logarithm gives

$$G_n := Q_n - A_n = \frac{\log \mathbb{E} M_n - \mathbb{E} \log M_n}{n - 1} \geq 0. \quad (7.3)$$

At a fixed order $n > 1$, equality holds precisely when M_n is constant almost surely. Constancy of one moment neither fixes the spectrum nor identifies a Hamiltonian. Even constancy of every spectral moment leaves the eigenvectors undetermined.

Proposition 7.1 (Von Neumann limit and its first correction). *For any ensemble of normalized density matrices on a common finite-dimensional Hilbert space,*

$$\lim_{n \rightarrow 1} A_n = \lim_{n \rightarrow 1} Q_n = \mathbb{E}S(\rho), \quad S(\rho) = -\text{tr}(\rho \log \rho). \quad (7.4)$$

Moreover, as $t \rightarrow 0$,

$$G_{1+t} = \frac{t}{2} \text{Var}[S(\rho)] + O(t^2). \quad (7.5)$$

The remainder is bounded uniformly over ensembles at fixed dimension.

Proof. Set $L_2(\rho) = \text{tr}[\rho(\log \rho)^2]$. Expansion of the normalized moment gives

$$M_{1+t} = 1 - tS + \frac{t^2}{2} L_2 + O(t^3), \quad \log M_{1+t} = -tS + \frac{t^2}{2} (L_2 - S^2) + O(t^3). \quad (7.6)$$

For n in a compact interval contained in $(0, \infty)$, the functions $x^n |\log x|^r$, extended by zero at $x = 0$, are uniformly bounded on $[0, 1]$. Summing over at most d eigenvalues bounds the required moment derivatives uniformly, including for rank-deficient states. The moment also has a positive lower bound near $n = 1$, so the logarithm has uniformly bounded derivatives there. These facts justify averaging the Taylor expansions. The linear coefficients of $\log \mathbb{E}M_{1+t}$ and $\mathbb{E} \log M_{1+t}$ are both $-\mathbb{E}S$. Their quadratic coefficients differ by $\text{Var}(S)/2$. Dividing their difference by t proves the claim. $\square$

Thus there is no nonzero annealed–quenched von Neumann entropy gap under these definitions. Entropy fluctuations survive as the slope of the finite-order Rényi gap. The statement holds for each finite dimension. A large-dimension saddle approximation followed by analytic continuation may require additional uniform estimates; such an approximation does not change the exact finite-dimensional identity. For an infinite-dimensional state space, the interchange of derivatives and averages must instead be justified by suitable integrability bounds.

7.2 Averaging a spectrum and averaging a state

Another nonnegative difference, often relevant to ensemble interpretations, is

$$\chi = S(\bar{\rho}) - \mathbb{E}S(\rho), \quad \bar{\rho} = \mathbb{E}\rho. \quad (7.7)$$

This is the Holevo quantity of the state ensemble. It compares the entropy of the averaged density operator with the mean entropy. It is not the $n \rightarrow 1$ limit of G_n . More explicitly,

$$\chi = \mathbb{E}D(\rho \| \bar{\rho}), \quad D(\rho \| \sigma) = \text{tr}[\rho(\log \rho - \log \sigma)] \geq 0, \quad (7.8)$$

where the support of ρ lies in that of $\bar{\rho}$ almost surely. Hence $\chi = 0$ precisely when the density matrices themselves agree almost surely. This is a statement about the specified state ensemble, not about whether an underlying theory is unique.

For comparison, let $\nu_\rho = \sum_{i=1}^d \delta_{\lambda_i(\rho)}$ be the eigenvalue counting measure. Entropy is a linear functional of this measure:

$$-\int_0^1 x \log x \, \mathbb{E}\nu_\rho(dx) = \mathbb{E}S(\rho). \quad (7.9)$$

Averaging the spectral density before performing the entropy integral therefore gives the mean entropy exactly. The eigenvalues of $\mathbb{E}\rho$ are generally not the averaged eigenvalue counting measure. A concavity argument for density operators cannot be transferred to this linear spectral integral.

An isospectral ensemble illustrates the distinction. Fix a density matrix ρ_0 and draw $\rho = U\rho_0U^\dagger$ with Haar-distributed U . All M_n , all Rényi entropies, and S are deterministic, so $G_n = 0$ and $\text{Var}(S) = 0$. Nevertheless $\bar{\rho} = I_d/d$ and $\chi = \log d - S(\rho_0)$, which is positive unless ρ_0 is maximally mixed. For a pure ρ_0 , $\chi = \log d$. The radiation states can therefore vary strongly while their entropies have no fluctuations at all. Conversely, a positive G_n can arise from random preparation within a fixed theory; it does not establish averaging over Hamiltonians.

7.3 An exactly evaluated two-state ensemble

Consider the qubit ensemble

$$\rho = \begin{cases} |0\rangle\langle 0|, & \text{with probability } p, \\ I_2/2, & \text{with probability } 1-p, \end{cases} \quad 0 \leq p \leq 1. \quad (7.10)$$

Its moment has the two values 1 and 2^{1-n} . All quantities in the preceding discussion can therefore be calculated without a saddle approximation:

$$Q_n = (1-p) \log 2, \quad A_n = -\frac{\log[p + (1-p)2^{1-n}]}{n-1}, \quad (7.11)$$

$$\text{Var}(S) = p(1-p)(\log 2)^2, \quad \chi = h_2\left(\frac{1+p}{2}\right) - (1-p) \log 2, \quad (7.12)$$

where $h_2(x) = -x \log x - (1-x) \log(1-x)$ is binary entropy. For $p = 1/2$, the common von Neumann limit is $\log 2/2$, whereas $\chi = 0.215762\dots$. At $n = 2$, $G_2 = \log(3/4) + \log 2/2 = 0.058892\dots$. The coefficient of $n-1$ in G_n is $(\log 2)^2/8 = 0.060057\dots$

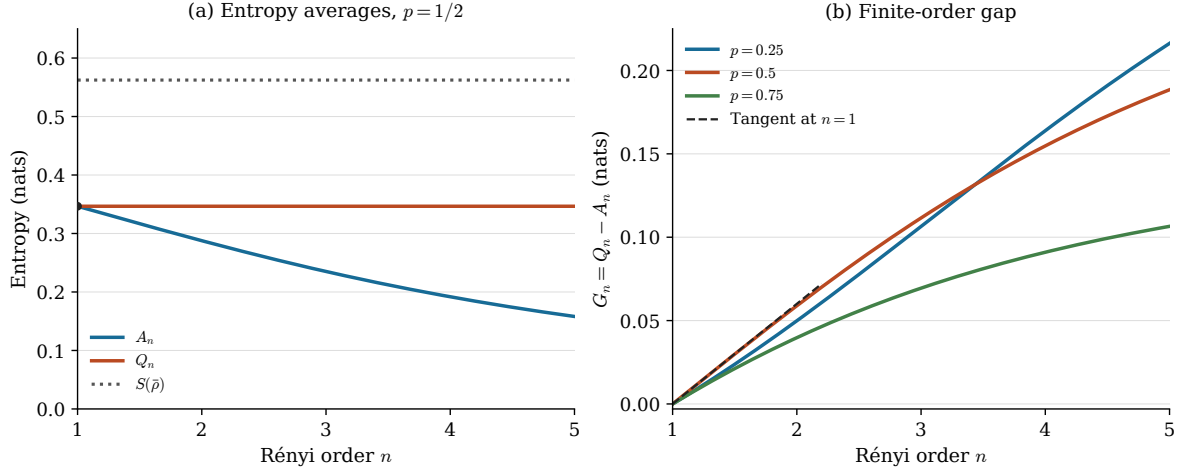


Figure 3. Exact averages for the qubit ensemble in (7.10). Left: A_n and Q_n at $p = 1/2$ approach the same von Neumann limit; the entropy $S(\bar{\rho})$ of the averaged state remains a different quantity. Right: the nonnegative gap $G_n = Q_n - A_n$ for three mixing probabilities. The dashed tangent at $p = 1/2$ is $(n-1)\text{Var}(S)/2$. Every curve is obtained from (7.11); no Page-time profile or gravitational fluctuation amplitude is assumed.

This example makes the observable distinction independent of any large-entropy limit. It also shows why the dependence of a fluctuation observable on Page time cannot be inferred from Jensen's

inequality. The probability distribution and its time dependence would have to be provided by a dynamical or state-preparation model. A positive finite Rényi gap, by itself, gives neither a universal peak position nor a universal amplitude.

7.4 Normalized replica covariances

The object that determines entropy fluctuations is the connected moment

$$C(n, \ell) = \mathbb{E}[M_n M_\ell] - \mathbb{E}M_n \mathbb{E}M_\ell. \quad (7.13)$$

Normalization immediately implies $C(1, \ell) = C(n, 1) = 0$. Differentiating with respect to the two independent replica indices gives

$$\partial_n \partial_\ell C(n, \ell)|_{n=\ell=1} = \text{Var}[S(\rho)]. \quad (7.14)$$

Both identities are exact, with differentiation justified as in Proposition 7.1. The covariance contains information absent from the one-point function $\mathbb{E}M_n$. Knowing the averaged spectral density, even exactly, does not determine it.

Proposition 3.1 gives an exact finite-dimensional example of this separation. Its two ensembles have the same $\mathbb{E}M_n$ at every positive real order, the same mean entropy, and the same averaged density matrix. Consequently all their A_n and their Holevo quantities agree. One ensemble has a fixed spectrum, whereas the other has two spectra with different entropies. The latter therefore has a nonzero entropy variance and a nonzero slope of G_n at one. Thus even complete knowledge of the one-point replica function cannot reconstruct this fluctuation observable.

