

From Hawking to Marchenko–Pastur: A Guide to the Page Curve and Beyond

Bernd Johannes Wuebben

New York, NY, USA

E-mail: wuebben@gmail.com

ABSTRACT: This tutorial develops the Page curve program from quantum entanglement and Hawking radiation to holographic entropy, replica wormholes, and random matrix theory. We derive the normalized random-state benchmark and its Marchenko–Pastur limit, explaining the roles of non-crossing partitions and spectral resolvents. We then distinguish averaging conditional entropies from calculating the entropy of a state with fluctuating energy or area. Gaussian broadening, finite reservoirs, and a quartic critical model illustrate how the preparation and the fluctuation law enter this distinction. Exact finite-dimensional examples separate averaged spectra from ensemble fluctuations and radiation data for one input from information about a quantum channel. Normalized annealed and quenched Rényi entropies have a common von Neumann limit, while their finite-order difference and connected replica moments measure additional ensemble information. Throughout, explicit calculations are distinguished from gravitational applications that require further assumptions. The resulting account provides a self-contained introduction to the established methods and to the remaining questions about microscopic information recovery.

KEYWORDS: Black hole information paradox, Page curve, replica wormholes, holographic entanglement, random matrices, free probability

Contents

Preface	1
I The Arena — Quantum Mechanics, Gravity, and the Paradox	3
1 Quantum Mechanics in One Hour	3
1.1 States, Hilbert spaces, and the superposition principle	3
1.2 Composite systems and entanglement	4
1.3 The density matrix	6
1.4 Von Neumann entropy and Rényi entropies	8
1.5 Unitarity and the no-information-loss principle	9
2 Black Holes and Hawking Radiation	11
2.1 Black holes in a nutshell	11
2.2 Black hole thermodynamics	12
2.3 Hawking radiation	14
2.4 The information paradox—sharp statement	15
3 Holography and Entanglement from Geometry	19
3.1 AdS/CFT: the holographic correspondence	19
3.2 Ryu–Takayanagi: entanglement from area	21
3.3 Quantum extremal surfaces: the upgrade	23
3.4 The island rule	25
3.5 What was achieved and what remains open	27
II The Gravitational Path Integral and Replica Wormholes	28
4 The Gravitational Path Integral	28
4.1 Path integrals: the basic idea	29
4.2 The genus expansion	30
4.3 Euclidean vs. Lorentzian: why we go to Euclidean signature	33
5 The Replica Trick and Replica Wormholes	36
5.1 The replica trick	36
5.2 The west-coast model	37
5.3 The master formula: permutations and topology	39
5.4 The two-saddle approximation and its failure	42
5.5 Summary: from permutations to distributions	44
III Free Probability and the Marchenko–Pastur Distribution	44

6	Free Probability and the Planar Random-State Limit	44
6.1	Classical vs. free probability	45
6.2	Non-crossing partitions, Catalan numbers, and Narayana numbers	46
6.3	Free cumulants and the moment-cumulant relation	47
6.4	The R -transform	48
6.5	The Marchenko–Pastur distribution	49
6.6	Analytic continuation: Carlson’s theorem	51
6.7	The west-coast scaling function	52
IV	Thermodynamic Fluctuations and Entropy Transitions	54
7	Entropy Averages and Gaussian Fluctuations	54
7.1	The fluctuating entropy difference	55
7.2	Canonical energy, thermodynamic entropy, and modular energy	55
7.3	The diagonal Page approximation	56
7.4	Gaussian formulas and their derivation	57
7.4.1	The averaged minimum	58
7.4.2	The averaged Page correction	58
7.4.3	Width, central deficit, and limiting cases	58
7.5	An order-one distinction from the planar spectrum	59
7.6	A useful identity beyond the Gaussian case	60
8	Thermodynamic Ensembles and Critical Fluctuations	60
8.1	Regular neutral AdS saddles	60
8.2	Fixed total energy and a finite reservoir	61
8.3	Negative heat capacity	63
8.4	Spinodal, coexistence, and continuous criticality	63
8.5	The finite-temperature critical crossover	65
8.6	How large must a critical reservoir be?	66
8.7	A normalized JT example	67
8.8	What these examples establish	68
9	Genus Corrections and the Topological Expansion	68
9.1	The genus-one permutation count	69
9.1.1	The closed-form formula	69
9.1.2	Small replica indices and direct enumeration	69
9.2	Normalized moments beyond the planar limit	70
9.2.1	Balanced dimensions	70
9.2.2	Why purity still has a correction	70
9.2.3	An independent entropy check	71
9.2.4	Unequal dimensions	71
9.3	The range of the genus expansion	71
9.4	Continuum handles, fluctuations, and QES corrections	72
9.4.1	The JT topology expansion	72
9.4.2	What the Akers–Penington result establishes	72
9.4.3	What can be compared between models	73

10 What Entropies and Spectra Determine	73
10.1 From a density matrix to an entropy	73
10.2 An averaged spectrum is not an ensemble of spectra	74
10.3 Two ensembles with exactly the same averaged spectrum	74
10.4 What a state design controls	76
10.5 Passing from moments to an entropy limit	77
10.6 A practical dictionary of the different questions	77
11 Rényi Averages and Entropy Fluctuations	78
11.1 The two Rényi averages	79
11.2 Averaging a spectrum and averaging a state	80
11.3 An exactly evaluated two-state ensemble	80
11.4 Normalized replica covariances	81
11.5 Concentration and its limits	82
11.6 Fluctuations within and between preparations	82
11.7 Application to gravitational ensembles	83
11.8 A four-replica calculation of purity fluctuations	83
11.9 Interpreting vanishing and nonvanishing differences	84
12 Comparing Gravitational Descriptions	85
12.1 What an entropy calculation establishes	85
12.2 The objects supplied by different approaches	86
12.2.1 Random-state and random-circuit models	86
12.2.2 Islands and replica wormholes	86
12.2.3 Microstate geometries and the fuzzball program	86
12.2.4 State-dependent interior operators	86
12.2.5 Geometric descriptions of entanglement	86
12.3 A comparison table with explicit status	87
12.4 Precision, normalization, and accessible observables	87
12.5 How to formulate a comparison that can be calculated	88
12.6 Reconstruction, decoding, and geometric descriptions	88
13 From Equilibrium States to Evaporation Channels	89
13.1 Entropy of a state and recovery of an input	89
13.2 An exact rising and falling entropy curve	90
13.3 An external reference distinguishes the channels	91
13.4 Consistency of successive radiation states	92
13.5 Fixed total energy does not fix the black-hole energy	92
13.6 A fluctuation width is not an evaporation timescale	93
13.7 What a dynamical comparison should specify	94
13.8 A two-qubit reference calculation	94
14 Further Calculations and Open Questions	95
14.1 Six concrete targets	95
14.2 How the checks fit together	96

V	Conclusion and Perspective	97
15	Conclusions	97
15.1	From entanglement to the replica spectrum	97
15.2	Thermodynamic preparation and entropy averaging	98
15.3	Spectra, ensembles, and information recovery	98
15.4	Further calculations	98
A	Notation and Conventions	99
A.1	Black hole variables	99
A.2	Hilbert space dimensions and the ratio γ	99
A.3	Combinatorial notation	99
A.4	Random matrix notation	99
A.5	Sign conventions and metric signature	99
B	The Non-Crossing Partition Lattice	99
B.1	Definition and first examples	100
B.2	The Narayana refinement	102
B.3	Lattice structure	103
B.3.1	The Möbius function	103
B.3.2	The Kreweras complement	104
B.4	Circle diagrams for $\text{NC}(4)$	104
B.5	The lattice diagram for $\text{NC}(3)$	104
B.6	The role of non-crossing partitions in gravity	105
C	JT Gravity: A Lightning Review	105
C.1	The action	105
C.2	The Schwarzian boundary mode	106
C.3	The disk partition function	106
C.4	The matrix integral dual	107
C.5	Why JT is the “hydrogen atom” of quantum gravity	107
D	Gaussian Broadening Integrals	107
D.1	The broadening integral	108
D.2	The Page curve as a broadened minimum	108
D.3	Non-Gaussian extensions	109
E	Random Matrix Theory: Marchenko–Pastur and Wishart Matrices	109
E.1	Wishart matrices	109
E.2	The Marchenko–Pastur distribution	109
E.3	Stieltjes transform derivation	110
E.3.1	Step 1: The resolvent identity	110
E.3.2	Step 2: The self-consistency relation	110
E.3.3	Step 3: Recovering the density	111
E.4	R -transform derivation	111
E.5	Edge universality and Tracy–Widom fluctuations	112

F	Selected Approaches to Black-Hole Information	112
F.1	Fuzzballs	112
F.2	Islands and the replica method	113
F.3	Random unitary circuits and fast scrambling	114
F.4	State-dependent operators	115
F.5	ER = EPR	115
F.6	Summary: what the proposals share and where they differ	116
G	Glossary of Key Terms	116

Preface

Purpose and prerequisites

The Page curve relates a question about quantum information to a calculation in gravity: how does the entropy of the radiation change while a black hole evaporates? Hawking’s calculation produces outgoing radiation entangled with degrees of freedom behind the horizon [38]. If this description is extrapolated through complete evaporation, the radiation remains mixed even when the collapsing matter was pure. Unitary evolution of a closed system preserves purity. The discrepancy is the black hole information paradox.

Modern calculations address this discrepancy using quantum extremal surfaces and replica wormholes [1, 4, 5, 27, 74, 76]. In controlled models they recover the turnover expected from a unitary description. Understanding this result requires several kinds of reasoning: the partial trace of a quantum state, black-hole thermodynamics, the saddle-point approximation, and the spectral statistics of random matrices. This tutorial develops the connections between these subjects.

We assume familiarity with linear algebra, probability, differentiation, and integration. Differential geometry is useful in the gravitational sections, but the combinatorial and spectral calculations can be followed independently. We introduce the quantum-information notation and give intermediate steps in the principal calculations. Examples progress from two qubits to large random matrices; gravitational interpretations follow the definitions of the mathematical objects.

Finite-dimensional examples are useful for two reasons. Normalization is unambiguous: each density matrix has trace one in every realization. They also distinguish questions sometimes conflated in a saddle-point treatment. Knowing the entropy differs from knowing the spectrum; knowing an averaged spectrum differs from knowing its fluctuations; and knowing the radiation state for one preparation differs from knowing the channel acting on arbitrary preparations. These distinctions remain relevant as dimensions grow.

What a Page curve establishes

For a pure state of radiation and a remaining black-hole system with effective dimensions k and L , respectively, the Schmidt decomposition gives

$$S(\rho_{\text{rad}}) = S(\rho_{\text{BH}}) \leq \log \min(k, L). \quad (0.1)$$

If the smaller subsystem is nearly maximally entangled with the larger one, this bound is nearly attained. Haar-random pure states provide the standard quantitative benchmark [71, 72]. The familiar rising and falling curve combines this typical-state statement with a model for the changing subsystem sizes. Unitarity alone gives neither typicality nor a unique intermediate-time trajectory. Complete evaporation into the specified radiation system does imply a pure final state when the initial state is pure and no inaccessible remnant or environment remains.

Replica calculations can establish more than this endpoint constraint. They compute moments of a reduced density matrix and, when spectral reconstruction or analytic continuation is controlled, its entropy. Their interpretation depends on the preparation, normalization, and whether the calculation represents a single theory or an ensemble average. A successful entropy calculation is therefore a substantial result without specifying every microscopic process that produces that entropy.

Established methods and the companion papers

The foundations used here belong to an extensive literature. Page’s random-state entropy formula [71], the Marchenko–Pastur law [58], and free-probability moment expansions [85, 95] provide the

random-matrix background. The gravitational replicas of Penington et al. [76] supply a setting for planar spectral calculations. The free-probability formulation of the weighted problem is developed by Wang [96], with related constructions in Cheng et al. [21]. Flat weights give the Marchenko–Pastur benchmark; planarity alone does not require that particular spectral law.

Thermodynamic corrections to subsystem entanglement were studied by Murthy and Srednicki [66]; smoothing of competing holographic entropy branches by fixed-area methods was developed by Marolf et al. [61]. We explain these ideas before discussing the additional comparisons in the author’s three companion papers. This tutorial contributes exposition and synthesis, rather than claiming the underlying methods as new results.

Paper I: spectral entropy and the diagonal approximation. *The Page Curve Through the Transition: An Order-One Correction to the Diagonal Approximation* [99] compares a conditional-entropy average with the entropy from the planar replica spectrum in its specified model. States associated with different energy or area sectors need not occupy orthogonal radiation subspaces. Averaging their conditional entropies need not give the entropy of their mixture. Under Paper I’s broad-distribution assumptions, the planar-minus-diagonal entropy tends to $-q \log q$, where q is the canonical weight of levels on the larger-black-hole side of the transition. We explain the approximation and its limits in Sections 6–8.

Paper II: fluctuation laws and critical scaling. *The Page Curve in Higher Dimensions: Phase Transitions, Critical Phenomena, and Finite-Size Broadening* [100] develops regular and critical fluctuations in a specified canonical random-state model. A quadratic stable saddle gives Gaussian fluctuations; at the charged AdS critical point a quartic saddle gives entropy fluctuations of order $S_c^{3/4}$, where S_c is the critical black-hole entropy. A reservoir heat capacity of order $S_c^{3/2}$ changes the limiting critical profile. A finite reservoir also changes the energy distribution even at fixed total energy. Applying these calculations to gravitational radiation additionally requires identifying the fluctuating entropy branches of that state.

Paper III: spectra, ensembles, and channels. *The Page Curve Is Not Enough: Universality, Ensemble Dependence, and Microscopic Ambiguity in Black Hole Information Recovery* [101] develops explicit finite-ensemble and channel comparisons. Equal averaged spectral measures can coexist with different ensemble fluctuations. Radiation trajectories for one input can agree between channels that differ on other inputs and on correlations with a reference. Normalized replica identities relate finite-order Rényi averages to these fluctuations. The examples delimit particular observables; they do not assign a universal list of predictions to entire proposals for quantum gravity.

Routes through the tutorial

The first three numbered sections introduce quantum states, Hawking radiation, and holographic entropy. Readers familiar with density matrices may begin with Section 2. The QES prescription in Section 3 motivates the replica calculation.

Sections 4 and 5 explain path integrals and replica sewing. The normalized random-state calculation makes the permutation sum concrete. Section 6 develops its non-crossing limit, resolvent, density, and entropy. This sequence can also be read as an introduction to the random-matrix part of the subject.

Sections 7 and 8 study conditional-entropy averages and their thermodynamic input. Gaussian calculations are followed by finite-reservoir and non-Gaussian examples. The Schwarzian example illustrates the role of the probability measure when energy is replaced by an entropy variable.

Section 9 distinguishes permutation counts, normalization corrections, and gravitational topology. Sections 10 and 11 compare spectral and ensemble information. Sections 12 and 13 apply the resulting questions to gravitational proposals and dynamical channels. Section 15 collects the conclusions. Appendices give notation, calculations, and supplementary background.

Complementary accounts include Harlow’s lectures on quantum information and holography [36], the island review by Almheiri et al. [2], the JT review by Mertens and Turiaci [64], and Raju’s discussion of the information problem [77]. Here we emphasize explicit spectral calculations and their state, ensemble, and dynamical interpretations.

Part I

The Arena — Quantum Mechanics, Gravity, and the Paradox

1 Quantum Mechanics in One Hour

This section is a rapid, self-contained tour of the quantum-mechanical concepts that will recur throughout the paper. The reader who is comfortable with Hilbert spaces, density matrices, and von Neumann entropy can skip ahead to Section 2; everyone else should read this section not for mathematical completeness—excellent textbooks abound [67, 80]—but for the *physical intuition* behind each formalism. We want to build a dictionary: states, entanglement, entropy, and unitarity. Every entry in that dictionary will reappear when we confront the black hole information paradox.

1.1 States, Hilbert spaces, and the superposition principle

A quantum system is described by a *Hilbert space* $\mathcal{H}$, a complex vector space equipped with an inner product. The pure states of the system are represented by rays in $\mathcal{H}$: unit vectors $|\psi\rangle \in \mathcal{H}$ with $\langle\psi|\psi\rangle = 1$, where an overall phase $e^{i\theta}|\psi\rangle$ describes the same physical state.

The *dimension* of $\mathcal{H}$ counts the number of perfectly distinguishable states accessible to the system. For a single qubit—a quantum two-level system—we have $\dim \mathcal{H} = 2$, with a basis $\{|0\rangle, |1\rangle\}$. For a spin- j particle, $\dim \mathcal{H} = 2j + 1$. For a system of n qubits, $\dim \mathcal{H} = 2^n$ —exponentially large in the number of constituents, a fact whose consequences pervade quantum information theory.

A crucial preview: Bekenstein–Hawking entropy as state-counting. Later, when we encounter a black hole with Bekenstein–Hawking entropy S_{BH} [12, 38], we will interpret this as a leading state-counting relation in a specified narrow energy window:

$$\log \dim \mathcal{H}_{\text{BH},E} \simeq S_{\text{BH}}(E). \quad (1.1)$$

The dimension refers to an effective microcanonical subspace, not the entire Hilbert space of a quantum field theory. Subleading corrections and the width of the energy window matter for a more precise count. A solar-mass black hole has $S_{\text{BH}} \sim 10^{77}$, giving an enormous effective number of states. For a pure joint state, the radiation entropy cannot exceed the logarithm of the remaining subsystem’s dimension. Near-saturation of this bound is an additional typicality statement, quantified by Page’s random-state calculation [71].

Superposition. The central structural axiom of quantum mechanics is the *superposition principle*: if $|0\rangle$ and $|1\rangle$ are both legitimate states, then so is any linear combination

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad |\alpha|^2 + |\beta|^2 = 1. \quad (1.2)$$

It is essential to grasp what this does *not* mean. The state $|\psi\rangle$ is *not* “the system is in state $|0\rangle$ with probability $|\alpha|^2$ and in state $|1\rangle$ with probability $|\beta|^2$.” That description corresponds to *classical ignorance*—a statistical mixture. The superposition $|\psi\rangle$ is a *definite* state of the system, as definite as $|0\rangle$ or $|1\rangle$ themselves. The distinction between quantum superposition and classical ignorance will become critical in Section 1.3, where we formalize both using density matrices, and again in Section 2, where the question of whether Hawking radiation is in a pure superposition or a classical mixture is precisely the information paradox.

The Born rule. Given a state $|\psi\rangle$ and an orthonormal basis $\{|i\rangle\}$ for $\mathcal{H}$, the probability of obtaining outcome i in a measurement is

$$p_i = |\langle i|\psi\rangle|^2. \quad (1.3)$$

For the superposition in Eq. (1.2), the Born rule gives $p_0 = |\alpha|^2$ and $p_1 = |\beta|^2$. The normalisation condition $|\alpha|^2 + |\beta|^2 = 1$ ensures $\sum_i p_i = 1$. We emphasise: the Born rule is a statement about the *statistics of measurement outcomes*, not about the ontological status of the state before measurement.

1.2 Composite systems and entanglement

When two quantum systems A and B are considered jointly, the Hilbert space of the composite system is the *tensor product*

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B. \quad (1.4)$$

If $\dim \mathcal{H}_A = d_A$ and $\dim \mathcal{H}_B = d_B$, then $\dim \mathcal{H} = d_A \cdot d_B$. The tensor product structure is the mathematical origin of all the phenomena—entanglement, monogamy, information scrambling—that make quantum mechanics qualitatively different from classical physics.

Product states and entangled states. A *product state* is one that can be written as $|\psi\rangle = |\phi\rangle_A \otimes |\chi\rangle_B$. In a product state, A and B are uncorrelated: measuring one reveals nothing about the other. A pure state that is *not* a product state is called *entangled*. A mixed state is separable if it is a convex combination of product states; a non-product mixed state need not be entangled.

The Bell state. The simplest and most important entangled state is the *Bell state* (also called an EPR pair), involving two qubits:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \quad (1.5)$$

where $|00\rangle \equiv |0\rangle_A \otimes |0\rangle_B$ and similarly for $|11\rangle$. Let us unpack the content of this state carefully, because the Bell state is the prototype for the entanglement structure that drives the information paradox.

- (i) **The total system is in a pure state.** We know everything there is to know about the two-qubit system: it is in the definite state $|\Phi^+\rangle$. Its entropy is zero.
- (ii) **Subsystem A is maximally uncertain.** If we restrict attention to qubit A alone, the Born rule gives

$$p(0_A) = \|\langle 0_A | \otimes \mathbb{1} | \Phi^+ \rangle\|^2 = \frac{1}{2}, \quad p(1_A) = \|\langle 1_A | \otimes \mathbb{1} | \Phi^+ \rangle\|^2 = \frac{1}{2}.$$

Qubit A , viewed in isolation, looks completely random. We will formalise this statement using the reduced density matrix in Section 1.3.

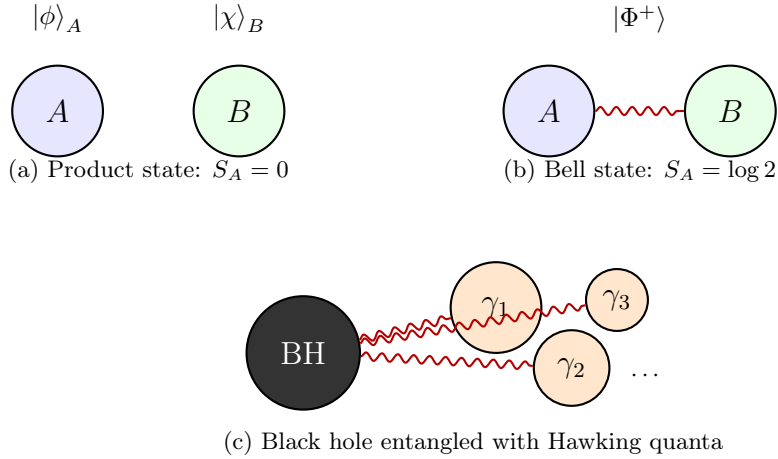


Figure 1. Entanglement illustrated. (a) In a product state, each subsystem is independently well-defined and has zero entropy. (b) In the Bell state $|\Phi^+\rangle$, the total system is pure, but each subsystem is maximally mixed with entropy $\log 2$ (wavy line denotes entanglement). (c) A black hole entangled with its Hawking radiation: each emitted quantum γ_i is entangled with the interior, making the radiation appear thermal when examined in isolation.

- (iii) **Measurements on A and B are perfectly correlated.** If we measure A and obtain $|0\rangle$, then a subsequent measurement on B will always yield $|0\rangle$, and similarly for $|1\rangle$. Correlations in a single fixed basis could still be mimicked by shared classical randomness; what no such classical account can reproduce are the correlations observed when A and B are measured in several non-commuting bases—this is the content of Bell’s theorem [13], to which we return below.
- (iv) **Entanglement produces ignorance from knowledge.** The total state is pure (zero entropy), yet either subsystem appears maximally mixed (maximum entropy $\log 2$). The ignorance about subsystem A is not due to any deficiency in our knowledge of the total system—it is an irreducible consequence of entanglement.

Point (iv) is perhaps the single most important conceptual message of this section. When a black hole is entangled with its Hawking radiation, the radiation *by itself* looks thermal—not because we have lost information, but because the information is stored in the correlations between the radiation and the black hole interior. Whether that information can eventually be recovered from the radiation alone is the information paradox.

Monogamy of entanglement. A qubit that is maximally entangled with one system *cannot* be entangled with any other system. More precisely, if A and B share a Bell state, then A is completely uncorrelated with any third system C . This is the *monogamy of entanglement*, and it is a theorem, not a choice [23].

Monogamy is what makes the black hole information paradox genuinely sharp. If the entropy of the accumulated radiation decreases after an emission, the new radiation has negative conditional entropy relative to the earlier radiation. A smooth horizon also requires the outgoing mode to have the appropriate near-vacuum correlations with interior modes. Treating these as independent subsystems

leads to competing entanglement requirements. Monogamy forbids both at once—the “AMPS firewall” argument [6] makes this tension precise.

EPR and nonlocality. The correlations in a Bell state are “nonlocal” in the sense that they cannot be reproduced by any *local hidden variable* theory—this is Bell’s theorem [13]. However, these correlations do not permit faster-than-light signalling; the reduced state of each subsystem is the same regardless of what measurement is performed on the other. The tension between nonlocal correlations and the impossibility of signalling is one of the deep conceptual puzzles of quantum mechanics, but it is not our primary concern in this paper. What matters for us is the entanglement structure itself and the entropy it generates for subsystems.

1.3 The density matrix

So far we have described quantum states as vectors $|\psi\rangle \in \mathcal{H}$. This description is complete when the system is in a *pure* state. But in practice—and this is the norm in black hole physics—we often deal with systems that are either (a) part of a larger entangled system, or (b) prepared with classical uncertainty about which quantum state they are in. In both cases, the appropriate description is the *density matrix* (or density operator).

Pure states. Given a pure state $|\psi\rangle$, the density matrix is the projector

$$\rho = |\psi\rangle\langle\psi|. \quad (1.6)$$

This is a Hermitian, positive semi-definite, trace-one operator on $\mathcal{H}$. A pure-state density matrix satisfies

$$\rho^2 = \rho, \quad \text{tr}(\rho^2) = 1, \quad S(\rho) = 0, \quad (1.7)$$

where $S(\rho)$ is the von Neumann entropy defined in Section 1.4. Purity ($\rho^2 = \rho$) is the density-matrix equivalent of “we know the state exactly.”

Mixed states. Suppose we have classical uncertainty: the system is in state $|\psi_i\rangle$ with probability p_i (with $p_i \geq 0$, $\sum_i p_i = 1$). The density matrix is

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|. \quad (1.8)$$

If the occupied vectors are not all the same ray, this gives a mixed state, with $\rho^2 \neq \rho$ and $\text{tr}(\rho^2) < 1$. The von Neumann entropy is strictly positive: $S(\rho) > 0$. The inequality $\text{tr}(\rho^2) < 1$ is a useful diagnostic of mixedness and is directly related to the second Rényi entropy (Section 1.4).

The reduced density matrix and the partial trace. Consider a bipartite pure state $|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$. The *reduced density matrix* of subsystem A is obtained by *tracing out* B :

$$\rho_A = \text{tr}_B(|\psi\rangle\langle\psi|) = \sum_j \langle j|_B |\psi\rangle\langle\psi| |j\rangle_B, \quad (1.9)$$

where $\{|j\rangle_B\}$ is any orthonormal basis for $\mathcal{H}_B$. The reduced density matrix ρ_A encodes *all* information that is accessible to an observer who can only make measurements on A .

Let us compute the reduced density matrix for the Bell state (1.5). Writing $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B)$, we have

$$\begin{aligned} \rho_A &= \text{tr}_B(|\Phi^+\rangle\langle\Phi^+|) \\ &= \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{\mathbb{I}_A}{2}, \end{aligned} \quad (1.10)$$

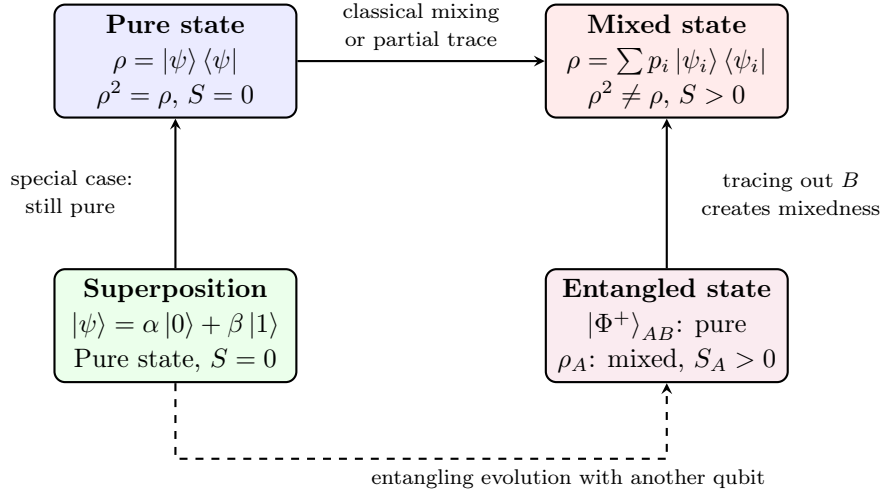


Figure 2. The taxonomy of quantum states. A superposition is a pure state with zero entropy. A mixture has positive entropy and arises either from classical ignorance or from tracing out part of an entangled system. The key insight: entanglement of a pure bipartite state forces each subsystem to be mixed. This mechanism is exactly what makes Hawking radiation appear thermal.

the *maximally mixed state* on qubit A . By symmetry, $\rho_B = \mathbb{I}_B/2$ as well.

Key property: equal subsystem entropies. For any pure state $|\psi\rangle_{AB}$ of a bipartite system, the reduced density matrices ρ_A and ρ_B have the same nonzero eigenvalues. Consequently,

$$S(\rho_A) = S(\rho_B). \quad (1.11)$$

This is a straightforward consequence of the Schmidt decomposition and will play a starring role in the Page curve: the entropy of the radiation equals the entropy of the remaining black hole, provided the total state is pure.

The maximally mixed state. On a d -dimensional Hilbert space, the maximally mixed state is

$$\rho_{\max} = \frac{\mathbb{I}}{d}, \quad S(\rho_{\max}) = \log d. \quad (1.12)$$

This is the state of maximum ignorance—every outcome in any orthonormal-basis measurement is equally likely. It is the density matrix describing one half of a maximally entangled state of two d -dimensional systems.

Pure vs. mixed vs. superposition: the critical distinction. Figure 2 summarises the relationships. A superposition $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ is a *pure* state—it has a density matrix $\rho = |\psi\rangle\langle\psi|$ with $\rho^2 = \rho$ and $S = 0$. A mixture $\rho = |\alpha|^2|0\rangle\langle 0| + |\beta|^2|1\rangle\langle 1|$ has $S > 0$ when both coefficients are nonzero. Physically, the superposition displays interference effects (e.g., in a double-slit experiment) while the mixture does not.

This distinction is the heart of the information paradox. Hawking’s original calculation [38] yields radiation described by a thermal density matrix—a *mixed* state with $S > 0$. If the black hole formed from a pure initial state, unitarity demands that the final state (after complete evaporation) be pure.

Resolving the paradox means explaining how a process that appears to produce mixed-state radiation can ultimately be consistent with purity.

1.4 Von Neumann entropy and Rényi entropies

We now introduce the entropy measures that quantify the amount of entanglement or ignorance encoded in a density matrix.

Von Neumann entropy. The von Neumann entropy of a density matrix ρ with eigenvalues $\{\lambda_i\}$ is

$$S_{\text{vN}}(\rho) = -\text{tr}(\rho \log \rho) = -\sum_i \lambda_i \log \lambda_i, \quad (1.13)$$

where $0 \log 0 \equiv 0$ by continuity. (Throughout this paper, “log” denotes the natural logarithm unless otherwise stated.) Key properties:

- (a) $S \geq 0$, with equality if and only if ρ is pure.
- (b) $S \leq \log d$ (where $d = \dim \mathcal{H}$), with equality if and only if ρ is maximally mixed.
- (c) For a bipartite pure state, $S(\rho_A) = S(\rho_B)$ (Eq. (1.11)).
- (d) *Strong subadditivity*: $S(\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC})$ [51].
- (e) *Araki–Lieb inequality*: $|S(\rho_A) - S(\rho_B)| \leq S(\rho_{AB})$ [7].

The von Neumann entropy is the quantum analogue of the Shannon entropy. When ρ describes one half of a bipartite pure state, $S(\rho)$ measures the *entanglement entropy*—the amount of quantum correlation between the two halves.

Rényi entropies. The Rényi entropies are a one-parameter family that generalises the von Neumann entropy:

$$S_n(\rho) = \frac{1}{1-n} \log \text{tr}(\rho^n), \quad n > 0, n \neq 1. \quad (1.14)$$

In the limit $n \rightarrow 1$, L’Hôpital’s rule recovers the von Neumann entropy:

$$\lim_{n \rightarrow 1} S_n(\rho) = S_{\text{vN}}(\rho). \quad (1.15)$$

Several special cases deserve mention.

- $S_{0+}(\rho) = \log(\mathbf{rank} \rho)$: the Hartley entropy, counting the number of nonzero eigenvalues (defined as the limit $n \rightarrow 0^+$ of Eq. (1.14)); the notation S_{0+} here is distinct from the JT topological entropy S_0 .
- $S_2 = -\log \text{tr}(\rho^2)$: the collision entropy. Here $\text{tr}(\rho^2)$ is the *purity* of ρ ; a pure state has $\text{tr}(\rho^2) = 1$ and $S_2 = 0$, while a maximally mixed state in d dimensions has $\text{tr}(\rho^2) = 1/d$ and $S_2 = \log d$. The second Rényi entropy is experimentally accessible [44] and computationally convenient.
- $S_\infty = -\log \lambda_{\text{max}}$: the min-entropy, determined by the largest eigenvalue.

Why the replica trick works. The quantity $\text{tr}(\rho^n)$ for *integer* $n \geq 2$ can be computed by considering n identical copies (“replicas”) of the system and sewing them together in a cyclic pattern. In the gravitational context, this procedure has a beautiful geometric interpretation: each replica is a copy of the spacetime, and $\text{tr}(\rho^n)$ is computed by a path integral over n -fold branched cover geometries. The sum over topologically distinct ways of connecting these replicas gives rise to *replica wormholes*, which

are the key to deriving the Page curve from semiclassical gravity. We will develop this idea in detail in Section 5.

The strategy is:

1. Compute $\text{tr}(\rho^n)$ for integer $n \geq 2$ using the replica method.
2. Analytically continue in n to a neighbourhood of $n = 1$.
3. Take the limit $n \rightarrow 1$ to extract S_{vN} .

This procedure is the *replica trick*. Integer values alone do not specify an analytic continuation without additional assumptions. For a known finite-dimensional density matrix, its spectrum directly defines $\text{tr} \rho^n$ for real $n > 0$. In a saddle-point calculation one must also control continuation and the order of limits. Carlson-type uniqueness conditions can be checked for the flat-weight spectral benchmark discussed in Section 6; they are not a general theorem about the full gravitational replica path integral.

The entanglement spectrum. The full set of eigenvalues $\{\lambda_i\}$ of a reduced density matrix is called the *entanglement spectrum*. The von Neumann entropy is a single number extracted from this spectrum; the Rényi entropies form a one-parameter family. But the spectrum itself contains strictly more information than any single entropy.

Section 6 derives the Marchenko–Pastur limit for a normalized flat-weight random state. In a gravitational application, identifying this law requires the corresponding preparation and weights. Later chapters distinguish a spectrum in one realization, an averaged spectral density, and correlations between eigenvalues in an ensemble. None of these spectra includes eigenvectors, and therefore none alone specifies every observable or the dynamics of a state.

1.5 Unitarity and the no-information-loss principle

The time evolution of an isolated quantum system is governed by a *unitary* operator:

$$|\psi(t)\rangle = U(t) |\psi(0)\rangle, \quad U^\dagger U = U U^\dagger = \mathbb{I}. \quad (1.16)$$

Unitarity is an extraordinarily powerful constraint. It preserves:

- (i) **Inner products:** $\langle \phi(t) | \psi(t) \rangle = \langle \phi(0) | \psi(0) \rangle$. Two states that start out distinguishable remain distinguishable forever.
- (ii) **Purity:** If $\rho(0) = |\psi(0)\rangle \langle \psi(0)|$ is pure, then $\rho(t) = U(t) \rho(0) U^\dagger(t)$ is also pure.
- (iii) **Entropy:** $S(\rho(t)) = S(\rho(0))$. The von Neumann entropy of the total system is a constant of the motion.
- (iv) **Information:** Since U is invertible, the initial state can in principle be recovered when the unitary is known and its inverse can be implemented: $|\psi(0)\rangle = U^\dagger(t) |\psi(t)\rangle$. No information is ever destroyed.

The preservation of information under unitary evolution is sometimes called the *no-information-loss principle*. In the context of the gravitational path integral, it is the statement that quantum gravity—whatever its ultimate formulation—must be a unitary theory.

The key constraint for black holes. Suppose a black hole forms from the collapse of a pure quantum state $|\psi_{\text{initial}}\rangle$ —say, a shell of photons in a definite quantum state. The black hole then evaporates completely via Hawking radiation. If quantum mechanics is unitary, then the final state of the radiation must also be pure:

$$|\psi_{\text{initial}}\rangle \xrightarrow{U} |\psi_{\text{radiation}}\rangle, \quad S(\rho_{\text{radiation}}) = 0. \quad (1.17)$$

This is a strong statement. It says that all the information that went into forming the black hole must come out in the Hawking radiation. The radiation may look thermal to any practical measurement, but it must be in a *globally pure* state—an astronomically complicated superposition whose correlations encode the initial state.

Why physicists take unitarity so seriously. One might ask: why not simply accept that black holes destroy information? Several lines of argument militate against this:

- (a) **Energy conservation violation.** Banks, Susskind, and Peskin [10] analyzed serious tensions between local nonunitary modifications and energy conservation. The conclusion depends on the form of the proposed evolution; it is not a theorem that every nonunitary quantum channel violates energy conservation.
- (b) **AdS/CFT.** In the AdS/CFT correspondence [54], the boundary conformal field theory is manifestly unitary. If the correspondence is exact, then the bulk gravitational theory—including black hole formation and evaporation—must also be unitary.
- (c) **Universal structure of quantum mechanics.** Unitary Hamiltonian evolution is the standard description of a closed system. Open systems admit nonunitary, trace-preserving channels while retaining consistent probabilities. Fundamental information loss would therefore require changing the closed-system dynamical postulate, rather than following merely from a different description of a subsystem.

The tension between Hawking’s apparently thermal radiation (mixed state, $S > 0$) and the demand of unitarity (pure final state, $S = 0$) is the black hole information paradox. Resolving this tension—and understanding the detailed mechanism by which information escapes—is the driving question of this paper.

Summary. We have assembled the following toolkit:

- Quantum states live in Hilbert spaces. A black hole’s entropy estimates the logarithm of the number of states in a suitable energy window.
- Composite systems have tensor-product Hilbert spaces. Entanglement is the rule, not the exception.
- The Bell state illustrates the core phenomenon: a globally pure state whose subsystems are maximally mixed. Entanglement creates ignorance.
- The density matrix describes both pure and mixed states. The partial trace over one subsystem produces the reduced density matrix.
- For a pure bipartite state, subsystem von Neumann entropy quantifies entanglement. Rényi entropies give additional spectral information; eigenvectors and channel data contain information beyond a spectrum.
- Unitarity preserves purity and information. A pure initial state must evolve to a pure final state.

Armed with this language, we are ready to introduce black holes and confront the paradox directly in Section 2.

2 Black Holes and Hawking Radiation

The previous section established the quantum-mechanical toolkit—density matrices, entanglement entropy, and the unitarity constraint—that we will need throughout this paper. We now turn to the gravitational arena in which all of those concepts are pushed to their breaking point: the physics of black holes. Our goal in this section is threefold. First, we introduce the classical black hole as a solution of Einstein’s equations, emphasizing the features—event horizons, the no-hair theorem—that make the information problem conceivable. Second, we develop the thermodynamic picture due to Bekenstein and Hawking that assigns entropy and temperature to a black hole. Third, we explain Hawking’s celebrated 1975 calculation of thermal radiation and trace its logic to its alarming conclusion: a violation of unitarity, the central principle of quantum mechanics. By the end of this section, the reader should understand the information paradox as a *precise* conflict between general relativity and quantum mechanics, and see how a Page curve tests the resulting entropy calculation.

2.1 Black holes in a nutshell

In an asymptotically flat classical spacetime, a black hole is the region that cannot send a signal to future null infinity. The boundary of that region is the **event horizon**, a null hypersurface that acts as a one-way membrane: signals can cross inward but never outward.

The Schwarzschild solution. The simplest black hole is the Schwarzschild solution, discovered in 1916 [82]. For a non-rotating, uncharged mass M , the metric in Schwarzschild coordinates reads

$$ds^2 = -\left(1 - \frac{r_s}{r}\right) dt^2 + \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 + r^2 d\Omega_2^2, \quad (2.1)$$

where $r_s = 2G_N M$ is the Schwarzschild radius (in units with $c = 1$; we set $\hbar = c = k_B = 1$ throughout this section) and $d\Omega_2^2$ is the round metric on S^2 . The coordinate singularity at $r = r_s$ is the event horizon; it can be removed by passing to Kruskal–Szekeres coordinates, confirming that nothing physically singular happens at the horizon itself. The true curvature singularity at $r = 0$ is physical and signals the breakdown of classical general relativity.

The event horizon as a causal boundary. The defining property of the horizon is *causal*: it is the boundary of the past of future null infinity, $\mathcal{H}^+ = \partial J^-(\mathcal{I}^+)$. An observer freely falling through the horizon notices nothing locally dramatic—tidal forces are small for a sufficiently massive black hole, scaling as M^{-2} . Yet an observer hovering outside sees the infalling object asymptotically redshift and freeze at the horizon, never quite crossing it. This tension between the local and global descriptions will reappear forcefully when we discuss the information paradox.

The no-hair theorem. One of the most striking features of classical black holes is their extreme simplicity. The *no-hair theorem* (more precisely, a collection of uniqueness results [19, 45, 46]) states, under suitable regularity, asymptotic-flatness, and connected-horizon assumptions, that a stationary black hole in four-dimensional Einstein–Maxwell theory is characterized by its mass M , electric charge Q , and angular momentum J . No other internal structure is visible from outside the horizon. The most general such solution is the Kerr–Newman metric.

This radical simplicity has a direct bearing on the information problem. If a black hole is formed from the gravitational collapse of, say, a library of books, all information about the contents of those books is hidden behind the horizon. The external spacetime is indistinguishable from one formed by the collapse of an equal mass of iron filings. Where did the information go? Classically, the answer is

straightforward: it fell behind the horizon and remains there. But as we shall see, Hawking’s discovery of black hole radiation reopens the question in a dramatic way.

Penrose diagrams. The causal structure of a black hole spacetime is most transparently displayed in a *Penrose (conformal) diagram* [75], which compactifies the infinite spacetime onto a finite region while preserving the light-cone structure. Figure 3 shows the Penrose diagram for a black hole formed from gravitational collapse and subsequently evaporating via Hawking radiation. The key features are: (i) the collapsing matter (shaded region) that forms the horizon; (ii) the event horizon itself, a null surface; (iii) the singularity at the top; (iv) future null infinity $\mathcal{I}^+$, where the Hawking radiation is collected; and (v) the endpoint of evaporation, after which the spacetime returns to flat Minkowski space (in the idealized picture).

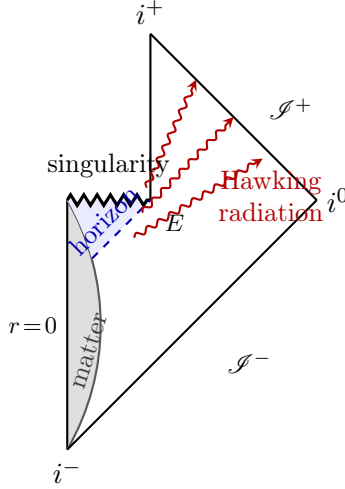


Figure 3. Penrose diagram for a black hole formed by gravitational collapse and evaporating via Hawking radiation. Light travels at 45° . The gray region is the collapsing matter (between the stellar surface and the $r=0$ axis). The dashed blue line is the event horizon; the zigzag line is the spacelike singularity. After complete evaporation at E , the spacetime returns to flat Minkowski geometry. Hawking quanta (wavy red arrows) propagate to future null infinity $\mathcal{I}^+$. The light-blue interior is causally disconnected from $\mathcal{I}^+$ —the heart of the information paradox.