For a Haar-distributed pure state on $\mathbb{C}^m \otimes \mathbb{C}^b$, let ρ be the reduction to $\mathbb{C}^m$. A useful exact comparison is the purity variance [10]:

$$\mathbb{E}M_2 = \frac{m+b}{mb+1}, \quad C(2, 2) = \frac{2(m^2-1)(b^2-1)}{(mb+1)^2(mb+2)(mb+3)}. \quad (7.15)$$

Appendix D derives this formula from normalized replica contractions. It vanishes if either subsystem is one-dimensional and, for $m = b \rightarrow \infty$, is $2m^{-4} + O(m^{-6})$. The finite normalization factors contribute at the leading order of the covariance, even though they only correct the leading mean purity at subleading order. A product of one-point approximations is therefore insufficient for computing this variance.

7.5 Concentration and its limits

For a specified order $n > 1$, concentration of M_n controls G_n . Write $M_n = \mu_n(1 + \delta_n)$, where $\mu_n = \mathbb{E}M_n$ and $\mathbb{E}\delta_n = 0$. If $|\delta_n| \leq \eta < 1$ almost surely, Taylor's theorem applied to $-\log(1 + \delta_n)$ gives the two-sided bound

$$\frac{\mathbb{E}\delta_n^2}{2(n-1)(1+\eta)^2} \leq G_n \leq \frac{\mathbb{E}\delta_n^2}{2(n-1)(1-\eta)^2}. \quad (7.16)$$

Under the same condition the quadratic approximation is

$$G_n = \frac{\text{Var}(M_n)}{2(n-1)\mu_n^2} + O\left(\frac{\mathbb{E}|\delta_n|^3}{(n-1)(1-\eta)^3}\right). \quad (7.17)$$

These estimates state an explicit condition controlling the logarithm. A small relative variance alone does not uniformly control rare events with M_n/μ_n very close to zero as the dimension increases. At

fixed dimension the lower bound $M_n \geq d^{1-n}$ supplies a bound, but its constant need not remain useful in a large-dimension limit.

Near $n = 1$, the numerator in the leading term of (7.17) itself vanishes quadratically: $\text{Var}(M_{1+t}) = t^2 \text{Var}(S) + O(t^3)$. This restores the linear behavior in (7.5); treating that numerator as independent of n would give a spurious singularity. Separately, an entropy variance bound has the direct probabilistic consequence

$$\mathbb{P}\{|S - \mathbb{E}S| \geq a\} \leq \frac{\text{Var}(S)}{a^2}, \quad a > 0. \quad (7.18)$$

This controls deviations of the entropy at the specified preparation or time. It does not imply concentration of the density operator, nor simultaneous agreement of an entire trajectory without further temporal control.

7.6 Fluctuations within and between preparations

An ensemble may contain more than one source of randomness. Let Θ denote a classical preparation parameter, such as an energy window or the number of available environment states. The law of total variance gives

$$\text{Var}(S) = \mathbb{E}[\text{Var}(S \mid \Theta)] + \text{Var}(\mathbb{E}[S \mid \Theta]). \quad (7.19)$$

The first term describes entropy fluctuations with the preparation held fixed; the second describes variation of the conditional mean between preparations. They need not have the same size or the same dependence on subsystem dimensions. Matching one of these terms between models does not match their total variance.

For example, keep the radiation dimension m fixed and draw a positive integer environment dimension B from a specified distribution. Conditional on $B = b$, draw a Haar bipartite state on $\mathbb{C}^m \otimes \mathbb{C}^b$. Every reduced state acts on the same radiation Hilbert space, so the preceding finite-dimensional identities apply even though the purifying space varies. The exact purity variance is

$$\begin{aligned} \text{Var}(M_2) = \mathbb{E} \left[\frac{2(m^2 - 1)(B^2 - 1)}{(mB + 1)^2(mB + 2)(mB + 3)} \right] \\ + \text{Var} \left(\frac{m + B}{mB + 1} \right). \end{aligned} \quad (7.20)$$

The second term is absent in a fixed-dimension comparison. This example supplies a precise probabilistic distinction between state fluctuations at fixed dimension and fluctuations of the available number of states. It does not identify an arbitrary gravitational ensemble with a mixture of Haar ensembles; that identification would require the conditional state law as well as the distribution of B .

Conditioning also clarifies how thermodynamic broadening and entropy fluctuations enter a comparison. A fluctuating energy can change the conditional mean entropy, while states within a narrow energy window can still have a nonzero conditional variance. A measured or computed total variance alone cannot separate these contributions. The preparation or conditioning must be stated before assigning a physical origin to either term. All statements here concern a classical probability distribution over density operators; replacing it by a coherent superposition is a different state construction.

7.7 Application to gravitational ensembles

When a path integral initially constructs an unnormalized positive operator W , the physical moment is

$$M_n = \frac{\text{tr } W^n}{(\text{tr } W)^n}. \quad (7.21)$$

The average of this ratio generally differs from replacing its denominator by $(\mathbb{E} \operatorname{tr} W)^n$. In a covariance calculation, fluctuations of the normalization can enter at the same order as the connected numerator. The normalization check $C(1, \ell) = 0$ must hold before using a replica derivative to infer entropy fluctuations.

Bousso and Miyaji [5] computed such fluctuations in the PSSY model, explicitly retaining the fluctuating normalization. Their calculation includes non-tubular and tubular contributions and finds a difference from fixed-dimension Haar fluctuations after the Page time, associated with fluctuations of the number of black hole states in the chosen energy window. This is an existing example of information beyond the mean Page curve. It is not a nonzero difference between the two normalized von Neumann limits in (7.4). Moreover, connected replica geometries need not introduce a handle: connectivity and genus are distinct properties, and the suppression of a covariance must be calculated in the specified model.

For a comparison between microscopic proposals, the ensemble measure, normalization prescription, state preparation, and observable must therefore be held fixed or related explicitly. A nonzero covariance can distinguish two specified preparations with different fluctuations. Neither its presence nor its absence alone decides whether gravity is described by one Hamiltonian. The finite Rényi gap and entropy variance provide useful additional data, while leaving the separate question of recovering an unknown input to a channel analysis.

8 Comparisons between gravitational descriptions

The examples above show that a radiation entropy, an averaged spectrum, a fluctuation, and a channel recovery test answer different questions. Applying these comparisons to gravity requires specifying the models on both sides. The broad names used for approaches to the information problem do not supply complete microscopic ensembles or evaporation maps. They can also refer to different parts of one description: a microscopic Hamiltonian, a family of states, a semiclassical calculation, or a prescription for interior operators.

8.1 What an entropy calculation establishes

A calculation of $S(R)$ in a specified state constrains that state's reduced density matrix. In gravity, the island prescription additionally relates the entropy to a generalized-entropy extremization problem [1, 9, 24]. When accompanied by a code-subspace description and suitable reconstruction assumptions, the calculation can support conclusions about recoverable interior observables. Ref. [25] contains an explicit Petz-map reconstruction argument in addition to its replica entropy computation. These additional structures are not encoded in the scalar function $S(R)$ alone.

This distinction prevents two opposite inferences. An entropy curve that agrees with a unitary random-state benchmark does not uniquely identify its preparation. Conversely, the existence of elementary replacement-channel examples does not invalidate a reconstruction argument that supplies more information than that curve. The examples in section 2 hold the specified radiation data fixed while changing the channel on other inputs. A gravity calculation that determines those additional inputs or their reference correlations can exclude the replacement channel.

Matching the full sequence $\mathbb{E} \operatorname{Tr} \rho_R^n$ likewise has a precise meaning. At fixed dimension it determines the averaged spectral measure, since that measure is supported on a compact interval. It does not determine a probability law on spectra. The finite-dimensional construction in section 3 demonstrates this difference without an asymptotic approximation. A two-trace correlator supplies additional data.

Even those spectral data leave operator matrix elements and the input dependence of the channel unspecified.

8.2 The objects supplied by different approaches

Random-state and random-circuit models. A Haar-induced state is a precisely defined statistical model. It gives finite-dimensional moments and fluctuations once the dimensions and symmetry class are fixed. A random circuit can also specify an actual channel and a time parameter. Its circuit architecture, depth, and required design order then affect the approach to a design [6]. Replacing a circuit by a Haar state at every time retains a useful benchmark but discards information about the temporal coupling of the states. Section 9 makes that distinction explicit.

Islands and replica wormholes. An island calculation is a prescription for entropies and related observables within a gravitational setup. In the model of ref. [25], the replica sum can be compared with a random-state spectral problem. In JT gravity, the matrix-integral description computes particular ensemble quantities [26]. Neither statement implies that every island calculation is an ensemble of the same microscopic Hamiltonians. Whether a boundary description is fixed or ensemble averaged is a question about its definition and factorization properties [17]. It cannot be settled by a von Neumann annealed–quenched gap under the normalized definitions of section 7, because that gap vanishes.

Microstate geometries and the fuzzball program. Microstate constructions seek a microscopic description of black-hole states and their effective geometric properties. Mathur’s analysis also emphasizes restrictions on attempts to restore purity through small corrections to a specified pair-creation process [19]. A radiation-spectrum comparison requires a chosen family of states, a probability distribution if an average is taken, and a coupling to the radiation. The name of the program alone does not determine an entanglement-spacing distribution or an entropy variance. In particular, a claim about integrability of a Hamiltonian does not by itself provide the local statistics of the reduced density matrix produced in a specified evaporation experiment.

State-dependent interior operators. State-dependent constructions concern the map between boundary observables and interior operators in a specified class of states [23]. To compare them through outgoing radiation, one must identify a boundary preparation and an observable whose prediction differs. A different representation of interior operators need not change the radiation density matrix. If the same microscopic state and outgoing algebra are used, all radiation entropies can agree without resolving how the interior is represented.

Geometric descriptions of entanglement. The ER=EPR proposal relates entanglement to geometric connections [16]. By itself it does not specify a normalized probability law for Hawking-radiation density matrices. A spectral or recovery comparison therefore requires a concrete realization of the proposal and a common set of observables. Its statements can coexist with a microscopic unitary theory and with a semiclassical island calculation. Treating these descriptions as mutually exclusive rows with independently fixed numerical predictions would impose more structure than their names provide.

8.3 A comparison table with explicit status

Table 2 collects calculations rather than assigning uncomputed predictions to broad programs. An equality in one row does not imply equality in another. In particular, the energy ensemble and accessible

Table 2. Specified comparisons and the information they add. The first four rows are calculated examples in this paper. The final two identify gravity calculations from the literature or further model-dependent questions; they are not predictions shared by every member of a named approach.

Comparison	Data that agree	Additional distinction
Isometry and replacement channel	All radiation states for one input	Reference mutual information and recovery fidelity differ exactly
Finite spectral ensembles	Averaged replica moments at every order and mean entropy	Purity and entropy covariances differ exactly
Fixed MP quantiles and Wishart states	Limiting density and mean entropy correction	Fixed versus fluctuating purity at finite dimension
Gaussian and quartic branch differences	Mean and variance	Central deficit and full averaged-minimum profile differ
Gravitational entropy fluctuations	A specified mean Page curve	Connected normalized replicas give model-dependent fluctuations; ref. [5] computes an example
Two proposed microscopic evaporation models	To be established for the chosen states and couplings	Compute a common operator, reference, or covariance observable before asserting a difference

algebra must be held fixed in a channel comparison, while normalization and matrix dimensions must be held fixed in a spectral comparison.

The last row is a research requirement. It is not an inference that all microscopic descriptions must eventually disagree on one of the listed quantities. Distinct descriptions of the same physics can agree on all accessible observables. Conversely, two models can already differ in a leading entropy curve if their state preparation, constraints, or dynamics differ. There is no general rule that physically meaningful differences begin only at a particular genus order.

8.4 Precision, normalization, and accessible observables

Three quantitative checks are needed before a small difference becomes a useful comparison. First, the error of each calculation must be smaller than the proposed distinction. Leading convergence to a common spectral density does not control a finite-size covariance; the Haar purity example illustrates the normalization terms required at that order. Second, the same ensemble operation must be computed on both sides. A ratio of averaged partition functions can differ from an average of normalized density-matrix moments. Third, the observable must belong to the algebra accessible in the proposed experiment. An operator involving an external reference requires that reference to be part of the state preparation.

Ensemble variance also differs from quantum measurement variance in a single member. For an observable O and states ρ_α , the variance in the average state decomposes as

$$\begin{aligned} \text{Tr}(\bar{\rho}O^2) - [\text{Tr}(\bar{\rho}O)]^2 &= \mathbb{E}_\alpha [\text{Tr}(\rho_\alpha O^2) - \text{Tr}(\rho_\alpha O)^2] \\ &\quad + \text{Var}_\alpha [\text{Tr}(\rho_\alpha O)], \quad \bar{\rho} = \mathbb{E}_\alpha \rho_\alpha. \end{aligned} \tag{8.1}$$

The two terms refer to repeated measurement within a preparation and to variation between preparations. Determining which variation a connected gravitational quantity represents requires an interpretation of the ensemble. A single fixed Hamiltonian can still support an ensemble of input states, random couplings in an experiment, or time samples. A nonzero variance is therefore not, by itself, evidence for an ensemble of Hamiltonians.

9 From equilibrium states to evaporation dynamics

Equilibrium distributions and instantaneous random-state models give well-defined entropy calculations. An evaporation process additionally specifies how successive radiation subsystems arise from one evolution. This distinction applies in AdS coupled to an absorbing bath as well as in asymptotically flat settings. Dynamical island calculations already study time-dependent evaporation [1, 2]; the issue here is which additional assumptions permit the equilibrium and random-state calculations of the preceding sections to describe such a history.

9.1 Consistency of successive radiation states

Use the finite-dimensional notation of Section 2. A sequential isometric model has maps

$$\begin{aligned} U_k : A &\longrightarrow B_k \otimes R_k, & W_{k+1} : B_k &\longrightarrow B_{k+1} \otimes b_{k+1}, \\ U_{k+1} &= (W_{k+1} \otimes I_{R_k})U_k, \end{aligned} \tag{9.1}$$

with the tensor factors reordered where necessary. Ancillas in a specified initial state may be included in B_k . The corresponding radiation channels $\mathcal{N}_k(X) = \text{tr}_{B_k}(U_k X U_k^\dagger)$ obey

$$\text{tr}_{b_{k+1}} \mathcal{N}_{k+1}(X) = \mathcal{N}_k(X) \quad \text{for every input operator } X. \tag{9.2}$$

This equation assumes that emitted modes are no longer acted upon; equivalently, their freely evolving observables can be described in an interaction picture. It is a constraint on channels and their joint outputs, not merely on the values of their entropies.

For example, assigning $|0\rangle\langle 0|$ to the first emitted qubit and a Bell state to the first two emitted qubits gives two valid density matrices, both pure. They cannot be the successive radiation states of this model: tracing the second qubit out of the Bell state gives $I/2$, not $|0\rangle\langle 0|$. Sampling an unrelated Haar state at each instant similarly does not establish equation (9.2). A dynamical random-state approximation must explain how the jointly distributed states are generated, even when each individual state has the expected marginal statistics. The constructions in Proposition 2.1 do satisfy this consistency condition because they release subsystems of a single prepared output.

Entropy inequalities give a weaker, useful necessary condition. Since $R_{k+1} = R_k b_{k+1}$,

$$S(R_{k+1}) - S(R_k) = S(b_{k+1}|R_k), \quad |S(R_{k+1}) - S(R_k)| \leq \ln \dim b_{k+1}. \tag{9.3}$$

The bound follows from subadditivity and the Araki–Lieb inequality. A decreasing radiation entropy requires negative conditional entropy of the new emission given the earlier radiation. In the Bell-pair example, the increment is $+\ln 2$ for the first q emissions and $-\ln 2$ for the last q . This accounts for the turnover through explicit early–late correlations, while remaining insufficient to decide whether those correlations encode the initial input.

9.2 Fixed total energy does not fix the black-hole energy

Energy conservation constrains the joint state rather than assigning zero variance to the energy of every subsystem. For subsystems described by an additive Hamiltonian, a fixed-total-energy space has the decomposition

$$\mathcal{H}_{E_{\text{tot}}} = \bigoplus_E \mathcal{H}_B(E) \otimes \mathcal{H}_R(E_{\text{tot}} - E). \quad (9.4)$$

A state in this space can have an arbitrary distribution p_E over the allowed black-hole energies. Orthogonality of the radiation energy sectors makes its reduced state block diagonal,

$$\rho_B = \bigoplus_E p_E \rho_{B,E}, \quad S(\rho_B) = - \sum_E p_E \ln p_E + \sum_E p_E S(\rho_{B,E}). \quad (9.5)$$

The first term is the entropy of the energy-sector probabilities, and the second is the mean entropy within sectors. Neither is fixed by specifying E_{tot} alone. Sampling assumptions and dynamics determine p_E .

The simplest illustration uses two qubits with excitation energy ε . The pure state

$$|\psi_p\rangle = \sqrt{p} |1\rangle_B |0\rangle_R + \sqrt{1-p} |0\rangle_B |1\rangle_R \quad (9.6)$$

has exact total energy ε , but

$$\text{Var}(E_B) = \varepsilon^2 p(1-p), \quad S(\rho_R) = -p \ln p - (1-p) \ln(1-p). \quad (9.7)$$

These are quantum fluctuations in one pure state; no average over Hamiltonians is involved. In an evaporating system, uncertainty in the remaining energy can likewise be correlated with the energy already emitted. Isolation does not justify replacing that distribution by a delta function at the mean black-hole energy.

The equilibrium counterpart further specifies the distribution. If states in a narrow total-energy shell are sampled with microcanonical weights, then the energy-sector probabilities are proportional to the product of the two densities of states. The stable quadratic expansion of this product, derived in Section 5, gives a variance of E_B/T equal to the effective heat capacity $(C_B^{-1} + C_R^{-1})^{-1}$. Fixing E_B itself removes this exchange fluctuation, whereas fixing $E_B + E_R$ does not. Because the two subsystem energies are anticorrelated, the variance of the difference between their thermodynamic entropies is four times the individual entropy variance at leading Gaussian order [33]. These thermodynamic entropies count states; they are distinct from the fine-grained entropies in equation (9.5). An outgoing flux need not be in equilibrium with the remaining black hole, so the equilibrium law cannot be assigned to evaporation without justifying its exchange and sampling assumptions.

9.3 A fluctuation width is not an evaporation timescale

Replacing an equilibrium parameter by a function of time does not determine either the radiation channel or its preparation law. Suppose, more specifically, that an instantaneous approximation has identified two entropy branches with mean difference $\Delta(t)$ and a distribution for their centered difference. Let its width in entropy units be $\sigma_\Delta(t)$, and assume that the distribution and the local entropy formula vary slowly while $\Delta(t)$ crosses zero. If $\dot{\Delta}(t_*) \neq 0$, linearization supplies the estimate

$$\tau_{\text{round}} \sim \frac{\sigma_\Delta(t_*)}{|\dot{\Delta}(t_*)|}, \quad \Delta(t_*) = 0. \quad (9.8)$$

Its coefficient depends on the convention used to define a crossover width. The estimate requires the fluctuation law and $\dot{\Delta}$ to change little over τ_{round} ; it also requires relaxation to justify the instantaneous description. If the crossing is tangent or relaxation is slow, the linear estimate fails. A thermodynamic variance alone supplies none of these dynamical conditions.