2.2 Black hole thermodynamics

The first hint that black holes are more than classical geometric objects came from a set of formal analogies—later understood to be exact identities—between the laws governing black hole mechanics and the laws of thermodynamics. These were developed in the early 1970s through the work of Bekenstein [12], Bardeen, Carter, and Hawking [11], and others.

The four laws of black hole mechanics

The analogy is summarized in Table 1.

The second law—Hawking’s area theorem [37]—states that, under the null energy condition and appropriate global causal assumptions, event-horizon area cannot decrease in classical evolution. This

Table 1. The four laws of black hole mechanics and their thermodynamic counterparts. The surface gravity κ plays the role of temperature, and the horizon area A plays the role of entropy.

Law	Thermodynamics	Black hole mechanics
Zeroth	T is constant in equilibrium	κ is constant on the horizon
First	$dE = T dS + \text{work}$	$dM = \frac{\kappa}{8\pi G_N} dA + \Omega dJ + \Phi dQ$
Second	$\delta S \geq 0$	$\delta A \geq 0$ (area theorem)
Third	$T = 0$ not reachable	$\kappa = 0$ not reachable

was the clue that first led Bekenstein to propose that a black hole carries entropy proportional to its area.

Bekenstein–Hawking entropy

Bekenstein [12] proposed that a black hole carries an entropy proportional to its horizon area, with a coefficient he could estimate only up to a factor of order one; Hawking’s temperature calculation [38] then fixed the coefficient, yielding the *Bekenstein–Hawking entropy*

$$S_{\text{BH}} = \frac{A}{4G_N}, \quad (2.2)$$

where A is the area of the event horizon and we work in units with $\hbar = c = k_B = 1$. For a Schwarzschild black hole of mass M , the horizon area is $A = 16\pi G_N^2 M^2$, giving

$$S_{\text{BH}} = 4\pi G_N M^2. \quad (2.3)$$

The Bekenstein–Hawking entropy is enormous. A solar-mass black hole has $S_{\text{BH}} \sim 10^{77}$, vastly exceeding the entropy of the star from which it formed. This is one of several hints that the number of states in a narrow black-hole energy window satisfies, to leading semiclassical order,

$$\log \dim \mathcal{H}_{\text{BH},E} \simeq S_{\text{BH}}(E), \quad (2.4)$$

corresponding to an exponentially large state count, despite the no-hair theorem telling us that the *external* description involves only three parameters. The tension between the simplicity of the exterior and the enormous number of microstates suggested by the entropy is a recurring theme.

Why area and not volume? In ordinary thermodynamic systems, entropy is an extensive quantity scaling with the volume. The Bekenstein–Hawking entropy, by contrast, scales with the *area* of the horizon. This is the first appearance of the **holographic principle**—the idea that the maximum entropy in a region of space scales with the boundary area rather than the bulk volume [89, 92]. We will see this principle realized concretely in the AdS/CFT correspondence (Section 3).

Hawking temperature

Having established the analogy between κ and temperature, the natural question is: what is the *actual* temperature? Hawking’s 1975 calculation [38] answered this definitively. A black hole radiates thermally at the **Hawking temperature**

$$T_H = \frac{\kappa}{2\pi}, \quad (2.5)$$

where κ is the surface gravity. For a Schwarzschild black hole, $\kappa = 1/(4G_N M)$, so

$$T_H = \frac{1}{8\pi G_N M}. \quad (2.6)$$

Note the inverse relationship with mass: *smaller black holes are hotter*. This has the dramatic consequence that Hawking radiation is a runaway process. As the black hole radiates and loses mass, it grows hotter, radiates faster, and eventually—in the absence of new physics—undergoes complete evaporation in a finite time.

For Einstein gravity at leading semiclassical order, the thermodynamic identifications are

$$T = T_H = \frac{\kappa}{2\pi}, \quad S = S_{\text{BH}} = \frac{A}{4G_N}. \quad (2.7)$$

These satisfy the first law in the form

$$dM = T_H dS_{\text{BH}} + \Omega dJ + \Phi dQ, \quad (2.8)$$

confirming the consistency of the entire thermodynamic framework.

A sanity check: black hole evaporation time. The Stefan–Boltzmann law gives the power radiated by a body of temperature T and area A as $P \sim \sigma T^4 A$. Using $T_H \propto M^{-1}$ and $A \propto M^2$, one finds $P \propto M^{-2}$, so $dM/dt \propto -M^{-2}$, giving an evaporation time

$$t_{\text{evap}} \sim G_N^2 M^3. \quad (2.9)$$

For a solar-mass black hole this is $\sim 10^{67}$ years [70], far longer than the age of the universe. The information paradox is therefore a conceptual rather than observational crisis—but no less sharp for that.

2.3 Hawking radiation

We now turn to the physical mechanism behind Hawking radiation and, crucially, the entanglement structure it creates. It is this entanglement structure that drives the information paradox.

The mechanism: particle creation in curved spacetime

Hawking’s original derivation [38] proceeds by studying a free quantum field on the background of a collapsing star that forms a black hole. The key insight is that the notion of “particle” is observer-dependent in curved spacetime. The vacuum state defined by an observer in the far past (before the black hole forms) is not the same as the vacuum defined by an observer in the far future (after the horizon has settled down). The two sets of mode functions are related by a **Bogoliubov transformation**:

$$\hat{b}_\omega = \int d\omega' (\alpha_{\omega\omega'} \hat{a}_{\omega'} + \beta_{\omega\omega'} \hat{a}_{\omega'}^\dagger), \quad (2.10)$$

where $\hat{a}_\omega$ and $\hat{b}_\omega$ are annihilation operators for the “in” and “out” mode decompositions, and $\alpha_{\omega\omega'}$ and $\beta_{\omega\omega'}$ are the Bogoliubov coefficients. A non-vanishing β coefficient means that the “in” vacuum contains “out” particles.

The central result of Hawking’s calculation is that the number spectrum of outgoing particles is thermal:

$$\langle \hat{N}_\omega \rangle = \frac{1}{e^{\omega/T_H} - 1} \quad (\text{bosons}), \quad (2.11)$$

a Planck distribution at the Hawking temperature T_H . (For fermions, the minus sign becomes a plus sign, giving a Fermi–Dirac distribution. The full result multiplies this occupation number by a frequency-dependent *greybody factor* Γ_ω , the transmission probability through the potential barrier outside the horizon. It modifies the spectrum measured at infinity while retaining the Hawking temperature in its occupation factor.) This thermal spectrum depends only on the mass M of the black hole (and its charge and spin, if present). At this level of the calculation it does not distinguish microscopic initial states with the same macroscopic parameters.

The entanglement structure of Hawking pairs

The thermal nature of the radiation has a deeper origin in *entanglement*. The Bogoliubov transformation creates particles in pairs: for each outgoing Hawking quantum at frequency ω escaping to infinity, there is a partner quantum at frequency ω that falls behind the horizon. The joint state of a single Hawking pair takes the form

$$|\Psi_\omega\rangle = \frac{1}{\sqrt{Z}} \sum_{n=0}^{\infty} e^{-n\omega/(2T_H)} |n\rangle_{\text{out}} \otimes |n\rangle_{\text{in}} , \quad (2.12)$$

where $Z = (1 - e^{-\omega/T_H})^{-1}$ is the normalization and $|n\rangle_{\text{out}}, |n\rangle_{\text{in}}$ are n -particle states of the outgoing and infalling modes, respectively. This is a two-mode squeezed state: tracing over the partner behind the horizon produces a thermal density matrix for the exterior mode,

$$\rho_{\text{out}} = \text{tr}_{\text{in}} |\Psi_\omega\rangle\langle\Psi_\omega| = \frac{1}{Z} \sum_n e^{-n\omega/T_H} |n\rangle\langle n| , \quad (2.13)$$

which is exactly a thermal (Gibbs) state.

Now consider the full radiation field, consisting of many such modes emitted over the black hole’s lifetime. Each outgoing mode is entangled with its partner behind the horizon, and these partners are hidden from any external observer. As the black hole evaporates and emits more and more Hawking quanta, the total entanglement entropy of the radiation grows:

$$S_{\text{vN}}(\rho_{\text{rad}}) \sim N_{\text{emitted}} \sim S_{\text{BH}}^{(\text{initial})} - S_{\text{BH}}(t) , \quad (2.14)$$

where N_{emitted} counts emitted quanta and the relation is parametric, not an equality of coefficients. Entropy production in the outgoing radiation depends on the species and greybody factors. Extrapolating the independent-pair approximation gives a monotonically increasing radiation entropy of order the initial black-hole entropy. The final evaporation stage is outside the regime of that semiclassical approximation.

This is the crux of the problem. If the black hole evaporates completely, the partners behind the horizon are simply gone—there is no subsystem left to purify the radiation. The radiation is left in a *mixed* state even though we assumed the initial state was pure.

2.4 The information paradox—sharp statement

We are now in a position to state the black hole information paradox precisely. The argument proceeds in three steps.

Step 1: Pure initial state. Consider a black hole formed from the gravitational collapse of matter in a pure quantum state $|\Psi_0\rangle$. (One can imagine collapsing a shell of photons in a carefully prepared quantum state.) The total system—black hole plus radiation—is described by a state in a Hilbert space $\mathcal{H}_{\text{BH}} \otimes \mathcal{H}_{\text{rad}}$. At early times, before any radiation has been emitted, the radiation Hilbert space is trivial and $S_{\text{vN}}(\rho_{\text{rad}}) = 0$.

Step 2: Hawking’s calculation. As the black hole evaporates, it emits thermal radiation. The entanglement between each outgoing Hawking quantum and its infalling partner causes $S_{\text{vN}}(\rho_{\text{rad}})$ to increase monotonically, as derived in (2.14).

Step 3: The final state. If the independent-pair description is extrapolated through complete evaporation, it predicts a mixed radiation state. In an idealized wave-packet description without greybody scattering, its schematic form is

$$\rho_{\text{rad}}^{(\text{final})} \approx \bigotimes_j \frac{e^{-\omega_j N_j / T_H(t_j)}}{Z_j}, \quad (2.15)$$

where j labels packets emitted at times t_j , N_j is their number operator, and Z_j normalizes each factor. The temperature changes during evaporation; a single Gibbs state at one temperature does not describe the whole history. The mixed result represents an evolution

$$|\Psi_0\rangle\langle\Psi_0| \longrightarrow \rho_{\text{rad}}^{(\text{final})} \neq |\Psi_f\rangle\langle\Psi_f| \quad (2.16)$$

from a pure state to a mixed state. Such an evolution *cannot* be described by any unitary operator.

Why this is a real paradox

One might be tempted to dismiss the information paradox as a technical subtlety or an artifact of approximations. It is neither. The paradox is a sharp conflict between two of the most well-tested principles in physics:

- (i) **Unitarity of quantum mechanics.** The evolution of any closed quantum system is unitary: pure states evolve to pure states, and information is conserved. This is not a minor technical assumption—it is the foundation of the entire quantum-mechanical framework, and fundamental modifications face consistency constraints, including tensions between locality and energy conservation [10]. Nonunitary subsystem channels themselves need not violate probability conservation.
- (ii) **Validity of semiclassical gravity in low-curvature regions.** Hawking’s calculation uses quantum field theory on a classical curved background. The approximation is applied *outside* the horizon, where curvature is low (for a massive black hole) and we have every reason to trust the semiclassical expansion for local observables. The issue is whether this approximation also controls the fine-grained entropy of all the radiation, especially after contributions from additional replica saddles become relevant. Low local curvature alone does not establish that the original entropy calculation remains uniformly valid throughout evaporation.

The paradox, then, is not “we don’t know the answer.” It is that two individually well-supported principles give contradictory predictions, and modifying either one comes at an enormous cost.

Proposed resolutions (preview)

Over the four decades following Hawking’s 1975 paper, a number of resolutions were proposed. We mention them briefly here, deferring detailed discussion to later sections:

- **Information is truly lost.** One simply accepts that unitarity is violated in quantum gravity. Hawking himself initially advocated this position [39] but later abandoned it.

- **Remnants.** The black hole does not evaporate completely but leaves behind a stable Planck-scale remnant that stores the missing information. This requires a Planck-mass object to have an exponentially large number of internal states, leading to serious difficulties with pair production and the species problem [32].
- **Information comes out in the radiation.** Unitarity is preserved, and the Hawking radiation is not truly thermal—subtle correlations encode the information of the initial state. This is the resolution most widely accepted today, and the Page curve provides the quantitative benchmark.
- **Fuzzballs.** The black hole interior is replaced by a horizon-scale microstate structure. Explicit smooth horizonless solutions realize some sectors; extrapolation to generic black-hole states and their complete evaporation remains a separate problem [62].

The Page curve as diagnostic

If a pure initial state evolves unitarily into radiation alone, the final radiation entropy is zero. If the entropy has grown at intermediate times, it must subsequently decrease. These facts do not fix its detailed history. The question is: what does the entropy curve $S_{\text{vN}}(\rho_{\text{rad}})$ look like as a function of time?

The answer was provided by Don Page in 1993 [71, 72]. Page modeled the joint state of black hole and radiation as a *random pure state* on a Hilbert space of dimension $d_{\text{BH}} d_{\text{rad}}$, with $d_{\text{BH}} = e^{S_{\text{BH}}}$; near-maximal entanglement of the smaller factor is then a consequence, not an assumption. Using the result for the average entanglement entropy of a random subsystem, he showed that

$$S_{\text{vN}}(\rho_{\text{rad}}) \approx \min(S_{\text{rad}}, S_{\text{BH}}(t)), \quad (2.17)$$

where S_{rad} is the coarse-grained (thermodynamic) entropy of the emitted radiation and $S_{\text{BH}}(t)$ is the remaining Bekenstein–Hawking entropy of the black hole. The key features of the **Page curve** are:

1. **Early times** ($t < t_{\text{Page}}$): the radiation entropy grows, tracking $S_{\text{rad}} \approx N_{\text{emitted}}$. The radiation subsystem is smaller than the black hole, and the entropy grows because each new Hawking quantum adds new entanglement.
2. **Page time** ($t = t_{\text{Page}}$): the entropy is near the crossing of the two coarse-grained entropies. In a fixed-total-dimension toy model this occurs at half the initial entropy. With irreversible entropy production, the crossing need not occur there.
3. **Late times** ($t > t_{\text{Page}}$): the entropy *decreases*, now tracking $S_{\text{BH}}(t)$. The radiation is the larger subsystem, and as the black hole shrinks, the maximum possible entanglement between the two subsystems decreases.
4. **Complete evaporation** ($t = t_{\text{evap}}$): the entropy returns to zero. The radiation is in a pure state, consistent with unitary evolution.

A simple evaporation model makes this dependence explicit. Suppose $dM/dt = -a/M^2$, with constant $a > 0$, so that $S_{\text{BH}}(t)/S_i = (1 - t/t_{\text{evap}})^{2/3}$, where S_i is the initial black-hole entropy. Let ξ be the ratio of radiation entropy production to the decrease in black-hole entropy:

$$S_{\text{rad}}(t) = \xi [S_i - S_{\text{BH}}(t)]. \quad (2.18)$$

For constant ξ , equating the branches gives

$$\frac{S_{\text{BH}}(t_P)}{S_i} = \frac{\xi}{1+\xi}, \quad \frac{t_P}{t_{\text{evap}}} = 1 - \left(\frac{\xi}{1+\xi} \right)^{3/2}. \quad (2.19)$$

The choice $\xi = 1$ gives $t_P/t_{\text{evap}} \simeq 0.6464$. An ideal massless blackbody flux in three spatial dimensions has $dS_{\text{rad}} = 4 dE_{\text{rad}}/(3T_H)$, while the first law has $-dS_{\text{BH}} = dE_{\text{rad}}/T_H$. Thus $\xi = 4/3$ in that idealization and $t_P/t_{\text{evap}} \simeq 0.5680$. Actual greybody factors and particle content change ξ and the mass-loss law [70]. Neither numerical time is a consequence of unitarity alone.

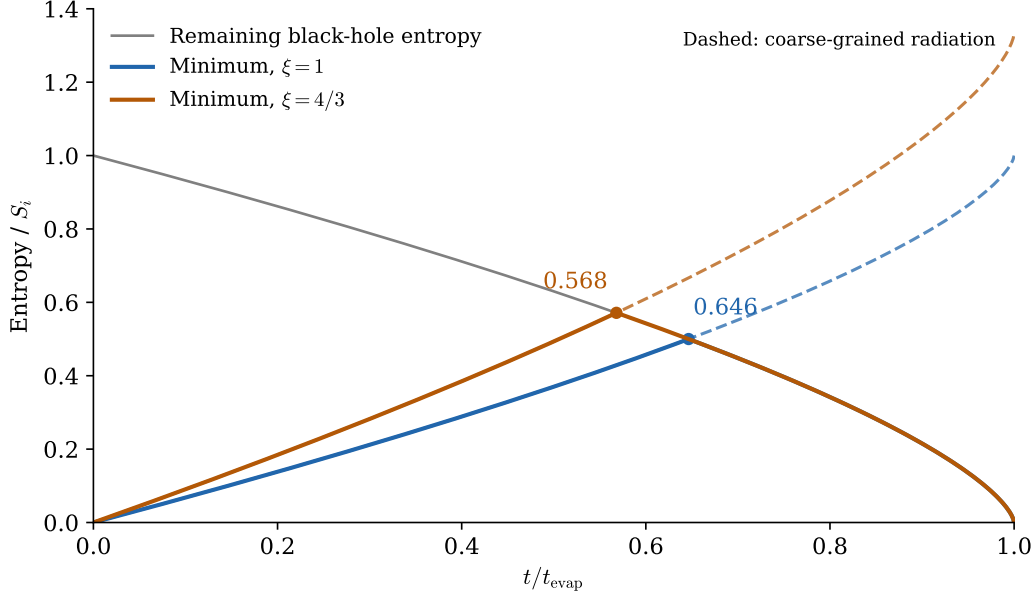


Figure 4. Typical-state entropy benchmarks for the specified mass-loss model. The black-hole entropy is $(1 - t/t_{\text{evap}})^{2/3}$ in units of its initial value. Solid curves are the minimum of this branch and $\xi[1 - (1 - t/t_{\text{evap}})^{2/3}]$; dashed curves are the corresponding coarse-grained radiation entropies. The choices $\xi = 1$ and $4/3$ give different crossing times. These are model curves, not numerical solutions of a gravitational evaporation problem. The late endpoint extrapolates beyond the semiclassical regime.

The four-decade gap

The significance of the Page curve is not merely qualitative. Equation (2.17) provides a *quantitative* benchmark once typicality and the effective subsystem entropies have been specified; unitarity supplies the endpoint constraints. Hawking’s original semiclassical calculation gives increasing radiation entropy. A central subsequent problem was to obtain a turnover directly from a controlled gravitational entropy calculation.

The challenge is clear: one needs a mechanism within the gravitational description itself that causes the entropy to turn over at the Page time. The low curvature at a macroscopic horizon motivates seeking this mechanism in a semiclassical treatment that includes information absent from Hawking’s original entropy calculation.

The development of *quantum extremal surfaces* [27], the island formula [4, 5, 74], and the *replica wormhole* derivations [1, 76] supplied such calculations in specific holographic and gravity-plus-bath

settings. The key insight, which we will develop in Sections 3–5, is that the gravitational path integral includes contributions from non-trivial spacetime topologies—saddle points connecting different replicas of the geometry—that Hawking’s original calculation did not include. In these models, including the relevant connected replica saddles produces the transition between growing and decreasing entropy branches.

Summary. Black holes are thermodynamic objects with entropy $S_{\text{BH}} = A/(4G_N)$ and temperature $T_H = \kappa/(2\pi)$. Hawking radiation is thermal, with each outgoing quantum entangled with a partner behind the horizon. The resulting monotonic growth of $S_{\text{vN}}(\rho_{\text{rad}})$ conflicts, upon complete evaporation into radiation alone, with a pure final state. The typical-state Page curve adds a quantitative benchmark for the intermediate entropy. The following sections explain gravitational calculations of this transition and the random-state benchmarks used to interpret them.

3 Holography and Entanglement from Geometry

Hawking’s entropy calculation conflicts with the pure final state required by unitary, complete evaporation from a pure preparation. The intermediate Page profile adds the typical-state assumptions of the preceding section. Holographic entropy gives a gravitational way to calculate that profile in controlled settings.

The key idea, developed over more than twenty years, is that *entanglement entropy is a geometric quantity*. In the right setting—the holographic correspondence between gravity and quantum field theory—the entanglement entropy of a boundary region is given at leading semiclassical order by the area of a bulk surface in units of $4G_N$. Upgrading this idea to include quantum corrections leads to the *quantum extremal surface* (QES) prescription and, from it, the *island rule*, which produces Page transitions in explicit gravity-plus-bath models.

We proceed in stages. Section 3.1 introduces the holographic correspondence (AdS/CFT). Section 3.2 presents the Ryu–Takayanagi formula relating entanglement entropy to area. Section 3.3 upgrades this to the quantum extremal surface prescription. Section 3.4 distills the island rule for the entropy of Hawking radiation. Finally, Section 3.5 takes stock of what has been achieved and what remains open.

3.1 AdS/CFT: the holographic correspondence

Anti-de Sitter space as a gravitational “box”

General relativity admits a family of maximally symmetric solutions classified by the sign of the cosmological constant Λ . Positive Λ gives de Sitter space (roughly, our universe at late times); zero gives Minkowski space; and negative Λ gives *anti-de Sitter* (AdS) space.

The defining property of AdS_{d+1} is constant negative curvature. A useful way to picture it: consider $(d+1)$ -dimensional spacetime whose spatial slices are hyperbolic, with a curvature radius ℓ set by $\Lambda = -d(d-1)/(2\ell^2)$. A massive particle released in AdS is pulled back toward the center by the curvature; it never reaches the boundary, turning around after a finite proper time. Crucially, however, light signals *can* reach the boundary in finite coordinate time. The upshot is that AdS behaves like a gravitational “box” with reflecting boundary conditions: energy is confined. Coupling the boundary to an external bath allows energy to leave. This makes it an ideal arena for quantum gravity: there is no S-matrix in AdS—the natural observables are correlation functions anchored to the boundary—and these, along with thermal equilibrium, can be defined without the infrared subtleties of asymptotically flat space.

The boundary of AdS_{d+1} —the surface at spatial infinity—is a d -dimensional manifold with a conformal structure. This boundary is not merely a mathematical convenience; it is where the dual quantum field theory lives.

The Maldacena conjecture

In 1997, Juan Maldacena proposed what is arguably the most consequential conjecture in modern theoretical physics [1998]. In its original form the statement is:

Type IIB string theory on $\text{AdS}_5 \times S^5$ is *exactly equivalent* to $\mathcal{N} = 4$ super-Yang–Mills theory on the four-dimensional boundary.

More broadly, the **AdS/CFT correspondence** asserts that a theory of quantum gravity in $(d + 1)$ -dimensional anti-de Sitter space is *dual* to a conformal field theory (CFT) living on the d -dimensional boundary. “Dual” here means *exactly equivalent*: every physical observable on one side has a precise counterpart on the other. The two descriptions are not approximations to each other; they are the *same* theory written in two different languages.

The term *holographic* captures the dimensional mismatch: a $(d + 1)$ -dimensional gravitational theory is fully encoded in a d -dimensional non-gravitational one, much as a three-dimensional image is encoded on a two-dimensional holographic plate. This resonates with the Bekenstein–Hawking entropy [12, 38], which already suggested that the number of degrees of freedom of a gravitational system scales with its boundary area, not its volume.

The holographic dictionary

The correspondence comes with a “dictionary” that translates between the two descriptions. Its key entries are:

1. **Bulk fields $\longleftrightarrow$ boundary operators.** A field ϕ propagating in the AdS bulk corresponds to a local operator $\mathcal{O}$ in the boundary CFT. The mass of the bulk field determines the scaling dimension of the boundary operator.
2. **Geometry $\longleftrightarrow$ entanglement.** The geometry of the bulk spacetime—its metric, topology, even its existence—is encoded in the entanglement structure of the boundary quantum state. This entry of the dictionary will be the focus of the rest of this section.
3. **Black hole $\longleftrightarrow$ thermal state.** A large black hole in the bulk corresponds to the boundary CFT at finite temperature. The Hawking temperature of the black hole equals the temperature of the CFT thermal state, and the Bekenstein–Hawking entropy gives the leading thermal entropy. A pure black-hole microstate is distinct from this thermal ensemble.
4. **Bulk partition function $\longleftrightarrow$ boundary partition function.** The gravitational path integral over asymptotically AdS geometries equals the CFT partition function. Schematically,

$$Z_{\text{gravity}}[\text{boundary conditions}] = Z_{\text{CFT}}[\text{sources}]. \quad (3.1)$$

This is the master equation of AdS/CFT.

Why this matters for the information paradox

The boundary CFT is a standard (albeit strongly coupled) quantum field theory. It has a Hilbert space, a Hamiltonian, and unitary time evolution. There is no room for information loss: time evolution is unitary by construction. Since the bulk gravity theory is *exactly* the same theory, gravity must also be unitary.

This gives a sharp resolution of the information paradox—at least in principle. If you form a black hole in AdS by throwing in matter, the boundary CFT evolves unitarily. With reflecting boundary conditions a large black hole can remain in equilibrium instead of evaporating completely. An evaporation experiment requires, for example, coupling the CFT to an external bath. The combined CFT–bath evolution is then the closed system to which unitarity applies.

The catch is that this argument is *structural*: it tells us that information is preserved, but not *how*. The boundary theory does not directly reveal what happens at the horizon, or how the interior degrees of freedom get out. To make progress on the “how,” we need tools that connect entanglement in the boundary theory to geometry in the bulk. That is the subject of the next subsection.

3.2 Ryu–Takayanagi: entanglement from area

The formula

In 2006, Shinsei Ryu and Tadashi Takayanagi proposed an elegant formula for the entanglement entropy of a boundary region [2006]. Let A be a spatial subregion of the boundary CFT, and let γ_A be the minimal-area surface in the bulk that is *homologous* to A —meaning γ_A is anchored to the boundary ∂A of region A and, together with A itself, bounds a bulk region. Then:

$$S(A) = \frac{\text{Area}(\gamma_A)}{4 G_N} \quad (\text{Ryu–Takayanagi formula}). \quad (3.2)$$

Here G_N is Newton’s constant in the bulk, and $S(A)$ is the von Neumann entropy of the reduced density matrix $\rho_A = \text{Tr}_{\bar{A}} |\Psi\rangle\langle\Psi|$.

This formula is striking: a *quantum information* quantity on the left (entanglement entropy, defined via the density matrix) is equal to a *classical geometric* quantity on the right (the area of a surface, computed from the metric). It is, in a precise sense, the quantitative version of the dictionary entry “geometry $\leftrightarrow$ entanglement.”

The factor of $4G_N$ in the denominator is the same factor that appears in the Bekenstein–Hawking entropy formula. This is not a coincidence: for a thermal state dual to a bulk black hole, the Bekenstein–Hawking entropy is a special case of the RT formula, obtained when the boundary region A is the entire boundary and the minimal surface γ_A wraps the black hole horizon. For a pure state on the entire boundary, by contrast, the entropy is zero; the state and homology condition are essential to this distinction.

Simple examples

To build intuition, consider two cases (illustrated in Figure 5):

Pure AdS. The bulk geometry is “empty” anti-de Sitter space, dual to the vacuum state of the CFT. For a connected boundary interval A in $\text{AdS}_3/\text{CFT}_2$, the minimal surface γ_A is a bulk geodesic—the shortest curve connecting the two endpoints of ∂A that dips into the interior. A short interval has a shallow geodesic (small area, small entropy); a longer interval has a deeper geodesic (larger area, larger entropy). This matches the universal logarithmic scaling of entanglement entropy in a two-dimensional CFT [17].

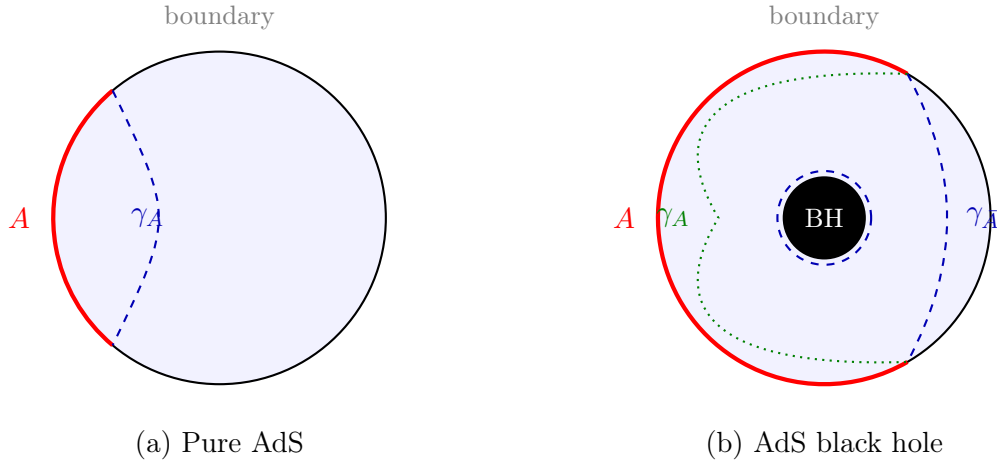


Figure 5. The Ryu–Takayanagi surface in two settings. (a) In pure AdS, the minimal surface γ_A homologous to boundary region A is a bulk geodesic connecting the endpoints of ∂A . (b) In a black hole background (the thermal state), a sufficiently large boundary region admits two candidates: a connected geodesic homologous to A (dotted green), and the complementary geodesic together with the horizon (both dashed blue). The RT prescription selects the candidate with smaller total area. When A is the entire boundary in the thermal state, the homologous surface is the horizon itself, and the RT formula reduces to the Bekenstein–Hawking entropy.

AdS with a black hole. Now consider a large black hole occupying the center of AdS. For a small boundary region A , the minimal surface is still a geodesic that barely enters the bulk. But as A grows to cover most of the boundary, a topological transition occurs: the minimal surface “jumps” to one that wraps around the black hole horizon. At this point the disconnected candidate gives $S(A) \simeq S(\bar{A}) + A_{\text{horizon}}/(4G_N)$ at leading order, with the same ultraviolet regulator on both boundary entropies. For the entire thermal system only the horizon contribution remains.

This topological transition—between two candidate surfaces—is a recurring theme. As we will see, a closely analogous transition underlies the Page curve itself.

Scope and limitations

The RT formula (3.2) applies to static geometries and holographic CFTs at leading order in $1/G_N$ (equivalently, large central charge). Its covariant generalization to dynamical spacetimes was given by Hubeny, Rangamani, and Takayanagi [2007], who replace the minimal surface with an *extremal* surface (one whose area is stationary under small deformations). We will refer to this covariant version as the HRT formula.

Crucially, both RT and HRT are *classical*: they compute the leading-order entanglement entropy and neglect quantum corrections (bulk entanglement, one-loop effects, etc.). For many purposes this is perfectly adequate. But for the information paradox, it is not enough. Bulk entropy can accumulate to the size of the area term over the evaporation history, so a contribution subleading at fixed time can participate in a leading competition at late times. This motivates the generalized-entropy prescription described next.

3.3 Quantum extremal surfaces: the upgrade

The need for quantum corrections

The RT formula computes entanglement entropy to leading order in $1/G_N$. In the holographic dictionary, $1/G_N$ plays the role of a large parameter (analogous to N^2 in a gauge theory with N colors), so the RT formula captures the “classical” or “planar” contribution. But quantum effects in the bulk—loop corrections, entanglement of bulk quantum fields—contribute at subleading order. For a complete answer, one needs to include these corrections.

The importance of this for black holes is immediate. The Bekenstein–Hawking entropy $S_{\text{BH}} = A_{\text{horizon}}/(4G_N)$ is an $O(1/G_N)$ quantity. Hawking’s monotonically rising radiation entropy is also $O(1/G_N)$. The “error” that leads to the information paradox is therefore a leading-order effect. Resolving it requires a formula that correctly competes the leading area term against the growing bulk entropy.

The first step beyond the classical formula was taken by Faulkner, Lewkowycz, and Maldacena [2013], who computed the leading quantum correction to holographic entanglement entropy: the bulk entanglement entropy S_{bulk} of the quantum fields, evaluated at the classical RT surface, is added to the area term. Engelhardt and Wall then promoted this one-loop correction to a full prescription that extremizes the generalized entropy itself.

The Engelhardt–Wall prescription

In 2015, Netta Engelhardt and Aron Wall proposed the **quantum extremal surface** (QES) prescription [2015]. It replaces the RT/HRT formula with:

$$S(A) = \min_{\chi} \text{ext}_{\text{extremal}} \left[\frac{\text{Area}(\chi)}{4G_N} + S_{\text{bulk}}(\Sigma_{\chi}) \right] \quad (\text{QES prescription}), \quad (3.3)$$

where:

- χ is a codimension-two surface in the bulk, anchored to ∂A on the boundary;
- Σ_{χ} is the bulk region on a Cauchy slice between χ and the boundary region A (the homology region; its domain of dependence is the “entanglement wedge”);
- $S_{\text{bulk}}(\Sigma_{\chi})$ is the von Neumann entropy of the bulk quantum fields restricted to Σ_{χ} ;
- “ext” means we find surfaces that *extremize* the generalized entropy (area term plus bulk entropy), and “min” selects the one with the smallest value among all extrema.

The two-step instruction—first extremize, then minimize—is essential. The functional $S_{\text{gen}}(\chi) = \text{Area}(\chi)/(4G_N) + S_{\text{bulk}}(\Sigma_{\chi})$ is called the **generalized entropy** of the surface χ . In general, there may be multiple surfaces at which S_{gen} is stationary. Each such surface is a quantum extremal surface. The prescription instructs us to pick the global *minimum* among these extrema—the one with the smallest generalized entropy.

The area and bulk entropy require a common renormalization prescription; the displayed formula uses Einstein gravity and suppresses higher-derivative entropy terms. It is the usual semiclassical prescription. Particular states and transitions can require refinements, including the leading QES corrections of Akers and Penington [3]; see Section 9.

When $S_{\text{bulk}} = 0$ (no bulk entropy contribution), the QES prescription reduces to HRT, or RT in a static setting: extremizing the area alone recovers the minimal surface, and we are back to Eq. (3.2).

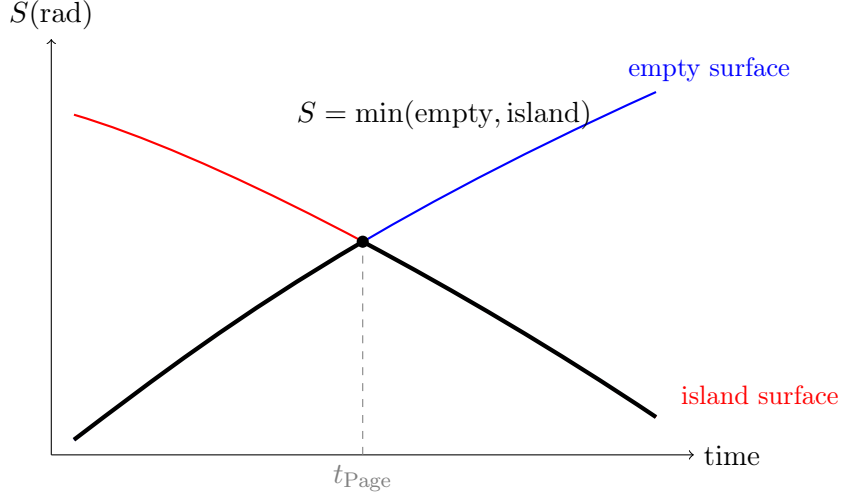


Figure 6. Competition between the two quantum extremal surfaces. The blue (empty surface) curve tracks Hawking’s monotonically growing result. The red (island surface) curve starts at the initial Bekenstein–Hawking entropy and decreases as the black hole evaporates. The QES prescription instructs us to take the global minimum (bold black curve), which transitions from the empty surface to the island surface at the *Page time*. The resulting curve is the Page curve.

Multiple extrema and the phase transition

The content of the QES prescription lies in the possibility of *multiple* quantum extremal surfaces for the same boundary region. This is where the resolution of the information paradox emerges.

Consider computing the entanglement entropy of the Hawking radiation emitted by an evaporating black hole. In an appropriate holographic setup (we will be more precise momentarily), the QES prescription identifies two competing extremal surfaces:

1. **The “empty” surface** (trivial QES). This surface is trivial in the sense that its area vanishes (or is negligible). The generalized entropy is then dominated by $S_{\text{bulk}}(\Sigma)$, which includes all the Hawking radiation. This contribution grows monotonically as the black hole evaporates, tracking Hawking’s original calculation.
2. **The “island” surface** (non-trivial QES). This surface lies near the black hole horizon. Its area term $\text{Area}(\chi)/(4G_N)$ gives a contribution close to the Bekenstein–Hawking entropy, which *decreases* as the black hole shrinks. The bulk entropy term is now evaluated on a union of the radiation region and an interior region—the “island” (Section 3.4 below). Including partners of outgoing modes can lower the bulk entropy of this union.

The QES prescription says: take the minimum. At early times, the empty surface has lower generalized entropy (the radiation entropy is still small), so it wins. At late times, the island surface has lower generalized entropy (the black hole has shrunk), so it wins. The crossover defines the **Page time** t_{Page} , and the envelope of the two curves is precisely the Page curve (Figure 6).

Schematically:

$$S(\text{rad}) = \begin{cases} S_{\text{bulk}}(\text{rad}) & t < t_{\text{Page}} \quad (\text{empty surface}), \\ \frac{A_{\text{horizon}}}{4G_N} + \dots & t > t_{\text{Page}} \quad (\text{island surface}). \end{cases} \quad (3.4)$$

Before the Page time, the entropy grows (tracking Hawking’s result). After the Page time, it decreases (tracking the shrinking horizon). The result is $S(\text{rad}) \approx \min(S_{\text{thermo}}, S_{\text{BH}})$, where S_{thermo} denotes the thermodynamic (coarse-grained) entropy of the radiation emitted so far and S_{BH} the instantaneous Bekenstein–Hawking entropy of the black hole—which is exactly the Page curve.

This is the central result:

The quantum extremal surface prescription, applied to an evaporating black hole, produces the Page curve. This was demonstrated independently in 2019 by Penington [2020] and by Almheiri, Engelhardt, Marolf, and Maxfield [2019].

This calculation gives a geometric entropy turnover in the specified semiclassical models. It goes beyond the original disconnected saddle without supplying a complete microscopic evaporation channel. The distinction between entropy and channel data is developed in Section 13.

3.4 The island rule

The QES prescription described above can be repackaged into a particularly intuitive form called the **island rule** [5]. This formulation makes explicit the role of the “island”—a region of the black hole interior that, after the Page time, is “owned by” the radiation.

Setup

Consider a black hole that is evaporating into a region we call the “radiation reservoir” (or simply “rad”). We want to compute $S(\text{rad})$, the entanglement entropy of the radiation. In a holographic setup, the radiation lives on a non-gravitational bath coupled to the gravitational region containing the black hole.

The formula

The island rule states:

$$S(\text{rad}) = \min_I \text{ext}_{\text{extremal}} \left[\frac{\text{Area}(\partial I)}{4G_N} + S_{\text{bulk}}(\text{rad} \cup I) \right] \quad (\text{island rule}). \quad (3.5)$$

Here:

- I is a spatial region in the gravitational sector (the “island”). The island may be empty ($I = \emptyset$) or non-empty.
- ∂I is the boundary of the island. When $I = \emptyset$, $\text{Area}(\partial I) = 0$.
- $S_{\text{bulk}}(\text{rad} \cup I)$ is the von Neumann entropy of the bulk quantum fields on the union of the radiation region and the island.
- The optimization is over all possible islands, including the trivial one.

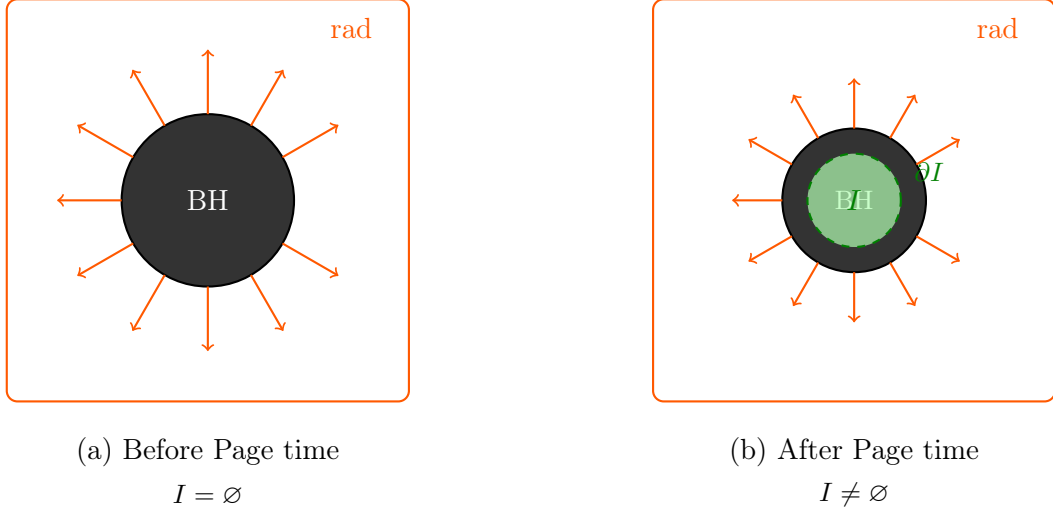


Figure 7. The island rule illustrated schematically. (a) Before the Page time, the island is empty ($I = \emptyset$). The entropy of the radiation is computed from the bulk entropy of the radiation alone, which grows monotonically. (b) After the Page time, a non-trivial island I appears inside the black hole. The radiation entropy is now computed from the generalized entropy: the area of ∂I (the QES, shown dashed) plus the bulk entropy of $\text{rad} \cup I$. Because the island “absorbs” the interior partners of the Hawking quanta, the bulk entropy term is reduced, and the total entropy decreases.

Two phases

The island rule identifies two competing saddle points:

No-island phase ($I = \emptyset$). When the island is empty, ∂I is also empty, and the formula reduces to

$$S(\text{rad}) = S_{\text{bulk}}(\text{rad}). \quad (3.6)$$

This is just the entropy of the Hawking radiation computed in the semiclassical approximation—Hawking’s original, monotonically growing answer.

Island phase ($I \neq \emptyset$). When a non-trivial island exists, it typically sits just inside the black hole horizon (Figure 7b). Its boundary ∂I is the quantum extremal surface, and the generalized entropy is approximately

$$S(\text{rad}) \approx \frac{A_{\text{horizon}}}{4G_N} + S_{\text{bulk}}(\text{rad} \cup I). \quad (3.7)$$

The bulk entropy of $\text{rad} \cup I$ is much smaller than $S_{\text{bulk}}(\text{rad})$ alone, because including the island brings in the interior partners of the Hawking quanta. These partners are entangled with the radiation, so including them in the same subsystem *reduces* the entanglement entropy. The dominant contribution is the area term, which decreases as the black hole evaporates.

The island rule selects whichever phase has lower entropy. Before the Page time, the no-island phase wins. After the Page time, the island phase wins. The result is the Page curve.

Interpretation: what the island means

The appearance of an island has a striking interpretation. After the Page time, a spatial region I inside the black hole is *encoded in the radiation*. In the language of quantum error correction (which underlies the modern understanding of holography), the island is part of the “entanglement wedge” of the radiation: an observer with access to the radiation can, in principle, reconstruct operators acting on the island.

This is deeply counterintuitive. In the evaporating case the island lies behind the horizon, spacelike separated from the late-time radiation region it purifies (in eternal-black-hole-plus-bath setups, islands can even extend outside the horizon). Yet the island rule tells us that, after the Page time, this interior region is reconstructable from the radiation. This describes the encoding of a bulk operator algebra, not a local signal propagating outward across a classical event horizon.

A caveat: formula, not mechanism

In its applicable regime the island rule computes semiclassical entropy and identifies an entanglement wedge associated with the radiation. Entanglement-wedge reconstruction gives operator-algebraic meaning to this region, conditional on the chosen semiclassical code subspace. The entropy formula alone does not specify a state-independent decoder or the entire dynamical process that accomplishes the encoding. Replica wormholes and entropy statistics supply additional calculations, developed in Sections 4 and 5.

3.5 What was achieved and what remains open

Let us take stock.

The achievement

The QES prescription and the island rule represent a genuine breakthrough. For the first time since Hawking’s 1975 paper [38], there exists a gravitational calculation that produces the Page curve:

1. The **Ryu–Takayanagi formula** [78] established that entanglement entropy is computed by the area of a minimal surface in the bulk.
2. The **quantum extremal surface prescription** [27] upgraded this to include bulk quantum corrections, introducing the generalized entropy.
3. **Penington** [2020] and **Almheiri et al.** [2019] showed that the QES prescription, applied to an evaporating black hole, produces the Page curve. The key is the exchange of dominance between two quantum extremal surfaces at the Page time.
4. The **island rule** [5] repackaged this into a formula directly for the radiation entropy, making the role of the island explicit.

In the stated holographic and bath-coupled models, these results derive the entropy turnover from a gravitational prescription. Reconstructing interior observables and identifying the full evaporation channel require additional information.

What remains open

Despite this progress, several deep questions remain:

1. **How does information escape?** The island rule tells us *that* information is encoded in the radiation, but not *how* it gets there. What is the physical mechanism at the horizon? Does an infalling observer notice anything? These questions remain open.
2. **Single theory vs. ensemble.** The gravitational path integral in JT gravity has an ensemble interpretation, whereas a fixed holographic CFT is a single theory. One must identify which quantities are averaged in a particular calculation; QES does not itself imply averaging over Hamiltonians. The relationship between single realizations and ensemble averages is a subject of active investigation [79, 101].
3. **What happens at the horizon?** Different proposals exist: firewalls [6], fuzzballs [62], smooth horizons with subtle encoding [73], ER=EPR [57]. The island rule does not by itself determine which horizon description applies. Whether two specified models yield the same observables must be calculated.
4. **Detailed transition structure.** The QES prescription predicts a sharp phase transition at the Page time: a discrete swap between two saddle points. For a finite-dimensional density matrix varying continuously with a parameter, entropy is continuous, although differentiability requires additional conditions. What controls the width and shape of this transition? What are the corrections to the sharp-swap approximation?

The second and fourth questions—the ensemble structure of the gravitational path integral and the detailed transition structure—are precisely the subject of the companion papers [99–101] and of the remainder of this tutorial: Papers I and II develop the detailed structure of the Page transition, while Paper III takes up the ensemble and microscopic-ambiguity questions. The gravitational path integral, replica wormholes, and free probability theory will provide the tools to resolve the transition and go beyond the leading-order Page curve. We turn to these tools in Part II.

Part II

The Gravitational Path Integral and Replica Wormholes

4 The Gravitational Path Integral

The previous sections assembled the physical ingredients: quantum mechanics, black holes, Hawking radiation, and the holographic principle. We now turn to the *computational* framework that makes the Page curve derivable within semiclassical gravity: the gravitational path integral. This is the tool that will, in the next section, produce replica wormholes and the leading entropy turnover in specified models.

The gravitational path integral is simultaneously one of the deepest and one of the most poorly defined objects in theoretical physics. It is deep because it encodes the quantum-mechanical sum over all possible spacetimes — not just over field configurations on a fixed background, but over

the geometries themselves. It is poorly defined because, unlike the path integral for a scalar field, a general nonperturbative definition for the gravitational theories considered here is not available. Nevertheless, when used with appropriate care — principally through the saddle-point approximation and its systematic corrections — it produces results of considerable power and precision.

This section develops the gravitational path integral in three stages. We begin with the basic idea of the path integral and explain what changes when the dynamical variable is the metric itself. We then introduce the genus expansion, which organizes two-dimensional topologies as a perturbative series. We then distinguish its relation to matrix integrals from the particular contraction combinatorics of a random-state model. Finally, we explain the Wick rotation to Euclidean signature and the thermal interpretation that makes the black hole path integral tractable.

4.1 Path integrals: the basic idea

Feynman’s sum over histories

The path integral formulation of quantum mechanics, due to Feynman [29], recasts the transition amplitude between two states as a sum over all possible histories connecting them. For a particle moving from position q_i at time t_i to position q_f at time t_f , the amplitude is

$$\langle q_f, t_f | q_i, t_i \rangle = \int \mathcal{D}q(t) e^{iS[q]/\hbar}, \quad (4.1)$$

where $S[q] = \int_{t_i}^{t_f} L(q, \dot{q}) dt$ is the classical action and the integral runs over *all* paths $q(t)$ with the specified boundary conditions. Each path contributes with a phase $e^{iS/\hbar}$; the classical path — the one that extremizes S — dominates because nearby paths constructively interfere, while paths far from the classical trajectory contribute rapidly oscillating phases that cancel.

The same logic extends to quantum field theory. For a scalar field ϕ on a fixed spacetime background, the generating functional (partition function) is

$$Z = \int \mathcal{D}\phi e^{iS[\phi]/\hbar}. \quad (4.2)$$

This is the starting point for all perturbative calculations in the Standard Model and in quantum field theory on curved spacetime.

Gravity: summing over geometries

The gravitational path integral makes a conceptual leap: it promotes the *spacetime geometry itself* to a dynamical variable. Rather than integrating over field configurations on a fixed background, we integrate over all metrics $g_{\mu\nu}$ on a manifold $\mathcal{M}$, weighted by the Einstein–Hilbert action:

$$Z = \int \mathcal{D}g_{\mu\nu} e^{iS_{\text{EH}}[g]/\hbar}, \quad S_{\text{EH}} = \frac{1}{16\pi G_N} \int_{\mathcal{M}} d^{d+1}x \sqrt{-g} (R - 2\Lambda) + S_{\text{bdy}}, \quad (4.3)$$

where R is the Ricci scalar, Λ is the cosmological constant, G_N is Newton’s constant, and S_{bdy} denotes the Gibbons–Hawking–York boundary term [31, 102] needed to make the variational problem well-posed on manifolds with boundary.

Three features distinguish the gravitational path integral from its field-theory counterpart:

1. **The integration variable is the geometry.** In ordinary quantum field theory, the background spacetime is fixed and the path integral sums over matter field configurations. In quantum gravity, the metric is itself integrated over. There is no fixed stage on which the drama unfolds; the stage is part of the show.

2. **The sum includes different topologies.** In particle physics, the path integral for a scalar field sums over all field configurations on $\mathbb{R}^{3,1}$. The topology of spacetime is fixed. In gravity, we must also sum over manifolds of different topology — spacetimes with handles, wormholes, and other non-trivial topological features. This “sum over topologies” has no analogue in the Standard Model and is the source of much of the richness (and difficulty) of quantum gravity.
3. **There is no known rigorous measure.** The measure $\mathcal{D}g_{\mu\nu}$ must account for the gauge redundancy of diffeomorphisms, the conformal factor problem (the Euclidean action is unbounded below for conformal deformations), and the absence of a background with respect to which “small fluctuations” can be defined. These problems are not merely technical; they reflect the deep conceptual challenge of defining quantum gravity non-perturbatively.

Despite these difficulties, the **saddle-point approximation** gives controlled calculations in suitable regimes, once boundary conditions, gauge fixing, and an integration contour have been specified.

The saddle-point approximation

The saddle-point (or stationary-phase) approximation to a path integral retains only the contributions from configurations that extremize the action — i.e., the classical solutions — together with Gaussian fluctuations around them. Schematically:

$$Z \approx \sum_{\text{saddles } \sigma} e^{-I[\sigma]} \cdot Z_{\sigma}^{(1)} \cdot (1 + O(G_N)), \quad (4.4)$$

where $I[\sigma]$ is the on-shell Euclidean action of the saddle σ (we explain the Euclidean formulation in Section 4.3), $Z_{\sigma}^{(1)}$ is the one-loop determinant from quadratic fluctuations around the saddle, and the remaining corrections are organized as a perturbative expansion in G_N .

The sum in Eq. (4.4) runs over saddles compatible with the boundary conditions and the integration contour. This is a crucial point: the gravitational path integral is not dominated by a single geometry, but by a *sum* over geometries, each weighted by e^{-I} . When there are multiple saddles, the one with the smallest action dominates, but the others contribute and can become important as external parameters (such as the temperature or the age of the black hole) are varied. Phase transitions occur precisely when the dominant saddle changes — the Hawking–Page transition [40] between thermal AdS and the AdS black hole being the classic example.

This is a semiclassical expansion, not a general nonperturbative definition. Away from degeneracies a dominant saddle can control the answer; near a transition several saddles can contribute comparably. In JT gravity the topological entropy S_0 separately controls the cost of handles, as explained below.

4.2 The genus expansion

The sum over topologies in the gravitational path integral admits a natural perturbative organization: the **genus expansion**. This expansion is most transparent in two-dimensional gravity, where orientable surfaces have a simple classification. Higher-dimensional topology does not admit the same general organization by one genus parameter.

Topology of surfaces

A closed, connected, orientable two-dimensional surface is completely characterized (up to homeomorphism) by a single integer: its **genus** g , the number of handles. The genus-zero surface is the sphere S^2 ; genus one is the torus T^2 ; genus two is a “pretzel” with two handles; and so on (Figure 9).

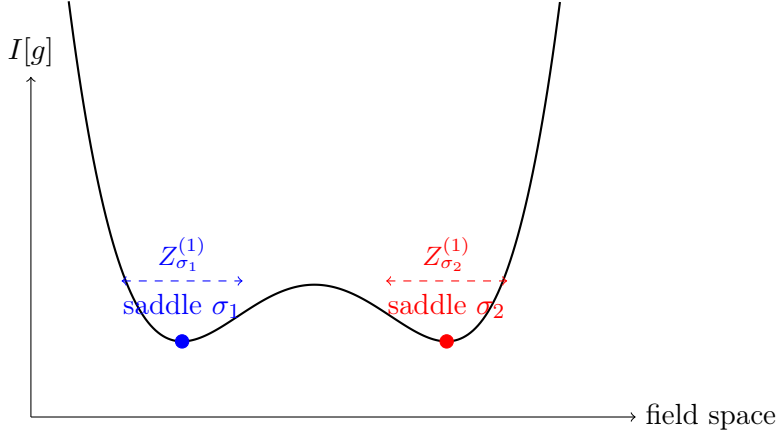


Figure 8. Schematic illustration of the saddle-point approximation. The gravitational path integral receives its dominant contributions from classical solutions (saddle points) of the Euclidean action. Gaussian fluctuations around each saddle produce the one-loop determinant $Z_{\sigma}^{(1)}$. The full answer is a sum over all saddles, weighted by $e^{-I[\sigma]}$.

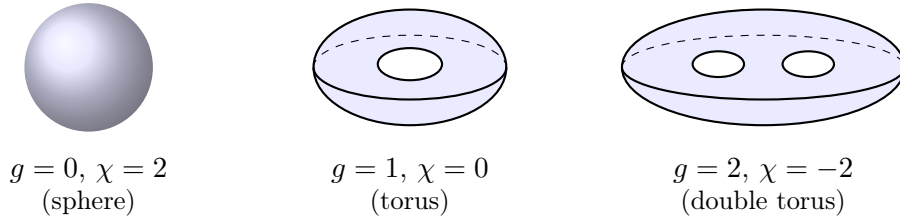


Figure 9. Orientable surfaces classified by genus g . The genus counts the number of handles: the sphere has none, the torus has one, and the double torus has two. The Euler characteristic $\chi = 2 - 2g$ decreases with increasing genus.

The topological invariant that appears naturally in the gravitational action is the **Euler characteristic**:

$$\chi = 2 - 2g. \quad (4.5)$$

For the sphere $\chi = 2$, for the torus $\chi = 0$, and for higher-genus surfaces $\chi < 0$. The Gauss–Bonnet theorem relates χ to the integral of the scalar curvature:

$$\frac{1}{4\pi} \int_{\mathcal{M}} d^2x \sqrt{g} R = \chi(\mathcal{M}) = 2 - 2g. \quad (4.6)$$

This is a topological invariant: it does not depend on the metric, only on the topology of the surface.

The genus expansion in JT gravity

Jackiw–Teitelboim (JT) gravity [47, 90] is a two-dimensional theory of gravity that serves as a powerful toy model for near-extremal black holes [56]; see Mertens and Turiaci [64] for a comprehensive review.

Its action is

$$I_{\text{JT}} = -S_0 \chi(\mathcal{M}) - \frac{1}{2} \int_{\mathcal{M}} d^2x \sqrt{g} \phi (R+2) - \int_{\partial\mathcal{M}} du \sqrt{h} \phi (K-1), \quad (4.7)$$

where ϕ is the dilaton field, K is the extrinsic curvature of the boundary, and S_0 is a topological coupling — the extremal entropy of the black hole.

The crucial feature is the topological term $-S_0 \chi$: each topology enters the path integral with a weight $e^{S_0 \chi}$. (For a surface with n boundaries the Euler characteristic is $\chi = 2 - 2g - n$ rather than $2 - 2g$, so, for example, the one-boundary disk carries a factor e^{S_0} ; since adding a handle lowers χ by two whatever the number of boundaries.) For a connected amplitude with a fixed number b of boundaries,

$$Z_{\text{conn},b}(\beta_1, \dots, \beta_b) \sim \sum_{g \geq 0} e^{(2-2g-b)S_0} Z_{g,b}(\beta_1, \dots, \beta_b). \quad (4.8)$$

Here $Z_{g,b}$ is the amplitude with the topological weight removed. Thus

$$\frac{e^{(2-2(g+1)-b)S_0} Z_{g+1,b}}{e^{(2-2g-b)S_0} Z_{g,b}} = e^{-2S_0} \frac{Z_{g+1,b}}{Z_{g,b}}. \quad (4.9)$$

At fixed genus and suitable fixed boundary parameters, large S_0 suppresses additional handles. The coefficients still matter: their growth makes this an asymptotic series, and other limits can compensate the topological suppression. A one-loop determinant on the disk is a correction at fixed topology; it does not add a handle.

Analogy with matrix models and large N

The structure of Eq. (4.8) is identical to the topological expansion of matrix integrals [91]. In a Hermitian matrix model with $N \times N$ matrices, Feynman diagrams can be drawn on surfaces of genus g , and the partition function admits the expansion

$$\ln Z_{\text{matrix}} = \sum_{g=0}^{\infty} N^{2-2g} F_g, \quad (4.10)$$

with the identification $N^2 \leftrightarrow e^{2S_0}$. This is not a coincidence. Saad et al. [79] showed, order by order in the genus expansion, that the perturbative series of JT gravity coincides with that of a specific double-scaled random matrix ensemble; the exact nonperturbative duality between the gravitational path integral and random matrix theory remains conjectural.

The correspondence relates the JT genus expansion to a matrix-integral expansion. A separate random-state problem describes the reduced density matrix of a bipartite Gaussian state. Its Wick contractions can also be drawn as ribbon graphs. At leading order, the surviving contractions are indexed by non-crossing partitions [85]. This is the connection used in Sections 5 and 6.

Planar contractions and non-crossing partitions

A ribbon graph records a cyclic order of edges at each vertex. Its faces supply factors of matrix dimension, and Euler's relation determines the power of that dimension. For the flat-weight Wishart moment problem, maximizing the number of faces selects permutations associated with non-crossing partitions. Their weights give the Marchenko–Pastur law. Changing the spectral weights changes the moment sum even when the contractions remain planar. In the gravitational problem these weights encode the preparation and disk amplitudes [21, 76, 96].

Calculation	Organization	Additional input
Connected JT amplitudes	Genus and boundaries	$Z_{g,b}(\beta_1, \dots)$
Gaussian matrix moments	Wick ribbon graphs	Matrix dimensions and covariance
Flat-weight planar moments	Non-crossing partitions	Uniform covariance
Weighted planar moments	Free moment expansions	Spectral weights

The shared combinatorial language is useful, but these are different amplitudes. In particular, a finite Wick sum at fixed moment order is not the complete JT sum over all handles. Section 9 returns to the distinction when discussing finite-size corrections.

Higher dimensions

In dimensions $d > 2$, the classification of manifolds by a single integer like the genus is no longer available — the homeomorphism problem for closed four-manifolds, for example, is algorithmically undecidable [59]. There is no general rule that a more complicated higher-dimensional topology has a larger Einstein action or an extra factor e^{-1/G_N} . The relative weights must be found from the admissible saddles and their actions in the theory and ensemble under consideration.

In the context of the replica trick and entropy calculations (developed in the next section), the relevant topologies in any dimension are “replica wormholes” — saddle points that connect different replicas of the spacetime. This Euclidean route to entanglement entropy is the same one by which the Ryu–Takayanagi formula itself is derived from the gravitational path integral [50]. The structure of the calculation, and the competition between connected and disconnected saddles, is qualitatively the same as in two dimensions, even though the detailed geometry is much richer.

4.3 Euclidean vs. Lorentzian: why we go to Euclidean signature

Throughout this section we have written the gravitational path integral in Euclidean signature, with a weight e^{-I} rather than e^{iS} . This subsection explains why.

Wick rotation

The passage from Lorentzian to Euclidean signature is achieved by the **Wick rotation**: the analytic continuation of the time coordinate to imaginary values,

$$t \longrightarrow -i\tau, \quad ds^2 = -dt^2 + d\vec{x}^2 \longrightarrow ds_E^2 = d\tau^2 + d\vec{x}^2. \quad (4.11)$$

Under this rotation, the Lorentzian action S becomes iI where I is the Euclidean action, and the oscillatory weight $e^{iS/\hbar}$ becomes the damped weight $e^{-I/\hbar}$:

$$e^{iS_{\text{Lor}}/\hbar} \longrightarrow e^{-I_{\text{Eucl}}/\hbar}. \quad (4.12)$$

For a stable bosonic action the Euclidean weight can improve convergence and makes steepest-descent estimates transparent. This improvement does not automatically hold in gravity, where the integration contour and negative modes require separate treatment. This is in contrast to the Lorentzian integral, where the oscillatory phase e^{iS} does not damp and the integral requires careful definition through $i\epsilon$ prescriptions.

In quantum field theory on flat space, the Wick rotation is a standard and well-understood procedure, related, under the appropriate positivity and regularity axioms, to Lorentzian correlation functions by Osterwalder–Schrader reconstruction [68, 69]. In gravity, the situation is more subtle.

The Euclidean action for gravity is not bounded below (the conformal factor problem), and not every Euclidean geometry has a Lorentzian continuation. Nevertheless, the Euclidean approach, used judiciously at the level of saddle-point geometries, has proven highly successful.

Thermal interpretation: periodicity and temperature

The most important consequence of the Wick rotation for black hole physics is the connection between **Euclidean time periodicity and temperature**.

Consider a quantum system at inverse temperature $\beta = 1/T$. Its thermal partition function is

$$Z(\beta) = \text{Tr}(e^{-\beta H}). \quad (4.13)$$

In the path-integral representation, this trace is computed by evolving in Euclidean time τ from 0 to β with **periodic boundary conditions**: the field configurations at $\tau = 0$ and $\tau = \beta$ are identified for bosonic variables; thermal fermions instead have antiperiodic boundary conditions. For a bosonic field,

$$Z(\beta) = \int_{\phi(\tau=0)=\phi(\tau=\beta)} \mathcal{D}\phi e^{-I_E[\phi]}. \quad (4.14)$$

The periodicity of Euclidean time,

$$\tau \sim \tau + \beta, \quad T = \frac{1}{\beta}, \quad (4.15)$$

is the Euclidean encoding of finite temperature. A system at temperature T is described by a Euclidean path integral on a geometry where the time direction is a circle of circumference $\beta = 1/T$.

The Euclidean black hole: the cigar geometry

This connection between periodicity and temperature acquires dramatic physical content when applied to the Schwarzschild black hole. Consider the Schwarzschild metric in the exterior region ($r > 2GM$):

$$ds^2 = -\left(1 - \frac{2GM}{r}\right) dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2 d\Omega^2. \quad (4.16)$$

Wick-rotating $t \rightarrow -i\tau$ gives the Euclidean Schwarzschild metric:

$$ds_E^2 = \left(1 - \frac{2GM}{r}\right) d\tau^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2 d\Omega^2. \quad (4.17)$$

Near the horizon $r = 2GM$, introduce the proper radial distance $\rho = 2\sqrt{2GM(r - 2GM)}$. Then the (τ, ρ) part of the metric becomes

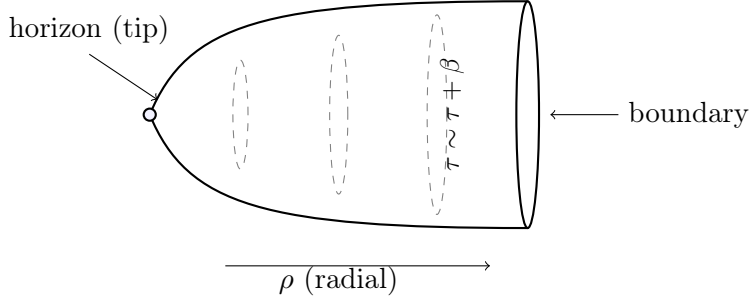
$$ds_E^2 \approx d\rho^2 + \frac{\rho^2}{(4GM)^2} d\tau^2. \quad (4.18)$$

This is the metric of flat $\mathbb{R}^2$ in polar coordinates, $ds^2 = d\rho^2 + \rho^2 d\theta^2$, provided we identify $\theta = \tau/(4GM)$. But flat space in polar coordinates is smooth at the origin $\rho = 0$ only if θ has period 2π . This requires

$$\frac{\tau}{4GM} \sim \frac{\tau}{4GM} + 2\pi \quad \implies \quad \tau \sim \tau + \underbrace{8\pi GM}_{\beta}, \quad (4.19)$$

and therefore the Hawking temperature is

$$T_H = \frac{1}{\beta} = \frac{1}{8\pi GM} = \frac{\hbar c^3}{8\pi G M k_B}. \quad (4.20)$$



$$\text{Smooth cap} \Rightarrow \beta = 8\pi GM \Rightarrow T_H = 1/(8\pi GM)$$

Figure 10. The Euclidean Schwarzschild black hole, often called the “cigar” geometry. Euclidean time τ is periodic, forming circles (dashed) whose radius shrinks to zero at the horizon, where the geometry smoothly caps off like the tip of a cigar. The requirement that this tip be smooth (no conical singularity) fixes $\beta = 8\pi GM$, yielding the Hawking temperature.

This is a striking result: the temperature of the Hawking radiation, originally derived by Hawking [38] through a difficult calculation of Bogoliubov coefficients in quantum field theory on a collapsing-star background, emerges here from a simple *regularity condition* on the Euclidean geometry. The requirement that the Euclidean continuation be smooth (free of conical singularities) at the horizon uniquely fixes the periodicity of Euclidean time, and hence the temperature.

The resulting geometry — shown in Figure 10 — is called the **cigar** (or “Euclidean cigar”). It has the topology of a disk $D^2 \times S^2$: the Euclidean time circle shrinks smoothly to zero size at the horizon, so the (τ, r) part of the geometry looks like a disk, capping off at its center. There is no interior region; the Euclidean black hole has no singularity or interior. The entire geometry is the smooth region outside the horizon, with the horizon itself realized as the smooth tip of the cigar.

The Euclidean action and black hole thermodynamics

The Euclidean action of the Schwarzschild black hole, evaluated on the cigar geometry with the Gibbons–Hawking–York boundary term properly included, gives [31]

$$I_{\text{cigar}} = \beta F = \beta M - S_{\text{BH}}, \quad (4.21)$$

where F is the free energy and

$$S_{\text{BH}} = \frac{A}{4G_N} \quad (4.22)$$

is the Bekenstein–Hawking entropy [12, 38]. Thus the Euclidean path integral evaluated on the black hole saddle gives $Z \approx e^{-I_{\text{cigar}}} = e^{S_{\text{BH}} - \beta M}$, which is the formal saddle contribution $e^{-\beta F}$. Asymptotically flat Schwarzschild has negative heat capacity and does not define a stable canonical equilibrium in an infinite heat bath. A cavity, a suitable finite reservoir, or a different stable branch is needed for an equilibrium fluctuation calculation (Section 8). The entire edifice of black hole thermodynamics — temperature, entropy, free energy — emerges naturally from the Euclidean gravitational path integral.

The price of going Euclidean

The Euclidean formulation is not without its difficulties. Two issues are particularly relevant for the developments in subsequent sections:

1. **Euclidean saddles without Lorentzian continuations.** Not every Euclidean saddle that contributes to the path integral has an obvious interpretation as a Lorentzian spacetime. The replica wormholes that will appear in Section 5 [1, 76] are a prime example: they are Euclidean geometries that connect n copies of the spacetime, and for non-integer n there is no literal n -fold cover. Continuation can instead be formulated using a quotient geometry with a conical defect. Even for integer n , the Lorentzian interpretation of these geometries is subtle and has been the subject of considerable recent work [24].
2. **The conformal factor problem.** The Euclidean Einstein–Hilbert action is not bounded below: conformal deformations $g_{\mu\nu} \rightarrow \Omega^2 g_{\mu\nu}$ can make the action arbitrarily negative. The Euclidean path integral, taken literally, does not converge. In practice, this is handled by specifying a contour of integration for the conformal factor, but the procedure introduces ambiguities that are not fully resolved.

Euclidean methods nevertheless provide well-tested semiclassical calculations of Hawking temperature, black-hole entropy, and transitions between admissible saddles. Their use requires the preparation, boundary conditions, and approximation to remain explicit.

Summary. The gravitational path integral motivates a sum over admissible geometries with quantum fluctuations around its saddles. In JT gravity, fixed-boundary amplitudes admit a genus expansion with an explicit topological weight. Matrix contractions provide a related planar organization, but the weights and observables must be specified separately. Euclidean regularity fixes the Schwarzschild temperature; it does not by itself establish canonical stability. The next section applies replica sewing to entropy and develops the normalized random-state benchmark.

5 The Replica Trick and Replica Wormholes

We develop the replica calculation first in a finite-dimensional random-state model, where normalization and every combinatorial coefficient can be checked exactly. We then explain its relation to the microcanonical, flat-weight limit of the west-coast model of replica wormholes. The distinction matters: the finite random-state formula is not a formula for every gravitational path integral. Within this benchmark, the complete planar sum resolves the transition missed by retaining only the fully disconnected and fully connected contractions. The calculation also supplies the preparation for the Marchenko–Pastur derivation in Section 6.

5.1 The replica trick

Our goal is to compute the von Neumann entropy of the radiation:

$$S_{\text{vN}} = -\text{Tr}(\rho \log \rho). \quad (5.1)$$

The difficulty is the logarithm. In a path integral formulation, ρ is computed as a functional integral over field configurations with specific boundary conditions. Taking the trace of ρ raised to an integer power is straightforward — it amounts to gluing boundary conditions cyclically — but taking the logarithm of an operator defined by a path integral has no natural geometric meaning.

The *replica trick* circumvents this obstacle by exploiting a simple identity. For any positive operator ρ with $\text{Tr } \rho = 1$,

$$S_{\text{vN}} = - \lim_{n \rightarrow 1} \frac{\partial}{\partial n} \text{Tr}(\rho^n). \quad (5.2)$$

The strategy is: (i) compute $\text{Tr}(\rho^n)$ for integer $n \geq 2$, where the calculation has a clean geometric interpretation; (ii) find a closed form in n ; (iii) analytically continue to $n \rightarrow 1$ and differentiate.

Why this works. For integer $n \geq 2$, write ρ in its eigenbasis: $\rho = \sum_{\alpha} \lambda_{\alpha} |\alpha\rangle\langle\alpha|$, so that $\text{Tr}(\rho^n) = \sum_{\alpha} \lambda_{\alpha}^n$. Each eigenvalue satisfies $0 \leq \lambda_{\alpha} \leq 1$, so λ_{α}^n is a perfectly well-defined analytic function of n for all $\text{Re } n > 0$. Differentiating at $n = 1$:

$$-\frac{\partial}{\partial n} \lambda_{\alpha}^n \Big|_{n=1} = -\lambda_{\alpha} \log \lambda_{\alpha}, \quad (5.3)$$

which, summed over α , gives S_{vN} . For a fixed finite-dimensional state this continuation is supplied by its spectrum, and $|\text{Tr } \rho^z| \leq 1$ for $\text{Re } z \geq 1$. An arbitrary interpolation of integer replica data need not be unique: a bound on the positive real axis alone does not control its growth in the complex plane. Section 6.6 states the additional analytic conditions used for the limiting Marchenko–Pastur moment integral.

The n -sheeted geometry. For integer n , the trace $\text{Tr}(\rho^n)$ has a direct path-integral interpretation. Consider a quantum field theory on a Euclidean spacetime M , and let ρ be the reduced density matrix of a spatial region A , obtained by tracing over the complement $\bar{A}$. In the path-integral language, ρ is represented by cutting the Euclidean manifold along A and leaving the two sides of the cut as “bra” and “ket” boundary conditions. To compute $\text{Tr}(\rho^n)$, one takes n copies of this cut manifold and glues them cyclically: the ket side of copy j is identified with the bra side of copy $j + 1$ (with $n + 1 \equiv 1$). The result is a partition function on an n -sheeted branched cover M_n of the original spacetime:

$$\text{Tr}(\rho^n) = \frac{Z_n}{Z_1^n}, \quad (5.4)$$

where $Z_n = \int \mathcal{D}\phi e^{-I[\phi; M_n]}$ is the Euclidean partition function on M_n and Z_1 is the single-copy partition function (the denominator ensures the correct normalization $\text{Tr } \rho = 1$).

Figure 11 illustrates this construction for $n = 3$. Three copies of the spacetime are arranged in a stack, each cut along the region A . The cyclic gluing $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$ produces a three-sheeted cover with a branch point at the boundary ∂A .

In gravity: dynamical topology. In a non-gravitational QFT, the manifold M_n is fixed and the path integral sums only over fields on that manifold. In a theory of gravity, the manifold itself is dynamical. A chosen gravitational path integral may include connected bulk geometries with the prescribed replicated boundary data. A stationary geometry connecting distinct replicas is called a *replica wormhole*. Whether a topology admits a contributing saddle, and with what weight, depends on the theory, boundary conditions, and integration contour. The flat-weight model below makes the relevant replica contractions explicit.

The rest of this section makes this idea precise in the simplest possible context: the west-coast model.

5.2 The west-coast model

The west-coast model, introduced by Penington, Shenker, Stanford, and Yang [76] (building on the JT gravity framework of Saad et al. [79]), is a solvable toy model that captures the essential physics of

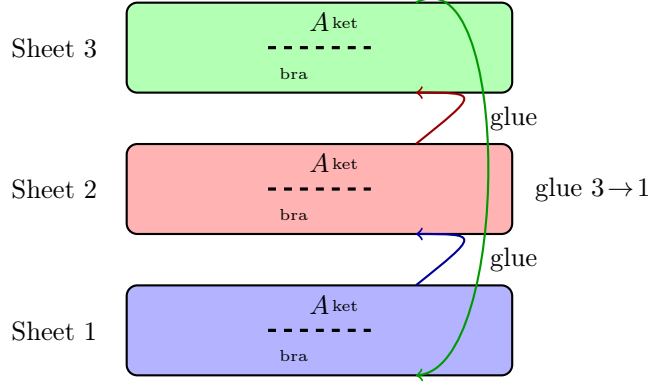


Figure 11. The n -sheeted replica geometry for $n = 3$. Each sheet represents one copy of the Euclidean spacetime, cut along the entangling region A . The cyclic identification (ket of sheet $j \rightarrow$ bra of sheet $j + 1$, with $3 + 1 \equiv 1$) produces the branched cover M_3 .

replica wormholes while remaining exactly computable. It consists of Jackiw–Teitelboim (JT) gravity on a disk with k flavors of end-of-the-world (EOW) branes.

The Hilbert space and the flat-weight approximation. For the finite-dimensional benchmark let $\mathcal{H}_{\text{BH}} = \mathbb{C}^L$ and $\mathcal{H}_{\text{rad}} = \mathbb{C}^k$. The gravitational interpretation requires a chosen microcanonical energy window and approximately constant state-preparation weights in that window. Its effective number of black-hole states is represented by L ; writing $L = e^{S_0}$ is a useful simplified counting convention, not an identity between the full JT Hilbert space and a finite space of that dimension. In a canonical JT preparation the energy density and brane weights must instead be retained.

We write $\gamma = k/L$. Scanning k at fixed L gives a family of states with an entanglement crossover at $\gamma = 1$. It does not specify a unitary evaporation channel. A dynamical model may have dimensions $k(t)$ and $L(t)$; its time dependence and energy constraints require separate assumptions.

The state. The combined state of the black hole and radiation is

$$|\psi\rangle = \frac{1}{\sqrt{Z}} \sum_{i=1}^k |\psi_i\rangle_{\text{BH}} \otimes |i\rangle_{\text{rad}}, \quad (5.5)$$

where $|\psi_i\rangle_{\text{BH}}$ is the state of the black hole interior conditioned on having emitted a brane of flavor i , and $Z = \sum_i \langle \psi_i | \psi_i \rangle$ is the normalization. This state is the gravity analog of a random bipartite pure state: if the states $|\psi_i\rangle$ were orthonormal, $|\psi\rangle$ would simply be a maximally entangled state of Schmidt rank k . The key point is that the overlaps $\langle \psi_j | \psi_i \rangle$ are *not* specified a priori — they are computed by the gravitational path integral.

Overlaps from gravity. The overlap $\langle \psi_j | \psi_i \rangle$ is computed by a Euclidean path integral on a manifold whose boundary conditions require brane flavor i on one side and flavor j on the other. At leading order (disk topology), only the $i = j$ boundary condition admits a solution:

$$\mathbb{E} \langle \psi_j | \psi_i \rangle \Big|_{\text{flat disk}} = L \delta_{ij}. \quad (5.6)$$

Replacing the overlaps and the normalization by these disk means gives

$$\rho_{\text{disc}} = \frac{1}{kL} \sum_{i,j} \mathbb{E} \langle \psi_j | \psi_i \rangle |i\rangle \langle j| = \frac{\mathbb{K}_k}{k}. \quad (5.7)$$

This is the disconnected approximation. In the isotropic Gaussian benchmark, unitary invariance also gives $\mathbb{E} \rho_{\text{rad}} = \mathbb{K}_k/k$ for the exactly normalized states. Neither statement identifies individual realizations with their average. In particular, for $k > L$ every realization has rank at most L , whereas its averaged matrix has rank k . Products of overlaps have connected contributions even when their individual means vanish. These contributions restore the nontrivial spectral moments; computing a moment of the averaged matrix would discard them.

The density matrix and its moments. Tracing over the black hole yields

$$\rho_{\text{rad}} = Z^{-1} \sum_{i,j} \langle \psi_j | \psi_i \rangle |i\rangle \langle j|.$$

The n th moment of ρ_{rad} involves a product of n such overlaps, summed over brane indices in the pattern dictated by the cyclic trace:

$$\text{Tr}(\rho_{\text{rad}}^n) = \frac{1}{Z^n} \sum_{i_1, \dots, i_n} \langle \psi_{i_2} | \psi_{i_1} \rangle \langle \psi_{i_3} | \psi_{i_2} \rangle \cdots \langle \psi_{i_1} | \psi_{i_n} \rangle. \quad (5.8)$$

Each product of overlaps is now computed by a gravitational path integral on n copies of the disk, with the crucial point being that the bulk geometry is free to connect these copies through wormholes. The different topological connections are labeled by permutations, as we explain next.

5.3 The master formula: permutations and topology

Let G be a $k \times L$ matrix of independent standard complex Gaussian entries and set

$$W = GG^\dagger, \quad T = \text{Tr } W, \quad \rho = \frac{W}{T}, \quad \tilde{\rho} = \frac{W}{kL}. \quad (5.9)$$

The matrix ρ is the reduction of a normalized Haar-random pure state. The matrix $\tilde{\rho}$ has unit trace only in expectation. Write $c_n = (1\,2\,\dots\,n)$, let $C(\sigma)$ count cycles, and define $(D)_n = D(D+1)\cdots(D+n-1)$. The exact normalized formula is

$$\boxed{\mathbb{E}_{\text{H}} \text{Tr } \rho^n = \frac{1}{(kL)_n} \sum_{\sigma \in S_n} k^{C(c_n^{-1}\sigma)} L^{C(\sigma)}.} \quad (5.10)$$

Here the expectation is over Haar states. The cycle orientation can be reversed by relabelling the permutations; it must be kept consistent with the chosen trace convention.

Derivation and the role of normalization. Expanding $\text{Tr } W^n$ produces n Gaussian entries and n conjugate entries. Wick's rule pairs them according to $\sigma \in S_n$. The black-hole indices form $C(\sigma)$ independent loops, each contributing L . The cyclic radiation trace leaves $C(c_n^{-1}\sigma)$ loops, each contributing k . Thus the numerator in Eq. (5.10) is $\mathbb{E} \text{Tr } W^n$.

The Gaussian vector of kL entries has independent radial and angular parts. Its squared norm T has the gamma distribution with shape kL and unit scale, so $\mathbb{E} T^n = (kL)_n$. The angular part is

a Haar vector, and $W = T\rho$ with T independent of ρ . Dividing $\mathbb{E} \text{Tr} W^n = \mathbb{E} T^n \mathbb{E}_H \text{Tr} \rho^n$ proves the formula. For the alternative normalization one obtains instead

$$\mathbb{E} \text{Tr} \tilde{\rho}^n = (kL)^{-n} \sum_{\sigma \in S_n} k^{C(c_n^{-1}\sigma)} L^{C(\sigma)}, \quad \mathbb{E}_H \text{Tr} \rho^n = \frac{(kL)^n}{(kL)_n} \mathbb{E} \text{Tr} \tilde{\rho}^n. \quad (5.11)$$

This second formula is convenient for organizing Wick contractions by genus. For fixed n and fixed positive k/L , the prefactor approaches one with relative corrections of order $(kL)^{-1}$, but these corrections cannot be omitted in a finite-size calculation.

A two-qubit normalization check. At $n = 2$, the identity and transposition give

$$\mathbb{E}_H \text{Tr} \rho^2 = \frac{k+L}{kL+1}, \quad \mathbb{E} \text{Tr} \tilde{\rho}^2 = \frac{1}{k} + \frac{1}{L}. \quad (5.12)$$

The normalized result is Lubkin's mean-purity formula [53]. For two qubits, $k = L = 2$, it is $4/5$, whereas the mean-normalized expression is 1. For $k = 1$ the normalized reduced state is necessarily pure and the first formula gives 1 exactly; the second gives $1 + 1/L$. These elementary checks identify the object being computed before any replica continuation. The Gaussian normalization of $\tilde{\rho}$ is not a harmless exact replacement for normalization of each state.

What carries over to gravity. In the microcanonical flat-weight approximation, the cycle factors reproduce the corresponding replica contractions of the west-coast model [76]. The factors $L^{C(\sigma)}$ are index-loop factors in the Gaussian derivation, not propagating graviton loops. General JT preparations assign energy-dependent partition functions to connected components. Their planar sums retain non-crossing combinatorics but need not give the same Marchenko–Pastur law. Additional continuum handles and moduli integrals are not enumerated by this finite Wick sum.

The genus of a permutation. Not all permutations contribute at the same order. The ribbon graph associated with a Wick contraction has a genus:

$$g(\sigma) = \frac{n+1 - C(\sigma) - C(c_n^{-1}\sigma)}{2}. \quad (5.13)$$

This formula has a precise topological origin: the permutation σ and the fixed cyclic permutation c_n together define a branched covering of the sphere, and $g(\sigma)$ is the genus of the covering surface (by the Riemann–Hurwitz formula). At fixed n and fixed $\gamma \in (0, \infty)$, each additional genus costs two powers of a dimension. If $L = e^{S_0}$ is adopted as the counting convention, this is the factor e^{-2gS_0} , up to fixed ratio factors. This is a statement about the Gaussian index expansion.

Rewriting Eq. (5.11) makes this expansion explicit:

$$\mathbb{E} \text{Tr}(\tilde{\rho}^n) = k^{1-n} \sum_{\sigma \in S_n} k^{-2g(\sigma)} \gamma^{n-C(\sigma)}, \quad (5.14)$$

where $\gamma = k/L$ and we have used $C(c_n^{-1}\sigma) = n+1 - C(\sigma) - 2g(\sigma)$ to write the per-permutation weight as $(kL)^{-n} k^{C(c_n^{-1}\sigma)} L^{C(\sigma)} = k^{1-n-2g(\sigma)} \gamma^{n-C(\sigma)}$, eliminating $C(c_n^{-1}\sigma)$ in favor of $g(\sigma)$ and $C(\sigma)$.

The genus-zero sector: non-crossing permutations. The genus-zero condition $g(\sigma) = 0$ is equivalent to

$$C(\sigma) + C(c_n^{-1}\sigma) = n+1. \quad (5.15)$$

Permutations satisfying this condition are precisely the *non-crossing permutations* $\text{NC}(n)$: each block is written as a cycle in the cyclic order inherited from c_n . This orientation is part of the definition.

These are the permutations that, when represented as connection patterns on n points arranged on a circle (in the order dictated by c_n), have no crossings. The number of non-crossing permutations of n elements is the *Catalan number*:

$$|\text{NC}(n)| = C_n = \frac{1}{n+1} \binom{2n}{n}. \quad (5.16)$$

The first few values are $C_1 = 1$, $C_2 = 2$, $C_3 = 5$, $C_4 = 14$, $C_5 = 42$, $C_6 = 132$, $C_7 = 429$.

Graphical interpretation. The connection between permutations and non-crossing partitions has a vivid graphical representation, shown in Figure 12. Place the numbers $1, 2, \dots, n$ on a circle in the order determined by the cyclic permutation c_n . For each cycle $(a_1 a_2 \dots a_m)$ of σ , draw arcs connecting $a_1 \rightarrow a_2 \rightarrow \dots \rightarrow a_m \rightarrow a_1$. If the cycles follow the prescribed orientation and their polygons have disjoint interiors, the permutation is non-crossing and contributes at genus zero. If crossings are unavoidable, the permutation has $g(\sigma) \geq 1$ and is suppressed by at least $L^{-2} = e^{-2S_0}$.