Similarly, a Marchenko–Pastur law at a fixed dimension ratio follows from a stated random-state or covariance model, not from time dependence of the ratio. Emission spectra, conserved charges and energy-dependent matrix elements can change the allowed states and their weights. In a gravitational application the relevant entropy branches must be derived from the generalized entropy prescription; they cannot always be replaced by logarithms of unweighted Hilbert space dimensions. The calculations of Sections 4 and 5 then provide conditional local models whose assumptions can be tested in the dynamical setting. They do not imply a universal sharp transition for isolated evaporation.

9.4 What a dynamical comparison should specify

A comparison between two evaporation models should fix the input space, the preparation, the emitted radiation algebra, and the conservation laws. It should state whether averages concern initial states, dynamics, or Hamiltonians, and identify any degrees of freedom that are traced out at the endpoint. Within that common setup, one can compare radiation moments at specified times, their connected fluctuations, and correlations between emissions. Reference-assisted recovery tests then address the dependence on unknown inputs, as in Section 2.3. Computational difficulty of a decoder is a further question beyond its mathematical existence [13].

The gravitational calculation has additional responsibilities: a specified coupling to a bath or asymptotic radiation region, a semiclassical state and its backreaction, and a justification of the entropy prescription in the regime being studied. These conditions are realized to different degrees in existing island models [1–3]. They should be evaluated model by model. Neither a fixed Hamiltonian nor fixed total energy eliminates all statistical fluctuations, and agreement of time-indexed equilibrium entropies does not establish agreement of quantum channels. These distinctions specify the additional calculations needed to connect the controlled examples in this paper to microscopic black-hole information recovery.

10 Further calculations

The explicit examples leave several concrete questions for gravitational models. Their resolution requires the following model-specific calculations.

10.1 Normalized two-trace fluctuations in a common preparation

The most direct extension of a computed mean Page curve is a connected correlator of normalized moments. One should specify the state preparation, energy window, boundary conditions, and random variables being averaged, then compute

$$C(n, \ell) = \mathbb{E}[\text{Tr}\rho_R^n \text{Tr}\rho_R^\ell] - \mathbb{E}[\text{Tr}\rho_R^n]\mathbb{E}[\text{Tr}\rho_R^\ell]. \quad (10.1)$$

The conditions $C(1, \ell) = C(n, 1) = 0$ are useful exact normalization checks. Differentiation at $(1, 1)$ gives the entropy variance when the limits are justified. The calculation of ref. [5] provides a gravitational example and a comparison with Haar-state fluctuations. Extending this comparison to another specified preparation would test how much of the covariance survives changes in the energy window, state-count fluctuations, or bath coupling.

The finite spectral construction in section 3 makes the missing information precise: even all one-trace averaged moments leave a two-trace covariance undetermined. A leading spectral calculation

therefore cannot replace this step. On the other hand, the covariance may itself be shared by several models. Establishing a difference requires its explicit value and controlled errors.

10.2 Generalized-entropy response near a critical point

The quartic thermodynamic law of Paper II suggests a calculation in a charged gravitational model with an actual entropy transition. First locate the competing generalized-entropy branches for the chosen state and radiation region. Next compute their joint response to the critical mode, including the covariance of the area and bulk-entropy terms. The relevant coefficient is the response of their difference, as in (5.14). A cancellation at that order would remove the $N^{3/4}$ contribution even though each branch fluctuates on that scale.

A finite bath then changes the fluctuation law through its entropy curvature. The $C_R \sim N^{3/2}$ crossover concerns the canonical critical model and its stated reservoir regularity assumptions. A physical application must check that the bath samples the required energy range and that remote configurations do not dominate the entropy moments. Finally, relaxation and evaporation times must be compared before interpreting the equilibrium distribution as a time-dependent radiation profile. These requirements make a critical enhancement a conditional prediction that can be tested in a chosen setup.

10.3 Spectral comparisons with matched conservation laws

The Wishart benchmark has no energy constraint. A useful next model would impose a common sector decomposition on both ensembles being compared, match the sector probabilities and dimensions, and evaluate both within-sector and between-sector fluctuations. For orthogonal radiation sectors, the entropy is $H(p) + \sum_{\alpha} p_{\alpha} S(\rho_{R,\alpha})$. Fluctuating sector probabilities introduce covariance terms absent from a fixed-dimension Haar calculation. Matching only the total dimension would miss them.

Such a construction would also clarify when the canonical planar correction of Paper I can enter a comparison. When radiation resolves the sector, the direct-sum identity applies; when sector contributions overlap in the radiation space, a spectral calculation is needed. One should determine this algebraic distinction from the state before choosing a random-matrix approximation. It need not be a small correction to the same model.

10.4 Reference correlations and recovery

A channel comparison can begin with a finite code subspace and an input entangled with a reference. The target quantities include $I(Q : R_t)$, the complementary correlations with the remaining black hole, and the error of a specified recovery map. These quantities address transfer of unknown quantum information, as in the decoding perspective of Hayden and Preskill [13]. They also permit a direct comparison with reconstruction statements associated with islands [24, 25].

A successful calculation should separate existence of a recovery map from the resources needed to implement it. An isometry has a perfect inverse on its image, but this fact gives no efficient circuit for that inverse. Conversely, a complicated decoder does not imply that the entropy curve is wrong. The chosen code subspace, accessible radiation algebra, and permitted operations determine the operational question.

10.5 Reproducibility of the worked comparisons

The figures in this paper are evaluated from the specified probability laws and finite-dimensional formulas. They contain no fitted gravity predictions. The accompanying calculations enumerate Haar

permutations, check the finite spectral examples, compare direct quadrature with closed fluctuation profiles, and evaluate the fixed-quantile sequence. The algebraic arguments establish the identities and limits; the numerical checks test their implementations and finite-size values. The figure programs and their dependencies are preserved with the source.¹

11 Conclusion

A Page curve constrains radiation entanglement, while identifying an information-recovery process requires information about the state preparation and channel. The distinction can be made exact. An isometry and a replacement channel reproduce every radiation state along one chosen preparation, although they have different reference correlations. Two finite-dimensional ensembles reproduce all averaged replica moments and the same mean von Neumann entropy, although their entropy fluctuations differ. Fixed Marchenko–Pastur spectra and Wishart spectra provide a corresponding comparison in the usual large-dimension regime.

These examples also identify quantities worth computing. Normalized two-trace correlations determine entropy fluctuations through a mixed replica derivative. The finite-order quenched-minus-annealed Rényi gap vanishes at order one, with a slope set by half the entropy variance. Entropy of an averaged state is a separate quantity. Keeping these operations distinct is necessary when comparing a gravitational path integral with an ensemble or a fixed preparation.

Thermodynamic broadening supplies additional information once the competing branches and their joint law are specified. Equal variances can give different crossover shapes, and a known fixed-law averaged-minimum profile can determine its scalar fluctuation distribution. Energy constraints, reservoir size, and cancellation of a common mode determine how that distribution enters a radiation entropy. A static distribution still requires an evolution law before it becomes an evaporation prediction.

The remaining gravitational question is therefore model specific: which additional spectral, operator, or reference-correlation data are fixed by the proposed description, and which predictions differ under a common preparation? Answering it would connect agreement with a Page curve to a more complete account of information recovery.

A Review of Free Probability and Non-Crossing Partitions

This appendix provides a self-contained summary of the free-probability tools used throughout the paper. The target reader is an information-paradox researcher who is familiar with random-matrix theory at the level of Wishart ensembles but has not encountered the combinatorial framework of free probability. Comprehensive treatments can be found in Nica and Speicher [21] and in the foundational papers of Voiculescu [30] and Speicher [29].

A.1 Non-crossing partitions

Let $[n] = \{1, 2, \dots, n\}$. A *partition* π of $[n]$ is a collection of non-empty, pairwise disjoint subsets (called *blocks*) whose union is $[n]$. The set of all partitions of $[n]$ is denoted $\mathcal{P}(n)$.

A partition $\pi \in \mathcal{P}(n)$ is *non-crossing* if there do not exist elements $a < b < c < d$ in $[n]$ such that a, c belong to one block and b, d belong to a different block. Equivalently, when the elements of $[n]$ are arranged around a circle, the convex hulls of the blocks do not intersect. The set of non-crossing partitions of $[n]$ is denoted $\text{NC}(n)$.

¹Reproduction commands are listed in the project build instructions. The computation programs are stored alongside the figure programs, with exact permutation checks separated from numerical quadrature.

Small examples. For $n = 1$ there is a single partition $\{\{1\}\}$, which is non-crossing. For $n = 2$ there are two partitions, $\{\{1\}, \{2\}\}$ and $\{\{1, 2\}\}$; both are non-crossing. For $n = 3$ there are five partitions: $\{\{1\}, \{2\}, \{3\}\}$, $\{\{1, 2\}, \{3\}\}$, $\{\{1, 3\}, \{2\}\}$, $\{\{1\}, \{2, 3\}\}$, and $\{\{1, 2, 3\}\}$. All five are non-crossing. For $n = 4$ there are 15 partitions in total, of which exactly 14 are non-crossing; the sole crossing partition is $\{\{1, 3\}, \{2, 4\}\}$.

Lattice structure. The set $\text{NC}(n)$, ordered by refinement (i.e. $\pi \leq \sigma$ if every block of π is contained in a block of σ), forms a lattice: any two elements have a unique meet and join. This lattice is self-dual and graded, with the rank of π equal to n minus the number of blocks [21].