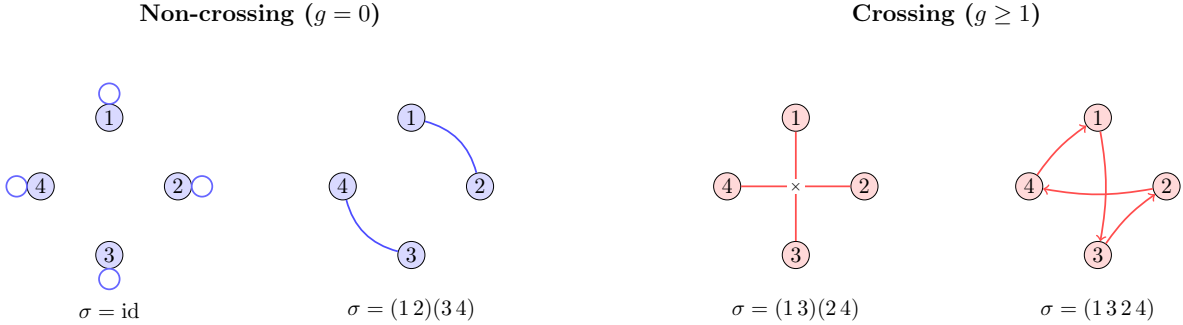


Figure 12. Permutations of $n = 4$ elements arranged on a circle in the order $(1, 2, 3, 4)$ determined by the cyclic permutation c_n . **Left:** Non-crossing permutations (genus 0): the identity $\sigma = \text{id}$ has four fixed points (self-loops), and $\sigma = (1\ 2)(3\ 4)$ has two non-crossing transpositions of adjacent elements. **Right:** Crossing permutations (genus ≥ 1): $\sigma = (1\ 3)(2\ 4)$ forces the arcs $1\text{--}3$ and $2\text{--}4$ to cross, and the 4-cycle $\sigma = (1\ 3\ 2\ 4)$ similarly produces unavoidable crossings. At genus zero, only non-crossing permutations contribute, and for $n = 4$ there are $C_4 = 14$ of them versus $4! = 24$ total.

The genus-zero moment. Restricting Eq. (5.11) to genus zero, and using the genus-zero constraint $C(c_n^{-1}\sigma) = n + 1 - C(\sigma)$ to simplify:

$$\mathbb{E} \text{Tr}(\tilde{\rho}^n) \Big|_{g=0} = k^{1-n} \sum_{\sigma \in \text{NC}(n)} \gamma^{n-C(\sigma)} = k^{1-n} M_n(\gamma) \quad (5.17)$$

where $M_n(\gamma) = \sum_{\sigma \in \text{NC}(n)} \gamma^{n-C(\sigma)}$ is the n th moment of the Marchenko–Pastur distribution with parameter γ [58]. (As a check, at $n = 2$ this gives $\text{Tr}(\rho^2)|_{g=0} = k^{-1}(1 + \gamma) = 1/k + 1/L$, the leading large-dimension purity.) The combination of the non-crossing combinatorics and these flat cycle weights determines the distribution. For normalized Haar states the precise limiting statement is $k^{n-1} \mathbb{E}_H \text{Tr} \rho^n \rightarrow M_n(\gamma)$ for fixed n , as developed in Section 6.

The non-crossing permutations can be further organized by cycle number using the *Narayana numbers* $N(n, j) = \frac{1}{n} \binom{n}{j} \binom{n}{j-1}$, which count the number of non-crossing permutations with exactly j

cycles:

$$M_n(\gamma) = \sum_{j=1}^n N(n, j) \gamma^{n-j}. \quad (5.18)$$

Each Narayana stratum $N(n, j)$ represents a family of partially connected topologies with the same number of bulk connected components, and the Narayana decomposition organizes the genus-zero sector into a hierarchy of increasing connectivity.

5.4 The two-saddle approximation and its failure

The leading saddle comparison underlying the island transition [4, 5, 74] motivates keeping two extreme contractions. In the flat-weight model these are the following terms of Eq. (5.11). The complete planar sum was already included in the resolvent analysis of Penington et al. [76].

1. **The identity** $\sigma = \text{id}$, corresponding to fully *disconnected* replicas. This permutation has $C(\text{id}) = n$ cycles (every element is a fixed point) and $C(c_n^{-1} \text{id}) = C(c_n^{-1}) = 1$ cycle. Its contribution is

$$\mathbb{E} \text{Tr}(\tilde{\rho}^n) \Big|_{\text{id}} = (kL)^{-n} \cdot k^1 \cdot L^n = k^{1-n}. \quad (5.19)$$

In units of the overall genus-zero prefactor k^{1-n} , this is the γ^0 term in the Narayana expansion (the $j = n$ stratum, with $N(n, n) = 1$).

2. **The cyclic permutation** $\sigma = c_n = (1 \ 2 \ \cdots \ n)$, corresponding to the fully *connected* wormhole. This permutation has $C(c_n) = 1$ cycle and $C(c_n^{-1} c_n) = C(\text{id}) = n$ cycles. Its contribution is

$$\mathbb{E} \text{Tr}(\tilde{\rho}^n) \Big|_{c_n} = (kL)^{-n} \cdot k^n \cdot L^1 = L^{1-n} = k^{1-n} \gamma^{n-1}. \quad (5.20)$$

This is the $j = 1$ stratum (Narayana number $N(n, 1) = 1$), with the highest power of γ .

These are the two extreme Narayana strata — $j = n$ (the identity) and $j = 1$ (the full cycle) — each of which contains a single permutation, since $N(n, n) = N(n, 1) = 1$. Together they give the *two-saddle approximation*:

$$\mathbb{E} \text{Tr}(\tilde{\rho}^n)_{\text{two-saddle}} = k^{1-n} (1 + \gamma^{n-1}) = k^{1-n} + L^{1-n}. \quad (5.21)$$

What is missing. For $n \geq 3$, there are $C_n - 2$ additional non-crossing permutations that are neither the identity nor the full cycle. These correspond to *partially connected* wormhole topologies: configurations where some — but not all — of the n replicas are connected through the bulk. In the flat-weight replica model these are additional planar contractions. At fixed γ , each contributes at the same order in the dimension expansion as the identity and cyclic contractions.

The exact genus-zero result is

$$\mathbb{E} \text{Tr}(\tilde{\rho}^n) \Big|_{g=0} = k^{1-n} \sum_{j=1}^n N(n, j) \gamma^{n-j} = k^{1-n} M_n(\gamma), \quad (5.22)$$

while the two-saddle approximation retains only the $j = 1$ and $j = n$ terms. The error is

$$\mathbb{E} \text{Tr}(\tilde{\rho}^n) \Big|_{g=0} - \mathbb{E} \text{Tr}(\tilde{\rho}^n)_{\text{two-saddle}} = k^{1-n} \sum_{j=2}^{n-1} N(n, j) \gamma^{n-j}, \quad (5.23)$$

which involves $\sum_{j=2}^{n-1} N(n, j) = C_n - 2$ permutations.

The error at the Page time. The failure is most dramatic at the Page time $\gamma = 1$ ($k = L$), where every non-crossing permutation contributes equally:

$$\left. \frac{\mathbb{E} \operatorname{Tr}(\tilde{\rho}^n)|_{g=0}}{\mathbb{E} \operatorname{Tr}(\tilde{\rho}^n)_{\text{two-saddle}}} \right|_{\gamma=1} = \frac{\sum_{j=1}^n N(n, j)}{1+1} = \frac{C_n}{2}. \quad (5.24)$$

This ratio grows rapidly with n . Using the asymptotic form of the Catalan numbers, $C_n \sim 4^n/(n^{3/2}\sqrt{\pi})$, the error grows as

$$\frac{C_n}{2} \sim \frac{4^n}{2n^{3/2}\sqrt{\pi}}, \quad (5.25)$$

which is *exponential* in n . This is a ratio of planar integer moments for $n \geq 2$. By itself it does not determine the error in a derivative at $n = 1$. That entropy error must be computed from the full limiting density, which is done in the next section. The large- n asymptotic displayed here is not a claim that the fixed- n planar expansion is uniform at arbitrarily large replica index.

Table 2 displays the ratio $C_n/2$ for small values of n , along with the number of missing saddles $C_n - 2$.

Table 2. Failure of the two-saddle approximation at the Page time $\gamma = 1$. The ratio $\mathbb{E} \operatorname{Tr}(\tilde{\rho}^n)|_{g=0} / \mathbb{E} \operatorname{Tr}(\tilde{\rho}^n)_{\text{two-saddle}} = C_n/2$ measures the multiplicative error from omitting partially connected wormholes. The number of missing genus-zero saddles is $C_n - 2$.

n	C_n	$C_n - 2$ (missing saddles)	$C_n/2$ (error ratio)	$4^n/(2n^{3/2}\sqrt{\pi})$
2	2	0	1.00	1.60
3	5	3	2.50	3.47
4	14	12	7.00	9.03
5	42	40	21.00	25.84
6	132	130	66.00	78.62
7	429	427	214.50	249.56

Several features of Table 2 deserve comment:

- For $n = 2$, the two-term sum exhausts the Wick numerator: there are only $C_2 = 2$ non-crossing permutations (the identity and the transposition), and both are included. The exact normalized moment still requires the denominator $kL(kL + 1)$. In addition, the logarithm of a mean purity and the mean logarithm of purity are different ensemble quantities.
- For $n = 3$, three partially connected saddles are already missing, and the error ratio is $5/2 = 2.5$. The two-saddle answer underestimates the planar mean-normalized third moment by a factor of 2.5 at $k = L$.
- For $n = 7$, the two-saddle approximation misses 427 of the 429 genus-zero topologies, and the error exceeds a factor of 200.
- The exact ratio $C_n/2$ approaches the asymptotic form $4^n/(2n^{3/2}\sqrt{\pi})$ from below. The convergence is somewhat slow (the ratio of exact to asymptotic is still ≈ 0.86 at $n = 7$) because the sub-leading corrections to the Catalan asymptotics are significant for small n .

Physical interpretation. The partially connected wormholes that the two-saddle approximation discards have a clear physical meaning. They are topologies in which *some* replicas are connected by wormholes while others remain disconnected. For example, at $n = 4$, the permutation $\sigma = (1\ 2)(3)(4)$ describes a configuration where replicas 1 and 2 are connected through a wormhole while replicas 3 and 4 remain independent. This permutation has $C(\sigma) = 3$ cycles and contributes at genus zero (one can verify that the arcs do not cross on the circle diagram).

The reason these partially connected topologies matter is that, near the Page transition, the fully disconnected and fully connected saddles are comparable in magnitude (both scale as L^{1-n} when $\gamma = 1$), and the partially connected saddles are *also* comparable. There is no parametric suppression of one saddle relative to another at $\gamma = 1$; all C_n genus-zero saddles contribute on an equal footing. Selecting just the dominant saddle replaces the crossover by a switch between the two extreme contractions. The complete planar spectral law resolves the associated entropy cusp. Keeping their sum at integer n and selecting a dominant entropy saddle are distinct approximations.

Consequences for the von Neumann entropy. For $n \geq 2$ the two extreme contractions give a useful asymptotic description when one dimension is much larger than the other. Adding those two terms and then setting $n = 1$ would count the same identity contraction twice: the result would be 2 rather than 1. Thus the two-term integer formula is not itself a normalized replica continuation near one.

The full planar sum avoids this ambiguity by specifying a probability density. Its entropy deficit at $k = L$ is $1/2$ nat in the large-dimension limit. That constant is the leading Page correction [71], and its replica resolvent derivation appears in Penington et al. [76]; see also the discussion of partially connected saddles in Marolf and Maxfield [60]. The Narayana organization in the companion paper [99] makes the contributions of each cycle count explicit. It does not introduce a new planar entropy law.

5.5 Summary: from permutations to distributions

The replica trace fixes a cyclic ordering; Wick contractions provide a permutation sum; and the leading contractions at fixed dimension ratio are non-crossing. For normalized Haar states, the rising factorial in Eq. (5.10) is part of the exact answer. The mean-normalized formula is useful bookkeeping for the planar limit, where its Narayana polynomial becomes the Marchenko–Pastur moment. This chain of statements explains why the flat-weight gravitational model has random-state spectral statistics, while leaving energy weights and continuum topology to their separate gravitational calculation.

Part III

Free Probability and the Marchenko–Pastur Distribution

6 Free Probability and the Planar Random-State Limit

Section 5 reduced the flat-weight replica calculation to a sum over non-crossing partitions. We now derive its limiting spectral density, the Marchenko–Pastur law, and then its entropy. The calculation is a concrete application of free probability, with all cycle weights specified. Non-crossing combinatorics alone would not determine this particular law.

The gravitational use of free probability has a broader established form. Wang [96] identifies a free multiplicative convolution in the west-coast model, with gravitational partition functions entering the free cumulants. Cheng et al. [21] obtain related laws for random tensor networks and competing cuts. The simple Marchenko–Pastur law discussed here requires the flat cycle weights of the finite-dimensional benchmark; more general matter spectra and energy-dependent gravitational weights give other free convolutions. We use the moment–cumulant formalism of Speicher [85] within Voiculescu’s free probability [95].

The result connects three familiar calculations: the planar Wick sum, the large-dimension spectrum of a reduced Haar state, and the leading Page entropy [52, 53, 58, 71]. The companion paper [99] develops the cycle-count organization and the effect of non-flat preparation weights. This section supplies the background needed to understand those distinctions.

6.1 Classical vs. free probability

We begin with a brief comparison of classical and free probability, emphasizing the single structural change — the restriction from all partitions to non-crossing partitions — that connects the two theories.

Classical independence. In classical probability, independence of X and Y implies factorization of their joint moments whenever those moments exist:

$$\mathbb{E}[X^m Y^n] = \mathbb{E}[X^m] \mathbb{E}[Y^n]. \quad (6.1)$$

The moment-cumulant relation expresses moments in terms of cumulants χ_m by summing over *all* partitions of the set $\{1, 2, \dots, n\}$:

$$M_n = \sum_{\pi \in \mathcal{P}(n)} \prod_{B \in \pi} \chi_{|B|}, \quad (6.2)$$

where $\mathcal{P}(n)$ denotes the set of all set partitions of $\{1, \dots, n\}$ and the product runs over the blocks B of the partition π . The number of partitions $|\mathcal{P}(n)|$ is the Bell number B_n ; the first few values are $B_1 = 1$, $B_2 = 2$, $B_3 = 5$, $B_4 = 15$, $B_5 = 52$.

Free independence. Free probability, introduced by Voiculescu [95], replaces the notion of independence with *free independence* (or *freeness*), a non-commutative analogue appropriate for operators on a Hilbert space. More precisely, unital subalgebras are free with respect to a normalized trace τ when $\tau(a_1 \cdots a_r) = 0$ for centered elements $\tau(a_j) = 0$ from alternating subalgebras. This condition replaces the factorization used for independent commuting variables. Its combinatorial consequence is the moment-cumulant relation (6.2) is replaced by a sum restricted to *non-crossing* partitions:

$$M_n = \sum_{\pi \in \text{NC}(n)} \prod_{B \in \pi} \kappa_{|B|} \quad (6.3)$$

where $\text{NC}(n)$ denotes the lattice of non-crossing partitions of $\{1, \dots, n\}$, and κ_m are the *free cumulants*, introduced by Speicher [85].

The gravitational connection. In the flat-weight replica calculation, the factors associated with each cycle factorize over the blocks of a non-crossing partition. This is exactly the algebraic form of a free moment–cumulant relation. The cycle weights then specify the free cumulants. General gravitational models must establish both the contributing contractions and their weights before this identification can be used. Planarity organizes many matrix-model limits [91], but it does not by itself fix their spectral density.

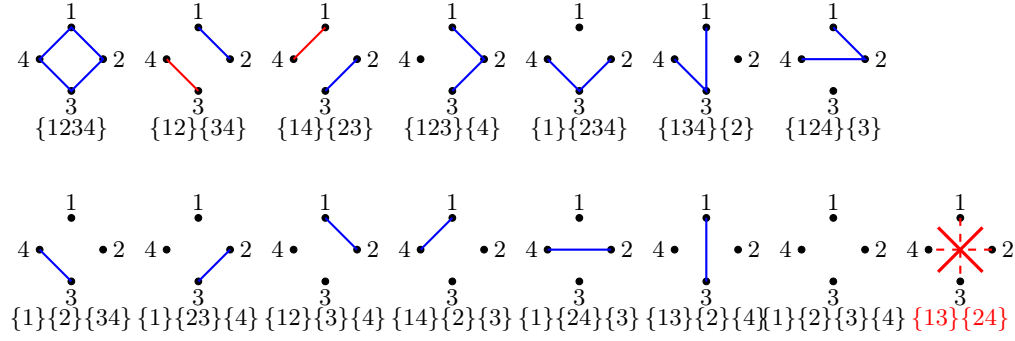


Figure 13. Non-crossing partitions of $\{1, 2, 3, 4\}$, arranged on circles. All $C_4 = 14$ non-crossing partitions are shown. The sole crossing partition $\{\{1, 3\}, \{2, 4\}\}$ is displayed last with dashed red arcs and an “X” to indicate that it is excluded from $\text{NC}(4)$. In the planar flat-weight replica model used here, these partitions label replica contraction patterns; each block labels a connected replica component.

6.2 Non-crossing partitions, Catalan numbers, and Narayana numbers

Definition. A partition π of $\{1, 2, \dots, n\}$ is *non-crossing* if there do not exist elements $a < b < c < d$ such that a and c belong to one block and b and d belong to a different block. Equivalently, when the elements are arranged on a circle and each block is represented by a convex hull, no two blocks cross.

Examples. For small n , the non-crossing partitions are:

- $n = 1$: $\{\{1\}\}$. One partition, one block.
- $n = 2$: $\{\{1, 2\}\}$ and $\{\{1\}, \{2\}\}$. Two partitions.
- $n = 3$: Five partitions: $\{\{1, 2, 3\}\}$, $\{\{1, 2\}, \{3\}\}$, $\{\{1, 3\}, \{2\}\}$, $\{\{1\}, \{2, 3\}\}$, $\{\{1\}, \{2\}, \{3\}\}$.
- $n = 4$: Fourteen partitions, listed in Figure 13.

The sole *crossing* partition for $n = 4$ that would otherwise be included among the 15 set partitions is $\{\{1, 3\}, \{2, 4\}\}$: placing the elements 1, 2, 3, 4 on a circle, the arcs connecting 1 to 3 and 2 to 4 necessarily cross.

Catalan numbers. The number of non-crossing partitions of $\{1, \dots, n\}$ is the n -th Catalan number,

$$C_n = \frac{1}{n+1} \binom{2n}{n}, \quad (6.4)$$

with values $C_0 = 1$, $C_1 = 1$, $C_2 = 2$, $C_3 = 5$, $C_4 = 14$, $C_5 = 42$, $C_6 = 132$. The Catalan numbers grow asymptotically as $C_n \sim 4^n / (n^{3/2} \sqrt{\pi})$, which is substantially slower than the Bell numbers B_n counting all partitions. The restriction to non-crossing partitions thus represents a dramatic simplification of the combinatorial structure.

Narayana numbers. A finer invariant of a non-crossing partition is its number of blocks. The number of non-crossing partitions of $\{1, \dots, n\}$ with exactly j blocks is the *Narayana number*

$$N(n, j) = \frac{1}{n} \binom{n}{j} \binom{n}{j-1}. \quad (6.5)$$

Table 3. Narayana numbers $N(n, j)$: the number of non-crossing partitions of $\{1, \dots, n\}$ with exactly j blocks. Row sums give the Catalan numbers C_n . In the planar flat-weight replica model used here, j counts the connected replica components of the contraction pattern.

$n \setminus j$	1	2	3	4	5	C_n
2	1	1				2
3	1	3	1			5
4	1	6	6	1		14
5	1	10	20	10	1	42

The Narayana numbers refine the Catalan numbers: $\sum_{j=1}^n N(n, j) = C_n$. They satisfy the symmetry

$$N(n, j) = N(n, n + 1 - j), \quad (6.6)$$

reflecting a duality between a partition and its Kreweras complement [48].

Table 3 lists the Narayana numbers for small n .

Gravitational interpretation. In the planar flat-weight replica model, a non-crossing partition of $\{1, \dots, n\}$ with j blocks labels a contraction pattern with j connected replica components. The west-coast interpretation associates a connected component with a gravitational amplitude joining the corresponding replicas. The Narayana number $N(n, j)$ counts these contraction patterns. It does not enumerate every geometry or saddle that may contribute to a gravitational amplitude with those boundary data.

6.3 Free cumulants and the moment-cumulant relation

We now make the free-probability structure precise. Recall the moment-cumulant relation (6.3):

$$M_n = \sum_{\pi \in \text{NC}(n)} \prod_{B \in \pi} \kappa_{|B|}. \quad (6.7)$$

This relation defines the free cumulants κ_m recursively from the moments M_n , or vice versa. The first few relations are:

$$M_1 = \kappa_1, \quad (6.8)$$

$$M_2 = \kappa_2 + \kappa_1^2, \quad (6.9)$$

$$M_3 = \kappa_3 + 3\kappa_2\kappa_1 + \kappa_1^3, \quad (6.10)$$

$$M_4 = \kappa_4 + 4\kappa_3\kappa_1 + 2\kappa_2^2 + 6\kappa_2\kappa_1^2 + \kappa_1^4. \quad (6.11)$$

The coefficients in these expressions are precisely the Narayana numbers from Table 3: for M_4 , the terms with one block (κ_4) appear $N(4, 1) = 1$ time, those with two blocks ($\kappa_3\kappa_1$ and κ_2^2) appear a total of $N(4, 2) = 6$ times, and so on.

Classical comparison. The classical moment-cumulant relation has the same form but sums over *all* partitions $\mathcal{P}(n)$. For instance, the classical relation for M_4 includes an additional term from the crossing partition $\{\{1, 3\}, \{2, 4\}\}$:

$$M_4^{\text{classical}} = \chi_4 + 4\chi_3\chi_1 + 3\chi_2^2 + 6\chi_2\chi_1^2 + \chi_1^4, \quad (6.12)$$

where the coefficient of χ_2^2 is 3 (from the three partitions into two pairs) rather than 2. The single additional contribution comes from the crossing partition $\{\{1, 3\}, \{2, 4\}\}$.

Free cumulants of the flat-weight model. The Narayana moment formula of Section 5 is reproduced by the free cumulants

$$\boxed{\kappa_m = \gamma^{m-1}, \quad \gamma = \frac{k}{L}}, \quad (6.13)$$

where k is the dimension of the radiation Hilbert space and L is the dimension of the black hole interior Hilbert space. The ratio γ compares the two dimensions. Calling its crossing of one a Page time additionally requires a specified time-dependent preparation.

The identification (6.13) is the central quantitative result of the genus-zero analysis. The free cumulants $\kappa_m = \gamma^{m-1}$ are precisely those of the *free Poisson distribution*, also known as the *Marchenko–Pastur distribution*. We shall derive this distribution explicitly in Section 6.5.

6.4 The R -transform

Free probability provides a generating-function technology, the R -transform, that packages the free cumulants into an analytic function and plays the role of the logarithm of the moment generating function in classical probability.

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

$$R(z) = \sum_{m=1}^{\infty} \kappa_m z^{m-1}. \quad (6.14)$$

For the flat-weight free cumulants (6.13), this becomes a geometric series:

$$\boxed{R(z) = \sum_{m=1}^{\infty} \gamma^{m-1} z^{m-1} = \frac{1}{1 - \gamma z}}. \quad (6.15)$$

Free convolution. The key property of the R -transform is its behavior under *free convolution* $\boxplus$: if X and Y are freely independent, then

$$R_{X+Y}(z) = R_X(z) + R_Y(z). \quad (6.16)$$

This is the free-probability analogue of the fact that the log-MGF (cumulant generating function) is additive under classical convolution. Free convolution describes limiting eigenvalue distributions of sums when the matrices are asymptotically free, as in suitable independent unitarily invariant ensembles. Classical independence of matrix entries or of the two matrices alone is insufficient. The additivity of R makes it useful once the freeness hypotheses hold.

From R -transform to spectral density. The R -transform determines the Cauchy (or Stieltjes) transform $G(z)$ of the eigenvalue density through the functional equation

$$G(R(z) + 1/z) = z. \quad (6.17)$$

Here $G(z)$ is defined as

$$G(z) = \int \frac{f(\lambda)}{z - \lambda} d\lambda, \quad (6.18)$$

where $f(\lambda)$ is the eigenvalue density. The density is recovered by the Stieltjes inversion formula:

$$f(\lambda) = -\frac{1}{\pi} \lim_{\epsilon \rightarrow 0^+} \text{Im } G(\lambda + i\epsilon). \quad (6.19)$$

For the gravitational R -transform (6.15), substituting $R(z) = 1/(1 - \gamma z)$ into (6.17) and solving for $G(z)$ yields a quadratic equation whose solution gives the Marchenko–Pastur distribution, as we now derive.

6.5 The Marchenko–Pastur distribution

Derivation from the R -transform. We seek the spectral density $f(\lambda)$ corresponding to the R -transform $R(z) = 1/(1 - \gamma z)$. The functional equation (6.17) states that G inverts the map $z \mapsto R(z) + 1/z$; evaluating that map at $z = G(w)$ must therefore return w :

$$w = \frac{1}{1 - \gamma G(w)} + \frac{1}{G(w)}. \quad (6.20)$$

Multiplying through by $G(1 - \gamma G)$ and collecting terms yields a quadratic equation for $G = G(w)$:

$$\gamma w G^2 - (w + \gamma - 1)G + 1 = 0. \quad (6.21)$$

The discriminant factorizes as $(w + \gamma - 1)^2 - 4\gamma w = (w - a_+)(w - a_-)$, where

$$a_{\pm} = (1 \pm \sqrt{\gamma})^2, \quad (6.22)$$

and the quadratic formula gives

$$G(w) = \frac{(w + \gamma - 1) - \sqrt{(w - a_+)(w - a_-)}}{2\gamma w}, \quad (6.23)$$

where the branch of the square root is chosen so that $G(w) \sim 1/w$ as $w \rightarrow \infty$.

For $\lambda \in [a_-, a_+]$, the square root in (6.23) becomes purely imaginary, and the Stieltjes inversion formula (6.19) gives the Marchenko–Pastur density:

$$f_{\text{MP}}(\lambda; \gamma) = \frac{\sqrt{(a_+ - \lambda)(\lambda - a_-)}}{2\pi\gamma\lambda}, \quad \lambda \in [a_-, a_+]. \quad (6.24)$$

For $\gamma > 1$, the density is supplemented by a point mass at the origin:

$$f(\lambda; \gamma) = \left(1 - \frac{1}{\gamma}\right) \delta(\lambda) + f_{\text{MP}}(\lambda; \gamma), \quad \gamma > 1. \quad (6.25)$$

The point mass corresponds to the $k - L$ exactly zero eigenvalues of the reduced density matrix when the radiation Hilbert space is larger than the black hole interior.

Physical interpretation. For a Haar-random pure state in $\mathbb{C}^k \otimes \mathbb{C}^L$, the empirical distribution of the rescaled radiation eigenvalues $k\lambda_i$ converges to this law as $k, L \rightarrow \infty$ at fixed $\gamma = k/L$. The same law follows for the planar flat-weight replica model [76]. Equality of this limiting density establishes equality of the corresponding limiting linear spectral statistics, subject to the integrability needed by the observable. It does not establish equality of finite-dimensional joint eigenvalue distributions, connected ensemble fluctuations, eigenvectors, or evaporation channels.

Marchenko–Pastur distributions also occur for specified Gram matrices of black-hole microstate families [9, 65]. Those Gram matrices and radiation density matrices are distinct operators; their common limiting law must be derived for each preparation.

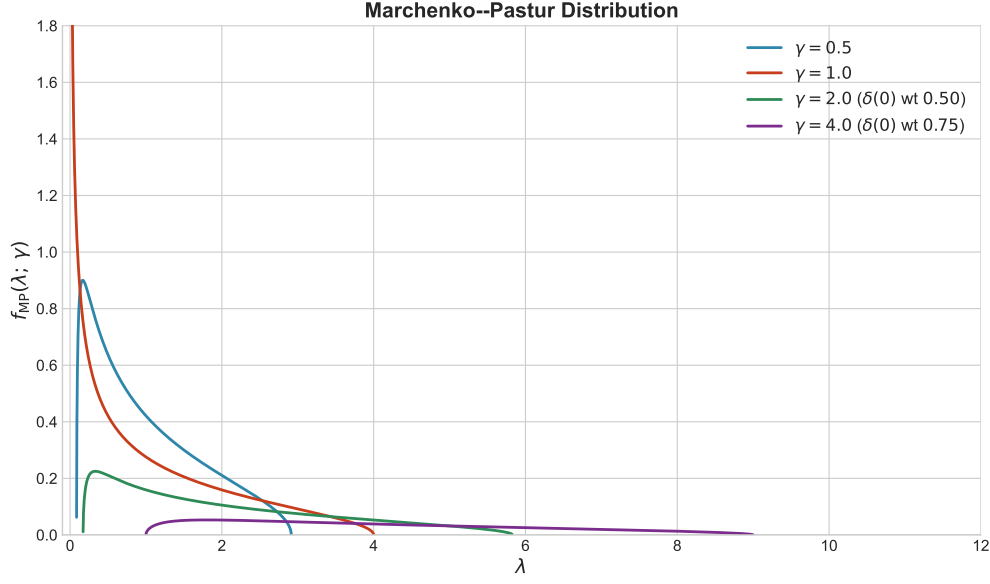


Figure 14. The Marchenko–Pastur density $f_{\text{MP}}(\lambda; \gamma)$ for $\gamma = 0.5, 1, 2, 4$. As γ increases, the support widens and shifts rightward. For every $\gamma \neq 1$ the density vanishes like a square root at both edges. For $\gamma > 1$, the lower edge $a_- = (1 - \sqrt{\gamma})^2$ remains positive and a point mass of weight $(1 - 1/\gamma)$ appears at $\lambda = 0$ (Eq. (6.25)). At the dimension crossing $\gamma = 1$, the lower edge touches zero ($a_- = 0$, $a_+ = 4$) and the density diverges as $\lambda^{-1/2}$ at the origin.

Special cases.

- **At equal dimensions** ($\gamma = 1$): the support is $[0, 4]$, and the density simplifies to

$$f_{\text{MP}}(\lambda; 1) = \frac{\sqrt{4 - \lambda}}{2\pi\sqrt{\lambda}}, \quad \lambda \in [0, 4]. \quad (6.26)$$

This is the quarter-circle law (up to a change of variables), the iconic distribution of random matrix theory.

- **Early time** ($\gamma \ll 1$): in the scaled variable of f_{MP} , the mass concentrates near $\lambda = 1$ (the support $[a_-, a_+]$ shrinks around 1). The physical eigenvalues of ρ_{rad} carry an additional normalization factor of $1/k$, so they cluster near $1/k$: a nearly maximally mixed state of the radiation.
- **Late time** ($\gamma \gg 1$): most eigenvalues are zero (the point mass dominates), and the nonzero part of the scaled density concentrates near $\lambda = \gamma$. In physical terms, the L nonzero eigenvalues cluster near $\gamma/k = 1/L$, giving entropy close to $\log L$. This becomes a nearly pure state only when the remaining factor itself has dimension near one.

Entropy from the density. For entropy it is convenient to use the smaller subsystem. Set $m = \min(k, L)$, $b = \max(k, L)$, and introduce the separate ratio $\eta = m/b \leq 1$; the raw ratio remains $\gamma = k/L$. Let λ now denote m times a nonzero eigenvalue. Its limiting law is $\text{MP}(\eta)$, with edges $(1 \pm \sqrt{\eta})^2$. No entropy information is lost, because for a pure global state the nonzero spectra of the two reduced density matrices coincide, so all spectral quantities are symmetric under $k \leftrightarrow L$. The

leading large-dimension von Neumann entropy is then

$$S_{\text{MP}} = \log \min(k, L) - \int_{a_-}^{a_+} f_{\text{MP}}(\lambda; \eta) \lambda \log \lambda \, d\lambda, \quad (6.27)$$

where the first term restores the eigenvalue normalization: writing $\lambda_{\text{phys}} = \lambda / \min(k, L)$ and using $\int f_{\text{MP}} \lambda \, d\lambda = M_1 = 1$, the sum $-\sum \lambda_{\text{phys}} \log \lambda_{\text{phys}}$ splits into exactly these two terms. The integral evaluates in closed form to $\eta/2$; the result is the west-coast scaling function discussed in Section 6.7, consistent with Eq. (6.35) below.

For an explicit evaluation, set $\lambda = 1 + \eta - 2\sqrt{\eta} \cos \theta$, $0 \leq \theta \leq \pi$. Then

$$\int \lambda \log \lambda f_{\text{MP}}(\lambda; \eta) \, d\lambda = \frac{2}{\pi} \int_0^\pi \sin^2 \theta \log(1 + \eta - 2\sqrt{\eta} \cos \theta) \, d\theta. \quad (6.28)$$

For $0 < \eta < 1$, expand the logarithm as $-2 \sum_{j \geq 1} \eta^{j/2} \cos(j\theta)/j$. Orthogonality with $\sin^2 \theta = (1 - \cos 2\theta)/2$ removes every term except $j = 2$, whose integral gives $\eta/2$. The endpoint $\eta = 1$ follows by integrability and passage to the limit. This derivation computes the constant deficit directly from the limiting spectrum.

6.6 Analytic continuation: Carlson's theorem

The replica trick computes $\text{Tr}(\rho^n)$ for integer $n \geq 2$ and then analytically continues to $n \rightarrow 1$ to extract the von Neumann entropy via

$$S_{\text{vN}} = - \lim_{n \rightarrow 1} \frac{\partial}{\partial n} \text{Tr}(\rho^n). \quad (6.29)$$

But analytic continuation from the integers is not unique in general — one can always add a function that vanishes at all positive integers (for instance, $\sin(\pi n)$ times any bounded function). A natural question is: *what justifies the specific analytic continuation given by the Marchenko–Pastur moment formula?*

The answer is provided by Carlson's theorem [18]:

Carlson's Theorem. Let $f(z)$ be holomorphic in the right half-plane $\text{Re}(z) > 0$, and suppose:

- (i) $f(n) = 0$ for all positive integers n ;
- (ii) $|f(z)| \leq C e^{c|z|}$ for some constants C, c with $c < \pi$.

Then $f(z) \equiv 0$.

The theorem implies that if two holomorphic functions agree at all positive integers and both satisfy the exponential growth bound with $c < \pi$, they must be identical. Thus, to justify the replica limit, one must verify that the Marchenko–Pastur moment formula

$$M_n(\eta) = \int_{a_-}^{a_+} \lambda^n f_{\text{MP}}(\lambda; \eta) \, d\lambda, \quad \mathbb{E}_{\text{H}} \text{Tr}(\rho^n) \sim m^{1-n} M_n(\eta), \quad (6.30)$$

satisfies the growth bound. The continuation is performed on the scaled moments $M_n(\eta)$; the prefactor $\min(k, L)^{1-n}$ is an explicit elementary function of n and requires no interpolation.

For a complex index z with $\text{Re } z > 0$, the compact support gives $|M_z(\eta)| \leq a_+^{\text{Re } z} \leq e^{|z| \log a_+}$. For real $n > 0$ this reads

$$|M_n(\eta)| \leq a_+^n = [(1 + \sqrt{\eta})^2]^n, \quad (6.31)$$

so the exponential growth rate is

$$c = \log a_+ = 2 \log(1 + \sqrt{\eta}). \quad (6.32)$$

Since $0 < \eta \leq 1$, the required bound is

$$c = 2 \log(1 + \sqrt{\eta}) \leq \log 4 < \pi. \quad (6.33)$$

Thus the moment integral has the required half-plane growth. At the lower edge zero, its derivative near $z = 1$ remains integrable because $\lambda \log \lambda f_{\text{MP}}(\lambda; 1)$ is integrable. An alternative, purely spectral justification is compact moment determinacy: polynomials are dense in the continuous functions on $[0, a_+]$, so all integer moments uniquely determine the measure. The entropy then follows by integration against that measure.

Within the stated bounded-exponential-growth class, the Marchenko–Pastur moment integral (6.30) is the *unique* analytic continuation of the scaled planar moments $M_n(\eta)$ from integer n to real n , and the replica limit $n \rightarrow 1$ is thereby justified.

6.7 The west-coast scaling function

We now have all the ingredients to extract the von Neumann entropy from the Marchenko–Pastur distribution. The result is the *west-coast scaling function*, which gives the planar entropy correction in the flat-weight model. It is the large-dimension limit of Page’s mean-entropy formula [71], rather than the exact finite-dimensional mean entropy.

The result. Define the dimensionless parameter

$$\Delta = \log \gamma = \log \frac{k}{L}, \quad (6.34)$$

which measures the logarithmic distance from the dimension crossing ($\Delta = 0$ when $k = L$). The von Neumann entropy of the radiation, computed from the Marchenko–Pastur distribution, takes the form

$$S_{\text{MP}} = \log \min(k, L) + F(\Delta), \quad (6.35)$$

where the *west-coast scaling function* is

$$F(\Delta) = -\frac{1}{2} e^{-|\Delta|}. \quad (6.36)$$

Properties. The west-coast scaling function has several notable features:

1. **Peak deficit:** $F(0) = -1/2$ nat. At the Page time, the entropy falls short of $\log \min(k, L) = \log k = \log L$ by exactly $1/2$ nat. This is the genus-zero analogue of the $O(1)$ correction computed by Page [71].
2. **Transition width:** The function decays as $e^{-|\Delta|}$, so the transition from “before Page time” to “after Page time” occurs over a window of width $O(1)$ in $\Delta = \log(k/L)$. This is an order-one window in the entropy difference, hence a relative width that vanishes when compared with a macroscopic entropy scale.
3. **Exponential tails:** For $|\Delta| \gg 1$, $F(\Delta) \rightarrow 0$ exponentially fast. Far from the Page time, the entropy is well approximated by the sharp Page curve $\log \min(k, L)$.
4. **Symmetry:** $F(\Delta) = F(-\Delta)$. This reflects the Narayana symmetry $N(n, j) = N(n, n+1-j)$ and, physically, the duality between radiation and black hole interior in the genus-zero sector.

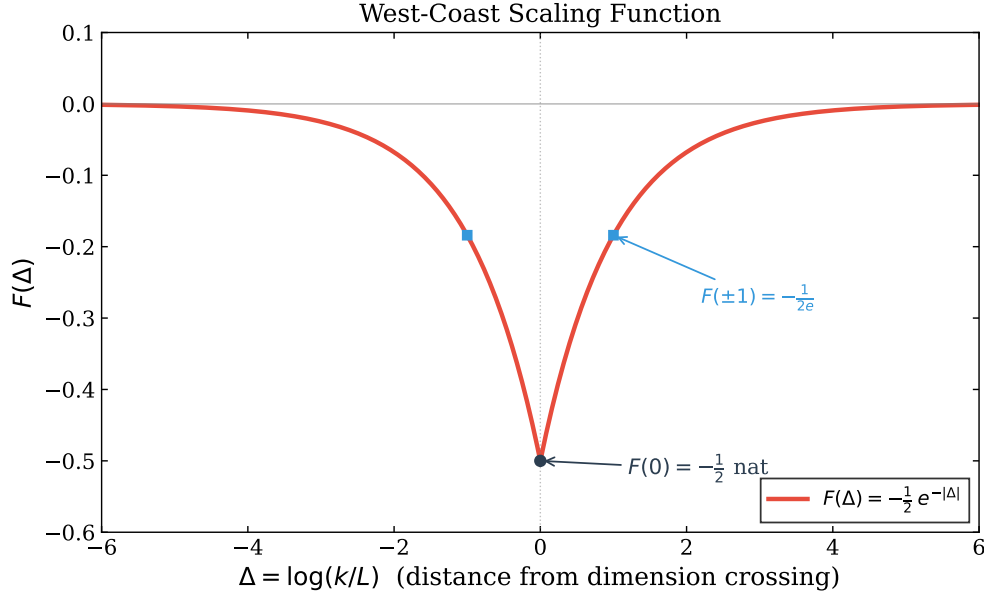


Figure 15. The west-coast scaling function $F(\Delta)$, where $\Delta = \log(k/L)$ measures the logarithmic distance from equal dimensions. The function reaches its minimum $F(0) = -1/2 \text{ nat}$ at $k = L$ and decays exponentially on both sides with e -folding scale $|\Delta| = 1$. The entire transition region has width $O(1)$ in Δ , corresponding to $O(1)$ in $\log(k/L)$.

The smooth Page curve. Combining (6.35) and (6.36), the flat-weight planar entropy is

$$S_{\text{MP}} = \min(S_{\text{rad}}, S_{\text{BH}}) - \frac{1}{2} e^{-|\Delta|}, \quad (6.37)$$

where $S_{\text{rad}} = \log k$ and $S_{\text{BH}} = \log L$ are the coarse-grained entropies. This replaces the sharp $\min(S_{\text{rad}}, S_{\text{BH}})$ of the idealized Page curve with a function whose first two derivatives match at $\Delta = 0$ when the branch mean is varied smoothly. The local rounding is quadratic, but this planar function need not be analytic to all orders at the crossing. At fixed L the curve rises and saturates; a falling branch additionally requires a decreasing black-hole dimension or another specified evolution.

Finite dimensions and the scope of the limit. For $m = \min(k, L)$ and $b = \max(k, L)$, the exact Haar mean entropy is

$$\mathbb{E}_{\text{H}} S(\rho) = H_{mb} - H_b - \frac{m-1}{2b}, \quad H_j = \sum_{r=1}^j \frac{1}{r}. \quad (6.38)$$

This is Page's result [71, 84]. Using $H_j = \log j + \gamma_E + 1/(2j) - 1/(12j^2) + O(j^{-4})$ gives

$$\mathbb{E}_{\text{H}} S(\rho) = \log m - \frac{m}{2b} + \frac{1}{2mb} + \frac{1}{12b^2} - \frac{1}{12m^2b^2} + O(b^{-4}), \quad (6.39)$$

with both dimensions large at fixed positive m/b . The leading correction to the Marchenko–Pastur entropy is therefore of order b^{-2} in this regime. At $m = b = L$ it is $7/(12L^2)$. For $m = b = 2$, the exact entropy is $1/3 \text{ nat}$, whereas the planar expression is $\log 2 - 1/2 \simeq 0.1931 \text{ nat}$. Small dimensions should be evaluated by the exact formula.

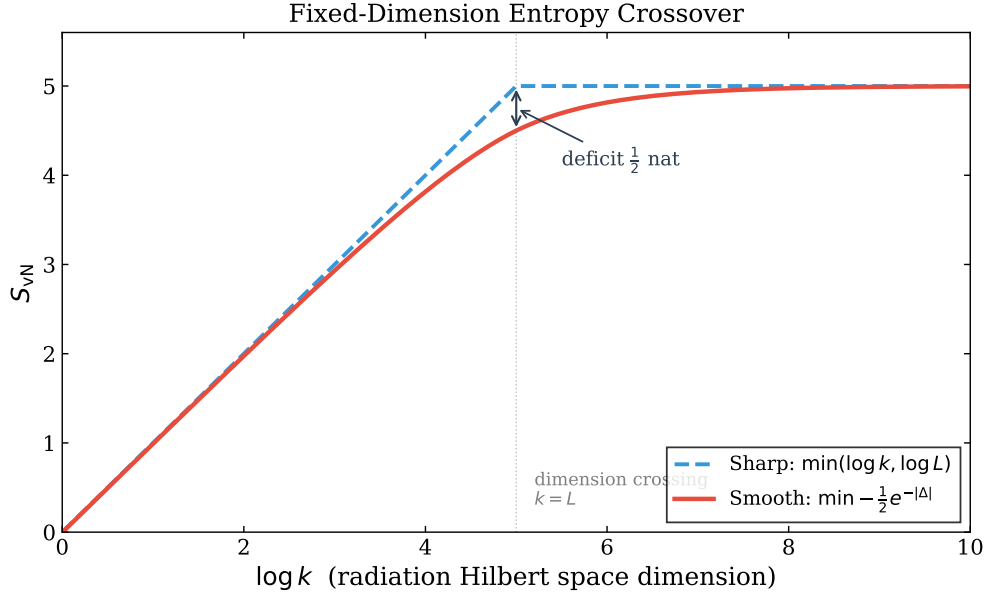


Figure 16. The fixed- L scan: comparison of the sharp dimension bound $\min(\log k, \log L)$ (dashed blue) with the smooth genus-zero Page curve $\min(\log k, \log L) - \frac{1}{2}e^{-|\Delta|}$ (solid red), where $\Delta = \log(k/L)$. The west-coast scaling function resolves the cusp at equal dimensions into a smooth transition. The maximum deficit is $1/2$ nat at $k = L$. This is a dimension scan; a time profile requires a dynamical preparation.

Other ensembles can have the same limiting density but different finite-size corrections or connected fluctuations. Weak convergence alone also needs supplementing by suitable tail control to pass to an unbounded spectral observable. The Gaussian/Haar benchmark supplies this control; the same conclusion for a gravitational preparation requires its own spectral analysis.

Summary. Flat cycle weights give the free cumulants $\kappa_m = \gamma^{m-1}$. Their R -transform determines the Marchenko–Pastur density, whose smaller-side form has ratio $\eta = \min(k, L)/\max(k, L)$. Integrating that density gives $S_{\text{MP}} = \log \min(k, L) - \eta/2$. The exact finite Haar answer has additional normalization and spectral corrections. General planar replica models use the same non-crossing organization with their own cycle weights.

Part IV

Thermodynamic Fluctuations and Entropy Transitions

7 Entropy Averages and Gaussian Fluctuations

The fixed-dimension Page calculation describes a random pure state on a tensor product whose dimensions are specified. Thermodynamic preparations introduce a second source of variation: the

number and weights of accessible states depend on energy. Averaging over that variation can round the leading minimum of two entropy branches on a scale larger than the order-one Page correction. Before performing the average, however, one must identify the random variable and the quantum state whose entropy is sought.

Enhanced corrections of this kind occur in chaotic energy eigenstates [66, 94]. In gravity, the diagonal approximation of Marolf et al. [61] averages fixed-area entropies, and Dong and Wang [26] obtained enhanced corrections by summing replica saddles in a chaotic holographic example. The Gaussian minimum formula derived below is established material from this literature. Its elementary derivation is included because it makes both the mechanism and the approximation explicit. Paper I [99] compares a specified level-resolved diagonal approximation with the planar spectrum; Paper II [100] studies regular and critical thermodynamic fluctuations within a canonical random-state model.

7.1 The fluctuating entropy difference

Let F_1 and F_2 be integrable random variables representing two candidate entropy branches. The exact identity

$$\mathbb{E} \min(F_1, F_2) = \frac{\mathbb{E} F_1 + \mathbb{E} F_2}{2} - \frac{1}{2} \mathbb{E} |F_1 - F_2| \quad (7.1)$$

shows which fluctuations matter. The deficit relative to the minimum of the mean branches is

$$B = \frac{1}{2} (\mathbb{E} |F_1 - F_2| - |\mathbb{E} F_1 - \mathbb{E} F_2|) \geq 0. \quad (7.2)$$

The inequality follows from convexity of the absolute value. In particular,

$$\sigma_{\text{gap}}^2 = \text{Var } F_1 + \text{Var } F_2 - 2 \text{Cov}(F_1, F_2). \quad (7.3)$$

A fluctuation shared equally by both branches cancels from their difference. A thermodynamic heat capacity therefore becomes an entropy-transition width only after the response of both branches has been specified. This point will be important for a finite reservoir in Section 8.

For example, if $F_1 = s$ is deterministic and $F_2 = Y$ is random with mean μ , write $\Delta = s - \mu$ and $\xi = Y - \mu$. Then

$$M(\Delta) = \mathbb{E} \min(\Delta, \xi), \quad B(\Delta) = \min(\Delta, 0) - M(\Delta) = \frac{1}{2} (\mathbb{E} |\Delta - \xi| - |\Delta|). \quad (7.4)$$

At the crossing of the mean branches,

$$B(0) = \frac{1}{2} \mathbb{E} |\xi| \leq \frac{1}{2} \sigma, \quad \sigma^2 = \mathbb{E} \xi^2. \quad (7.5)$$

Neither a Gaussian distribution nor gravity is needed for these statements. They describe an average of a minimum; an identification with a quantum entropy is an additional physical statement.

7.2 Canonical energy, thermodynamic entropy, and modular energy

Consider a system with a finite canonical partition function $Z(\beta) = \text{Tr } e^{-\beta H}$, with Boltzmann's constant set to one. Differentiating $\log Z$ gives the exact fluctuation identity

$$\text{Var } E = \partial_\beta^2 \log Z = T^2 C_{\text{can}}, \quad C_{\text{can}} = \frac{d\langle E \rangle}{dT}. \quad (7.6)$$

Other conserved parameters are held fixed in the derivative. The canonical modular energy of an energy level is

$$K(E) = \beta E + \log Z = -\log x(E), \quad x(E) = e^{-\beta E} / Z. \quad (7.7)$$

When levels are sampled with probabilities $x(E)$, $\langle K \rangle$ is the canonical von Neumann entropy and $\text{Var } K = C_{\text{can}}$ exactly.

The thermodynamic entropy $S(E)$, used to model a density of states, is a different function. In a black-hole model it may be the Bekenstein–Hawking entropy along a family of equilibrium geometries. Its canonical energy law is

$$p_\beta(E) dE \propto e^{S(E)-\beta E} dE.$$

Suppose this distribution is controlled by a single stable, nondegenerate saddle at E_* . With $C_B = dE_*/dT > 0$, the first law and a Taylor expansion give

$$S(E) - \beta E = S(E_*) - \beta E_* - \frac{(E - E_*)^2}{2T^2 C_B} + \dots, \quad (7.8)$$

$$S(E) - S(E_*) = \frac{E - E_*}{T} + \dots. \quad (7.9)$$

Consequently,

$$\text{Var } S(E) \simeq C_B, \quad \text{Var } K = C_{\text{can}} \simeq C_B. \quad (7.10)$$

The approximation requires small relative fluctuations, controlled higher terms in the action and measure, and negligible weight in competing phases. It is not an exact identity between the geometric entropy variance and a heat capacity.

An exact way to track the difference is to set $I(E) = \beta E - S(E)$. Then

$$(K - \langle K \rangle) - (S - \langle S \rangle) = I - \langle I \rangle. \quad (7.11)$$

At a regular saddle the fluctuations of I are order one, whereas those of S and K can be order $\sqrt{C_B}$. Their leading broad distributions therefore agree, while their means, higher cumulants, and subleading widths need not. This distinction is visible even in JT gravity.

For the disk Schwarzian density in the convention used in Section 8.7, the leading geometric entropy width is

$$\sigma_{S,\text{lead}} = \sqrt{4\pi^2/\beta}, \quad \sigma_{S,\text{lead}}|_{\beta=2\pi} = \sqrt{2\pi}. \quad (7.12)$$

The exact variance of K is instead $4\pi^2/\beta + 3/2$. Direct integration at $\beta = 2\pi$ gives a geometric entropy standard deviation 2.305939; the leading expression is 2.506628. These are different approximations to different observables, as detailed below.

For a neutral AdS black hole at large radius in bulk dimension $d + 1$,

$$C_B \simeq (d - 1)S_*, \quad r_+/\ell \gg 1, \quad (7.13)$$

and the regular leading width is

$$\sigma_S \simeq \sqrt{(d - 1)S_*}. \quad (7.14)$$

Large entropy and large radius are separate limits. At fixed r_+/ℓ , the coefficient C_B/S_* is generally different from $d - 1$. Nor does a fixed total energy imply a fixed black-hole energy: a finite reservoir permits energy exchange and changes the fluctuation law.

7.3 The diagonal Page approximation

Let $s = \log k$ and $y = \log L$ denote the logarithms of two large subsystem dimensions. The leading fixed-dimension Page formula is

$$S_{\text{fixed}}(s, y) = \min(s, y) - \frac{1}{2}e^{-|s-y|}. \quad (7.15)$$

It follows from the Marchenko–Pastur limit of the normalized Haar calculation [71, 76]. Finite dimensions have the normalization corrections discussed in Section 5; (7.15) is not their exact entropy formula.

For a real random variable Y with probability law P , define

$$S_{\text{diag}}(s; P) := \int \left[\min(s, y) - \frac{1}{2} e^{-|s-y|} \right] P(dy). \quad (7.16)$$

This definition gives a conditional Page average. Its two terms are

$$S_{\text{diag}} = E_{\text{min}} + E_{\text{corr}}, \quad E_{\text{min}} = \mathbb{E} \min(s, Y), \quad E_{\text{corr}} = -\frac{1}{2} \mathbb{E} e^{-|s-Y|}. \quad (7.17)$$

If a shifted law $Y = \mu + \xi$ is held fixed while s varies, this is a convolution in $\Delta = s - \mu$. Independence of a putative radiation variable and Y does not by itself turn this average into a gravitational entropy. One must specify what Y represents and how sectors enter the reduced state.

Here is a canonical construction for which the approximation can be compared with a definite spectrum. Let $x_j = e^{-\beta E_j} / Z$ and take independent complex Gaussian coefficients g_{ij} of unit variance. Define

$$|\Psi\rangle = \frac{1}{\sqrt{\mathcal{N}}} \sum_{i=1}^k \sum_j \sqrt{x_j} g_{ij} |i\rangle_R |j\rangle_B, \quad \mathcal{N} = \sum_{i,j} x_j |g_{ij}|^2. \quad (7.18)$$

The reference space is common to all black-hole levels. Its dimension k is a control parameter; this definition imposes no evaporation law. The variable in the level-resolved approximation is $Y = K_j = -\log x_j$, sampled with probability x_j , rather than the geometric entropy evaluated at E_j . In the planar model the probability law ν of rescaled reference eigenvalues $u = k\lambda$ is specified by

$$R_\nu(w) = \mathbb{E}_P \frac{1}{1 - aw}, \quad a = kx_j = e^{s-K_j},$$

$$S_{\text{planar}} = s - \int u \log u \nu(du). \quad (7.19)$$

Equation (7.19) describes the planar limit of the normalized Gaussian construction; it is not an exact finite-dimensional identity. Thus an averaged Page kernel and a spectral entropy are two computable quantities in the same specified model. For equiprobable x_j , this is the Marchenko–Pastur benchmark. Nonuniform weights generally give a different spectrum.

There is a different answer when a reduced state is an orthogonal direct sum, $\rho_R = \bigoplus_\alpha p_\alpha \rho_{R,\alpha}$. The following finite-dimensional identity is exact:

$$S(\rho_R) = H(p) + \sum_\alpha p_\alpha S(\rho_{R,\alpha}), \quad H(p) = -\sum_\alpha p_\alpha \log p_\alpha. \quad (7.20)$$

For instance, mixing two pure states supported on orthogonal subspaces gives entropy $\log 2$, whereas their average conditional entropy is zero. If those two pure states coincide, the mixture instead has zero entropy. The sector probabilities and conditional entropies alone do not distinguish these cases; the supports matter. This elementary example explains why an average of conditional entropies requires a careful interpretation.

7.4 Gaussian formulas and their derivation

Suppose $Y \sim \mathcal{N}(\mu, \sigma^2)$ in the explicitly defined average (7.16). A Gaussian with support on the entire real line is an effective fluctuation law. Its physical use assumes that the mean entropy is far from any lower bound compared with σ . We write $\phi(z) = (2\pi)^{-1/2} e^{-z^2/2}$ and $\Phi(z) = \int_{-\infty}^z \phi(t) dt$.

7.4.1 The averaged minimum

Writing $Y = \mu + \sigma Z$, with Z standard normal, gives

$$\begin{aligned} E_{\min} &= s - \mathbb{E}(s - Y)^+ \\ &= s - \sigma \mathbb{E}(\Delta/\sigma - Z)^+, \quad x^+ = \max(x, 0). \end{aligned} \quad (7.21)$$

To evaluate the remaining integral, use $\phi'(z) = -z\phi(z)$:

$$\begin{aligned} \mathbb{E}(\Delta - \sigma Z)^+ &= \int_{-\infty}^{\Delta/\sigma} (\Delta - \sigma z) \phi(z) dz \\ &= \Delta \Phi(\Delta/\sigma) + \sigma \phi(\Delta/\sigma). \end{aligned} \quad (7.22)$$

Therefore

$$E_{\min}(\Delta, \sigma) = \mu + \Delta \Phi(-\Delta/\sigma) - \sigma \phi(\Delta/\sigma). \quad (7.23)$$

The slope is a tail probability and the curvature is minus the fluctuation density:

$$\frac{\partial E_{\min}}{\partial s} = \Phi(-\Delta/\sigma), \quad -\frac{\partial^2 E_{\min}}{\partial s^2} = \frac{1}{\sigma} \phi(\Delta/\sigma). \quad (7.24)$$

The mean-branch crossing has slope one half. A useful operational width is the interval over which the slope decreases from 0.9 to 0.1: $2\Phi^{-1}(0.9)\sigma \simeq 2.563103\sigma$. This is an entropy interval; converting it to a duration requires a dynamical relation between s and time.

7.4.2 The averaged Page correction

Split the exponential average at $Y = s$. The contribution from $Y < s$ is $e^{-s}\mathbb{E}[e^Y \mathbf{1}_{Y < s}]$; the other is $e^s\mathbb{E}[e^{-Y} \mathbf{1}_{Y > s}]$. Completing the square in each Gaussian integral yields

$$\begin{aligned} E_{\text{corr}}(\Delta, \sigma) &= -\frac{1}{2}e^{\sigma^2/2} \left[e^{-\Delta} \Phi\left(\frac{\Delta - \sigma^2}{\sigma}\right) \right. \\ &\quad \left. + e^{\Delta} \Phi\left(\frac{-\Delta - \sigma^2}{\sigma}\right) \right]. \end{aligned} \quad (7.25)$$

The large exponential is compensated by a small normal tail. Numerical implementations should combine the exponents with a logarithmic normal CDF, or use a scaled complementary error function, to avoid overflow.

7.4.3 Width, central deficit, and limiting cases

The full diagonal approximation in this Gaussian example is

$$S_{\text{diag}} = E_{\min} + E_{\text{corr}}. \quad (7.26)$$

At $\Delta = 0$, the deficit of the averaged minimum is exactly $\sigma/\sqrt{2\pi}$. The full diagonal entropy differs from the fixed kernel at the mean by

$$S_{\text{diag}}(\mu) - S_{\text{fixed}}(\mu, \mu) = -\frac{\sigma}{\sqrt{2\pi}} + \frac{1}{2} - \frac{1}{2}e^{\sigma^2/2} \text{erfc}(\sigma/\sqrt{2}). \quad (7.27)$$

The final term is of order σ^{-1} as $\sigma \rightarrow \infty$. Thus the averaged minimum determines the leading broad-distribution deficit, while averaging the finite Page correction changes it at a smaller order.

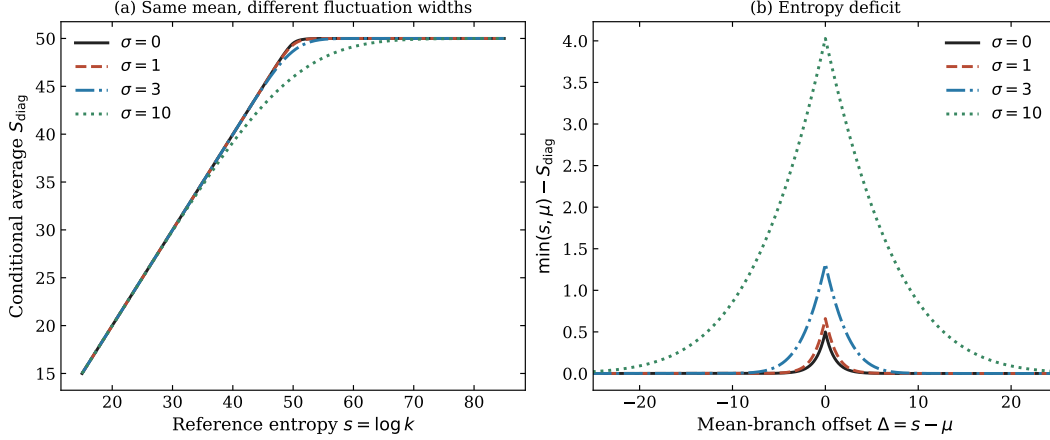


Figure 17. Gaussian conditional Page averages defined by (7.16), with $\mu = 50$ and $\sigma = 0, 1, 3, 10$. The plotted variable is the reference entropy $s = \log k$. All curves use the same mean branch. Increasing the width of its Gaussian fluctuation rounds the transition; these curves do not independently establish the entropy of a gravitational state. The right panel displays the central deficit relative to the minimum of the mean branches.

For $u = \Delta/\sigma$, the minimum has the exact rescaled form $E_{\min} - \mu = \sigma[u\Phi(-u) - \phi(u)]$. This is a scaling identity for Gaussian fluctuations. A quartic critical law or a bimodal mixture has a different profile. In the opposite limit $\sigma \rightarrow 0$, the distribution collapses at μ and the full expression tends to $\min(s, \mu) - \frac{1}{2}e^{-|s-\mu|}$. Its intrinsic crossover remains order one. The fluctuation-induced width is order σ when $\sigma \gg 1$; for $\sigma \lesssim 1$, the fixed-kernel smoothing cannot be neglected.

7.5 An order-one distinction from the planar spectrum

The full diagonal approximation should not be confused with the planar entropy in (7.19), even when both share the same leading broadening. Set $\delta = K - s$ and $q = \Pr(\delta > 0)$. For a family of canonical laws with $\mathbb{E}_{Pa} < \infty$, suppose

$$\Pr(|\delta| \leq A) \longrightarrow 0 \quad \text{for every fixed } A > 0, \quad q \longrightarrow q_* \in (0, 1). \quad (7.28)$$

The broad-distribution result of Paper I [99] is

$$S_{\text{planar}} - S_{\text{diag}} \longrightarrow -q_* \log q_*. \quad (7.29)$$

This assertion concerns the specified planar spectral model. It is not a general theorem about averaging gravitational entropies. At the centered crossing of a broad symmetric law, $q_* = 1/2$, giving $\frac{1}{2} \log 2$. The leading deficit grows with the width, so the relative error decreases, but the additive error does not vanish.

The spectrum gives a physical explanation. Canonical levels whose individual probabilities are much smaller than $1/k$ collectively contribute an approximately uniform spectral component of total weight q_* . Its entropy contains $-q_* \log q_*$ in addition to $q_* \log k$. The diagonal Page average retains the latter but omits the former. Related spectral components already appear in the canonical replica analysis of Penington et al. [76]; the companion result evaluates the limiting difference. The Gaussian correction in (7.25) belongs to the diagonal approximation and does not supply this missing order-one spectral contribution.

7.6 A useful identity beyond the Gaussian case

For any real X and fixed threshold a ,

$$\min(X, a) = a - (a - X)^+, \quad (7.30)$$

and hence

$$\mathbb{E} \min(X, a) = a - \mathbb{E}(a - X)^+. \quad (7.31)$$

For $X \sim \mathcal{N}(\mu, \sigma^2)$, the half-line integral is

$$\mathbb{E}(a - X)^+ = (a - \mu)\Phi((a - \mu)/\sigma) + \sigma\phi((a - \mu)/\sigma). \quad (7.32)$$

This is also the Gaussian integral in Bachelier's option-pricing formula [8]; no financial interpretation is required here. For a general centered fluctuation with density p_ξ , the same identity gives $-M''(\Delta) = p_\xi(\Delta)$. Thus the curvature of an averaged minimum directly records its fluctuation density.

For example, equally likely values $\xi = \pm a$ give a deficit $a/2$ at the center and two distinct changes of slope, rather than a rounded Gaussian corner. A quartic density gives another explicit profile. These examples show why variance alone is insufficient outside a Gaussian regime. The next section develops the thermodynamic distributions that can supply such inputs, while keeping the quantum-entropy interpretation separate.

8 Thermodynamic Ensembles and Critical Fluctuations

The Gaussian integral of the preceding section does not depend on spacetime dimension. Its physical input does: a state preparation, a probability law, and a fluctuating difference between candidate entropy branches. This section works through several examples. Neutral AdS black holes provide regular stable saddles; finite reservoirs impose energy constraints; a charged critical point replaces the Gaussian law by a quartic one. JT gravity supplies an explicitly normalized distribution for which the distinction between geometric entropy and modular energy can be checked directly.

The thermodynamic formulas are standard. The particular finite-size critical profile and its use in the canonical random-state model follow Paper II [100]. They should be read with the distinction between an averaged minimum, the full diagonal Page average, and a planar spectral entropy established in Section 7.

8.1 Regular neutral AdS saddles

For a Schwarzschild–AdS black hole in bulk dimension $d + 1 \geq 4$,

$$\begin{aligned} S_* &= \frac{\Omega_{d-1} r_+^{d-1}}{4G}, & T_H &= \frac{d - 2 + dr_+^2/\ell^2}{4\pi r_+}, \\ C_B &= T_H \frac{dS_*/dr_+}{dT_H/dr_+} = (d - 1)S_* \frac{dr_+^2/\ell^2 + d - 2}{dr_+^2/\ell^2 - d + 2}. \end{aligned} \quad (8.1)$$

Here ℓ is the AdS length, Ω_{d-1} the area of the unit $(d - 1)$ -sphere, and the derivative holds ℓ fixed. The heat capacity is positive for $r_+ > \ell\sqrt{(d - 2)/d}$. To avoid both metastability and phase coexistence, first take $r_+/\ell > 1$ fixed and let $G \rightarrow 0$. The black-hole phase then dominates thermal AdS in the usual Hawking–Page comparison [40, 97].

If the density-of-states model is $e^{S(E)} dE/E_0$, with E_0 a fixed energy unit, the canonical distribution near this saddle is Gaussian at leading order. For a deterministic competing branch and a fluctuating thermodynamic entropy, the variance is

$$\sigma_S^2 \simeq C_B. \quad (8.2)$$

For the canonical modular energy $K = \beta E + \log Z$ the corresponding exact identity is $\text{Var } K = C_{\text{can}}$, and its leading semiclassical value is again C_B . In either case the Gaussian diagonal average, with mean μ and $\Delta = s - \mu$, is

$$S_{\text{diag}}(s) = \int_{\mathbb{R}} S_{\text{fixed}}(s, \mu + x) \frac{e^{-x^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}} dx. \quad (8.3)$$

This equation uses the full fixed-dimension kernel defined in (7.15); dropping its exponential term gives the averaged minimum.

The frequently used large-radius estimate is

$$\sigma_S \simeq \sqrt{(d-1)S_*}, \quad r_+/\ell \gg 1. \quad (8.4)$$

There are two independent approximations here. The Gaussian law requires a well-controlled thermodynamic saddle, while replacing C_B/S_* by $d-1$ requires a large geometric radius:

$$\frac{C_B}{S_*} = (d-1) \left[1 + \frac{2(d-2)}{d} \frac{\ell^2}{r_+^2} + O(\ell^4/r_+^4) \right]. \quad (8.5)$$

Increasing S_* by reducing G at fixed radius improves the first approximation but leaves the geometric correction unchanged.

For a Gaussian law the central deficit of the averaged minimum is

$$B(0) = \sqrt{\frac{C_B}{2\pi}} \simeq \sqrt{\frac{(d-1)S_*}{2\pi}} \quad \text{at large radius.} \quad (8.6)$$

It grows in absolute entropy units while decreasing relative to S_* . At $r_+/\ell = 2$, the ratios C_B/S_* in bulk dimensions four, five, and six are 26/11, 27/7, and 92/17, respectively. These differ appreciably from their large-radius limits 2, 3, and 4. Figure 18 shows both the geometric dependence and the large-radius comparison. The dimensions are compared at equal S_* ; this is not a comparison at fixed mass across different theories.

8.2 Fixed total energy and a finite reservoir

Fixing the black-hole energy and fixing the total energy of a black hole plus reservoir impose different constraints. Let $E_B + E_R = E_{\text{tot}}$ and suppose the two systems can exchange energy. Their density-of-states distribution is

$$p(E_B) dE_B \propto \exp[S_B(E_B) + S_R(E_{\text{tot}} - E_B)] dE_B. \quad (8.7)$$

At a stationary point the temperatures agree. Put $\epsilon = E_B - E_{B,*}$, and let C_B and C_R be the heat capacities at the common temperature T . The entropy expansion is

$$S_B + S_R = S_{\text{tot},*} - \frac{\epsilon^2}{2T^2} \left(\frac{1}{C_B} + \frac{1}{C_R} \right) + O(\epsilon^3). \quad (8.8)$$

If this quadratic maximum controls the distribution, then

$$\text{Var } E_B \simeq T^2 C_{\text{eff}}, \quad C_{\text{eff}} = (C_B^{-1} + C_R^{-1})^{-1}, \quad C_B^{-1} + C_R^{-1} > 0. \quad (8.9)$$

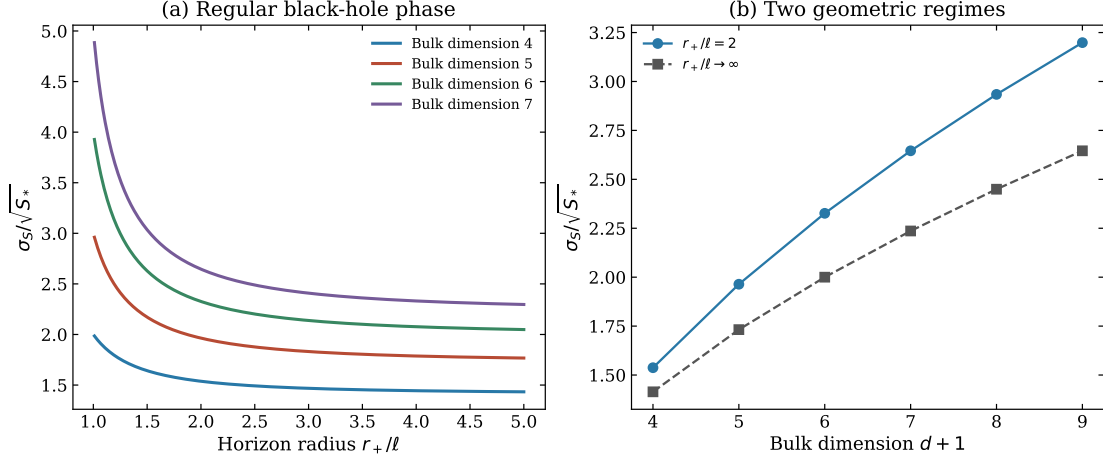


Figure 18. Leading regular-phase entropy width $\sigma_S/\sqrt{S_*} = \sqrt{C_B/S_*}$ for neutral AdS black holes. Left: its dependence on $r_+/\ell > 1$ in several bulk dimensions. Right: values at $r_+/\ell = 2$ and the separate large-radius limit, as functions of the bulk dimension $d+1$. The curves describe a single stable black-hole phase; no extrapolation through a spinodal is made.

The final inequality is the local stability condition. Positive capacities make C_{eff} their harmonic combination. As $C_R/C_B \rightarrow \infty$ with $C_B > 0$, the canonical result is recovered:

$$\text{Var } S_B \simeq C_{\text{eff}} \longrightarrow C_B. \quad (8.10)$$

For a genuinely small reservoir, its full entropy function may be needed; a heat capacity at one energy does not guarantee a Gaussian distribution.

The first law also gives $\delta S_B = \epsilon/T$ and $\delta S_R = -\epsilon/T$ to linear order. It follows that

$$\text{Var } S_B \simeq C_{\text{eff}}, \quad \text{Var}(S_B - S_R) \simeq 4C_{\text{eff}}. \quad (8.11)$$

The second relation, not the first, governs the averaged minimum if the competing branches are the two subsystem entropies. Energy conservation makes their fluctuations anticorrelated. Treating one branch as fixed would miss a factor of two in the standard deviation.

A numerical example makes the distinction explicit. Take two regular systems with $C_B = C_R = 100$, at equal mean entropies. Then $C_{\text{eff}} = 50$. The entropy standard deviation of either system is $\sqrt{50}$, whereas the standard deviation of their difference is $\sqrt{200}$. The leading central deficit of their averaged minimum is

$$B(0) = \frac{\sqrt{200}}{\sqrt{2\pi}} = \frac{10}{\sqrt{\pi}} \simeq 5.641896. \quad (8.12)$$

It is compatible with an exactly fixed total energy. This is the same thermodynamic mechanism that gives an enhanced correction for an equal bipartition in the energy-eigenstate analysis of Murthy and Srednicki [66].

The total entropy $S_B + S_R$, by contrast, is stationary at equilibrium. Its fluctuation starts quadratically and is order one in the Gaussian approximation. It must not be substituted for the fluctuating branch difference. More generally, if that difference is $g(E_B)$ under the energy constraint, its leading variance is $g'(E_{B,*})^2 \text{Var } E_B$. A common mode can cancel, and a vanishing linear derivative requires a new expansion.

Table 4. Distinct constraints in a regular quadratic regime. Each variance concerns the indicated observable. An actual entanglement calculation must also specify the sector structure of the state.

Preparation	Fluctuating quantity	Leading variance
Fixed E_B	$S_B(E_B)$	0 from energy exchange
Infinite reservoir, $C_B > 0$	S_B	C_B
Fixed $E_B + E_R$, stable finite bath	S_B	C_{eff}
Same finite bath, two entropy branches	$S_B - S_R$	$4C_{\text{eff}}$

These equilibrium statements do not determine the time-dependent state of an evaporating black hole. To apply them dynamically, one needs a preparation and an approximation justifying the instantaneous distribution, together with the relevant entropy branches. Unitarity or fixed total energy alone does not select a sharp or broad Page transition.

8.3 Negative heat capacity

For an asymptotically flat Schwarzschild black hole in bulk dimension $d + 1 \geq 4$,

$$C_B = -(d - 1)S_B < 0. \quad (8.13)$$

Indeed $T_H \propto r_+^{-1}$ while $S_B \propto r_+^{d-1}$. Such a branch is unstable in contact with an infinite reservoir of fixed temperature: increasing its energy lowers its temperature, encouraging further inflow. The unrestricted asymptotically flat canonical integral also fails because the black-hole density of states grows too rapidly.

The finite-reservoir stability condition in (8.9) is different. A positive reservoir with $C_R < |C_B|$ can make the combined entropy locally concave along the allowed energy exchange. For example, $C_B = -100$ and $C_R = 25$ give $C_{\text{eff}} = 100/3 > 0$. This is a local thermodynamic statement; it does not prove global stability or construct a gravitational equilibrium boundary condition. Cavity thermodynamics illustrates why boundary conditions matter [103]. Negative heat capacity therefore rules out an unrestricted canonical treatment of that branch, but does not force every allowed fixed-total-energy state to have negligible subsystem fluctuations.

8.4 Spinodal, coexistence, and continuous criticality

Three distinct phenomena are often placed under the heading of a thermodynamic transition. They require different probability laws.

The neutral AdS spinodal. The minimum temperature of the neutral black-hole family occurs at

$$r_0 = \ell \sqrt{\frac{d-2}{d}}, \quad T_0 = \frac{\sqrt{d(d-2)}}{2\pi\ell}. \quad (8.14)$$

Set $N_0 = S(r_0)$, $x = r_+/r_0 - 1$, and $t_0 = (T - T_0)/T_0$. The action $I = M/T - S$ has the local form

$$I(x) - I(0) = (d-1)N_0 \left(\frac{x^3}{6} - t_0 x \right) + O(N_0[x^4 + |t_0|x^2 + t_0^2|x|]). \quad (8.15)$$

For $t_0 > 0$, the stable and unstable stationary radii lie near $x_{\pm} = \pm\sqrt{2t_0}$. The stable saddle's response obeys

$$C_B \sim (d-1)N_0(2t_0)^{-1/2}, \quad (8.16)$$

and its Gaussian entropy width is

$$\sigma_S \sim \sqrt{(d-1)N_0} (2t_0)^{-1/4}. \quad (8.17)$$

The barrier is of order $N_0 t_0^{3/2}$. A well-localized Gaussian therefore requires $t_0 \gg N_0^{-2/3}$. The apparent divergence cannot be extrapolated past that condition. The cubic local action is not confining on the real line, so it does not define a replacement equilibrium distribution by itself. A distribution conditioned to remain in a metastable basin needs an explicit basin prescription or dynamical construction.

Hawking–Page coexistence. The Hawking–Page temperature is instead $T_{\text{HP}} = (d-1)/(2\pi\ell)$, at $r_+ = \ell$. It exceeds T_0 [40]. The black-hole branch there has finite heat capacity $C_B = (d-1)^2 S_{\text{HP}}$. Between T_0 and T_{HP} it is metastable relative to thermal AdS, even when its local heat capacity is positive.

When both phases are retained, the choice of phase supplies another fluctuation. If their probabilities are w and $1-w$, with conditional means μ_B, μ_A and variances V_B, V_A of a chosen observable, then

$$V = wV_B + (1-w)V_A + w(1-w)(\mu_B - \mu_A)^2. \quad (8.18)$$

For modular energy at coexistence, the mean separation is extensive. A mixture with both weights of order one can therefore have width of order the entropy, even though each phase has an ordinary square-root width. It is bimodal. Replacing it by a Gaussian of the same variance loses its two changes of slope in the averaged minimum. For two equally weighted narrow components separated by D , the central deficit is $D/4$. The coexistence window has relative temperature width of order S_{HP}^{-1} , with finite-size shifts depending on phase prefactors.

The charged continuous critical point. For a spherical RN–AdS black hole, choose the charge parameter q so that

$$T_H(r) = \frac{1}{4\pi r} \left(d-2 + \frac{dr^2}{\ell^2} - \frac{(d-2)q^2}{r^{2d-4}} \right), \quad S(r) = \frac{\Omega_{d-1} r^{d-1}}{4G}. \quad (8.19)$$

The fixed-charge small–large black-hole transition has a critical endpoint [20, 49]. At that endpoint,

$$r_c^2 = \frac{(d-2)^2 \ell^2}{d(d-1)}, \quad q_c^2 = \frac{r_c^{2d-4}}{(2d-3)(d-1)}, \quad T_c = \frac{(d-2)^2}{\pi(2d-3)r_c}. \quad (8.20)$$

Here $T'_H(r_c) = T''_H(r_c) = 0$. Let $N = S(r_c)$, $R = r/r_c$, and hold the critical geometry fixed while $G \rightarrow 0$, so $N \rightarrow \infty$. In the density-of-states model $e^{S(E)} dE/E_0$ on the physical RN outer-horizon family, the first law gives

$$\frac{dI_c}{dR} = N(d-1)R^{d-2} \left(\frac{T_H(R)}{T_c} - 1 \right), \quad I_c = \beta_c M - S. \quad (8.21)$$

Since $T_H/T_c - 1 = (2d-3)(R-1)^3/6 + \dots$,

$$I_c(R) - I_c(1) = N a_d (R-1)^4 + O(N(R-1)^5), \quad a_d = \frac{(d-1)(2d-3)}{24}. \quad (8.22)$$

The first restoring term is quartic and positive. Within this RN family the critical saddle is the global minimum; applying the construction to a larger theory requires checking additional phases and instabilities.

Put

$$X = (a_d N)^{1/4} (R-1), \quad L_N = (d-1) a_d^{-1/4} N^{3/4}. \quad (8.23)$$

The energy measure contains the Jacobian $M'(R)dR$, smooth and positive at $R = 1$. It is constant to leading order on this shrinking neighborhood. The resulting normalized limiting density is

$$p_0(X) = \frac{2}{\Gamma(1/4)} e^{-X^4}. \quad (8.24)$$

The centered S and K divided by L_N both tend to X in this model. Their common leading entropy standard deviation and central averaged-minimum deficit are

$$\sigma \sim L_N \sqrt{\frac{\Gamma(3/4)}{\Gamma(1/4)}}, \quad B(0) \sim \frac{L_N \sqrt{\pi}}{2\Gamma(1/4)}. \quad (8.25)$$

For example, $\int_0^\infty x^m e^{-x^4} dx = \frac{1}{4} \Gamma((m+1)/4)$ immediately gives both moments. Their ratio is $B(0)/\sigma = 2^{-5/4}$, distinct from the Gaussian value $1/\sqrt{2\pi}$. The width is finite at every finite N , grows as $N^{3/4}$, and decreases relative to the extensive entropy as $N^{-1/4}$. In the canonical random-state model, the broad-distribution planar correction from (7.29) remains order one.

8.5 The finite-temperature critical crossover

Let $t = (T - T_c)/T_c$ at $q = q_c$. The exact action relation

$$I_\beta(R) - I_\beta(1) = \frac{I_c(R) - I_c(1) - t[S(R) - N]}{1 + t} \quad (8.26)$$

shows that the correct thermal scaling variable is $h = L_N t$. At fixed h , the limiting density becomes

$$p_h(X) = \frac{e^{-X^4 + hX}}{\mathcal{Z}(h)}, \quad \mathcal{Z}(h) = \int_{\mathbb{R}} e^{-x^4 + hx} dx. \quad (8.27)$$

Write $\mu_h = \mathbb{E}_h X$ and $v_h = \text{Var}_h X$. The centered entropy profile at $\Delta = zL_N$ is

$$\frac{B(zL_N)}{L_N} \longrightarrow \frac{1}{2} [\mathbb{E}_h |z - X + \mu_h| - |z|], \quad \frac{\sigma}{L_N} \longrightarrow \sqrt{v_h}. \quad (8.28)$$

The averaged entropy is smooth through this finite crossover for $|t| = O(N^{-3/4})$. Its deficit relative to the sharp minimum contains a cusp at $z = 0$ because the reference minimum itself has a cusp.

For large $|h|$, the maximum of the density is near $X_* = (h/4)^{1/3}$, with Gaussian variance $(12X_*^2)^{-1}$. Equivalently, writing $b_d = (2d - 3)/6$, the regular saddle outside the critical window obeys

$$C_B \sim \frac{(d-1)N}{3b_d^{1/3}} |t|^{-2/3}, \quad N^{-3/4} \ll |t| \ll 1, \quad (8.29)$$

and hence

$$\sigma \sim \sqrt{\frac{d-1}{3b_d^{1/3}}} N^{1/2} |t|^{-1/3}. \quad (8.30)$$

The familiar critical response exponent describes this outer Gaussian regime. The quartic law supplies its finite-size continuation, rather than an unlimited width at $t = 0$.

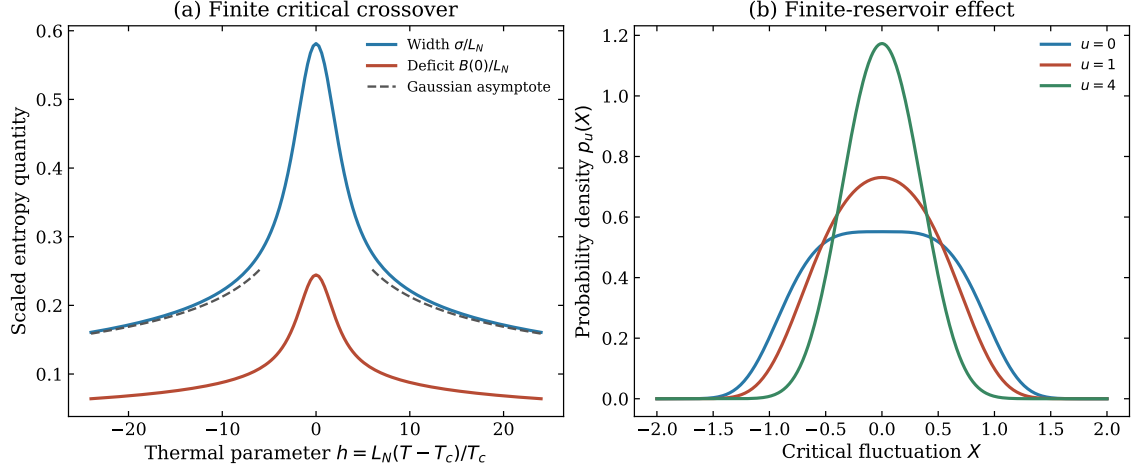


Figure 19. Controlled critical crossover from the normalized density $p_h(X) \propto e^{-X^4+hX}$. Left: the finite width σ/L_N and central deficit $B(0)/L_N$; the dashed curve is the Gaussian large- $|h|$ width, displayed only away from $h = 0$. Right: the critical density modified by a reservoir, $p_u(X) \propto e^{-X^4-uX^2}$, for several u . These are limiting distributions of the specified thermodynamic model, not extrapolations of a divergent Gaussian variance.

8.6 How large must a critical reservoir be?

At the critical point, expand a reservoir entropy around equilibrium:

$$S_R(E_{R,c} - \epsilon) = S_R(E_{R,c}) - \frac{\epsilon}{T_c} - \frac{\epsilon^2}{2T_c^2 C_R} + \mathcal{R}_R(\epsilon). \quad (8.31)$$

On the canonical critical scale, $\epsilon/T_c = L_N X + o(L_N)$. If the remainder tends to zero on bounded X intervals, the allowed energy domain includes these intervals, and the composite distribution has sufficiently controlled tails to pass to its first two moments, the combined law is

$$p_u(X) = \frac{e^{-X^4-uX^2}}{Z_u}, \quad u = \lim_{N \rightarrow \infty} \frac{L_N^2}{2C_R} \geq 0. \quad (8.32)$$

The tail requirement is separate from a Taylor expansion: rare distant energies must not carry a finite share of the scaled variance.

The canonical quartic limit requires $u = 0$, or $C_R \gg N^{3/2}$. A reservoir with $C_R \asymp N^{3/2}$ changes the shape at leading order. In the reservoir-limited regime $C_R \rightarrow \infty$ but $C_R \ll N^{3/2}$, the appropriate scale becomes $\sqrt{C_R}$. Provided the same local and tail control hold on that narrower scale, the law becomes Gaussian and $\sigma_B \sim \sqrt{C_R}$. An ordinary extensive reservoir, with capacity of order N , is therefore insufficient to reproduce the canonical critical width. If the two competing branches are S_B and S_R , their leading difference is $2L_N X$, so both its width and central deficit are twice the one-fluctuating-branch values.

This is a concrete consequence of the state preparation. It also supplies a useful test for applications: identify the thermodynamic mode, determine the constraint, and compute its coupling to the branch difference before assigning a critical exponent to an entropy transition.

Table 5. Moments of the geometric entropy shift x from the normalized disk JT density (8.34). The final column is the distinct canonical modular-energy variance.