A.2 Catalan numbers

The cardinality of $\text{NC}(n)$ is the n -th Catalan number:

$$C_n = |\text{NC}(n)| = \frac{1}{n+1} \binom{2n}{n}. \quad (\text{A.1})$$

The first several values are

$$C_1 = 1, \quad C_2 = 2, \quad C_3 = 5, \quad C_4 = 14, \quad C_5 = 42, \quad C_6 = 132. \quad (\text{A.2})$$

The generating function is

$$C(z) = \sum_{n=0}^{\infty} C_n z^n = \frac{1 - \sqrt{1 - 4z}}{2z}, \quad (\text{A.3})$$

satisfying the quadratic relation $C(z) = 1 + zC(z)^2$. The asymptotic growth is

$$C_n \sim \frac{4^n}{n^{3/2}\sqrt{\pi}} \quad (n \rightarrow \infty). \quad (\text{A.4})$$

For comparison, the total number of partitions of $[n]$ is the Bell number B_n , which grows much faster: $B_4 = 15$, $B_5 = 52$, $B_6 = 203$. The distinction between C_n and B_n is the combinatorial signature of the difference between free and classical probability.

A.3 Narayana numbers

The Catalan number C_n counts all non-crossing partitions of $[n]$ without regard to the number of blocks. A finer count is provided by the Narayana numbers:

$$N(n, k) = \frac{1}{n} \binom{n}{k} \binom{n}{k-1}, \quad k = 1, 2, \dots, n, \quad (\text{A.5})$$

which count the number of non-crossing partitions of $[n]$ with exactly k blocks. They satisfy

$$\sum_{k=1}^n N(n, k) = C_n. \quad (\text{A.6})$$

For example, when $n = 4$: $N(4, 1) = 1$, $N(4, 2) = 6$, $N(4, 3) = 6$, $N(4, 4) = 1$, summing to $C_4 = 14$. The symmetry $N(n, k) = N(n, n+1-k)$ reflects the self-duality of the non-crossing partition lattice.

In the uniform-weight replica model, the Narayana polynomial groups planar permutations by their cycle number. Its interpretation in terms of replica geometries depends on that model [18, 25].

A.4 The Möbius function on the non-crossing partition lattice

For a locally finite partially ordered set, the Möbius function is defined by $\mu(\pi, \pi) = 1$ and $\sum_{\pi \leq \tau \leq \sigma} \mu(\pi, \tau) = 0$ when $\pi < \sigma$. For the interval from the discrete partition 0_n to a non-crossing partition π , its value factorizes over the blocks:

$$\mu_{\text{NC}}(0_n, \pi) = \prod_{B \in \pi} (-1)^{|B|-1} C_{|B|-1}. \quad (\text{A.7})$$

In particular, $\mu(0_n, 1_n) = (-1)^{n-1} C_{n-1}$. General interval values follow from the recursion; counting only how many blocks of π lie in each block of σ does not specify the requisite interval factors. Möbius inversion states that if $g(\sigma) = \sum_{\pi \leq \sigma} f(\pi)$, then $f(\sigma) = \sum_{\pi \leq \sigma} g(\pi) \mu(\pi, \sigma)$. Applied to non-crossing partitions, this recovers free cumulants from moments [21, 29].

A.5 Free cumulants and the moment–cumulant relation

Let a be a non-commutative random variable with moments $M_n = \varphi(a^n)$, where φ is a tracial state. The *free cumulants* κ_n are defined implicitly by the moment–cumulant formula

$$M_n = \sum_{\pi \in \text{NC}(n)} \prod_{B \in \pi} \kappa_{|B|}, \quad (\text{A.8})$$

where the sum is over all non-crossing partitions of $[n]$ and the product is over the blocks B of π . This formula can be inverted using the Möbius function (A.7) to express κ_n in terms of the moments.

The free cumulants should be contrasted with the *classical cumulants* c_n , which are defined by the analogous formula with the sum taken over *all* partitions $\mathcal{P}(n)$:

$$M_n = \sum_{\pi \in \mathcal{P}(n)} \prod_{B \in \pi} c_{|B|}. \quad (\text{A.9})$$

The restriction from $\mathcal{P}(n)$ to $\text{NC}(n)$ is the defining algebraic difference between classical and free probability. For random variables that are free (the non-commutative analogue of independence), mixed free cumulants vanish, just as mixed classical cumulants vanish for independent random variables [21].

A.6 The R -transform

The generating function of the free cumulants is the R -transform:

$$R(z) = \sum_{n=1}^{\infty} \kappa_n z^{n-1}. \quad (\text{A.10})$$

It is the free-probability analogue of the logarithm of the moment generating function in classical probability. The key property is *additivity under free convolution*: if a and b are freely independent, then

$$R_{a+b}(z) = R_a(z) + R_b(z). \quad (\text{A.11})$$

The R -transform is related to the Cauchy transform $G(z) = \sum_{n=0}^{\infty} M_n z^{-n-1}$ by the functional equation

$$G(R(z) + 1/z) = z, \quad (\text{A.12})$$

which provides a practical route to computing the eigenvalue distribution from the free cumulants [30].

A.7 The free Poisson (Marchenko–Pastur) distribution

The distribution of central importance for the Page curve is the *free Poisson distribution*, also known as the Marchenko–Pastur (MP) law. It is characterized by geometric free cumulants:

$$\kappa_m = \gamma^{m-1}, \quad m = 1, 2, 3, \dots, \quad (\text{A.13})$$

where $\gamma > 0$ is the ratio parameter. Substituting into the R -transform (A.10) gives a geometric series:

$$R(z) = \sum_{n=1}^{\infty} \gamma^{n-1} z^{n-1} = \frac{1}{1 - \gamma z}. \quad (\text{A.14})$$

The moments follow from the moment–cumulant relation (A.8):

$$M_n = \sum_{\pi \in \text{NC}(n)} \gamma^{n-|\pi|} = \sum_{k=1}^n N(n, k) \gamma^{n-k}, \quad (\text{A.15})$$

recovering the Narayana polynomial encountered in the main text.

The eigenvalue density of the Marchenko–Pastur law with ratio parameter γ is

$$f_{\text{MP}}(x; \gamma) = \frac{1}{2\pi\gamma x} \sqrt{(\lambda_+ - x)(x - \lambda_-)}, \quad \lambda_{\pm} = (1 \pm \sqrt{\gamma})^2, \quad (\text{A.16})$$

supported on $[\lambda_-, \lambda_+]$, with an additional point mass $(1 - 1/\gamma) \delta(x)$ when $\gamma > 1$. At $\gamma = 1$ (the Page transition), the lower edge touches zero and the density reduces to

$$f_{\text{MP}}(x; 1) = \frac{1}{2\pi x} \sqrt{x(4 - x)}, \quad x \in [0, 4]. \quad (\text{A.17})$$

The Marchenko–Pastur distribution arises as the limiting eigenvalue distribution of Wishart matrices $W = XX^\dagger/L$, where X is a $k \times L$ matrix with independent centered complex Gaussian entries normalized by $\mathbb{E}|X_{ij}|^2 = 1$, and $k/L \rightarrow \gamma > 0$. In the black-hole context, k and L are the radiation and black-hole Hilbert-space dimensions, and $\gamma = k/L$ parametrizes the Page transition [22, 32].

A.8 Why this matters for replica wormholes

The uniform-weight model of ref. [25] relates a planar replica sum to the moments in (A.15). The restriction to non-crossing partitions organizes the sum, and the specified weights determine its value. Both ingredients are necessary. For example, the first three moments of a law with free cumulants r_1, r_2, r_3 are

$$M_1 = r_1, \quad M_2 = r_1^2 + r_2, \quad M_3 = r_1^3 + 3r_1r_2 + r_3. \quad (\text{A.18})$$

Changing r_3 changes M_3 while leaving the first two moments unchanged whenever the modified cumulants belong to a probability law. For example, take a free compound Poisson law of rate one. One nonnegative jump distribution gives masses $1/2, 1/2$ to $0, 2$; another gives masses $1/3, 1/2, 1/6$ to $0, 1, 3$. Their first two moments are $1, 2$, but their third moments are $4, 5$. Since the jump moments are the free cumulants, the resulting laws have $M_1 = 1$, $M_2 = 3$, and respectively $M_3 = 11, 12$. The non-crossing combinatorics is the same in both calculations.

In particular, planarity by itself does not imply the Marchenko–Pastur law, and a gravitational genus expansion does not prove free independence of unspecified radiation modes. Canonically weighted models have different free cumulants; their free-probability formulation is developed in ref. [31]. The order-one comparison in (5.17) uses that more general spectral problem.

Likewise, a one-trace moment formula does not determine the covariance of two traces. The latter uses two boundary cycles and, for normalized states, the joint normalization factors. Appendix D works out this distinction at finite dimension. It cannot be recovered by subtracting C_n^2 from C_{2n} in an ordinary one-circle non-crossing count.

B Normalized moments and the spectral comparison

This appendix supplies the finite-dimensional normalization, entropy integral, and deterministic approximation used in Sections 4 and 6. The random-state identities are standard. Their purpose here is to place the comparison of ensembles on a common normalization and to separate statements about a limiting spectrum from statements about its fluctuations.

B.1 The symmetric-projector formula

Let $D = mb$ and $P_\psi = |\psi\rangle\langle\psi|$ for a Haar-random unit vector in $\mathbb{C}^D$. Unitary invariance and support on the symmetric subspace imply

$$\mathbb{E}_H P_\psi^{\otimes k} = \frac{P_{\text{sym},k}}{\binom{D+k-1}{k}} = \frac{1}{(D)_k} \sum_{\pi \in S_k} U_\pi. \quad (\text{B.1})$$

Indeed, the symmetric tensor representation is irreducible, so the invariant operator is proportional to its identity; taking the trace fixes the constant. Here U_π permutes the k copies of the full bipartite space. Factoring it as $U_\pi^R \otimes U_\pi^B$ and contracting the radiation factors with a permutation τ gives

$$\mathbb{E}_H \prod_{a=1}^r \text{Tr} \rho_R^{k_a} = \frac{1}{(mb)_K} \sum_{\pi \in S_K} m^{c(\tau\pi)} b^{c(\pi)}, \quad K = \sum_a k_a, \quad (\text{B.2})$$

where τ has disjoint cycles of lengths $k_1, \dots, k_r$. This form follows because a permutation on copies of $\mathbb{C}^d$ has trace $d^{c(\pi)}$. An inverse in the convention for the cyclic permutation changes the indexing of the sum but not its value.