β/π	$\langle x \rangle$	$\text{sd}(x)$	γ_1	$\text{Var } K$
1	13.566324	3.401121	0.050218	14.066371
2	7.280839	2.305939	0.134644	7.783185
4	4.119496	1.524513	0.262853	4.641593

8.7 A normalized JT example

For the disk Schwarzian density, in units with Schwarzian coupling one,

$$\rho_0(E) = \frac{e^{S_0}}{2\pi^2} \sinh(2\pi\sqrt{2E}), \quad E \geq 0. \quad (8.33)$$

This is the density per unit dE [79]. It is the disk density, not a nonperturbative completion of JT gravity. Its canonical probability is proportional to $\rho_0(E)e^{-\beta E}dE$. Define the geometric entropy shift $x = S_{\text{BH}} - S_0 = 2\pi\sqrt{2E}$. Since $dE = x dx/(4\pi^2)$, its normalized density is

$$p_\beta(x) = \frac{x \sinh x e^{-\beta x^2/(8\pi^2)}}{Z_x(\beta)}, \quad x \geq 0, \quad (8.34)$$

$$Z_x(\beta) = \int_0^\infty x \sinh x e^{-\beta x^2/(8\pi^2)} dx.$$

The factor x is the Jacobian. Omitting it changes the mean, variance, and the resulting conditional entropy average.

The disk partition function is proportional to $e^{S_0} \beta^{-3/2} e^{2\pi^2/\beta}$. Therefore

$$\langle E \rangle = \frac{2\pi^2}{\beta^2} + \frac{3}{2\beta}, \quad C_{\text{can}} = \frac{4\pi^2}{\beta} + \frac{3}{2} = \text{Var } K. \quad (8.35)$$

These identities are exact within the disk density. They are not formulas for $\text{Var } x$. For the latter, the distribution (8.34) can be integrated directly. Define its skewness by

$$\gamma_1 = \frac{\mathbb{E}(x - \langle x \rangle)^3}{(\text{Var } x)^{3/2}}. \quad (8.36)$$

The values in Table 5 come from the saved quadrature calculation accompanying this section. Neither the centered geometric distribution nor the centered modular distribution depends on S_0 ; the extremal degeneracy shifts entropy means.

Figure 20 isolates the effect of this non-Gaussian shape on a defined conditional average. It compares a fixed kernel at the mean geometric entropy, its average against a Gaussian with the same mean and variance, and its average against the full density (8.34). Matching the variance avoids attributing a width error to skewness. All three curves use the same conditional Page kernel. None is labeled the exact planar JT entropy.

At high temperature the action has a well-separated saddle, and the leading geometric width $\sqrt{4\pi^2/\beta}$ captures the dominant scale. At moderate temperature the lower boundary and non-Gaussian shape matter. For the canonical random-state model, one should push the same energy law forward to $K = \beta E + \log Z$ and then compute its diagonal or planar entropy as appropriate. Improving the probability distribution in a conditional average is distinct from replacing that average by a spectral calculation. It is also distinct from adding higher gravitational topologies.

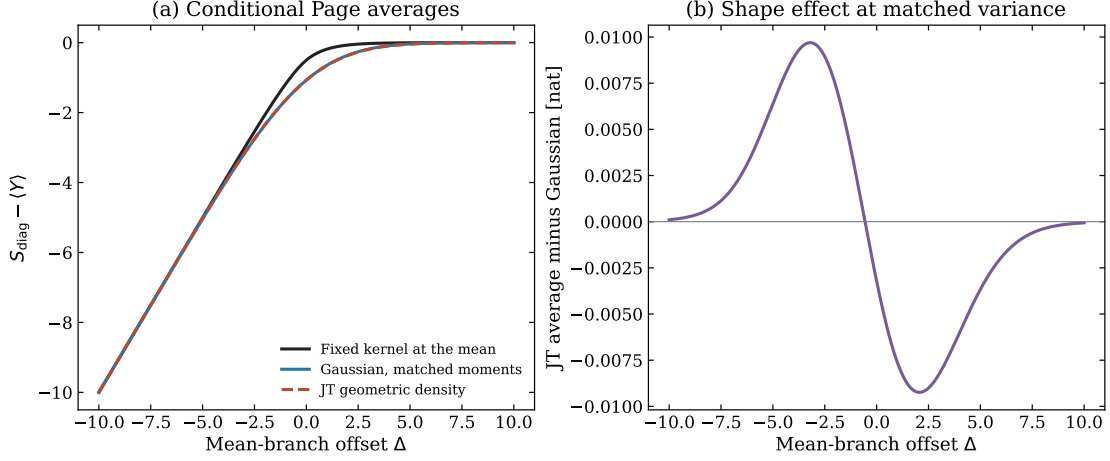


Figure 20. Conditional Page averages for the geometric entropy $Y = S_0 + x$, at $\beta = 2\pi$ and $S_0 = 20$. Left: the fixed kernel, a Gaussian average with matched moments, and numerical integration against the normalized JT density. Right: the difference between the two averages, isolating distribution shape at the same variance. The JT standard deviation is 2.305939 and skewness is 0.134644. These are diagonal conditional averages using $Y = S_{\text{BH}}$; the canonical random-state calculation instead uses $Y = K$ and its planar spectrum is a separate calculation.

8.8 What these examples establish

The regular Gaussian profile follows from a quadratic saddle and a specified fluctuating branch difference. Its width often scales as the square root of an extensive heat capacity, with coefficients depending on geometry and constraints. A finite total energy still permits fluctuations of subsystems, and their anticorrelation changes the relevant entropy difference.

At a charged continuous critical point the quartic action gives a finite $N^{3/4}$ width in the canonical model, a thermal window of order $N^{-3/4}$, and a reservoir crossover at capacity of order $N^{3/2}$. A spinodal instead marks the failure of a metastable saddle, while coexistence produces a mixture of distinct phases. These distributions should be treated separately.

Finally, thermodynamic entropy, canonical modular energy, averaged conditional entropy, and planar spectral entropy are distinct quantities. Their leading broad limits can agree without their order-one corrections agreeing. Keeping those distinctions explicit allows the subsequent discussion of higher moments and replica topology to address identifiable observables.

9 Genus Corrections and the Topological Expansion

The Marchenko–Pastur density is a leading large-dimension result. To go beyond it, one must specify both the ensemble and the normalization of the observable. This section develops the first correction in the complex Gaussian/Haar benchmark and then compares its organization with the continuum JT topology expansion. Three operations will be kept distinct: counting nonplanar Wick contractions, normalizing a random state, and adding handles to a gravitational surface. All can affect a replica calculation, but their coefficients are not interchangeable.

9.1 The genus-one permutation count

9.1.1 The closed-form formula

For $c_n = (1\,2\cdots n)$, recall the genus of a contraction,

$$g(\sigma) = \frac{n+1 - C(\sigma) - C(c_n^{-1}\sigma)}{2}. \quad (9.1)$$

The associated connected ribbon graph has one face, n edges, and $C(\sigma) + C(c_n^{-1}\sigma)$ vertices, so this equation is Euler's formula. Genus zero selects the C_n non-crossing permutations. Genus one requires

$$C(\sigma) + C(c_n^{-1}\sigma) = n - 1. \quad (9.2)$$

Let $g_j(n)$ denote the number of permutations of genus j . The classical genus-one count is

$$g_1(n) = \binom{2n-2}{n-1} \frac{(n-1)(n-2)}{12}. \quad (9.3)$$

This is a count of factorizations of a long cycle, or equivalently rooted bicolored one-face maps; see Cori and Hetyei [25], Goupil and Schaeffer [34]. The Harer–Zagier polygon-gluing problem [35] is related but counts a different family, so its coefficients should not be substituted for Eq. (9.3). The count is not a new gravitational result.

One convenient way to extract these numbers is their generating function

$$\sum_{n \geq 3} g_1(n) z^n = \frac{z^3}{(1-4z)^{5/2}}. \quad (9.4)$$

Indeed, the binomial series gives the coefficient $4^{n-3}(5/2)_{n-3}/(n-3)!$, which simplifies to Eq. (9.3). Here the generating-function identity encodes the combinatorial enumeration; the subsequent coefficient extraction is elementary.

9.1.2 Small replica indices and direct enumeration

The first values are

$$g_1(1), \dots, g_1(8) = 0, 0, 1, 10, 70, 420, 2310, 12012. \quad (9.5)$$

For example, at $n = 3$ five permutations are non-crossing and the remaining permutation $(1\,3\,2) = c_3^{-1}$ has genus one. At $n = 5$, the full partition of $5! = 120$ permutations is $42 + 70 + 8$ at genera zero, one, and two. The accompanying computation enumerates all permutations through $n = 8$ and checks both Euler's formula and the closed form.

Dividing the two closed forms gives

$$\frac{g_1(n)}{C_n} = \frac{n(n-1)(n-2)(n+1)}{24(2n-1)} \sim \frac{n^3}{48}. \quad (9.6)$$

Thus genus-one permutations first outnumber planar permutations at $n = 5$. Their numerical count is not their contribution to a moment: the dimension weights must still be included.

Table 6. Permutation counts and their ratios. A zero in the genus-one column concerns the Wick numerator; it does not imply that the normalized moment has no relative correction of order L^{-2} .

n	C_n	$g_1(n)$	$g_1(n)/C_n$	$n^3/48$
1	1	0	0	0.021
2	2	0	0	0.167
3	5	1	0.200	0.563
4	14	10	0.714	1.333
5	42	70	1.667	2.604
6	132	420	3.182	4.500
7	429	2310	5.385	7.146
8	1430	12012	8.400	10.667

9.2 Normalized moments beyond the planar limit

9.2.1 Balanced dimensions

Set $k = L$, with L an integer dimension, and use the normalized state of Eq. (5.9). The exact finite formula becomes

$$\mathbb{E}_H \text{Tr } \rho^n = L^{1-n} \frac{L^{2n}}{(L^2)_n} \sum_{g=0}^{\lfloor (n-1)/2 \rfloor} g_g(n) L^{-2g}. \quad (9.7)$$

The bound on g follows because each of the two cycle counts is at least one. The numerator is a finite polynomial for each integer n . The corresponding mean-normalized moment has the same expression without $L^{2n}/(L^2)_n$. In particular its genus-one contribution is exactly

$$\delta_1[\mathbb{E} \text{Tr } \tilde{\rho}^n] = L^{1-n} g_1(n) L^{-2}. \quad (9.8)$$

The normalization factor has its own expansion,

$$\frac{L^{2n}}{(L^2)_n} = \prod_{j=0}^{n-1} \left(1 + \frac{j}{L^2}\right)^{-1} = 1 - \frac{n(n-1)}{2L^2} + O(L^{-4}), \quad (9.9)$$

for fixed n . Multiplying the numerator and denominator expansions gives

$$L^{n-1} \mathbb{E}_H \text{Tr } \rho^n = C_n + \frac{g_1(n) - \frac{1}{2}n(n-1)C_n}{L^2} + O(L^{-4}). \quad (9.10)$$

Both terms in the correction are required. The negative normalization term can exceed the positive contribution from nonplanar contractions.

9.2.2 Why purity still has a correction

At $n = 1$, the exact moment is one. At $n = 2$, $g_1(2) = 0$, but the normalization factor is $L^2/(L^2 + 1)$. Hence

$$\mathbb{E}_H \text{Tr } \rho^2 = \frac{2L}{L^2 + 1} = \frac{2}{L} - \frac{2}{L^3} + O(L^{-5}). \quad (9.11)$$

The leading relative correction is $-L^{-2}$. This provides a direct counterexample to identifying absence of a genus-one contraction with absence of a correction at that order. For $n = 3$ the exact formula is

$$\mathbb{E}_H \text{Tr } \rho^3 = \frac{5L^2 + 1}{(L^2 + 1)(L^2 + 2)} = \frac{5}{L^2} - \frac{14}{L^4} + O(L^{-6}). \quad (9.12)$$

The numerator coefficient $g_1(3) = 1$ and the denominator coefficient $-3C_3 = -15$ combine to give -14 .

For clarity, the resulting order statements are

$$\frac{\delta \mathbb{E} \operatorname{Tr} \rho^2}{2/L} = O(L^{-2}), \quad \frac{\delta \mathbb{E} \operatorname{Tr} \rho^3}{5/L^2} = O(L^{-2}). \quad (9.13)$$

They refer to ensemble mean moments. The average of a Rényi entropy is not obtained by taking the logarithm of its mean moment; the distinction is examined in Section 11.

9.2.3 An independent entropy check

The von Neumann entropy involves a derivative at replica index one, not the value of one integer moment. A closed form for $g_1(n)$ at integer n does not by itself justify differentiating a chosen interpolation of it. In the Haar model the exact Page formula (6.38) provides an independent calculation:

$$\mathbb{E}_{\text{H}} S(\rho) = H_{L^2} - H_L - \frac{L-1}{2L} = \log L - \frac{1}{2} + \frac{7}{12L^2} + O(L^{-4}). \quad (9.14)$$

Thus the entropy does receive a correction at the same dimension order. At $L = 2$ the exact result is $1/3$ nat, while the leading expression is $\log 2 - 1/2$. Equations (9.11) and (9.14) check different observables, with different coefficients, in the same normalized ensemble.

9.2.4 Unequal dimensions

At fixed $\gamma = k/L$, the genus- g coefficient is a polynomial in γ rather than the unweighted count $g_g(n)$. For example, the mean-normalized expansion is

$$\mathbb{E} \operatorname{Tr} \tilde{\rho}^n = k^{1-n} \sum_g k^{-2g} \sum_{\substack{\sigma \in S_n \\ g(\sigma)=g}} \gamma^{n-C(\sigma)}. \quad (9.15)$$

The normalized answer multiplies this by $(kL)^n/(kL)_n$. Counting permutations without their cycle numbers is sufficient only at $\gamma = 1$. More general matrix ensembles change the contraction weights or even the available contractions, so their finite-size corrections require an independent calculation.

9.3 The range of the genus expansion

For fixed n the permutation numerator is finite, whereas its normalized rational expression has an infinite expansion in inverse dimension. Neither fact implies uniform control as n grows with L . At balanced dimensions, the ratio of the genus-one to planar contribution to the mean-normalized moment is

$$\frac{g_1(n)}{C_n L^2} \sim \frac{n^3}{48 L^2}. \quad (9.16)$$

A necessary check for retaining only genus zero is therefore $n^3/L^2 \ll 1$. At n of order $L^{2/3}$, the first omitted term alone is no longer parametrically small. This estimate does not establish a complete uniform asymptotic expansion; it identifies a scale at which a fixed-replica argument cannot simply be extrapolated.

Remark 9.1 (The Catalan/genus-one crossover). At a fixed small replica index, large dimension suppresses genus one even when $g_1(n) > C_n$. If the replica index is allowed to grow, the combinatorial factor competes with the suppression. Both the index and dimension must be specified when claiming that a correction is small.

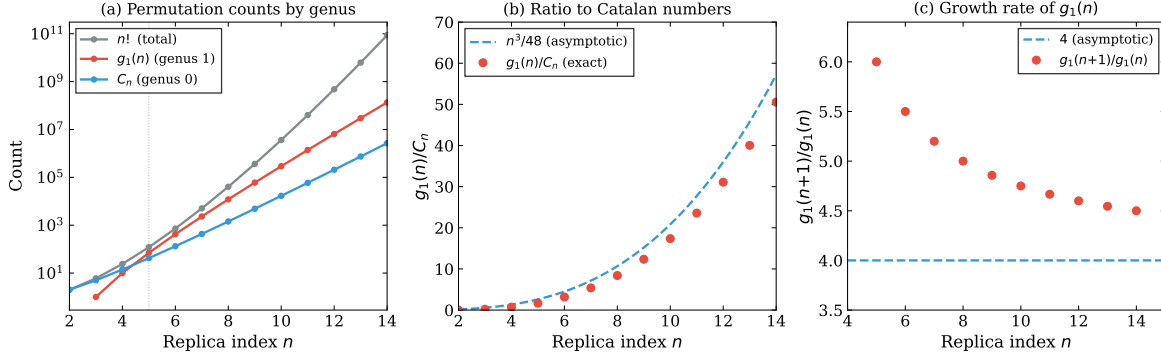


Figure 21. Genus-one permutation counting. (a) The counts $n!$, $g_1(n)$, and C_n . (b) The exact ratio $g_1(n)/C_n$ and its large- n approximation $n^3/48$. (c) The ratio $g_1(n+1)/g_1(n) = 2(2n-1)/(n-2)$ approaches four. These are counts. Their contributions to moments include dimension weights and, for normalized states, the rising-factorial denominator.

9.4 Continuum handles, fluctuations, and QES corrections

9.4.1 The JT topology expansion

For connected JT amplitudes with b asymptotic boundaries, the perturbative organization is [79]

$$Z_b^{\text{conn}}(\beta_1, \dots, \beta_b) \sim \sum_{g \geq 0} e^{(2-2g-b)S_0} Z_{g,b}(\beta_1, \dots, \beta_b). \quad (9.17)$$

The coefficients include integrals over moduli spaces of hyperbolic surfaces, with boundary wavefunctions fixed by the preparation. Unlike the Wick genus at fixed integer n , the continuum genus is unbounded at a fixed boundary number. In the JT matrix-integral description the perturbative coefficients match a double-scaled spectral expansion. Defining a nonperturbative theory requires additional choices; a formal topology series alone does not fix a unique completion.

Connectedness and genus are also different notions. A genus-zero surface can connect two or more boundaries and thus contribute to a connected ensemble correlator. A connected fluctuation does not automatically carry the relative weight of an added handle. Likewise, a one-loop determinant about a fixed classical saddle is an expansion in fluctuations of fields on that saddle, not a change in its topology. These distinctions are needed before importing powers of e^{-S_0} into an entropy-variance calculation.

9.4.2 What the Akers–Penington result establishes

Akers and Penington [3] identify conditions in which the usual minimal-QES formula needs leading-order corrections. The issue concerns the bulk state between competing surfaces, characterized by smooth min- and max-entropies. It is not the genus-one coefficient in Eq. (9.3).

A useful illustration is their mixture argument. For $\rho = p\rho_1 + (1-p)\rho_2$, entropy concavity and the classical label bound imply

$$pS(\rho_1) + (1-p)S(\rho_2) \leq S(\rho) \leq pS(\rho_1) + (1-p)S(\rho_2) + h_2(p), \quad (9.18)$$

where $h_2(p) = -p \log p - (1-p) \log(1-p)$. For fixed p , this uncertainty is at most $\log 2$. A naive saddle prescription that instead differs from the weighted component entropies by order $1/G_N$ contradicts

these bounds. The required correction is then leading in the semiclassical expansion, even though no genus-one permutation has been invoked. This example explains why a state-sensitive QES correction cannot be identified with a universal exponentially suppressed handle correction.

9.4.3 What can be compared between models

The permutation counts are mathematical constants in every genus. The physical weights assigned to replica components, the continuum moduli integrals, normalization fluctuations, and state preparation are model data. A comparison between two theories must calculate these data for the same observable and regime. The existence of a topological expansion, or an agreement of leading thermodynamic entropies, does not establish equality of their planar density or inequality of their first subleading coefficient. The finite examples in the following sections demonstrate what spectra and connected observables can distinguish without presupposing those uncomputed gravity comparisons.

10 What Entropies and Spectra Determine

The random-state calculation gives a precise answer to a specified statistical question. It does not, by itself, identify the dynamics that prepared the state. This section explains how much information is retained by the different quantities used in that calculation. The distinction is familiar from probability: a mean does not specify a distribution, and a one-point distribution does not specify correlations. Here the random objects are density matrices and their spectra.

We develop the distinction through elementary examples, including the finite ensembles of Paper III [101]. These examples are useful because every state and every probability can be written down. They do not assume a semiclassical approximation, an entropy expansion, or a particular microscopic model of a black hole. Their role is to identify which further calculations are required before an entropy result can answer a question about ensembles or information recovery.

10.1 From a density matrix to an entropy

For a density matrix ρ on $\mathbb{C}^d$, let $\lambda_1, \dots, \lambda_d$ be its eigenvalues, including zeros. The spectral counting measure and its moments are

$$\nu_\rho = \sum_{i=1}^d \delta_{\lambda_i}, \quad M_n(\rho) = \text{tr } \rho^n = \int_0^1 x^n \nu_\rho(dx). \quad (10.1)$$

The measure has mass d , while its first moment is one. The distinction between these normalizations matters when comparing a density of entanglement eigenvalues with a probability density. Entropy is another linear functional of the counting measure:

$$S(\rho) = - \int_0^1 x \log x \nu_\rho(dx), \quad S_n(\rho) = \frac{\log M_n(\rho)}{1-n}. \quad (10.2)$$

Thus a spectrum determines all these entropies, but a single entropy does not determine the spectrum. Even the full spectrum determines a matrix only up to a unitary change of eigenvectors.

For an explicit example, choose $0 < p < 1$, $p \neq 1/2$, and set

$$\rho = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix}, \quad H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \rho' = H \rho H^\dagger. \quad (10.3)$$

Every spectral entropy agrees. However, for the same physical observable $Z = \text{diag}(1, -1)$,

$$\text{tr}(\rho Z) = 2p - 1, \quad \text{tr}(\rho' Z) = 0. \quad (10.4)$$

This is a change of state while the measurement basis is held fixed. Rotating both the state and every observable would instead be a change of representation. A radiation spectrum omits precisely this relation between the eigenvectors and the accessible observables.

The analogy with thermodynamics has a limited but useful meaning. An entropy is a coarse description of a state, just as an equation of state is a coarse description of a material. Agreement of the coarse data is compatible with disagreement of other data. It is not a theorem that all microscopic theories agree on the coarse data: different Hamiltonians, constraints, or preparations can already give different leading entropies. The task is to state the data held fixed in each comparison.

10.2 An averaged spectrum is not an ensemble of spectra

Let ρ now be random, with a specified probability distribution over normalized density matrices on the same space. Then

$$\bar{\nu} = \mathbb{E}\nu_\rho, \quad \mathbb{E}M_n = \int x^n \bar{\nu}(dx), \quad \mathbb{E}S = - \int x \log x \bar{\nu}(dx). \quad (10.5)$$

These are one-point spectral data. They describe the expected number of eigenvalues in any interval, not how eigenvalues occur together in one realization. The two-point measure $\mathbb{E}[\nu_\rho(dx)\nu_\rho(dy)]$ retains the latter information. For example,

$$\mathbb{E}[M_n M_\ell] = \iint x^n y^\ell \mathbb{E}[\nu_\rho(dx)\nu_\rho(dy)] \quad (10.6)$$

cannot generally be recovered from $\bar{\nu}$.

This distinction persists even if all positive integer moments are known exactly. Measures on $[0, 1]$ are determined by their moments: polynomials are uniformly dense among continuous functions, and equality of their integrals therefore implies equality of the measures. For a fixed d , the zeroth moment is supplied by $\bar{\nu}([0, 1]) = d$. Consequently all $\mathbb{E}M_n$ determine $\bar{\nu}$. They do not specify a probability measure on the space of spectra. To see the missing information directly, it is enough to use four-dimensional states.

10.3 Two ensembles with exactly the same averaged spectrum

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

$$\begin{aligned} L_u &= (c + u, c - u, c + u, c - u), \\ L_v &= (c + v, c - v, c + v, c - v), \\ L_{uv} &= (c + u, c - u, c + v, c - v). \end{aligned} \quad (10.7)$$

All entries are positive and every row sums to one. Ensemble $\mathcal{A}$ chooses the diagonal matrix with spectrum L_u or L_v with equal probabilities. Ensemble $\mathcal{B}$ always chooses the spectrum L_{uv} . In both ensembles we then conjugate the state by an independent Haar unitary in $U(4)$. The conjugation leaves every spectral statistic unchanged and makes both averaged states equal to $I_4/4$. It can be omitted if only spectral comparisons are of interest.

For any function g defined on the four eigenvalue values, introduce

$$a_g(t) = g(c + t) + g(c - t). \quad (10.8)$$

The statistic $\text{tr } g(\rho)$ has values $2a_g(u)$ and $2a_g(v)$ in $\mathcal{A}$, and the single value $a_g(u) + a_g(v)$ in $\mathcal{B}$. This proves the following exact comparison.

Proposition 10.1 (Equal averaged spectra and unequal fluctuations). *The ensembles defined by (10.7) satisfy*

$$\mathbb{E}_{\mathcal{A}} \text{tr } g(\rho) = \mathbb{E}_{\mathcal{B}} \text{tr } g(\rho) = a_g(u) + a_g(v). \quad (10.9)$$

For real-valued functions g, h , their covariances are

$$\begin{aligned} \text{Cov}_{\mathcal{A}}(\text{tr } g(\rho), \text{tr } h(\rho)) &= [a_g(u) - a_g(v)][a_h(u) - a_h(v)], \\ \text{Cov}_{\mathcal{B}}(\text{tr } g(\rho), \text{tr } h(\rho)) &= 0. \end{aligned} \quad (10.10)$$

In particular, their averaged spectral measures, all averaged replica moments, and mean von Neumann entropies agree, while their purity and entropy variances differ.

To verify the covariance, subtract the common mean. In $\mathcal{A}$ the two centered values of $\text{tr } g(\rho)$ are $\pm[a_g(u) - a_g(v)]$. The same sign occurs for h in each realization, so the product is the right-hand side of (10.10) in both realizations. Every spectral statistic in $\mathcal{B}$ is constant. This derivation also shows that the difference concerns correlations between eigenvalues, rather than an error in matching the one-point density.

For purity, $a_g(t) = 2(c^2 + t^2)$, and hence

$$\text{Var}_{\mathcal{A}}(M_2) = 4(u^2 - v^2)^2, \quad \text{Var}_{\mathcal{B}}(M_2) = 0. \quad (10.11)$$

For entropy define $f(t) = -(c+t)\log(c+t) - (c-t)\log(c-t)$. Then

$$\text{Var}_{\mathcal{A}}(S) = [f(u) - f(v)]^2, \quad \text{Var}_{\mathcal{B}}(S) = 0. \quad (10.12)$$

Since $f'(t) = \log[(c-t)/(c+t)] < 0$ for $0 < t < c$, the first variance is positive when $u < v$. There is no large-dimension limit in this statement. Each matrix can be purified by an environment of dimension four, so the example also defines ensembles of bipartite pure states.

A numerical example with rational eigenvalues. Take $u = 0$ and $v = 1/8$. The three spectra become

$$L_0 = (1/4, 1/4, 1/4, 1/4), \quad L_{1/8} = (3/8, 1/8, 3/8, 1/8), \quad L_{0,1/8} = (1/4, 1/4, 3/8, 1/8). \quad (10.13)$$

Their purities are $1/4$, $5/16$, and $9/32$, respectively. Averaging the first two therefore gives the third, whereas their purity variance is $1/1024$. Their common mean entropy is

$$\mathbb{E}S = \frac{5}{2} \log 2 - \frac{3}{8} \log 3, \quad \text{Var}_{\mathcal{A}}S = \frac{1}{4} \left(\frac{3}{4} \log 3 - \log 2 \right)^2. \quad (10.14)$$

The entropy of each averaged matrix is $\log 4$. Thus even the entropy of the averaged state and its difference from the mean entropy agree between these ensembles. Section 11 will relate their different entropy variances to finite-order Rényi averages.

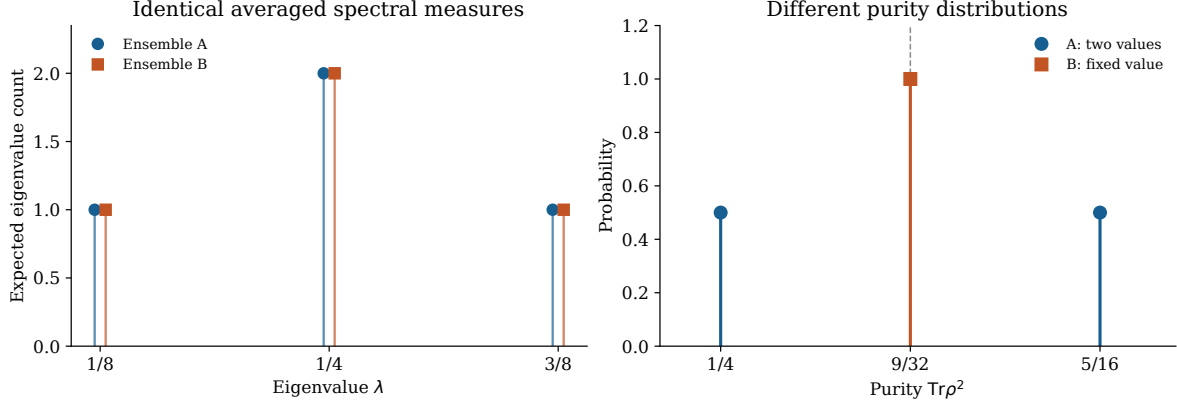


Figure 22. The exact four-dimensional example (10.13). Left: the averaged eigenvalue counting measures agree; the markers are displaced horizontally only for visibility. Right: ensemble $\mathcal{A}$ has two equally probable purity values, whereas $\mathcal{B}$ has the constant purity $9/32$. Their means coincide, but their variances are $1/1024$ and zero. These are exact finite distributions, not random samples.

10.4 What a state design controls

Random circuits are often used to justify a random-state approximation. The precise intermediate notion is a design. Let $\mathcal{E}$ be a probability distribution on normalized pure states of $\mathbb{C}^k \otimes \mathbb{C}^L$, and define its n -copy moment operator by

$$\Omega_{\mathcal{E}}^{(n)} = \mathbb{E}_{\mathcal{E}}(|\psi\rangle\langle\psi|)^{\otimes n}. \quad (10.15)$$

An exact projective state t -design reproduces the Haar moment operators through order t . For approximate matching, specify the error in trace norm:

$$\epsilon_n = \|\Omega_{\mathcal{E}}^{(n)} - \Omega_{\text{H}}^{(n)}\|_1. \quad (10.16)$$

This definition is part of the assumption. A theorem about approximate unitary designs stated in another norm must be translated before its error can be substituted here. A unitary design acting on a fixed pure input supplies a state design; one individual draw from the unitary ensemble is not itself that design.

The replica permutation identity makes the consequence transparent. Let $T_{R,n}$ cyclically permute the n copies of the radiation factor and act trivially on the remaining factors. Then

$$\text{tr } \rho_R^n = \text{tr}[(|\psi\rangle\langle\psi|)^{\otimes n} T_{R,n}], \quad \|T_{R,n}\|_{\infty} = 1. \quad (10.17)$$

Subtracting the two ensemble averages and using $|\text{tr } AB| \leq \|A\|_1 \|B\|_{\infty}$ gives

$$|\mathbb{E}_{\mathcal{E}} M_n - \mathbb{E}_{\text{H}} M_n| \leq \epsilon_n. \quad (10.18)$$

A product $M_{n_1} \cdots M_{n_r}$ uses disjoint permutation cycles and requires $n_1 + \cdots + n_r$ copies. In particular, two-copy matching controls mean purity, while four-copy matching controls its second moment. Because $0 \leq M_2 \leq 1$,

$$|\text{Var}_{\mathcal{E}}(M_2) - \text{Var}_{\text{H}}(M_2)| \leq \epsilon_4 + 2\epsilon_2. \quad (10.19)$$

The second term bounds the difference between the squares of the means. An exact two-design need not satisfy the four-copy condition.

The scale of the desired observable also matters. The natural scaled spectral moment is $k^{n-1}M_n$, so its error is at most $k^{n-1}\epsilon_n$. Recovering a Marchenko–Pastur moment as $k, L \rightarrow \infty$ therefore requires control at that scale, in addition to the corresponding Haar limit. Recovering a small covariance can require substantially greater precision than recovering a leading mean. The local random-circuit results of Brandão, Harrow, and Horodecki [16] give design statements under specified circuit assumptions, with depth depending on design order. A scrambling-time estimate alone supplies neither arbitrary-order design matching nor these dimension-dependent error bounds.

10.5 Passing from moments to an entropy limit

At fixed dimension, the compact support of the unscaled eigenvalues makes the full moment problem well behaved. In a large-dimension calculation the useful variables are instead $x_i = k\lambda_i$, with empirical probability measure $\mu_k = k^{-1} \sum_i \delta_{x_i}$. Now

$$S(\rho_R) = \log k - \int x \log x \mu_k(dx). \quad (10.20)$$

A limiting law for μ_k must control this integral, including any spectral tail. Weak convergence alone tests bounded continuous functions; $x \log x$ is unbounded at infinity. It is therefore useful to state one sufficient condition explicitly.

Suppose μ_k converges weakly in probability to a deterministic measure μ , and $\sup_k \mathbb{E} \int x^2 \mu_k(dx) < \infty$. For $K > e$,

$$\int_{x>K} x \log x \mu_k(dx) \leq \frac{\log K}{K} \int x^2 \mu_k(dx). \quad (10.21)$$

The expected right-hand side tends to zero as $K \rightarrow \infty$. The integrand extends continuously to zero and is bounded below, so a bounded truncation converges in probability and in mean. Removing the truncation proves convergence of the entropy correction in (10.20). The same tail argument controls the limiting integral. This is one route from a spectral law to entropy; it displays the extra estimate that finitely many moment matches do not automatically provide.

For the flat-weight Haar/Wishart model, the Marchenko–Pastur law and Page’s mean entropy are established benchmarks [58, 71]. Their appearance in two calculations is evidence of common state statistics in the stated regime. It does not imply agreement of eigenvectors, finite-size covariances, or channel matrix elements. Conversely, the presence of non-crossing partitions alone does not force the Marchenko–Pastur law: different free cumulant weights yield different moment sequences. Both the combinatorics and the weights must be specified.

10.6 A practical dictionary of the different questions

Table 7 summarizes implications established by the definitions. Its rows are not a ranking of microscopic theories. They indicate which data are discarded when a more detailed object is replaced by a summary.

The last row will be derived in Section 13. It is an important positive result: sufficiently complete reference-assisted data do determine a channel. The limitations established here concern particular incomplete data, not an impossibility of learning quantum dynamics. Likewise, extra spectral moments and covariances can distinguish specified models when their predictions differ. What cannot be inferred is a universal separation of broad gravitational programs before those predictions have been calculated.

Table 7. Data about a state, an ensemble, and a channel. A stronger input in one row need not supply the data of another row.

Given data	What they determine	What remains unspecified
One density matrix	All expectation values in that state	Its preparation and channel action on other inputs
Its full spectrum	Every spectral entropy	Eigenvectors relative to physical observables
Averaged spectral measure	All mean linear spectral statistics	Covariances between those statistics
Finite-order state design	Corresponding copy moments	Higher-order errors, general entropy limits, and time correlations
Full joint output for a maximally entangled reference input	The channel on every input	The physical implementation and complexity of a decoder

Exercises. For the spectra in (10.13), compute the common third and fourth mean moments and verify that $\mathbb{E}_{\mathcal{A}} M_2^2 - \mathbb{E}_{\mathcal{B}} M_2^2 = 1/1024$. Next omit the Haar conjugation and compare the two averaged matrices. The averaged spectral measure still agrees, but equality of averaged matrices no longer follows automatically. Finally, use (10.10) to calculate $\text{Cov}(M_2, S)$ and explain its sign: the more concentrated spectrum has higher purity and lower entropy. These checks distinguish the role of spectral randomness from the role of random eigenvectors.

11 Rényi Averages and Entropy Fluctuations

The terms annealed and quenched originate in problems where a logarithm of a partition function is taken after or before an average over disorder. The same ordering issue arises for replica moments. To use the analogy, however, we must specify the normalized random quantity inside the logarithm. The derivations below develop the finite-dimensional identities used in Paper III [101].

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 $0 \log 0 = 0$.

11.1 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

$$A_n = \frac{\log \mathbb{E} M_n}{1 - n}, \quad (11.1)$$

$$Q_n = \frac{\mathbb{E} \log M_n}{1 - n} = \mathbb{E} S_n(\rho). \quad (11.2)$$

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

$$G_n := Q_n - A_n = \frac{\log \mathbb{E} M_n - \mathbb{E} \log M_n}{n - 1} \geq 0. \quad (11.3)$$

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

$$\lim_{n \rightarrow 1} A_n = \lim_{n \rightarrow 1} Q_n = \mathbb{E} S(\rho), \quad S(\rho) = -\text{tr}(\rho \log \rho). \quad (11.4)$$

Moreover, as $t \rightarrow 0$,

$$G_{1+t} = \frac{t}{2} \text{Var}[S(\rho)] + O(t^2). \quad (11.5)$$

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

$$M_{1+t} = 1 - tS + \frac{t^2}{2} L_2 + O(t^3), \quad \log M_{1+t} = -tS + \frac{t^2}{2} (L_2 - S^2) + O(t^3). \quad (11.6)$$

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.

11.2 Averaging a spectrum and averaging a state

Another nonnegative difference, often relevant to ensemble interpretations, is

$$\chi = S(\bar{\rho}) - \mathbb{E}S(\rho), \quad \bar{\rho} = \mathbb{E}\rho. \quad (11.7)$$

This is the Holevo quantity of the state ensemble [42]. 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,

$$\chi = \mathbb{E}D(\rho\|\bar{\rho}), \quad D(\rho\|\sigma) = \text{tr}[\rho(\log \rho - \log \sigma)] \geq 0, \quad (11.8)$$

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:

$$- \int_0^1 x \log x \, \mathbb{E}\nu_\rho(dx) = \mathbb{E}S(\rho). \quad (11.9)$$

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.

11.3 An exactly evaluated two-state ensemble

Consider the qubit ensemble

$$\rho = \begin{cases} |0\rangle\langle 0|, & \text{with probability } p, \\ I_2/2, & \text{with probability } 1-p, \end{cases} \quad 0 \leq p \leq 1. \quad (11.10)$$

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:

$$Q_n = (1-p) \log 2, \quad A_n = - \frac{\log[p + (1-p)2^{1-n}]}{n-1}, \quad (11.11)$$

$$\text{Var}(S) = p(1-p)(\log 2)^2, \quad \chi = h_2\left(\frac{1+p}{2}\right) - (1-p) \log 2, \quad (11.12)$$

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 $(\log 2)^2/8 = 0.060057\dots$

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.

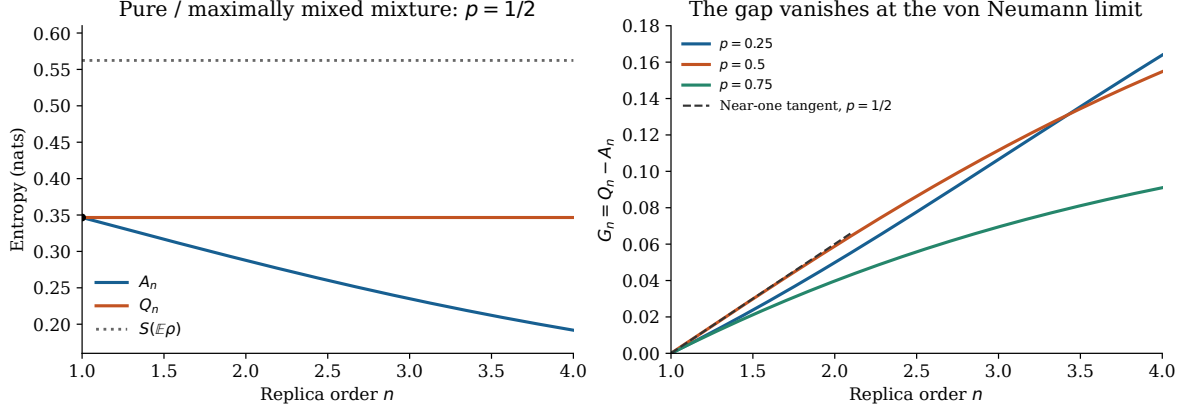


Figure 23. Exact averages for the qubit ensemble in (11.10). Left: A_n and Q_n at $p = 1/2$ approach the same von Neumann limit; the entropy $S(\bar{\rho})$ of the averaged state remains a different quantity. Right: the nonnegative gap $G_n = Q_n - A_n$ for three mixing probabilities. The dashed tangent at $p = 1/2$ is $(n - 1)\text{Var}(S)/2$. Every curve is obtained from (11.11); no Page-time profile or gravitational fluctuation amplitude is assumed.

11.4 Normalized replica covariances

The object that determines entropy fluctuations is the connected moment

$$C(n, \ell) = \mathbb{E}[M_n M_\ell] - \mathbb{E}M_n \mathbb{E}M_\ell. \quad (11.13)$$

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

$$\partial_n \partial_\ell C(n, \ell)|_{n=\ell=1} = \text{Var}[S(\rho)]. \quad (11.14)$$

Both identities are exact, with differentiation justified as in Proposition 11.1. 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 10.1 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 [33]:

$$\mathbb{E}M_2 = \frac{m+b}{mb+1}, \quad C(2, 2) = \frac{2(m^2-1)(b^2-1)}{(mb+1)^2(mb+2)(mb+3)}. \quad (11.15)$$

We derive this formula below 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.

11.5 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

$$\frac{\mathbb{E}\delta_n^2}{2(n-1)(1+\eta)^2} \leq G_n \leq \frac{\mathbb{E}\delta_n^2}{2(n-1)(1-\eta)^2}. \quad (11.16)$$

Under the same condition the quadratic approximation is

$$G_n = \frac{\text{Var}(M_n)}{2(n-1)\mu_n^2} + O\left(\frac{\mathbb{E}|\delta_n|^3}{(n-1)(1-\eta)^3}\right). \quad (11.17)$$

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 (11.17) itself vanishes quadratically: $\text{Var}(M_{1+t}) = t^2 \text{Var}(S) + O(t^3)$. This restores the linear behavior in (11.5); treating that numerator as independent of n would give a spurious singularity. Separately, an entropy variance bound has the direct probabilistic consequence

$$\mathbb{P}\{|S - \mathbb{E}S| \geq a\} \leq \frac{\text{Var}(S)}{a^2}, \quad a > 0. \quad (11.18)$$

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.

11.6 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

$$\text{Var}(S) = \mathbb{E}[\text{Var}(S | \Theta)] + \text{Var}(\mathbb{E}[S | \Theta]). \quad (11.19)$$

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

$$\begin{aligned} \text{Var}(M_2) = \mathbb{E} \left[\frac{2(m^2 - 1)(B^2 - 1)}{(mB + 1)^2(mB + 2)(mB + 3)} \right] \\ + \text{Var} \left(\frac{m + B}{mB + 1} \right). \end{aligned} \quad (11.20)$$

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.

11.7 Application to gravitational ensembles

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

$$M_n = \frac{\text{tr } W^n}{(\text{tr } W)^n}. \quad (11.21)$$

The average of this ratio generally differs from replacing its denominator by $(\mathbb{E} \text{tr } W)^n$. 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 [15] 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 (11.4). 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.

11.8 A four-replica calculation of purity fluctuations

It is instructive to derive the Haar purity variance explicitly. This also checks which replica data a fluctuation requires. In the following calculation the radiation and environment dimensions are k and L , and $D = kL$. For a Haar vector on $\mathbb{C}^D$, its r -copy average is

$$\mathbb{E}(|\psi\rangle\langle\psi|)^{\otimes r} = \frac{1}{(D)_r} \sum_{\pi \in S_r} U_\pi, \quad (D)_r = D(D+1) \cdots (D+r-1). \quad (11.22)$$

This follows from unitary invariance: the average is the identity on the symmetric subspace, divided by its dimension. The symmetrizer is $r!^{-1} \sum_{\pi} U_\pi$, and its trace is $\binom{D+r-1}{r} = (D)_r / r!$, giving the displayed normalization. The identity is a finite-dimensional statement, before a planar approximation.

Let $C(\pi)$ count the cycles of a permutation, including fixed points. The radiation contraction for M_2^2 uses $\tau = (12)(34)$, so

$$\mathbb{E}M_2^2 = \frac{1}{(kL)_4} \sum_{\pi \in S_4} k^{C(\tau\pi)} L^{C(\pi)}. \quad (11.23)$$

The two disjoint cycles represent the two traces. This differs from a single four-cycle, which would calculate $\mathbb{E}M_4$. Enumerating the 24 permutations gives the following polynomial:

$$\begin{aligned} \sum_{\pi \in S_4} k^{C(\tau\pi)} L^{C(\pi)} &= k^2 L^4 + 2k^3 L^3 + k^4 L^2 + 4kL^3 \\ &\quad + 10k^2 L^2 + 4k^3 L + 2kL. \end{aligned} \quad (11.24)$$

For example, the identity permutation contributes $k^2 L^4$, while $\pi = \tau$ contributes $k^4 L^2$. The coefficients sum to 24 when $k = L = 1$, as they must. Dividing by $(kL)_4$, subtracting $[(k+L)/(kL+1)]^2$, and collecting terms gives

$$\text{Var}(M_2) = \frac{2(k^2 - 1)(L^2 - 1)}{(kL + 1)^2(kL + 2)(kL + 3)}. \quad (11.25)$$

This reproduces (11.15) with $m = k, b = L$. For two qubits, the mean purity is $4/5$ and its variance is $3/175$. When either subsystem has dimension one, the variance vanishes exactly. Both limits provide simple checks on the normalization.

At $k = L$, the large-dimension expansion is

$$\mathbb{E}M_2 = \frac{2}{k} - \frac{2}{k^3} + O(k^{-5}), \quad \text{Var}(M_2) = \frac{2}{k^4} + O(k^{-6}). \quad (11.26)$$

The correction to the mean is small compared with the mean itself. Nevertheless, its product with the leading mean contributes to the square of the mean at order k^{-4} , the very order of the variance. It is therefore inconsistent to retain connected terms of this size while discarding the normalization corrections needed in the subtraction. A leading spectral density is not sufficient for a subleading fluctuation calculation.

11.9 Interpreting vanishing and nonvanishing differences

There are several distinct reasons why a difference can vanish. If ρ is deterministic, all ensemble fluctuations vanish. If its spectrum is deterministic but its eigenvectors vary, every G_n still vanishes, although the Holevo quantity can be positive. If only one moment is deterministic, G_n vanishes at that order without fixing all other spectral observables. Finally, the normalized von Neumann annealed and quenched limits agree in every finite-dimensional ensemble, including ones with large entropy variance. These are mathematically different statements.

The converse implications are equally limited. A nonzero G_2 certifies fluctuations of purity in the specified ensemble; it does not identify their origin. A nonzero entropy variance certifies variation of entropy between preparations; it does not imply that every measurement distinguishes those preparations. A nonzero Holevo quantity certifies that the density matrices vary, even when every spectral entropy is fixed. None of these statements alone determines whether the Hamiltonian was fixed, randomly drawn, or incompletely specified.

Exercises. For the qubit ensemble (11.10), calculate $C(n, \ell)$ directly and verify $C(n, \ell) = p(1 - p)(1 - 2^{1-n})(1 - 2^{1-\ell})$. Differentiate at $n = \ell = 1$ and recover its entropy variance. Next use (10.10) to calculate $C(n, \ell)$ for the two four-dimensional ensembles. Their $\mathbb{E}M_n$ agree at every order, but one covariance is zero and the other generally is not. Finally compare an equally weighted ensemble of $|0\rangle$

Table 8. Three nonnegative quantities and their distinct meanings. All statements refer to normalized states in a common finite dimension.

Quantity	Vanishing condition	Information retained
G_n for fixed $n > 1$	M_n is almost surely constant	Variation of that particular moment
$\text{Var}(S)$	S is almost surely constant	Entropy fluctuations and the slope of G_n at one
$S(\mathbb{E}\rho) - \mathbb{E}S(\rho)$	ρ is almost surely constant	Variation of full density matrices, including eigenvectors

and $|1\rangle$ with the deterministic state $I_2/2$. Both averaged matrices are $I_2/2$, while the mean entropies are zero and $\log 2$. This last example makes clear why specifying an averaged matrix also fails to specify an ensemble.

12 Comparing 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.

12.1 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 [4, 27, 74]. When accompanied by a code-subspace description and suitable reconstruction assumptions, the calculation can support conclusions about recoverable interior observables. Ref. [76] 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 13 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 $\mathbb{E}\text{Tr}\rho_R^n$ 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 10.3 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.

12.2 The objects supplied by different approaches

12.2.1 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 [16]. 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 13 makes that distinction explicit.

12.2.2 Islands and replica wormholes

An island calculation is a prescription for entropies and related observables within a gravitational setup. In the model of ref. [76], the replica sum can be compared with a random-state spectral problem. In JT gravity, the matrix-integral description computes particular ensemble quantities [79]. 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 [60]. It cannot be settled by a von Neumann annealed–quenched gap under the normalized definitions of section 11, because that gap vanishes.

12.2.3 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 [63]. Explicit smooth supergravity solutions represent important families of microstates in particular charge sectors; identifying them with every black-hole microstate, or extrapolating their construction to a generic nonextremal evaporating hole, requires further work [14, 62]. 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.

12.2.4 State-dependent interior operators

State-dependent constructions concern the map between boundary observables and interior operators in a specified class of states [73]. To compare them through outgoing radiation, one must identify a boundary preparation and an observable whose prediction differs. This is a question about which algebra is represented, on which states, and with what accuracy. It does not prescribe a unique probability distribution of outgoing states. 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.

12.2.5 Geometric descriptions of entanglement

The ER=EPR proposal relates entanglement to geometric connections [57]. By itself it does not specify a normalized probability law for Hawking-radiation density matrices. Proposals relating complexity to geometric growth concern additional observables and a definition of circuit cost [87]. They do not specify an entanglement spectrum from an entropy alone. A spectral or recovery comparison therefore

Table 9. Specified comparisons and the information they add. The first two 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
Gravitational entropy fluctuations	A specified mean Page curve	Connected normalized replicas give model-dependent fluctuations; ref. [15] 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

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.

12.3 A comparison table with explicit status

Table 9 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 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.

12.4 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

$$\begin{aligned} \text{Tr}(\bar{\rho}O^2) - [\text{Tr}(\bar{\rho}O)]^2 &= \mathbb{E}_\alpha [\text{Tr}(\rho_\alpha O^2) - \text{Tr}(\rho_\alpha O)^2] \\ &\quad + \text{Var}_\alpha [\text{Tr}(\rho_\alpha O)], \quad \bar{\rho} = \mathbb{E}_\alpha \rho_\alpha. \end{aligned} \quad (12.1)$$

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.

12.5 How to formulate a comparison that can be calculated

A useful comparison begins with a common preparation, not a list of names for research programs. Suppose two models are claimed to describe the same experiment. Specify a radiation algebra $\mathcal{A}_R$, an initial-state space A , conserved charges, and a coupling to the outgoing modes. If an average is involved, specify the probability measure and which quantities are held fixed. One may then ask whether the two models agree on

$$\mathbb{E} \text{tr} \rho_R^n, \quad \text{Cov}(\text{tr} \rho_R^n, \text{tr} \rho_R^\ell), \quad \text{tr}(\rho_R O), \quad (\text{id}_Q \otimes \mathcal{N})(\rho_{QA}). \quad (12.2)$$

These are, respectively, averaged spectral data, spectral correlations, operator data, and reference-assisted channel data. There is no universal ordering of their practical discriminatory power. Two models may agree on every spectral quantity and differ in an operator matrix element. They may agree on the full state for one input and differ on another input. They may instead disagree already in the leading entropy because they prepare different energy sectors.

For an example of the required precision, consider a proposed difference δV between two purity variances of order k^{-4} . An entropy calculation accurate only up to order one cannot resolve that difference. Even a moment approximation with absolute error $o(k^{-1})$ may be insufficient. If the comparison uses approximate state designs, the bound (10.19) requires control of the relevant two-copy and four-copy errors relative to $|\delta V|$. If it instead uses a gravitational saddle expansion, every contribution at that order, including normalization and the subtraction of disconnected terms, must be included. The name of the approximation does not supply its error estimate.

A further distinction concerns a family of states and a family of Hamiltonians. A single Hamiltonian admits an ensemble of microcanonical states, a thermal density matrix, different pure inputs, and time-dependent preparations. Each can lead to different spectral fluctuations. Conversely, an ensemble of Hamiltonians can have observables whose fluctuations vanish in a specified limit. The physical interpretation of a covariance therefore comes from the definition of the ensemble and the factorization properties of the theory, not from the covariance's sign alone.

12.6 Reconstruction, decoding, and geometric descriptions

The information problem contains several related tasks. An entropy calculation determines a scalar property of a reduced state. Entanglement wedge reconstruction concerns the representation of a bulk algebra in boundary degrees of freedom. A decoding problem asks for a channel that recovers a specified input with a stated fidelity. An implementation problem asks whether that decoder can be carried out with accessible resources. These tasks can support one another, but the necessary assumptions should remain visible when passing between them.

In particular, a mathematical decoder need not be computationally simple. The Hayden–Preskill setup specifies an input, its reference, the prior entanglement of the black hole, and access to early and late radiation [41]. The resulting recovery question contains more data than a pure-input Page curve. Likewise, the Petz reconstruction analysis of Penington et al. [76] does not rest on the entropy graph alone. The elementary counterexample in the next section therefore identifies what the graph omits; it does not contradict a calculation that supplies these additional channel or code-subspace data.

Geometric traversability is a separate issue. A traversable-wormhole protocol specifies a coupling that changes the physical evolution; Gao, Jafferis, and Wall provide a concrete example[30]. It is not a general consequence that implementing any decoder makes an Einstein–Rosen bridge traversable. Similarly, a conjectural relation between complexity and geometric growth does not by itself fix a universal recovery time. A comparison must state the protocol, the allowed interactions, and the notion of recovery before these times can be equated.

These distinctions leave substantial room for useful calculations. The normalized fluctuation calculation of Bousso and Miyaji is an existing example where a gravitational ensemble supplies information beyond the mean Page curve[15]. Explicit microstate emission models can be compared with statistical benchmarks when their state sampling and radiation coupling are specified. A boundary construction can be tested on a common code subspace through reference correlations. The tutorial’s finite examples provide checks on such calculations; they are not replacements for carrying them out in gravity.

A comparison exercise. Suppose two calculations report the same $\mathbb{E} \text{tr} \rho_R^n$ for every integer $n \geq 1$, in the same finite radiation dimension. List what follows and what does not. The averaged spectral measures and mean entropies agree. Purity variances need not agree, by Proposition 10.1. Matrix elements in a fixed basis need not agree, by (10.4). Finally, even equality of all radiation matrix elements for one pure input does not determine a channel on other inputs. The next section derives that last statement and explains the positive role of full Choi tomography.

13 From Equilibrium States to Evaporation Channels

The island calculations recover a turnover of the fine-grained radiation entropy in controlled semiclassical settings [4, 5, 74, 76]. 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.

13.1 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

$$S(R_k) = -\text{tr}(\rho_{R_k} \ln \rho_{R_k}), \quad S(R_k) = S(B_k) \leq \min\{\ln d_{R_k}, \ln d_{B_k}\}. \quad (13.1)$$

Approach to this bound is a typicality statement, as in the random-state calculation of Page [71]. 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 13.1 (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

$$\mathcal{V}(X) = VXV^\dagger, \quad \mathcal{C}(X) = \text{tr}(X) |\Omega\rangle\langle\Omega|. \quad (13.2)$$

Both maps are channels. For every k ,

$$\mathcal{V}_k(\psi_0) = \mathcal{C}_k(\psi_0) = \text{tr}_{b_{k+1}\cdots b_N} |\Omega\rangle\langle\Omega|, \quad (13.3)$$

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 (13.3) 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.

13.2 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 $q + j$:

$$|\Omega\rangle = \bigotimes_{j=1}^q \frac{|0\rangle_j |0\rangle_{q+j} + |1\rangle_j |1\rangle_{q+j}}{\sqrt{2}}. \quad (13.4)$$

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

$$S(R_k) = \min(k, 2q - k) \ln 2, \quad 0 \leq k \leq 2q. \quad (13.5)$$

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

$$W|\psi\rangle_A = |\Omega\rangle_{R_N} \otimes |\psi\rangle_E, \quad \mathcal{C}(X) = \text{tr}_E(WXW^\dagger), \quad (13.6)$$

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 (13.6) also identifies why this construction is not a model of complete black-hole evaporation with a proven absence of remnants or hidden sectors.

13.3 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 [41, 81]. Let

$$|\Phi_D\rangle_{QA} = \frac{1}{\sqrt{D}} \sum_{i=1}^D |i\rangle_Q |i\rangle_A, \quad |\Psi_V\rangle = (I_Q \otimes V)|\Phi_D\rangle. \quad (13.7)$$

The two final joint states are

$$\rho_{QR_N}^{\mathcal{V}} = |\Psi_V\rangle\langle\Psi_V|, \quad \rho_{QR_N}^{\mathcal{C}} = \frac{I_Q}{D} \otimes |\Omega\rangle\langle\Omega|. \quad (13.8)$$

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

$$I(Q : R_N)_{\mathcal{V}} = 2 \ln D, \quad I(Q : R_N)_{\mathcal{C}} = 0. \quad (13.9)$$

This is a different input preparation from the pure-state experiment in equation (13.3). 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

$$F_V(\mathcal{N}) = \langle\Psi_V|(\text{id}_Q \otimes \mathcal{N})(|\Phi_D\rangle\langle\Phi_D|)|\Psi_V\rangle. \quad (13.10)$$

For the two channels,

$$F_V(\mathcal{V}) = 1, \quad F_V(\mathcal{C}) = \frac{1}{D^2}. \quad (13.11)$$

Indeed, substituting the second state in equation (13.8) 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 [81]. 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

$$J_{\mathcal{N}} = (\text{id}_Q \otimes \mathcal{N})(|\Phi_D\rangle\langle\Phi_D|) = \frac{1}{D} \sum_{i,j} |i\rangle\langle j|_Q \otimes \mathcal{N}(|i\rangle\langle j|). \quad (13.12)$$

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

$$\mathcal{N}(\omega) = D \operatorname{Tr}_Q[(\omega^{\mathsf{T}} \otimes I_R) J_{\mathcal{N}}], \quad (13.13)$$

where the transpose is taken in the basis defining $|\Phi_D\rangle$. This is the standard channel-state correspondence [22]; 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 $J_{\mathcal{N}}$.

13.4 Consistency of successive radiation states

Use the finite-dimensional notation of Section 13. A sequential isometric model has maps

$$\begin{aligned} U_k : A &\longrightarrow B_k \otimes R_k, & W_{k+1} : B_k &\longrightarrow B_{k+1} \otimes b_{k+1}, \\ U_{k+1} &= (W_{k+1} \otimes I_{R_k}) U_k, \end{aligned} \quad (13.14)$$

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

$$\operatorname{tr}_{b_{k+1}} \mathcal{N}_{k+1}(X) = \mathcal{N}_k(X) \quad \text{for every input operator } X. \quad (13.15)$$

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 (13.15). 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 13.1 do satisfy this consistency condition because they release subsystems of a single prepared output.

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

$$S(R_{k+1}) - S(R_k) = S(b_{k+1}|R_k), \quad |S(R_{k+1}) - S(R_k)| \leq \ln \dim b_{k+1}. \quad (13.16)$$

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.

13.5 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

$$\mathcal{H}_{E_{\text{tot}}} = \bigoplus_E \mathcal{H}_B(E) \otimes \mathcal{H}_R(E_{\text{tot}} - E). \quad (13.17)$$

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,

$$\rho_B = \bigoplus_E p_E \rho_{B,E}, \quad S(\rho_B) = - \sum_E p_E \ln p_E + \sum_E p_E S(\rho_{B,E}). \quad (13.18)$$

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

$$|\psi_p\rangle = \sqrt{p} |1\rangle_B |0\rangle_R + \sqrt{1-p} |0\rangle_B |1\rangle_R \quad (13.19)$$

has exact total energy ε , but

$$\text{Var}(E_B) = \varepsilon^2 p(1-p), \quad S(\rho_R) = -p \ln p - (1-p) \ln(1-p). \quad (13.20)$$

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 8.2, 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 [100]. These thermodynamic entropies count states; they are distinct from the fine-grained entropies in equation (13.18). 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.

13.6 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

$$\tau_{\text{round}} \sim \frac{\sigma_\Delta(t_*)}{|\dot{\Delta}(t_*)|}, \quad \Delta(t_*) = 0. \quad (13.21)$$

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 6 and 7 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.

13.7 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 13.3. Computational difficulty of a decoder is a further question beyond its mathematical existence [41].

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 [4, 5]. 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.

13.8 A two-qubit reference calculation

The general channel construction can be checked with two emitted qubits. Let the input be one qubit and define the isometry

$$V|0\rangle = |\Phi^+\rangle, \quad V|1\rangle = |\Phi^-\rangle, \quad |\Phi^\pm\rangle = \frac{|00\rangle \pm |11\rangle}{\sqrt{2}}. \quad (13.22)$$

The two Bell states are orthogonal, so $V^\dagger V = I_2$. Compare $\mathcal{V}(X) = V X V^\dagger$ with $\mathcal{C}(X) = \text{tr}(X) |\Phi^+\rangle\langle\Phi^+|$. For input $|0\rangle$, both produce the same complete radiation state. After the first emission the state is $I_2/2$, and after the second it is $|\Phi^+\rangle\langle\Phi^+|$. Their radiation entropies are therefore exactly 0, $\log 2$, 0 at the three cuts.

The channels differ for other inputs. For $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$,

$$V|+\rangle = |00\rangle, \quad \mathcal{C}(|+\rangle\langle+|) = |\Phi^+\rangle\langle\Phi^+|. \quad (13.23)$$

Measuring Z on the first emitted qubit gives expectation one for $\mathcal{V}$ and zero for $\mathcal{C}$. The common history of the chosen input did not determine even this simple response to a different preparation.

For a reference-entangled input $(|00\rangle_{QA} + |11\rangle_{QA})/\sqrt{2}$, the isometric output is

$$|\Psi_V\rangle_{QR} = \frac{|0\rangle_Q |\Phi^+\rangle_R + |1\rangle_Q |\Phi^-\rangle_R}{\sqrt{2}}. \quad (13.24)$$

Its entropies satisfy $S(Q) = S(R) = \log 2$ and $S(QR) = 0$, so $I(Q : R) = 2 \log 2$. The replacement channel gives $(I_Q/2) \otimes |\Phi^+\rangle\langle\Phi^+|$, with $S(QR) = S(Q) = \log 2$ and $S(R) = 0$, so $I(Q : R) = 0$. The larger radiation entropy in the first case is a consequence of successful reference entanglement, not a failure of information recovery. A purity criterion appropriate to an unentangled input must therefore not be imposed on this different experiment.

Table 10. Static data and additional requirements for a dynamical interpretation. The requirements are conditional on the intended observable; they are not claims that every evaporation model has the same fluctuation law.

Static calculation	What can be retained	Additional dynamical input
Haar/MP state statistics	A conditional state benchmark	A justified state law and consistent successive marginals
Energy-sector distribution	A law under its preparation assumptions	Emission dynamics and exchange or sampling assumptions
Entropy crossover width	Width in entropy units	Branch velocity, relaxation, and variation of the law during the crossing
Pure-input entropy trajectory	Entropies of that preparation	Channel action on an input space and reference correlations
A mathematical decoder	Recovery with its stated error	Access to radiation, allowed operations, and computational cost

For completeness, the Choi reconstruction formula can be checked at the level of matrix units. Substituting $J_{\mathcal{N}} = D^{-1} \sum_{ij} |i\rangle\langle j| \otimes \mathcal{N}(|i\rangle\langle j|)$ into (13.13) gives

$$D \operatorname{tr}_Q[(\omega^\top \otimes I) J_{\mathcal{N}}] = \sum_{ij} \omega_{ij} \mathcal{N}(|i\rangle\langle j|) = \mathcal{N}(\omega). \quad (13.25)$$

Thus complete tomography includes the off-diagonal input operators that a single pure-input history leaves untested. The use of a maximally entangled input also explains why the factor D and the transpose appear. Measuring only the eigenvalues or entropy of the Choi state would discard some of these matrix elements and would not in general reconstruct the channel.

14 Further Calculations and Open Questions

The preceding sections separate several calculations that can be pursued without assuming that all microscopic descriptions share one universal hierarchy. The most direct extensions begin with a normalized state model and then add one piece of information that its mean entropy omits. The examples also provide elementary consistency checks for more difficult gravitational computations.

14.1 Six concrete targets

1. Normalized spectral correlations in a common preparation. A useful extension of a one-point replica calculation is $C(n, \ell)$ for independently variable replica indices. The normalization checks $C(1, \ell) = C(n, 1) = 0$ must be satisfied, and the mixed derivative at one must equal the entropy variance. The purity calculation in Section 11.8 is a finite-dimensional reference. For a gravitational model, one should specify the energy window and separate fluctuations in the number of available states from fluctuations conditional on that number. Bousso and Miyaji’s PSSY calculation already

implements a nontrivial part of this program [15]; it is a starting point for comparisons, not an uncomputed prediction.

2. Finite-order Rényi averages with controlled continuation. The objects to compare are A_n and Q_n at $n > 1$, or the covariance that controls their near-one difference. Their normalized von Neumann limits agree. A calculation of a nonzero limit should therefore first check its normalization, its definition of entropy, and the order of its large-dimension and replica limits. A controlled result would state the approximation error on an interval near one and identify any nonuniform behavior. The qubit figure in Section 11 provides a direct check on the sign and slope conventions.

3. Conditional spectra in fluctuating energy sectors. Thermodynamics specifies sector weights only after the preparation is chosen. It does not specify whether sector contributions are visible as orthogonal blocks in the radiation algebra. A useful calculation would hold the energy law fixed while deriving the overlaps and the conditional state law. One can then compare the entropy of the resulting state with the average of conditional entropies, rather than identifying them by assumption. The diagonal and planar distinctions in Paper I [99] give a controlled example of this issue. Its extension to another gravity model requires its own spectral equation and state normalization.

4. Regular and critical reservoir limits. At a regular stable saddle, a finite reservoir gives the effective capacity $(C_B^{-1} + C_R^{-1})^{-1}$. Near a critical saddle the quadratic approximation can fail, and a non-Gaussian law must be derived before an entropy average is evaluated. Paper II [100] provides conditional finite-reservoir and quartic calculations. Extending them to evaporation requires a dynamical justification of local equilibration, the energy-sector law, and the branch variables. One should compute their covariance as well as their marginal variances: the difference of two correlated entropy branches can have a very different width from either branch separately.

5. Consistent evaporation channels. A time-indexed family of density matrices should be replaced or supplemented by maps satisfying (13.15). A tractable circuit with energy or charge conservation is one possible starting point. One can test its single-time moments, correlations between emissions, and recovery of an input entangled with a reference. The absence of an unobserved endpoint system is an additional physical assumption that must be established if final radiation purity is used as evidence of complete recovery. The Bell-pair construction demonstrates why that assumption cannot be read off an entropy graph.

6. A matched comparison of two gravitational descriptions. Choose an observable, a preparation, and an approximation accuracy before comparing two descriptions. The target may be a connected replica moment, an operator matrix element, or a recovery fidelity. The comparison becomes informative when both calculations use the same radiation algebra and conserved charges and have errors smaller than the claimed difference. If the descriptions are dual accounts of the same physics, agreement may be the appropriate result. If they are different models, a discrepancy must be traced to an identified assumption rather than assigned to the label of a research program.

14.2 How the checks fit together

For a spectral calculation, begin with positivity and normalization of the proposed state. Identify whether the calculated object is a moment of a normalized state, a ratio of averages, or an average of ratios. Then specify whether the result is exact at finite dimension or only at a fixed order of an asymptotic expansion. If the question concerns fluctuations, compute the relevant products of traces

and their subtractions at the same accuracy. If it concerns entropy, justify the replica derivative or the spectral integral, including the required limit interchange.

For a dynamical calculation, add the input space and sequential output maps. Check compatibility of their reduced states and the treatment of already emitted modes. State whether the test uses a fixed pure input, a distribution of inputs, or an input entangled with a reference. An entropy curve, an average fidelity, and full channel tomography are different data sets; choosing one should match the physical question. Finally, separate the existence of a recovery channel from the resources needed to implement it.

These checks do not reduce the importance of the Page curve. They explain how its information can be extended. A controlled entropy calculation, a normalized covariance, and a reference-assisted channel test each add a definite statement. Combining them in a specified gravitational model is a substantive research objective, with the finite examples serving as transparent tests of what each calculation has and has not determined.

Part V

Conclusion and Perspective

15 Conclusions

15.1 From entanglement to the replica spectrum

A pure state of a composite system can have mixed reduced states. Hawking radiation entangled with interior partners therefore need not violate unitarity while the black hole remains present. The difficulty arises when the semiclassical description is extrapolated through complete evaporation, leaving mixed radiation and no remaining purifier. The Bell-pair and partial-trace calculations isolate the quantum structure of that argument.

Ryu–Takayanagi surfaces, their covariant generalization, and quantum extremal surfaces relate entropy to bulk area and bulk entanglement [27, 43, 78]. In the evaporating models of Almheiri et al. [4, 5], Penington [74], competing entropy branches produce a Page transition. Replica wormholes give a path-integral derivation [1, 76]. These are results for specified gravitational preparations and observables.

Random-state models make a substantial part of the spectral calculation explicit. Exact normalized Haar moments distinguish the random trace from its mean. Their large-dimension expansion is organized by permutations, and its leading flat-weight sector by non-crossing partitions. The resolvent gives the Marchenko–Pastur distribution, whose entropy recovers the large-dimension limit of Page’s result [58, 71]. Free probability supplies the moment calculus; the gravitational weighted problem requires additional spectral input [21, 96]. A common planar organization does not require all preparations to have the same spectrum.

The finite-dimensional calculation also clarifies subleading terms. Absence of a crossing permutation does not imply absence of a normalized observable’s correction at the corresponding order: the normalization itself contributes. Adding a handle to a gravitational surface differs from evaluating a fluctuation determinant on a fixed surface. Neither operation can be identified without a derivation with the leading QES corrections studied by Akers and Penington [3].

15.2 Thermodynamic preparation and entropy averaging

For conditional states ρ_x with probabilities p_x , the average $\sum_x p_x S(\rho_x)$ need not equal $S(\sum_x p_x \rho_x)$. Orthogonal sectors contribute a classical mixing entropy; overlapping sectors require spectral overlap information. A diagonal approximation therefore needs a definition of the conditional states and of what happens to their labels when they are averaged. Paper I's comparison with planar spectral entropy is an application of this distinction [99].

Once an averaged-minimum model is specified, its smoothing can be computed. A Gaussian branch difference gives the familiar normal-density and normal-distribution functions. A stable canonical saddle can produce this law, but identifying its variance with a black-hole heat capacity uses a local thermodynamic expansion. If both entropy branches fluctuate, their covariance contributes to the variance of their difference.

At a quartic critical saddle the scale and profile change. A finite reservoir can compete with the quartic term and restore a quadratic contribution. Paper II makes these model regimes explicit [100], complementing fixed-energy entanglement calculations [66] and fixed-area holographic analysis [61]. Fixed total energy does not fix each subsystem energy. An ensemble name alone does not determine a radiation-entropy width without identifying the state and observable.

15.3 Spectra, ensembles, and information recovery

The final chapters give concrete examples of different information supplied by different observables. Two ensembles can have identical averaged spectral measures, all averaged moments, and mean entropy, while their entropy and purity fluctuations differ. A one-point spectral average records the frequency of individual eigenvalues, but not their grouping into realizations. Connected moments contain information about this grouping.

For normalized finite-dimensional states, annealed and quenched Rényi entropies have the same von Neumann limit. Their finite-order difference is still informative: close to replica number one, its leading coefficient is half the entropy variance. Entropy of the averaged density matrix is a third quantity. These distinct averaging operations answer distinct questions.

At the dynamical level, knowing the radiation state for one preparation need not identify a channel. Agreement on that input leaves the action on other inputs and coherent superpositions undetermined. Reference-system correlations reveal the difference. Conversely, a full normalized Choi state determines a finite-dimensional channel when its input and output spaces and basis convention are fixed. This specifies which stronger data settle the identification problem.

These constructions do not establish equality of predictions between two particular theories of gravity. They specify the comparisons needed for such a conclusion. Fuzzballs, islands, random dynamics, state-dependent reconstruction, and ER=EPR address different aspects of the information problem and are not automatically mutually exclusive. A quantitative comparison fixes a common preparation and observable and then tests whether the predictions differ within their stated approximations.

15.4 Further calculations

The remaining gravitational questions can now be formulated precisely. One can compare a diagonal conditional-entropy average with a planar spectral calculation for the same preparation. One can compute normalized connected replicas, including normalization corrections, to determine fluctuations in a specified ensemble. One can derive the joint fluctuations of competing generalized entropies instead of inferring them from one heat capacity. Finally, one can test whether a sequence of equilibrium

approximations comes from a consistent evaporation channel and which reference correlations that channel preserves.

The tutorial develops the calculations needed to formulate these questions and finite models in which several can be answered completely. Their gravitational applications require the corresponding preparation, spectral weights, and dynamics. Keeping these inputs explicit allows the Page curve to serve both as a quantitative entropy calculation and as a starting point for further tests of information recovery.

A Notation and Conventions

This appendix collects the notation and conventions used throughout the paper. We work exclusively in natural units,

$$\hbar = c = k_B = 1, \quad (\text{A.1})$$

so that energy, mass, temperature, and inverse length all share the same dimension, and the Bekenstein–Hawking entropy is the dimensionless ratio $A/4G_N$. Entropies are measured in *nats* (base- e), not bits (base-2); the conversion is $1 \text{ bit} = \ln 2 \text{ nats}$. All logarithms are natural unless stated otherwise.

A.1 Black hole variables

A.2 Hilbert space dimensions and the ratio γ

The Haar benchmark uses finite-dimensional subsystems. In gravity these dimensions represent specified effective state counts, for example in a chosen energy window; the full JT Hilbert space is not finite-dimensional.

A.3 Combinatorial notation

A.4 Random matrix notation

A.5 Sign conventions and metric signature

We adopt the “mostly plus” metric signature $(-, +, +, +)$. The Riemann tensor is defined by

$$R^\alpha{}_{\beta\gamma\delta} V^\beta = (\nabla_\gamma \nabla_\delta - \nabla_\delta \nabla_\gamma) V^\alpha, \quad (\text{A.2})$$

the Ricci tensor by $R_{\mu\nu} = R^\alpha{}_{\mu\alpha\nu}$, and the Einstein equations read

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G_N T_{\mu\nu}. \quad (\text{A.3})$$

In two-dimensional JT gravity these reduce to the dilaton equation of motion reviewed in [Appendix C](#).

B The Non-Crossing Partition Lattice

Non-crossing partitions, introduced by Kreweras [\[48\]](#) and developed into a central pillar of free probability by Speicher [\[85, 86\]](#), organize the planar flat-weight replica contractions of [Section 6](#). This appendix provides a self-contained development.

Symbol	Definition
S_{BH}	Bekenstein–Hawking entropy of the black hole, $S_{\text{BH}} = A/(4G_N)$. Treated as a function of time during evaporation.
S_0	Extremal (topological) entropy: the zero-temperature piece of S_{BH} . In JT gravity, $S_0 = \phi_0/(4G_N)$ where ϕ_0 is the constant (topological) piece of the dilaton; the boundary value ϕ_b instead sets the Schwarzian coupling C (Appendix C).
S_{rad}	Coarse-grained (thermodynamic) entropy of the Hawking radiation.
Δ	Entropy-branch difference: $\Delta \equiv S_{\text{rad}} - S_{\text{BH}}$ in the illustrative evaporation model. Other preparations define their branch entropies locally.
t_P	Page time: the instant at which $S_{\text{rad}} = S_{\text{BH}}$, i.e. $\Delta = 0$.
β	Inverse Hawking temperature, $\beta = 1/T_H$.
C_B (or C_V)	Geometric heat capacity of the black hole at constant volume (or at fixed charge, as appropriate). For a Schwarzschild black hole in four dimensions, $C_V = -8\pi G_N M^2$.
σ	Standard deviation of the explicitly specified entropy variable or difference of branch entropies. At a stable regular canonical saddle, linearization gives $\text{Var } S_{\text{BH}} \simeq C_B$. For a finite reservoir, the corresponding energy variance is $T^2 C_{\text{eff}}$, where $C_{\text{eff}} = (C_B^{-1} + C_R^{-1})^{-1} > 0$. Correlations enter the variance of a difference of branch entropies.

Table 11. Black-hole thermodynamic variables. Geometric heat capacities, canonical energy variances, and entropy-coordinate variances agree only under the ensemble and linearization assumptions stated in the text.

B.1 Definition and first examples

Definition. A *partition* of the set $[n] = \{1, 2, \dots, n\}$ is a collection $\pi = \{B_1, B_2, \dots, B_j\}$ of non-empty, pairwise disjoint subsets (*blocks*) whose union is $[n]$. The partition π is *non-crossing* if, whenever $a < b < c < d$ and $a, c \in B_i$ and $b, d \in B_k$, we have $i = k$. Equivalently, when the elements of $[n]$ are placed in order around a circle and the elements of each block are connected by convex hulls, no two blocks cross.

We denote the set of all non-crossing partitions of $[n]$ by $\text{NC}(n)$. Its cardinality is the Catalan number C_n . For small n the enumeration proceeds as follows.

$n = 1$: There is a single partition $\{1\}$; it is trivially non-crossing. $|\text{NC}(1)| = C_1 = 1$.

Symbol	Definition
$\mathcal{H}_{\text{rad}}$	Hilbert space of the radiation subsystem; dimension k .
$\mathcal{H}_{\text{BH}}$	Hilbert space of the finite black-hole factor; dimension L .
k	$\dim \mathcal{H}_{\text{rad}}$. Fixed in each static calculation; written $k(t)$ when a dynamical model specifies its time dependence.
L	$\dim \mathcal{H}_{\text{BH}}$. The flat-weight scan can vary k at fixed L ; an evaporation model may instead have $L(t)$. The convention $L = e^{S_0}$ is a simplified state count, not a general JT identity.
γ	Ratio $\gamma \equiv k/L$, as in Section 6. Free cumulants have a replica-order subscript, κ_m ; surface gravity is denoted κ in the main text and κ_s in this appendix. The Page transition occurs at $\gamma = 1$.
η	Symmetric ratio $\eta = \min(k, L)/\max(k, L) \leq 1$, used for the smaller-side spectral density and entropy.
D	In Haar moment formulas, $D = kL$ and $(D)_n$ is its rising factorial. In Section 13, D instead denotes the channel input dimension. Thermodynamic N is defined locally.

Table 12. Hilbert space dimension variables.

Symbol	Definition
C_n	Catalan number: $C_n = \frac{1}{n+1} \binom{2n}{n}$. Counts the non-crossing partitions of $\{1, \dots, n\}$, i.e. $C_n = \text{NC}(n) $. The first values are $C_0 = 1$, $C_1 = 1$, $C_2 = 2$, $C_3 = 5$, $C_4 = 14$, $C_5 = 42$.
$N(n, j)$	Narayana number: $N(n, j) = \frac{1}{n} \binom{n}{j} \binom{n}{j-1}$ for $1 \leq j \leq n$. Counts non-crossing partitions of $\{1, \dots, n\}$ with exactly j blocks. Satisfies $\sum_{j=1}^n N(n, j) = C_n$.
c_n	Cyclic permutation $(1\ 2 \cdots n)$ imposed by an n th trace moment.
$\text{NC}(n)$	Lattice of non-crossing partitions of $\{1, \dots, n\}$.
$\text{Kr}(\pi)$	Kreweras complement of $\pi \in \text{NC}(n)$.
μ_{NC}	Möbius function on the lattice $\text{NC}(n)$.

Table 13. Combinatorial notation.

$n = 2$: The two partitions of $\{1, 2\}$ are $\{\{1, 2\}\}$ and $\{\{1\}, \{2\}\}$. Both are non-crossing. $|\text{NC}(2)| = C_2 = 2$.

Symbol	Definition
W	Wishart matrix: $W = XX^\dagger$, where X is a $k \times L$ matrix with i.i.d. mean-zero entries of variance $1/L$ (Eq. (E.1)). This is $1/L$ times the unit-variance Wishart matrix denoted W in Section 5.
$\rho_{\text{MP}}(\lambda)$	Marchenko–Pastur distribution with ratio parameter $\gamma = k/L$.
$\lambda_\pm$	Marchenko–Pastur edges: $\lambda_\pm = (1 \pm \sqrt{\gamma})^2$.
$G(z)$	Stieltjes transform of a spectral distribution: $G(z) = \int \frac{\rho(\lambda)d\lambda}{z-\lambda}$.
$R(z)$	R -transform: $R(z) = G^{(-1)}(z) - 1/z$, where $G^{(-1)}$ is the functional inverse of G .
$\Phi(z), \phi(z)$	Standard normal CDF and PDF, respectively.

Table 14. Random matrix and probability notation.

$n = 3$: Of the five partitions of $\{1, 2, 3\}$, all are non-crossing (there are not enough elements to create a crossing). $|\text{NC}(3)| = C_3 = 5$. The partitions, grouped by number of blocks, are:

- 1 block: $\{1, 2, 3\}$.
- 2 blocks: $\{1, 2\}\{3\}$, $\{1, 3\}\{2\}$, $\{1\}\{2, 3\}$.
- 3 blocks: $\{1\}\{2\}\{3\}$.

$n = 4$: Of the 15 set partitions of $\{1, 2, 3, 4\}$, exactly 14 are non-crossing. The single crossing partition is $\{\{1, 3\}, \{2, 4\}\}$, whose blocks interleave. $|\text{NC}(4)| = C_4 = 14$.

$n = 5$: $|\text{NC}(5)| = C_5 = 42$.

B.2 The Narayana refinement

The Narayana number $N(n, j)$ counts the non-crossing partitions of $[n]$ with exactly j blocks:

$$N(n, j) = \frac{1}{n} \binom{n}{j} \binom{n}{j-1}, \quad 1 \leq j \leq n. \quad (\text{B.1})$$

These satisfy the Catalan sum rule

$$\sum_{j=1}^n N(n, j) = C_n = \frac{1}{n+1} \binom{2n}{n}, \quad (\text{B.2})$$

and the *Narayana duality*

$$N(n, j) = N(n, n+1-j), \quad (\text{B.3})$$

reflecting the bijection that sends each non-crossing partition to its Kreweras complement (see below).

$n \setminus j$	1	2	3	4	5	6	7	C_n
1	1							1
2	1	1						2
3	1	3	1					5
4	1	6	6	1				14
5	1	10	20	10	1			42
6	1	15	50	50	15	1		132
7	1	21	105	175	105	21	1	429

Table 15. Narayana numbers $N(n, j)$. The manifest duality $N(n, j) = N(n, n + 1 - j)$ is visible as the left–right symmetry of each row. The row sums give Catalan numbers C_n .

B.3 Lattice structure

The set $\text{NC}(n)$ carries a natural partial order: $\pi \leq \sigma$ if every block of π is contained in a block of σ (i.e. π is a *refinement* of σ). Under this order, $\text{NC}(n)$ is a lattice: any two elements have a unique meet (greatest lower bound) and join (least upper bound). The minimum element $\hat{0}$ is the discrete partition $\{\{1\}, \{2\}, \dots, \{n\}\}$ and the maximum element $\hat{1}$ is the single-block partition $\{\{1, 2, \dots, n\}\}$.

This lattice is *self-dual* and *graded*: the rank function $\text{rk}(\pi) = n - |\pi|$ (where $|\pi|$ is the number of blocks) satisfies $\text{rk}(\hat{0}) = 0$ and $\text{rk}(\hat{1}) = n - 1$. The number of elements at rank r is $N(n, n - r)$.

B.3.1 The Möbius function

The Möbius function of the non-crossing partition lattice takes a simple closed form. For the interval $[\hat{0}, \hat{1}] = \text{NC}(n)$, one has [48]

$$\mu_{\text{NC}}(\hat{0}, \hat{1}) = (-1)^{n-1} C_{n-1}. \quad (\text{B.4})$$

More generally, for $\pi \leq \sigma$ in $\text{NC}(n)$ with the interval $[\pi, \sigma]$ decomposing as a product of intervals of sizes $n_1, \dots, n_r$,

$$\mu_{\text{NC}}(\pi, \sigma) = \prod_{i=1}^r (-1)^{n_i-1} C_{n_i-1}. \quad (\text{B.5})$$

The n_i are the sizes of the non-crossing factors in the interval isomorphism; they are not in general the sizes of the original blocks of π . For example,

$$\mu_{\text{NC}}(\pi, \hat{1}) = \prod_{B \in \text{Kr}(\pi)} (-1)^{|B|-1} C_{|B|-1}. \quad (\text{B.6})$$

For $\pi = \{\{1, 2\}, \{3, 4\}\}$ the upper interval has two elements, so its Möbius value is -1 . Multiplying the factors associated with the two original two-element blocks would instead give $+1$ and would be wrong. This factorization is essential to the moment–cumulant formulas of free probability: the free cumulants κ_n are related to moments m_n by a Möbius inversion on $\text{NC}(n)$.

B.3.2 The Kreweras complement

Let $\pi \in \text{NC}(n)$. Interleave a second copy of $[n]$, labelled $\bar{1}, \bar{2}, \dots, \bar{n}$, between the original elements on the circle: $1, \bar{1}, 2, \bar{2}, \dots, n, \bar{n}$. The *Kreweras complement* $\text{Kr}(\pi)$ is the largest non-crossing partition of $\{\bar{1}, \dots, \bar{n}\}$ such that $\pi \cup \text{Kr}(\pi)$ is non-crossing on the $2n$ points.

The map Kr is an anti-automorphism of the lattice $\text{NC}(n)$: it reverses the partial order, $\pi \leq \sigma \implies \text{Kr}(\pi) \geq \text{Kr}(\sigma)$, and satisfies $|\text{Kr}(\pi)| + |\pi| = n + 1$. In particular, a partition with j blocks maps to one with $n + 1 - j$ blocks, which is the origin of the Narayana duality (B.3).

B.4 Circle diagrams for $\text{NC}(4)$

Figure 24 displays all 14 non-crossing partitions of $\{1, 2, 3, 4\}$ as circle diagrams, organized by the number of blocks j .