For a single trace, the first three moments are

$$\begin{aligned} \mathbb{E} \text{Tr} \rho_R &= 1, \\ \mathbb{E} \text{Tr} \rho_R^2 &= \frac{m+b}{mb+1}, \\ \mathbb{E} \text{Tr} \rho_R^3 &= \frac{m^2 + 3mb + b^2 + 1}{(mb+1)(mb+2)}. \end{aligned} \quad (\text{B.3})$$

The third numerator follows from six permutations: the identity and the inverse three-cycle supply the m^2 and b^2 terms, the three transpositions supply $3mb$, and the remaining three-cycle supplies one. These expressions reduce to one whenever either subsystem has dimension one, as required for a pure reduced state.

For purity $P = \text{Tr} \rho_R^2$, the four-copy formula gives

$$\mathbb{E} P^2 = \frac{(mb+4)(m+b)^2 + 2(mb+1)}{(mb+1)(mb+2)(mb+3)}. \quad (\text{B.4})$$

Subtracting $[(m+b)/(mb+1)]^2$ yields Eq. (6.13). The full permutation count and the normalization cancellation are displayed in Appendix D; the result agrees with the known purity cumulants [10]. No approximation of the random denominator is required.

B.2 Planar terms and finite-dimensional corrections

For the single-cycle contraction, the cycle inequality is

$$c(\pi) + c(\tau_k \pi) \leq k + 1. \quad (\text{B.5})$$

It follows from the triangle inequality for transposition length $|\pi| = k - c(\pi)$. Equality selects permutations associated with non-crossing partitions. If $c(\pi) = j$, their contribution after multiplication by m^{k-1} is γ^{k-j} , and their number is $k^{-1} \binom{k}{j} \binom{k}{j-1}$. This proves the leading Narayana sum in Eq. (4.4) directly from the normalized finite-dimensional expression.

At $m = b$, these terms give C_k for each fixed k . The nonplanar numerator and the normalization denominator both contribute to the first correction. For example,

$$m \mathbb{E} \text{Tr} \rho_R^2 = 2 - \frac{2}{m^2} + O(m^{-4}), \quad m^2 \mathbb{E} \text{Tr} \rho_R^3 = 5 - \frac{14}{m^2} + O(m^{-4}). \quad (\text{B.6})$$

The purity numerator contains only the two permutations already present at planar order. Its nonzero correction arises from normalization. For the third moment, the additional numerator term and normalization combine to give the coefficient -14 . Counting a subset of additional permutations therefore does not by itself give the correction to a normalized observable. A fortiori, the vanishing of an integer-index coefficient at $k = 1$ does not determine its derivative at $k = 1$ and cannot establish a vanishing von Neumann entropy correction.

A compactly supported limiting measure is uniquely determined by all its moments: equality on polynomials extends to equality on continuous functions by uniform polynomial approximation. This is a statement about the measure. It does not identify an arbitrary analytic function of a complex replica index from its integer values. Here the entropy is obtained from the identified measure by integration, with the tail control in Eq. (6.6) when the finite-dimensional measures are not uniformly supported. No Carlson-type continuation assumption is needed for this argument.

B.3 Evaluation of the Marchenko–Pastur entropy integral

For $0 < \gamma \leq 1$, put $r = \sqrt{\gamma}$ and $x = 1 + r^2 - 2r \cos \theta$, $0 \leq \theta \leq \pi$. The density then satisfies

$$d\mu_\gamma(x) = \frac{2 \sin^2 \theta}{\pi x} d\theta, \quad \int x \log x d\mu_\gamma(x) = \frac{2}{\pi} \int_0^\pi \sin^2 \theta \log(1 + r^2 - 2r \cos \theta) d\theta. \quad (\text{B.7})$$

For $r < 1$, expand

$$\log(1 + r^2 - 2r \cos \theta) = -2 \sum_{j \geq 1} \frac{r^j \cos(j\theta)}{j}. \quad (\text{B.8})$$

Orthogonality gives $\int_0^\pi \sin^2 \theta \cos(j\theta) d\theta = 0$ except at $j = 2$, where the integral is $-\pi/4$. Therefore

$$\int x \log x d\mu_\gamma(x) = \frac{r^2}{2} = \frac{\gamma}{2}. \quad (\text{B.9})$$

The endpoint $r = 1$ follows by continuity of the integral; the logarithmic singularity at $\theta = 0$ is integrable after multiplication by $\sin^2 \theta$. Together with normalization of ρ_R , this proves the entropy deficit used in the spectral comparison. For $\gamma > 1$, the zero eigenvalues contribute nothing to the entropy, and exchanging the two subsystem dimensions yields Eq. (4.9).

The same measure can evaluate Rényi statistics at any specified real index where its moment integral is finite. A negative index at $\gamma = 1$ must account for the hard-edge density, which behaves as $x^{-1/2}$. In particular, convergence of positive integer moments should not be used without additional assumptions to infer statistics that are singular at zero. This provides another reason to state which observable is being compared rather than referring to a spectrum as a single universal datum.

B.4 Quantile normalization and convergence

Write $Q_\gamma = F_\gamma^{-1}$. For fixed $\gamma \in (0, 1]$, Q_γ is continuous on $[0, 1]$ with bounded values in $[x_-, x_+]$. Thus, for every continuous function f on that interval,

$$\frac{1}{m} \sum_{i=1}^m f\left(Q_\gamma\left(\frac{i-1/2}{m}\right)\right) \longrightarrow \int_0^1 f(Q_\gamma(u)) du = \int f(x) d\mu_\gamma(x). \quad (\text{B.10})$$

Taking $f(x) = x$ gives $\bar{q}_m \rightarrow 1$. Replacing $q_{i,m}$ by $q_{i,m}/\bar{q}_m$ consequently preserves the weak limit and every fixed positive moment. The scaled eigenvalues remain in a fixed compact interval for sufficiently large m , so the choice $f(x) = x \log x$, continuously extended at zero, also applies. This establishes the fixed-spectrum part of Proposition 6.2 without random sampling.

Purification introduces no further spectral restriction when $b_m \geq m$: the Schmidt rank of every state constructed in Eq. (6.9) is at most m . At a fixed pair of dimensions, conjugating by U can change measurements in a specified basis but leaves every spectral statistic constant. The vanishing purity variance therefore holds exactly, rather than as a numerical smallness assertion. The entropy variance also vanishes exactly in this ensemble. In the Haar ensemble it is positive for $m, b > 1$, since entropy is a continuous nonconstant function on a pure-state space with full-support Haar measure. Its size is a separate question from equality of the limiting entropy.

The construction describes a sequence of ensembles at specified bipartitions. It does not assert that arbitrary choices of these states can be joined into an energy-conserving evaporation process with a fixed Hamiltonian. Such a dynamical embedding is additional physical information, exactly as the relation between an equilibrium spectral law and a radiation channel is additional information. The conclusion needed here concerns identifiability: neither the limiting density nor its entropy fixes the probability law of spectral fluctuations.

B.5 Reproducibility of the comparison

Figure 2 uses $\gamma = 1/2$, $b = 2m$, and $m = 2, 4, 8, \dots, 256$. The fixed spectrum is computed by inverting the cumulative integral in Eq. (B.7) at the midpoint probabilities in Eq. (6.9), followed by exact normalization of the numerical eigenvalue array. The Wishart mean entropy uses Eq. (4.8), and its purity variance uses Eq. (6.13). Thus the figure combines a deterministic quadrature with exact finite-dimensional formulas; it contains no Monte Carlo estimate.

Independent rational-arithmetic sums over S_k , for $k \leq 4$, verify Eqs. (B.3) and (B.4) at several dimension pairs, including dimensions one. The second-moment normalization, positivity of the constructed spectra, unit trace, and convergence of the quantile entropy are checked alongside the figure data.² These checks supplement the permutation and Riemann-sum arguments; the propositions do not depend on numerical fitting.

C Fluctuation profiles and finite-reservoir limits

This appendix gives the probability calculations used in section 5. The RN–AdS critical scaling and its reservoir interpolation are imported from the thermodynamic model of Paper II [33]; the identities below make their use here self-contained.

²The self-contained computation is in `figures/code/compute05_spectral_comparison.py`; the plot is produced by `figures/code/fig05_ensemble_comparison.py`.

C.1 An integrated-tail representation

For a centered symmetric integrable variable Y with density f , write $d = |\Delta|$. Splitting the absolute-value integral at $-\Delta$ gives

$$B_Y(\Delta) = \int_d^\infty (y - d)f(y) dy. \quad (\text{C.1})$$

It follows that B_Y is nonnegative, even, and decreases with d . For a standard Gaussian, $\int_d^\infty y\phi(y) dy = \phi(d)$, which yields (5.4). The deficit is largest at the crossing and vanishes as $d \rightarrow \infty$. These properties belong to the symmetric fixed-law averaged-minimum model; they are not a statement about the ensemble variance of a von Neumann entropy.

If Y has a finite second moment, Tonelli's theorem also gives

$$\int_{-\infty}^\infty B_Y(\Delta) d\Delta = 2 \int_0^\infty f(y) \int_0^y (y - d) dd dy = \frac{\mathbb{E}Y^2}{2}. \quad (\text{C.2})$$

The two deficits in figure 1 therefore have the same integrated area $1/2$, despite different central heights and tails. The area and peak convey different information about the law.

C.2 Quartic moments and standardization

For X with the density in (5.6), substituting $u = x^4$ in the positive-half-line integral yields, for $r > -1$,

$$\mathbb{E}|X|^r = \frac{\Gamma((r+1)/4)}{\Gamma(1/4)}. \quad (\text{C.3})$$

Thus $v = \mathbb{E}X^2 = \Gamma(3/4)/\Gamma(1/4)$, $\mathbb{E}|X| = \sqrt{\pi}/\Gamma(1/4)$, and $\mathbb{E}X^4 = 1/4$. For $Y = X/\sqrt{v}$,

$$f_Y(y) = \frac{2\sqrt{v}}{\Gamma(1/4)} e^{-v^2 y^4}, \quad B_Y(0) = \frac{\sqrt{\pi}}{2\sqrt{\Gamma(1/4)\Gamma(3/4)}} = 2^{-5/4}, \quad (\text{C.4})$$

where the last equality uses $\Gamma(1/4)\Gamma(3/4) = \pi\sqrt{2}$. An explicit profile before standardization is

$$B_X(z) = \frac{\Gamma(1/2, |z|^4) - |z|\Gamma(1/4, |z|^4)}{2\Gamma(1/4)}, \quad B_Y(\Delta) = v^{-1/2} B_X(\sqrt{v}\Delta). \quad (\text{C.5})$$

Here $\Gamma(a, z) = \int_z^\infty u^{a-1} e^{-u} du$. Figure 1 uses the tail integral directly; the incomplete-gamma expression provides an independent evaluation.

C.3 Reservoir assumptions and response of two branches

At the critical point, let $\epsilon = E_B - E_{B,c}$. On the canonical scale of Paper II, $\epsilon/T_c = L_N X + o(L_N)$, and the black-hole action is $X^4 + o(1)$. The reservoir entropy expansion is

$$S_R(E_{R,c} - \epsilon) = S_R(E_{R,c}) - \epsilon/T_c - \epsilon^2/(2T_c^2 C_R) + \mathcal{R}_R(\epsilon). \quad (\text{C.6})$$

If $\mathcal{R}_R(T_c L_N x) = o(1)$ uniformly on compact sets, the allowed energies contain every fixed compact set on that scale, and the normalized laws are tight, the limiting density is $Z_u^{-1} e^{-x^4 - ux^2}$ with $u = \lim L_N^2/(2C_R)$ finite. Convergence of entropy variances additionally requires uniform integrability of the squared scaled energy and entropy variables. A local saddle expansion does not exclude remote regions carrying small probability and large moments. When transferring to a new branch difference, its own remainder must satisfy the moment condition in (5.14).

In the regime $C_R \rightarrow \infty$ and $C_R/L_N^2 \rightarrow 0$, the natural scale is $\epsilon/(T_c\sqrt{C_R})$. The same local and tail assumptions must hold on this narrower scale. The quartic term is then $O(C_R^2/L_N^4) = o(1)$, while the reservoir term is $y^2/2$. The resulting Gaussian limit has $\text{Var}S_B \sim C_R$. A reservoir whose capacity stays bounded need not admit a Gaussian approximation and requires its complete density of states.

At a regular common-temperature equilibrium, $\delta S_B = \epsilon/T$ and $\delta S_R = -\epsilon/T$ to first order. For $F_1 = S_B$ and $F_2 = S_R$, their difference responds with coefficient $2/T$, while their sum is stationary. More generally, the coefficient is

$$(\partial_{E_B} - \partial_{E_R})(F_1 - F_2)|_*. \quad (\text{C.7})$$

This derivative supplies information that neither marginal heat capacity contains.

The stationary sum gives a useful separate example. In the Gaussian exchange approximation, let $Z = \epsilon/(T\sqrt{C_{\text{eff}}})$ be standard normal. The total entropy loss is $L = Z^2/2$, so

$$\mathbb{E}L = \frac{1}{2}, \quad \text{Var}L = \frac{1}{2}. \quad (\text{C.8})$$

These constants concern a quadratic entropy cost. They do not imply $\text{Var}S_B = 1/2$, $\text{Var}(S_B - S_R) = 1/2$, or a universal microcanonical Page-transition width. Those quantities fluctuate linearly in the exchange mode and have the capacity-dependent variances given in the main text.

D Normalized Haar replica covariances

This appendix derives the finite-dimensional purity covariance used in Section 7. The result is standard in the theory of random bipartite states [10]. The derivation is included to exhibit the normalization terms that enter a connected replica calculation and to provide an exactly solvable comparison for more general gravitational ensembles.

D.1 The Haar moment identity

Let $|\psi\rangle$ be Haar distributed on $\mathbb{C}^m \otimes \mathbb{C}^b$, set $D = mb$, and define $\rho = \text{tr}_B |\psi\rangle\langle\psi|$. For positive integers $q_1, \dots, q_r$, put $q = \sum_{j=1}^r q_j$ and let $\tau \in S_q$ have disjoint cycles of lengths $q_1, \dots, q_r$. If $c(\pi)$ denotes the number of cycles of a permutation, including fixed points, then

$$\mathbb{E} \prod_{j=1}^r \text{tr} \rho^{q_j} = \frac{1}{(D)_q} \sum_{\pi \in S_q} m^{c(\tau\pi)} b^{c(\pi)}, \quad (D)_q = D(D+1) \cdots (D+q-1). \quad (\text{D.1})$$

Composition is from right to left. Replacing the convention for the cyclic permutation by its inverse gives the same summed result.

To prove the identity, unitary invariance gives

$$\mathbb{E}(|\psi\rangle\langle\psi|)^{\otimes q} = \frac{\sum_{\pi \in S_q} U_\pi}{(D)_q}, \quad (\text{D.2})$$

where U_π permutes tensor factors. Indeed, the Haar average is the normalized projector onto the symmetric tensor power, whose dimension is $(D)_q/q!$, while the projector is $q!^{-1} \sum_{\pi} U_\pi$. The product of reduced traces is obtained by contracting this operator with U_τ on the m -dimensional factors and the identity on the b -dimensional factors. A permutation trace contributes one factor of the subsystem dimension per cycle, producing (D.1). This proof fixes both the numerator and the rising-factorial denominator.

For a single moment of order two, $\tau = (12)$. The two permutations give $mb^2 + m^2b = D(m + b)$, and hence

$$\mathbb{E} \operatorname{tr} \rho^2 = \frac{m + b}{D + 1}. \quad (\text{D.3})$$

The formula applies to either ordering of the subsystem dimensions. A pure bipartite state has the same nonzero reduced eigenvalues on both sides, which accounts for the symmetry under $m \leftrightarrow b$.

D.2 The fourth replica moment

For $\mathbb{E}(\operatorname{tr} \rho^2)^2$, take $\tau = (12)(34)$ and enumerate the 24 permutations in S_4 . Table 3 groups them by the two cycle counts appearing in (D.1). Thus the finite computation can be checked without a choice of planar diagram convention.

Table 3. Cycle counts for $\tau = (12)(34)$. The contribution of a row is its multiplicity times $m^{c(\tau\pi)}b^{c(\pi)}$.

$c(\pi)$	$c(\tau\pi)$	Multiplicity
1	1	2
1	3	4
2	2	10
2	4	1
3	1	4
3	3	2
4	2	1

The resulting polynomial is

$$\begin{aligned} N_{22} &= m^2b^4 + m^4b^2 + 2m^3b^3 + 4mb^3 + 4m^3b + 10m^2b^2 + 2mb \\ &= D[D(m + b)^2 + 4(m^2 + b^2) + 10D + 2]. \end{aligned} \quad (\text{D.4})$$

Consequently

$$\mathbb{E}(\operatorname{tr} \rho^2)^2 = \frac{D(m + b)^2 + 4(m^2 + b^2) + 10D + 2}{(D + 1)(D + 2)(D + 3)}. \quad (\text{D.5})$$

Subtracting the square of (D.3) gives

$$\operatorname{Var}(\operatorname{tr} \rho^2) = \frac{2(m^2 - 1)(b^2 - 1)}{(D + 1)^2(D + 2)(D + 3)}. \quad (\text{D.6})$$

This agrees with the second cumulant computed by Giraud [10]. For $m = b = 2$, the mean purity is $4/5$, its second moment is $23/35$, and its variance is $3/175$. These numbers also follow directly from the seven rows of Table 3.

For $m = b$ large, the cancellation is particularly informative:

$$\mathbb{E}M_2^2 = \frac{4}{m^2} - \frac{6}{m^4} + O(m^{-6}), \quad (\mathbb{E}M_2)^2 = \frac{4}{m^2} - \frac{8}{m^4} + O(m^{-6}). \quad (\text{D.7})$$

The variance $2m^{-4} + O(m^{-6})$ depends on both subleading coefficients. Keeping the leading term of each expectation would correctly reproduce neither the coefficient nor the covariance itself. Likewise, retaining only permutations linking the two replica groups in an unnormalized numerator need not implement the subtraction for normalized states.

D.3 An exact two-qubit Rényi gap

The same Haar ensemble permits a nontrivial exact Rényi calculation when $m = b = 2$. Write the two eigenvalues as $(1 + u)/2$ and $(1 - u)/2$, with $0 \leq u \leq 1$. The induced Haar eigenvalue distribution is $3u^2 du$; equivalently, the purity $P = (1 + u^2)/2$ has density $3\sqrt{2P - 1}$ on $[1/2, 1]$ [10]. This density is normalized and reproduces $\mathbb{E}P = 4/5$ and $\mathbb{E}P^2 = 23/35$. The two order-two entropies are

$$A_2 = -\log(4/5), \quad Q_2 = -3 \int_0^1 u^2 \log\left(\frac{1 + u^2}{2}\right) du = \frac{\pi}{2} - \frac{4}{3}. \quad (\text{D.8})$$

For completeness, integration by parts gives

$$\int_0^1 u^2 \log\left(\frac{1 + u^2}{2}\right) du = -\frac{2}{3} \int_0^1 \frac{u^4}{1 + u^2} du = \frac{4}{9} - \frac{\pi}{6}. \quad (\text{D.9})$$

Thus $G_2 = \pi/2 - 4/3 + \log(4/5) = 0.014319 \dots > 0$. The fixed-order gap is compatible with the equality of the two von Neumann limits: they probe different aspects of the moment function. This example also illustrates the distinction between an exact Rényi gap and its small-fluctuation approximation. Substituting the purity mean and variance into the quadratic expression (7.17) gives $3/224$, rather than the exact value. Higher centered moments account for the difference.

D.4 Why normalization affects the leading covariance

A Gaussian construction makes the issue explicit. Let X be an $m \times b$ complex matrix with independent standard complex Gaussian entries, with $\mathbb{E}|X_{ia}|^2 = 1$, and put $W = XX^\dagger$. Its trace $T = \text{tr } W$ has the Gamma distribution of shape D and unit scale. The normalized matrix $\rho = W/T$ has the reduced Haar state distribution and is independent of T . This follows by writing the D complex Gaussian coordinates as an independent radius and uniform direction. In particular, $\mathbb{E}T^q = (D)_q$.

The operator $\tilde{\rho} = W/D$ has mean trace one but fluctuating trace. It satisfies

$$\widetilde{M}_2 := \text{tr } \tilde{\rho}^2 = \left(\frac{T}{D}\right)^2 M_2, \quad \mathbb{E}\widetilde{M}_2 = \frac{m + b}{D}. \quad (\text{D.10})$$

Independence of the radius and direction, followed by (D.4), yields

$$\text{Var}(\widetilde{M}_2) = \frac{4(m^2 + b^2) + 10D + 2}{D^3}. \quad (\text{D.11})$$

For $m = b$, this is $18m^{-4} + 2m^{-6}$, compared with the normalized leading coefficient $2m^{-4}$. The distinction persists at the first nonzero covariance order. It is not an effect that can be discarded because $T/D \rightarrow 1$ in probability.

The same issue is visible before computing purity. For the physical state, $M_1 = 1$ identically, and therefore $C(1, \ell) = 0$ for every ℓ . For $\tilde{\rho}$, the trace variance instead equals $\text{Var}(T/D) = 1/D$. Any connected-moment expression proposed for normalized states must satisfy the former identity. Subtracting the normalization fluctuations only after a large-dimension truncation can miss the cancellations that enforce it. In gravitational applications, these cancellations concern the normalization of each state, even if the path integral was initially normalized only after ensemble averaging [5].

D.5 From moment covariances to entropy variance

The double replica derivative in (7.14) can also be understood directly from the spectral counting measure. Define its covariance measure by

$$K(dx, dy) = \mathbb{E}[\nu_\rho(dx)\nu_\rho(dy)] - \mathbb{E}\nu_\rho(dx)\mathbb{E}\nu_\rho(dy). \quad (\text{D.12})$$

Then

$$C(n, \ell) = \int_{[0,1]^2} x^n y^\ell K(dx, dy), \quad \text{Var}(S) = \int_{[0,1]^2} x \log x y \log y K(dx, dy). \quad (\text{D.13})$$

The measure includes the contribution with the same eigenvalue in both factors. Removing that diagonal term would compute a different correlation function. Fixed trace implies the sum rule $\int x K(dx, dy) = 0$, interpreted as an identity of signed measures in y . This is the spectral form of $C(1, \ell) = 0$.

Equations (D.12) and (D.13) show precisely which information an averaged density of eigenvalues omits: it determines linear statistics in mean, while their variances require a two-point correlation. At fixed dimension, the differentiations are justified by the boundedness argument in Proposition 7.1. In an asymptotic calculation, controlling the covariance and its replica derivatives requires more than convergence of the one-point density. In particular, integer moments at a finite set of orders do not by themselves specify the continuation or its derivatives at one.

For reproducibility, the seven multiplicities in Table 3 exhaust S_4 , while (D.1) supplies an independent exact check of trace normalization and purity covariance for any positive integer m, b . The accompanying calculation uses rational arithmetic to check these identities for several unequal as well as equal subsystem dimensions. The qubit curves in Figure 3 are evaluations of the separate closed-form example (7.11); they do not use these Haar asymptotics.

References

- [1] Almheiri, A., Engelhardt, N., Marolf, D., and Maxfield, H., 2019, The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole, *JHEP* 12, 063.
- [2] Almheiri, A., Mahajan, R., Maldacena, J., and Zhao, Y., 2020, The Page curve of Hawking radiation from semiclassical geometry, *JHEP* 03, 149.
- [3] Almheiri, A., Hartman, T., Maldacena, J., Shaghoulian, E., and Tajdini, A., 2021, The entropy of Hawking radiation, *Rev. Mod. Phys.* 93, 035002.
- [4] Bianchi, E. and Donà, P., 2019, Typical entanglement entropy in the presence of a center: Page curve and its variance, *Phys. Rev. D* 100, 105010, arXiv:1904.08370.
- [5] Bousso, R. and Miyaji, M., 2024, Fluctuations in the entropy of Hawking radiation, *Phys. Rev. D* 109, 026006, arXiv:2307.13920.
- [6] Brandão, F. G. S. L., Harrow, A. W., and Horodecki, M., 2016, Local random quantum circuits are approximate polynomial-designs, *Commun. Math. Phys.* 346, 397–434, arXiv:1208.0692.
- [7] Choi, M.-D., 1975, Completely positive linear maps on complex matrices, *Linear Algebra and its Applications* 10, 285–290, doi:10.1016/0024-3795(75)90075-0.
- [8] Collins, B. and Śniady, P., 2006, Integration with respect to the Haar measure on unitary, orthogonal and symplectic group, *Commun. Math. Phys.* 264, 773–795.
- [9] Engelhardt, N. and Wall, A. C., 2015, Quantum extremal surfaces: holographic entanglement entropy beyond the classical regime, *JHEP* 01, 073.
- [10] Giraud, O., 2007, Distribution of bipartite entanglement for random pure states, *J. Phys. A: Math. Theor.* 40, 2793–2801, doi:10.1088/1751-8113/40/11/014, arXiv:quant-ph/0611285.
- [11] Hawking, S. W., 1976, Breakdown of predictability in gravitational collapse, *Phys. Rev. D* 14, 2460.

- [12] Hawking, S. W. and Page, D. N., 1983, Thermodynamics of black holes in anti-de Sitter space, *Commun. Math. Phys.* 87, 577.
- [13] Hayden, P. and Preskill, J., 2007, Black holes as mirrors: quantum information in random subsystems, *JHEP* 09, 120.
- [14] Lloyd, S. and Pagels, H., 1988, Complexity as thermodynamic depth, *Annals of Physics* 188, 186–213.
- [15] Lubkin, E., 1978, Entropy of an n-system from its correlation with a k-reservoir, *Journal of Mathematical Physics* 19, 1028–1031.
- [16] Maldacena, J. and Susskind, L., 2013, Cool horizons for entangled black holes, *Fortschr. Phys.* 61, 781.
- [17] Marolf, D. and Maxfield, H., 2020, Transcending the ensemble: baby universes, spacetime wormholes, and the order and disorder of black hole information, *JHEP* 08, 044.
- [18] Marolf, D., Wang, S., and Wang, Z., 2020, Probing phase transitions of holographic entanglement entropy with fixed area states, *Journal of High Energy Physics* 2020(12), 084.
- [19] Mathur, S. D., 2009, The information paradox: a pedagogical introduction, *Class. Quant. Grav.* 26, 224001.
- [20] Murthy, C. and Srednicki, M., 2019, Structure of chaotic eigenstates and their entanglement entropy, *Physical Review E* 100, 022131.
- [21] Nica, A. and Speicher, R., 2006, *Lectures on the Combinatorics of Free Probability*, Cambridge University Press.
- [22] Page, D. N., 1993, Average entropy of a subsystem, *Phys. Rev. Lett.* 71, 1291.
- [23] Papadodimas, K. and Raju, S., 2014, State-dependent bulk-boundary maps and black hole complementarity, *Phys. Rev. D* 89, 086010.
- [24] Penington, G., 2020, Entanglement wedge reconstruction and the information paradox, *JHEP* 09, 002.
- [25] Penington, G., Shenker, S., Stanford, D., and Yang, Z., 2022, Replica wormholes and the black hole interior, *JHEP* 03, 205, arXiv:1911.11977.
- [26] Saad, P., Shenker, S., and Stanford, D., 2019, JT gravity as a matrix integral, arXiv:1903.11115.
- [27] Schumacher, B., 1996, Sending entanglement through noisy quantum channels, *Phys. Rev. A* 54, 2614–2628, doi:10.1103/PhysRevA.54.2614, arXiv:quant-ph/9604023.
- [28] Sen, S., 1996, Average entropy of a subsystem, *Phys. Rev. Lett.* 77, 1–3, arXiv:hep-th/9601132.
- [29] Speicher, R., 1994, Multiplicative functions on the lattice of non-crossing partitions and free convolution, *Math. Ann.* 298, 611.
- [30] Voiculescu, D., 1991, Limit laws for random matrices and free products, *Invent. Math.* 104, 201.
- [31] Wang, J., 2023, Beyond islands: a free probabilistic approach, *Journal of High Energy Physics* 2023(10), 040.
- [32] Wuebben, B. J., 2026, The Page curve through the transition: an order-one correction to the diagonal approximation, SSRN Working Paper No. 6372958.
- [33] Wuebben, B. J., 2026, The Page curve in higher dimensions: phase transitions, critical phenomena, and finite-size broadening, SSRN Working Paper No. 6372978.