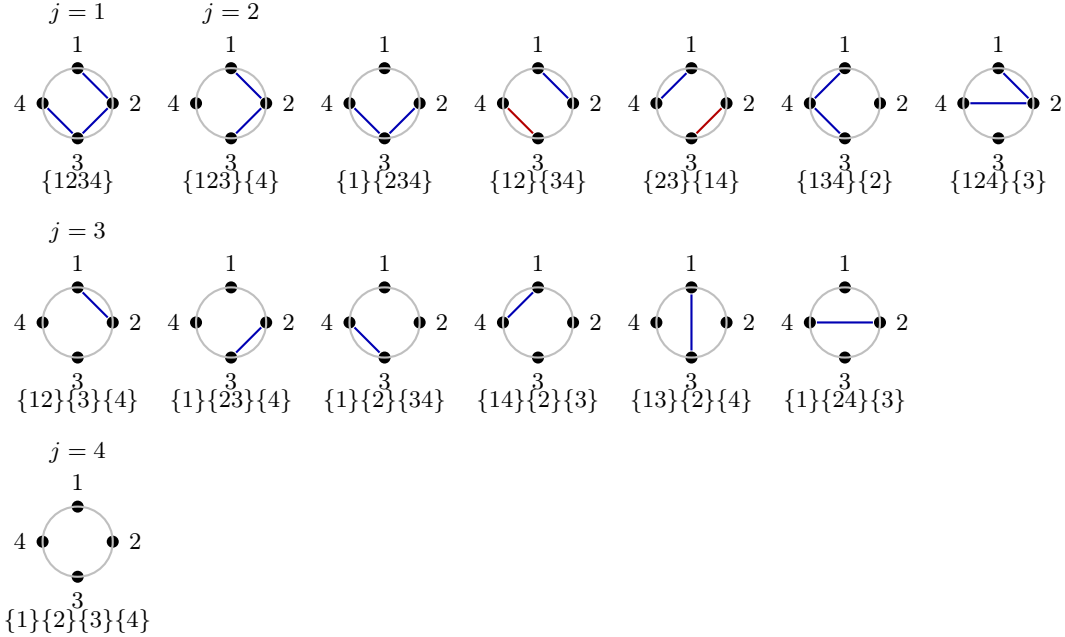


Figure 24. All 14 non-crossing partitions of $\{1, 2, 3, 4\}$, displayed as circle diagrams. The elements 1, 2, 3, 4 are placed clockwise starting from the top of each circle; elements within a block are connected by edges (chords or arcs). The partitions are organized by number of blocks j : one partition with $j = 1$ (Narayana number $N(4, 1) = 1$), six with $j = 2$ ($N(4, 2) = 6$), six with $j = 3$ ($N(4, 3) = 6$), and one with $j = 4$ ($N(4, 4) = 1$). The single crossing partition $\{1, 3\}\{2, 4\}$, whose chords would intersect, is absent.

B.5 The lattice diagram for $\text{NC}(3)$

The Hasse diagram of $\text{NC}(3)$ provides a compact visualization of the partial order. There are $C_3 = 5$ elements, arranged in three ranks.

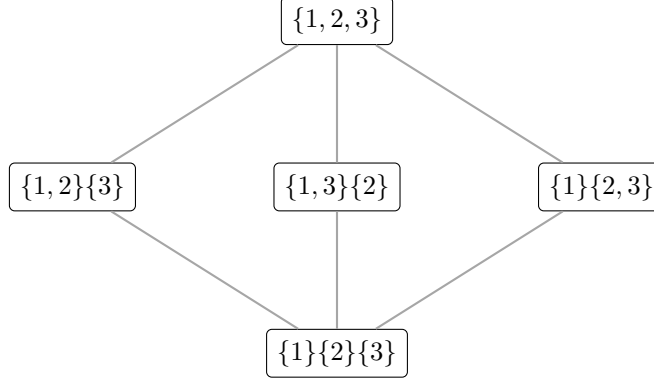


Figure 25. Hasse diagram of the non-crossing partition lattice $\text{NC}(3)$. The minimum element $\hat{0} = \{1\}\{2\}\{3\}$ is at the bottom; the maximum element $\hat{1} = \{1, 2, 3\}$ is at the top. The lattice is self-dual: the Kreweras complement exchanges $\hat{0} \leftrightarrow \hat{1}$ and maps rank-1 elements to rank-1 elements.

B.6 The role of non-crossing partitions in gravity

In the flat-weight replica calculation, the leading Wick contractions are labelled by non-crossing permutations, equivalently non-crossing partitions [85, 95]. Non-crossing pairings use a doubled set of points under the usual bijection; they are not all pair partitions of the original n points. The *crossing* partition $\{1, 3\}\{2, 4\}$ defines a genus-one Wick ribbon graph, with relative dimension suppression L^{-2} at fixed ratio. This is not by itself a count of continuum JT handles. The Marchenko–Pastur distribution emerges precisely through its n -th moment

$$m_n = \sum_{\pi \in \text{NC}(n)} \gamma^{n-|\pi|} \quad (\text{B.7})$$

summed over all non-crossing partitions, where $|\pi|$ counts the number of blocks and $\gamma = k/L$ is the Hilbert-space ratio. This is the content of the free-probability moment–cumulant relation with free cumulants $\kappa_m = \gamma^{m-1}$: each block $B \in \pi$ contributes $\kappa_{|B|} = \gamma^{|B|-1}$, and the exponents add up to $\sum_{B \in \pi} (|B| - 1) = n - |\pi|$. For $n = 2$, for example, the two non-crossing partitions $\{12\}$ and $\{1\}\{2\}$ give $m_2 = \gamma + 1$, the second moment of the mean-one Marchenko–Pastur law.

C JT Gravity: A Lightning Review

Jackiw–Teitelboim (JT) gravity [47, 90] is a model of dilaton gravity in two spacetime dimensions that has become the “hydrogen atom” of quantum gravity: rich enough to exhibit a gravitational path integral with wormholes, ensemble averaging, and replica calculations, yet simple enough to be exactly solvable. This appendix provides the essential ingredients needed in the main text.

C.1 The action

The JT gravity action consists of a topological term and a dynamical dilaton piece:

$$I_{\text{JT}} = -\frac{S_0}{2\pi} \left[\frac{1}{2} \int_{\mathcal{M}} \sqrt{g} R + \int_{\partial\mathcal{M}} \sqrt{h} K \right] - \left[\frac{1}{2} \int_{\mathcal{M}} \sqrt{g} \phi (R + 2) + \int_{\partial\mathcal{M}} \sqrt{h} \phi_b (K - 1) \right]. \quad (\text{C.1})$$

The first bracket is the two-dimensional Einstein–Hilbert action (a topological invariant by the Gauss–Bonnet theorem), weighted by S_0 . On a surface of genus g with n boundaries, it contributes $-S_0(2 - 2g - n)/(2\pi) \times 2\pi = S_0(2g + n - 2)$ — more precisely, $e^{S_0(1-2g)}$ per connected surface with one boundary. The second bracket is the dynamical content: the dilaton field ϕ enforces $R = -2$ (constant negative curvature, i.e. locally AdS_2) in the bulk, and the boundary condition $\phi_b \rightarrow \infty$ selects asymptotically AdS_2 geometries.

C.2 The Schwarzian boundary mode

After fixing $R = -2$, the bulk geometry is rigid (locally a quotient of $\mathbb{H}^2$), and the remaining degrees of freedom are *boundary gravitons* — wiggles of the boundary curve in the ambient hyperbolic disk. These are described by a single function $f(\tau)$ parameterizing the boundary trajectory, with the effective action [56, 88]

$$I_{\text{Sch}}[f] = -C \int_0^\beta d\tau \{f, \tau\}, \quad \{f, \tau\} \equiv \frac{f'''}{f'} - \frac{3}{2} \left(\frac{f''}{f'} \right)^2, \quad (\text{C.2})$$

Here $f(\tau) = \tan[\pi u(\tau)/\beta]$ is a projective coordinate on the thermal circle, with $u(\tau + \beta) = u(\tau) + \beta$. The symbol $\{f, \tau\}$ denotes the Schwarzian derivative and $C = \phi_r/(8\pi G_N)$ is the “Schwarzian coupling” (with ϕ_r the renormalized boundary value of the dilaton). The path integral over f can be performed exactly, either by localization [88] or by recognizing the Schwarzian theory as the boundary (edge-mode) theory of two-dimensional $\text{SL}(2, \mathbb{R})$ BF theory.

C.3 The disk partition function

The genus-zero, single-boundary contribution to the JT gravity partition function is the exact Schwarzian path integral. With Schwarzian coupling C , the disk-level density of states is

$$\rho_{\text{disk}}(E) = e^{S_0} \frac{C \sinh(2\pi\sqrt{2CE})}{2\pi^2}, \quad (\text{C.3})$$

a density per unit dE , in the measure convention of Saad et al. [79]. An overall C -dependent constant can alternatively be absorbed into the normalization of the measure; it does not affect a canonical probability at fixed C . This density of states grows exponentially, consistent with the Bekenstein–Hawking entropy. Its Laplace transform, $Z(\beta) = \int_0^\infty dE \rho(E) e^{-\beta E}$ with β the inverse temperature, is Gaussian after the substitution $E = u^2$: using $\int_0^\infty \sinh(a\sqrt{E}) e^{-\beta E} dE = \frac{\sqrt{\pi} a}{2\beta^{3/2}} e^{a^2/(4\beta)}$ with $a = 2\pi\sqrt{2C}$, one obtains the disk partition function

$$Z_{\text{disk}}(\beta) = e^{S_0} \frac{C^{3/2}}{\sqrt{2\pi}} \frac{e^{2\pi^2 C/\beta}}{\beta^{3/2}}. \quad (\text{C.4})$$

For the entropy coordinate $x = 2\pi\sqrt{2CE}$, $E = x^2/(8\pi^2 C)$ and $dE = x dx/(4\pi^2 C)$. Therefore the normalized canonical distribution in that coordinate is

$$p_\beta(x) = \frac{x \sinh x e^{-\beta x^2/(8\pi^2 C)}}{\int_0^\infty u \sinh u e^{-\beta u^2/(8\pi^2 C)} du}, \quad x \geq 0. \quad (\text{C.5})$$

The factor x is the Jacobian. It cannot be dropped when computing entropy fluctuations or the conditional average in Section 8.

The factor e^{S_0} comes from the topological term (genus $g = 0$, one boundary); the remaining factor is the exact one-loop partition function of the Schwarzian theory.

C.4 The matrix integral dual

Saad et al. [79] match the perturbative JT amplitudes, order by order in topology, to a double-scaled Hermitian matrix integral. Schematically,

$$\left\langle \prod_{i=1}^b Z(\beta_i) \right\rangle_{\text{JT}} \sim \left\langle \prod_{i=1}^b \text{Tr} e^{-\beta_i H} \right\rangle_{\text{matrix}}, \quad (\text{C.6})$$

where the equality is between the corresponding perturbative expansions. The one-boundary disk density does not uniquely specify the nonperturbative completion. Such a completion requires a definition of the matrix ensemble beyond its formal genus series.

In this formulation the path integral computes ensemble quantities rather than directly specifying one fixed Hamiltonian. Connected amplitudes encode connected ensemble correlations. A surface with two boundaries can already be connected at genus zero, so connected fluctuations must not all be assigned the suppression of an additional handle. The effective counting parameter e^{S_0} organizes topology; it should not be read as the dimension of the full continuous-energy JT Hilbert space.

C.5 Why JT is the “hydrogen atom” of quantum gravity

The analogy is apt in several respects:

1. **Exact solvability.** Just as the hydrogen atom can be solved exactly in quantum mechanics, JT amplitudes can be computed recursively to arbitrary perturbative order through the matrix-integral construction. A specified matrix ensemble can additionally supply nonperturbative data.
2. **Near-extremal applications.** For black holes with an appropriate near-AdS₂ throat and a controlled dimensional reduction, JT gravity captures a leading low-energy sector. Extra light modes, charges, and boundary conditions can require additional fields and change its range of validity [56].
3. **A laboratory for ideas.** Many central calculations in the Page curve program — replica wormholes, islands, the QES formula, the genus expansion of entropy — can be exhibited explicitly in JT gravity before attempting the (much harder) generalization to higher dimensions.
4. **Limitations are instructive.** JT gravity is two-dimensional, has no propagating gravitons, and its partition function is an ensemble average rather than a single-system quantity. Understanding precisely *where* JT gravity fails to capture higher-dimensional physics is as important as understanding where it succeeds.

D Gaussian Broadening Integrals

The conditional entropy approximation of the main text averages the minimum $\min(S_{\text{rad}}, S_{\text{BH}})$ by its average over Gaussian fluctuations of the horizon area. The required integrals — the expectations $\mathbb{E}[\max(K - X, 0)]$ and $\mathbb{E}[\min(X, K)]$ for a normal random variable X — have simple closed forms, collected in this appendix.

D.1 The broadening integral

Let $X \sim \mathcal{N}(\mu, \sigma^2)$ with $\sigma > 0$, and let K be a fixed threshold. Then

$$\mathbb{E}[\max(K - X, 0)] = (K - \mu) \Phi(z) + \sigma \phi(z), \quad z \equiv \frac{K - \mu}{\sigma}, \quad (\text{D.1})$$

where Φ is the standard normal CDF and ϕ is the standard normal PDF.

Derivation. The expectation is

$$\begin{aligned} \mathbb{E}[\max(K - X, 0)] &= \int_{-\infty}^K (K - x) \frac{1}{\sigma} \phi\left(\frac{x - \mu}{\sigma}\right) dx \\ &= (K - \mu) \int_{-\infty}^K \frac{1}{\sigma} \phi\left(\frac{x - \mu}{\sigma}\right) dx + \int_{-\infty}^K (\mu - x) \frac{1}{\sigma} \phi\left(\frac{x - \mu}{\sigma}\right) dx \\ &= (K - \mu) \Phi\left(\frac{K - \mu}{\sigma}\right) + \sigma \int_{-\infty}^{(K - \mu)/\sigma} (-u) \phi(u) du \\ &= (K - \mu) \Phi\left(\frac{K - \mu}{\sigma}\right) + \sigma \phi\left(\frac{K - \mu}{\sigma}\right), \end{aligned} \quad (\text{D.2})$$

where the third line substitutes $u = (x - \mu)/\sigma$ and the final step uses $\phi'(u) = -u \phi(u)$, so that the integrand of the second term is a total derivative. The final line agrees with the boxed formula (D.1).

Because ϕ is even and $\Phi(-z) = 1 - \Phi(z)$, the same result can be written with the opposite sign convention $\bar{z} \equiv (\mu - K)/\sigma = -z$:

$$\mathbb{E}[\max(K - X, 0)] = (K - \mu) \Phi(-\bar{z}) + \sigma \phi(\bar{z}), \quad \bar{z} = \frac{\mu - K}{\sigma}, \quad (\text{D.3})$$

with $\bar{z} > 0$ when the mean lies above the threshold.

From the identity $\min(X, K) = K - \max(K - X, 0)$, one obtains the companion formula

$$\mathbb{E}[\min(X, K)] = K - (K - \mu) \Phi(z) - \sigma \phi(z), \quad z = \frac{K - \mu}{\sigma}. \quad (\text{D.4})$$

This is the form used for the conditional average of a sharp minimum.

D.2 The Page curve as a broadened minimum

For a conditional model with fixed radiation branch entropy $K = S_{\text{rad}}$, let $X = S_{\text{BH}}$ have mean μ and variance σ^2 . Define $\langle \Delta \rangle = K - \mu$. The sharp conditional average is then exactly, within the assumed Gaussian distribution,

$$S_{\text{diag}}^{\text{sharp}} := \mathbb{E} \min(S_{\text{rad}}, S_{\text{BH}}) = S_{\text{rad}} - \langle \Delta \rangle \Phi(\langle \Delta \rangle / \sigma) - \sigma \phi(\langle \Delta \rangle / \sigma). \quad (\text{D.5})$$

It tends to the minimum of the two means as $\sigma \rightarrow 0$, and at equal means has deficit $\sigma/\sqrt{2\pi}$. The symbol S_{diag} records that this is an average of conditional entropies. It is not, without an additional argument, the entropy of a quantum mixture of overlapping states.

The Gaussian distribution follows from a regular, locally stable saddle in the chosen ensemble. It does not require or establish statistically independent Planck-area cells. Canonically, $\text{Var } E = T^2 C_{\text{can}}$ is exact, whereas $\delta S_{\text{BH}} \simeq \delta E/T$ and $\text{Var } S_{\text{BH}} \simeq C_B$ use linearization and saddle equivalence of the heat capacities. The thermal modular variable $K_{\text{th}} = \beta E + \log Z$ instead satisfies $\text{Var } K_{\text{th}} = C_{\text{can}}$ exactly.

If two subsystems exchange energy at fixed total energy, a regular saddle gives $\text{Var } E_B \simeq T^2 (C_B^{-1} + C_R^{-1})^{-1}$. Both branch entropies then fluctuate. At equal temperatures, $\delta(S_B - S_R) \simeq 2\delta E_B/T$, so their difference has variance $4(C_B^{-1} + C_R^{-1})^{-1}$, rather than the variance of either branch alone. More generally the covariance term in $\text{Var}(S_B - S_R) = \text{Var } S_B + \text{Var } S_R - 2\text{Cov}(S_B, S_R)$ must be retained.

D.3 Non-Gaussian extensions

Small non-Gaussian corrections about a regular saddle may be organized by an Edgeworth expansion when the relevant cumulants possess a controlled scaling. For a standardized variable with skewness ζ_3 and excess kurtosis ζ_4 , its first terms are

$$f(x) \sim \phi(x) \left[1 + \frac{\zeta_3}{6} H_3(x) + \frac{\zeta_4}{24} H_4(x) + \frac{\zeta_3^2}{72} H_6(x) + \cdots \right], \quad (\text{D.6})$$

where the H_j are probabilists' Hermite polynomials. This is an asymptotic expansion in the region where the saddle approximation is controlled, not a globally positive probability density for arbitrary cumulants. The cumulants come from the specified probability law, not automatically from a permutation genus coefficient.

At a critical point where the quadratic saddle term vanishes, a quartic limiting density can be the leading law rather than a small Edgeworth correction. If the entropy coordinate satisfies $X - X_c = N^{3/4}u$ and the exponent tends to au^4 with $a > 0$, the width scales as $N^{3/4}$. The Gaussian formulas above must then be replaced by integrals against the normalized quartic density; Section 8 derives the corresponding thermodynamic assumptions.

E Random Matrix Theory: Marchenko–Pastur and Wishart Matrices

This appendix provides a self-contained derivation of the Marchenko–Pastur (MP) distribution from several complementary perspectives. The MP law is the spectral law underlying the flat-weight random-state benchmark.

E.1 Wishart matrices

Definition. Let X be a $k \times L$ random matrix whose entries X_{ij} are independent, identically distributed complex Gaussian variables with mean zero and variance $1/L$. Thus the matrix W below is $1/L$ times the unit-variance Wishart matrix in Section 5. Broader sample covariance ensembles have the same leading law under appropriate moment and independence hypotheses. The *Wishart matrix* (or *sample covariance matrix*) is

$$W = XX^\dagger. \quad (\text{E.1})$$

It is a $k \times k$ positive semidefinite Hermitian matrix. Almost surely, when $k \leq L$ it has k positive eigenvalues; when $k > L$ it has L positive eigenvalues and $k - L$ zero eigenvalues.

The physical context is as follows. A Haar-random pure state $|\Psi\rangle \in \mathcal{H}_{\text{rad}} \otimes \mathcal{H}_{\text{BH}}$ has a reduced density matrix $\rho_{\text{rad}} = \text{tr}_{\text{BH}} |\Psi\rangle\langle\Psi|$ whose eigenvalues are of order $1/k$ (they sum to one). After rescaling by k , the eigenvalue distribution of $k\rho_{\text{rad}}$ — a mean-one spectrum, equal in law to that of $kW/\text{Tr } W$ — converges, in the limit $k, L \rightarrow \infty$ with $\gamma = k/L$ fixed, to the MP law [58, 71].

E.2 The Marchenko–Pastur distribution

The Marchenko–Pastur distribution with ratio parameter $\gamma > 0$ is

$$\rho_{\text{MP}}(\lambda) = \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{2\pi\gamma\lambda} \mathbf{1}_{[\lambda_-, \lambda_+]}(\lambda) + \max\left(1 - \frac{1}{\gamma}, 0\right) \delta(\lambda), \quad (\text{E.2})$$

where the edges are

$$\lambda_{\pm} = (1 \pm \sqrt{\gamma})^2. \quad (\text{E.3})$$

For $\gamma < 1$ (i.e. $k < L$, on the smaller-radiation side) the distribution is supported on $[\lambda_-, \lambda_+]$ with $\lambda_- = (1 - \sqrt{\gamma})^2 > 0$, and there is no atom at zero. For $\gamma > 1$ (the larger-radiation side) the continuous part is still supported on $[\lambda_-, \lambda_+]$ with $\lambda_- > 0$ — a gap $(0, \lambda_-)$ opens between zero and the continuous support — and in addition a point mass of weight $1 - 1/\gamma$ appears at $\lambda = 0$, reflecting the $k - L$ zero eigenvalues of the rank- L matrix W . At the dimension crossing $\gamma = 1$, $\lambda_- = 0$ and the point mass vanishes.

E.3 Stieltjes transform derivation

The Stieltjes transform of a probability measure μ on $\mathbb{R}$ is

$$G(z) = \int \frac{d\mu(\lambda)}{z - \lambda}, \quad z \in \mathbb{C} \setminus \text{supp}(\mu). \quad (\text{E.4})$$

The measure can be recovered from G via the Stieltjes inversion formula:

$$d\mu(\lambda) = -\frac{1}{\pi} \lim_{\epsilon \rightarrow 0^+} \text{Im} G(\lambda + i\epsilon) d\lambda. \quad (\text{E.5})$$

We now derive the self-consistent equation for $G(z)$ in the Wishart case. The resolvent of $W = XX^\dagger$ is $\mathbf{G}(z) = (zI_k - W)^{-1}$, and $G(z) = \frac{1}{k} \text{tr} \mathbf{G}(z)$.

E.3.1 Step 1: The resolvent identity

Using the Woodbury identity (or directly from the definitions), one establishes for the $k \times k$ resolvent:

$$(zI_k - XX^\dagger)^{-1} = \frac{1}{z} I_k + \frac{1}{z} X (zI_L - X^\dagger X)^{-1} X^\dagger. \quad (\text{E.6})$$

Taking $\frac{1}{k} \text{tr}$ of both sides:

$$G(z) = \frac{1}{z} + \frac{1}{z} \cdot \frac{1}{k} \text{tr} [X^\dagger X (zI_L - X^\dagger X)^{-1}]. \quad (\text{E.7})$$

Now $X^\dagger X$ is an $L \times L$ matrix, and its non-zero eigenvalues coincide with those of $XX^\dagger$. Denote the Stieltjes transform of the empirical spectral distribution of $X^\dagger X$ by $\tilde{G}(z)$. Then

$$\frac{1}{L} \text{tr} (zI_L - X^\dagger X)^{-1} = \tilde{G}(z), \quad (\text{E.8})$$

and

$$\frac{1}{L} \text{tr} [X^\dagger X (zI_L - X^\dagger X)^{-1}] = -1 + z \tilde{G}(z). \quad (\text{E.9})$$

Substituting back, using $\frac{1}{k} \text{tr} = \frac{1}{\gamma} \cdot \frac{1}{L} \text{tr}$ with $\gamma = k/L$:

$$G(z) = \frac{1}{z} \left[1 + \frac{1}{\gamma} (-1 + z \tilde{G}(z)) \right] = \frac{1 - 1/\gamma}{z} + \frac{1}{\gamma} \tilde{G}(z). \quad (\text{E.10})$$

E.3.2 Step 2: The self-consistency relation

Since $XX^\dagger$ and $X^\dagger X$ share the same non-zero eigenvalues, their Stieltjes transforms are related by

$$\tilde{G}(z) = \frac{1 - \gamma}{z} + \gamma G(z), \quad (\text{E.11})$$

(accounting for the $|k - L|$ extra zero eigenvalues in the larger matrix; this is the inverse of (E.10)). These two equations are the same eigenvalue identity and do not close the system. Self-averaging

alone would not close it either: many deterministic spectral measures satisfy the same relation. The additional input is the limiting Wick moment calculation, Eq. (B.7), which supplies the free cumulants $\kappa_m = \gamma^{m-1}$. Their generating function gives

$$G(z) = \frac{1}{z - R(G(z))}, \quad R(w) = \frac{1}{1 - \gamma w}, \quad (\text{E.12})$$

where $R(w) = \sum_{m \geq 1} \kappa_m w^{m-1}$. The relation between R and G follows from the non-crossing moment recursion; it supplies the probabilistic information absent from the eigenvalue identity. Substituting $R(G) = 1/(1 - \gamma G)$ and clearing denominators,

$$G \left(z - \frac{1}{1 - \gamma G} \right) = 1 \implies z G (1 - \gamma G) - G = 1 - \gamma G, \quad (\text{E.13})$$

which rearranges to the **Marchenko–Pastur equation**:

$$\boxed{\gamma z G^2 - (z + \gamma - 1) G + 1 = 0.} \quad (\text{E.14})$$

This is a quadratic in G , solved by

$$G(z) = \frac{(z + \gamma - 1) - \sqrt{(z + \gamma - 1)^2 - 4\gamma z}}{2\gamma z}, \quad (\text{E.15})$$

where the branch of the square root is chosen so that $G(z) \sim 1/z$ as $z \rightarrow \infty$ (the normalization condition $\int d\mu = 1$).

E.3.3 Step 3: Recovering the density

The discriminant of (E.14) is

$$\Delta(z) = (z + \gamma - 1)^2 - 4\gamma z = (z - \lambda_+)(z - \lambda_-), \quad (\text{E.16})$$

with $\lambda_{\pm} = (1 \pm \sqrt{\gamma})^2$; expanding either side gives $z^2 - 2(1 + \gamma)z + (1 - \gamma)^2$. For real $\lambda \in (\lambda_-, \lambda_+)$, $\Delta(\lambda) < 0$ and the square root is purely imaginary. Applying the Stieltjes inversion formula (E.5):

$$\rho(\lambda) = -\frac{1}{\pi} \text{Im} G(\lambda + i0^+) = \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{2\pi\gamma\lambda}, \quad (\text{E.17})$$

which reproduces the continuous part of (E.2). For $\gamma > 1$, the branch of (E.15) fixed by $G \sim 1/z$ at infinity acquires a simple pole at $z = 0$ with residue $1 - 1/\gamma$, reproducing the point mass at zero.

E.4 R -transform derivation

The R -transform, introduced by Voiculescu [95] as the free-probability analogue of the log-characteristic function, provides an alternative route to spectral distributions. It is defined by

$$R(w) = G^{(-1)}(w) - \frac{1}{w}, \quad \text{equivalently} \quad G(R(w) + 1/w) = w, \quad (\text{E.18})$$

where $G^{(-1)}$ denotes the functional inverse of G (so that $G(G^{(-1)}(w)) = w$).

The key property is *additivity under free convolution*: if a and b are freely independent random variables, then

$$R_{a+b}(w) = R_a(w) + R_b(w). \quad (\text{E.19})$$

For a Marchenko–Pastur distribution with ratio γ , the R -transform is simple:

$$R_{\text{MP}}(w) = \frac{1}{1 - \gamma w}. \quad (\text{E.20})$$

Indeed, evaluating the defining relation (E.18) at $w = G(z)$ gives

$$z = R(G) + \frac{1}{G} = \frac{1}{1 - \gamma G} + \frac{1}{G}, \quad (\text{E.21})$$

multiplying through by $G(1 - \gamma G)$ yields $zG(1 - \gamma G) = G + (1 - \gamma G)$, which is again the quadratic (E.14), $\gamma z G^2 - (z + \gamma - 1)G + 1 = 0$.

The R -transform route is especially powerful when computing the spectral distribution of sums of freely independent random matrices, as arises in the analysis of the genus-zero contribution to the Page curve.

E.5 Edge universality and Tracy–Widom fluctuations

The MP distribution describes the *bulk* eigenvalue density. At a regular nonzero spectral edge, the density vanishes as a square root: $\rho(\lambda) \sim |\lambda - \lambda_{\pm}|^{1/2}$. For complex Gaussian Wishart matrices, the edge fluctuations have the *Tracy–Widom distribution* [93].

Specifically, for the largest eigenvalue λ_{max} of a complex Gaussian Wishart matrix at a fixed positive finite ratio:

$$\frac{\lambda_{\text{max}} - \lambda_+}{\sigma_{\text{TW}} \min(k, L)^{-2/3}} \xrightarrow{d} F_2, \quad (\text{E.22})$$

where F_2 is the Tracy–Widom distribution for the Gaussian Unitary Ensemble (GUE); the fluctuation scale is $\min(k, L)^{-2/3}$, and σ_{TW} is a known $O(1)$ function of γ .

The Tracy–Widom distribution is *not* Gaussian: it is asymmetric, with a right tail that decays as $\exp(-\frac{4}{3}x^{3/2})$ and a left tail that decays as $\exp(-|x|^3/12)$. These statements concern eigenvalues of a specified random matrix ensemble. They do not identify eigenvalue-edge fluctuations with fluctuations of horizon area or with competition between island saddles. Such an identification would require an additional derivation linking the observables and their limits.

At the special point $\gamma = 1$ (equal dimensions), the lower edge reaches the origin: $\lambda_- \rightarrow 0$ while $\lambda_+ \rightarrow 4$. For exactly balanced dimensions, or a fixed dimension difference, the smallest-eigenvalue limit at this hard edge is not of the Airy-kernel (Tracy–Widom) type: they are governed by the Bessel kernel, with limiting distributions expressible in Painlevé form.

F Selected Approaches to Black-Hole Information

The approaches below supply different kinds of information: candidate microstates, an entropy prescription, a dynamical model, interior operators, and a geometric interpretation of entanglement. They need not be mutually exclusive. This appendix explains the cited constructions and the questions they address; it is not an exhaustive survey of the current status of each research program. Section 12 explains how to formulate a quantitative comparison in a common setting.

F.1 Fuzzballs

Proponents. Mathur [62], Bena and Warner [14], and collaborators.

Core idea. The fuzzball proposal asserts that the classical black hole geometry is an effective description that averages over an exponentially large number of horizon-free, horizonless solutions

(“fuzzballs” or “microstate geometries”) in string theory. Smooth horizonless microstate geometries are a controlled subset of the proposed microscopic states. The full fuzzball proposal need not describe every state by a smooth classical geometry, and a general count of the known smooth families has not been shown to exhaust the black-hole entropy.

What has been constructed. Explicit microstate geometries have been constructed for certain supersymmetric (BPS) black holes, particularly in five dimensions:

- Two-charge systems: explicit families associated with profile data [62].
- Three-charge systems: large families of “bubbling” geometries carrying the same asymptotic charges as the BMPV black hole [14].
- Non-BPS extensions: partial constructions using “almost-BPS” and “scaling” solutions.

What remains open.

- No complete construction of microstate geometries for the Schwarzschild or Kerr black holes.
- The “typical” microstate geometry (the one that dominates the thermodynamic ensemble) has not been identified; the known solutions are atypical, carrying special quantum numbers.
- The cited geometric constructions do not specify a normalized evaporation ensemble from which the radiation spectrum and its transition profile could be computed. That is additional dynamical input.
- It is unclear whether the “fuzzball complementarity” (the experience of an infalling observer) can be reconciled with a horizonless geometry.

F.2 Islands and the replica method

Proponents. Penington [74]; Almheiri, Engelhardt, Marolf, and Maxfield [4]; Almheiri, Mahajan, Maldacena, and Zhao [5]; Penington, Shenker, Stanford, and Yang [76].

Core idea. The island formula generalizes the quantum extremal surface (QES) prescription [27] to settings with gravity coupled to a non-gravitational bath. The fine-grained entropy of the radiation is computed as

$$S(\text{rad}) = \min_{\mathcal{I}} \left[\frac{\text{Area}(\partial\mathcal{I})}{4G_N} + S_{\text{bulk}}(\text{rad} \cup \mathcal{I}) \right], \quad (\text{F.1})$$

where $\mathcal{I}$ is the “island” — a region of the black hole interior that is entanglement-wedge reconstructable from the radiation. In the standard evaporating examples, before the Page time the dominant extremum is at $\mathcal{I} = \emptyset$ (no island), and $S(\text{rad})$ grows. After the crossing, a non-trivial island dominates and, for the specified evaporation dynamics, gives the decreasing branch.

The replica wormhole derivation [76] provides a gravitational path integral justification: the island saddle is the dominant contribution to $\text{tr}(\rho^n)$ for n near 1, and the analytic continuation to $n \rightarrow 1$ yields the island formula.

What has been established.

- The island formula reproduces the Page curve in JT gravity coupled to a bath [5, 76].
- In higher-dimensional models (Randall–Sundrum braneworlds), similar island calculations have been performed.

- The JT matrix-integral formulation organizes amplitudes by genus. The relation between this expansion and a particular replicated entropy calculation depends on its boundary conditions and normalization.

What remains open.

- JT matrix-integral replica calculations compute ensemble quantities. Applying an ensemble result to a typical member needs a concentration argument; QES methods in fixed holographic theories have a broader scope than that particular JT ensemble interpretation.
- Extending a bath calculation to a fully gravitating setup requires specifying the radiation observable and its algebra; the tensor-factor treatment of the external bath cannot simply be carried over.
- The Page curve is one entropy statistic, not the first spectral moment (which is $\text{Tr } \rho = 1$). Rényi moments and the spectral density already contain nontrivial planar information. Their finite-size corrections require the specified model and normalization.
- The role of ensemble averaging vs. self-averaging is a central unresolved issue.

F.3 Random unitary circuits and fast scrambling

Proponents. Hayden and Preskill [41]; Sekino and Susskind [83].

Core idea. The black hole is modelled as a quantum information processor that scrambles information via random unitary dynamics. The Hayden–Preskill protocol considers an old black hole (past the Page time) into which a small diary is thrown. If the black hole dynamics is approximated by a Haar-random unitary, the diary’s information is recoverable from the radiation after collecting a number of additional qubits set by the diary size and desired accuracy. Relating this information-theoretic statement to a physical waiting time requires a scrambling model and an emission rate.

Key results.

- Hayden–Preskill: after the Page time, a Haar-random unitary acting on the black hole allows recovery of a diary thrown into the black hole from only $O(1)$ qubits of radiation beyond the diary size: after c extra quanta are collected, the recovery error is exponentially small in c [41].
- Sekino–Susskind: black holes motivate the fast-scrambling conjecture, with a scrambling time of order $\beta \log S$ [83].
- Approximate unitary t -designs: local random quantum circuits are proven to form approximate unitary t -designs in circuit depth polynomial in t and in the number of qubits [16]. Applying such a circuit theorem to evolution generated by a particular chaotic Hamiltonian requires a separate argument. This motivates — but does not prove — the Haar-random approximation used in random matrix calculations.

What remains open.

- The Haar-random unitary model treats the black hole as a “black box”; it does not provide a geometric or gravitational description of the interior.
- The connection between the scrambling time and the emergence of the island saddle in the gravitational path integral is understood heuristically but not rigorously.
- Finite-order design information controls only the corresponding copy moments. Additional precision and tail estimates are needed to control a spectral or entropy limit.

F.4 State-dependent operators

Proponents. Papadodimas and Raju [73].

Core idea. Papadodimas and Raju proposed that the operators describing the black hole interior (the “mirror operators”) must be *state-dependent*: they are defined relative to a particular equilibrium state $|\Psi\rangle$ of the CFT and cannot be written as fixed, state-independent operators on the Hilbert space. This is motivated by the observation that a state-independent construction of interior operators would violate the basic counting of degrees of freedom (the “AMPS” or “firewall” argument [6]).

Key features.

- The mirror operators $\tilde{O}$ are constructed as functions of the CFT Hamiltonian and the state $|\Psi\rangle$, acting on a “code subspace” of states near $|\Psi\rangle$.
- Within this code subspace, the mirror operators satisfy the expected bulk effective field theory algebra, including the correct commutation relations.
- The construction is formulated in a large- N regime; powers of $1/N$ should not be identified with e^{-S_0} : it applies to low-point correlation functions in the large- N limit.

What remains open.

- State dependence is philosophically contentious: it challenges the standard quantum-mechanical framework in which observables are represented by fixed operators.
- Operator-algebraic approaches provide another framework for black-hole observables and entropy [98]. Their relation to a particular mirror-operator construction requires matching the states, algebras, and approximations; it is not an automatic reformulation of state dependence.
- Whether state-dependent operators can be consistently defined beyond perturbation theory in $1/N$ is unknown.
- The proposal does not directly address the Page curve; its primary concern is the smoothness of the black hole interior rather than the entropy of the radiation.

F.5 ER = EPR

Proponents. Maldacena and Susskind [57].

Core idea. The ER = EPR conjecture proposes that any two quantum systems connected by Einstein–Podolsky–Rosen (EPR) entanglement are also connected by an Einstein–Rosen (ER) bridge (a non-traversable wormhole). In the context of black hole evaporation, this motivates a possible geometric interpretation of correlations between radiation and the remaining system. Assigning a semiclassical wormhole to each individual Hawking quantum is not an established consequence. Geometric descriptions of more general entanglement patterns remain conjectural outside controlled examples such as the thermofield double.

Key features.

- The thermofield double state $|TFD\rangle$ of two copies of a CFT is dual to the two-sided eternal AdS black hole (the original ER bridge), providing a concrete example of ER = EPR [55].

- *Complexity*: The conjecture connects entanglement to geometry through the concept of quantum computational complexity [87]. The interior volume of the wormhole grows linearly in time, which is conjectured to be dual to the linear growth of complexity of the boundary state (the “complexity = volume” or “complexity = action” proposals).
- *Traversable wormholes*: Gao et al. [30] showed that a double-trace deformation coupling the two boundaries of the eternal black hole can render the wormhole traversable. This provides an operational protocol for transmitting information through the wormhole, which in the boundary description corresponds to quantum teleportation.

What remains open.

- The unrestricted ER = EPR conjecture extends beyond the controlled examples described above; those examples do not establish its general form.
- The “Planck-scale wormholes” connecting individual Hawking quanta are far below the semiclassical regime and cannot be described by classical geometry. What “wormhole” means in this regime is unclear.
- The relationship between ER = EPR and the island/replica approach is suggestive but not precise: the replica wormholes that appear in the n -replica calculation are Euclidean and topological, while the ER bridges are Lorentzian.
- The unrestricted conjecture does not specify a unique radiation channel or spectrum for comparison with another model.

F.6 Summary: what the proposals share and where they differ

These approaches address different questions and need not be mutually exclusive. Microstate geometries concern microscopic states; islands concern entropy and reconstruction; random circuits supply dynamics in specified models; mirror operators concern interior observables; and ER = EPR concerns the interpretation of entanglement. A common motivation of unitary black-hole physics does not mean that each approach has independently calculated the same radiation spectrum. Quantitative comparisons require matched preparations, observables, and approximation orders, as discussed in Section 12.

G Glossary of Key Terms

The following terms appear frequently throughout the paper. Definitions are given in the context of black hole physics and the Page curve program; mathematical terms receive their standard definitions.

AdS/CFT The Anti-de Sitter/Conformal Field Theory correspondence [54]: a duality between a theory of quantum gravity in $(d + 1)$ -dimensional Anti-de Sitter spacetime and a conformal field theory living on its d -dimensional boundary.

Annealed average

The quantity $\ln \overline{Z}$, computed by first averaging Z and then taking the logarithm. Contrasted with the quenched average $\overline{\ln Z}$, where the logarithm is taken before averaging. Their difference depends on fluctuations of Z . For normalized replica moments $Z = X_n = \text{Tr } \rho^n$, the corresponding entropy averages have the same finite-dimensional limit at $n = 1$.

Approach	Construction discussed	Information supplied
Microstate geometries	Explicit horizonless solutions for selected charge sectors	Candidate microscopic states; a complete evaporation ensemble is additional input
Islands and replicas	QES and replica calculations in specified bath-coupled models	Entropy and, with reconstruction arguments, a recoverable bulk algebra
Random unitaries	Haar and circuit models with a prescribed input and output split	Moment averages and recovery bounds for the chosen ensemble
Mirror operators	Interior operators near a selected equilibrium state	A local code-subspace operator algebra in the stated large- N regime
ER = EPR	Two-sided AdS geometry for the thermofield double; broader conjectural extensions	A geometric interpretation in controlled examples, not a unique evaporation channel

Table 16. Examples discussed in this appendix and the type of information each supplies. These constructions address different questions; quantitative comparisons require the common preparation and observable described in Section 12.

Area operator

The area observable associated with a specified surface. If $S_{\text{BH}} = \hat{A}/(4G_N)$, then $\text{Var } \hat{A} = 16G_N^2 \text{Var } S_{\text{BH}}$. The further approximation $\text{Var } S_{\text{BH}} \simeq C_B$ requires a stable regular canonical saddle and linearization in energy.

Bekenstein–Hawking entropy

$S_{\text{BH}} = A/(4G_N)$: the entropy of a black hole, proportional to the area A of the event horizon [12, 38].

Catalan number

$C_n = \frac{1}{n+1} \binom{2n}{n}$: counts the non-crossing partitions of $\{1, \dots, n\}$, among many other combinatorial objects.

Dilaton

A scalar field ϕ appearing in the JT gravity action that enforces constant negative curvature in the bulk. Its constant (topological) piece ϕ_0 sets the extremal entropy $S_0 = \phi_0/(4G_N)$, while its boundary value ϕ_b sets the Schwarzian coupling C .

Entanglement entropy

The von Neumann entropy $S = -\text{tr}(\rho \ln \rho)$ of the reduced density matrix of a subsystem. For a pure global state, measures its bipartite entanglement; for a mixed global state it also includes ordinary mixedness.

ER bridge

An Einstein–Rosen bridge: a non-traversable wormhole connecting two asymptotic regions of a black hole spacetime.

Extremal surface

A codimension-two bulk surface with stationary area under local deformations, subject

to the prescribed anchoring and homology conditions. The HRT prescription uses such surfaces at leading semiclassical order.

Free cumulant

The analogue of a classical cumulant in free probability. Related to moments by summation over non-crossing partitions (Möbius inversion on $\text{NC}(n)$).

Free probability

A non-commutative probability theory, developed by Voiculescu [95], in which the notion of independence is replaced by *free independence*. Classical independence corresponds to tensor products; free independence corresponds to free products.

Fuzzball

A proposed microscopic black-hole state in string theory; smooth horizonless microstate geometries provide a controlled subset of candidate realizations [62].

Genus expansion

The perturbative expansion of the gravitational path integral (or a matrix integral) organized by the genus g of the contributing surfaces. For connected one-boundary JT amplitudes, the convention is $Z = \sum_g e^{(1-2g)S_0} Z_g$; with b boundaries the topological weight is $e^{(2-2g-b)S_0}$.

Haar measure

The unique unitarily invariant probability measure on the unitary group $U(N)$. Used to define “random” unitary operators and “random” pure states.

Hawking radiation

Thermal radiation emitted by a black hole due to quantum effects near the horizon [38]. The temperature is the Hawking temperature $T_H = \kappa_s/(2\pi)$, where κ_s is the surface gravity.

Island

A region $\mathcal{I}$ of the black hole interior (or its entanglement wedge) that contributes to the fine-grained entropy of the radiation via the island formula (F.1).

JT gravity

Jackiw–Teitelboim gravity: a two-dimensional dilaton gravity theory. See Appendix C.

Kreweras complement

An anti-automorphism of the non-crossing partition lattice that maps a partition with j blocks to one with $n + 1 - j$ blocks. See Appendix B.

Marchenko–Pastur distribution

The limiting eigenvalue distribution of Wishart matrices $W = XX^\dagger$ (entries of X i.i.d. with mean zero and variance $1/L$) in the limit $k, L \rightarrow \infty$ with $\gamma = k/L$ fixed [58]. See Appendix E.

Narayana number

$N(n, j) = \frac{1}{n} \binom{n}{j} \binom{n}{j-1}$: counts non-crossing partitions of $[n]$ with exactly j blocks.

Non-crossing partition

A partition of $[n]$ such that no two blocks “cross” when elements are arranged on a circle. See Appendix B.

Page curve

The radiation entropy as a function of time. The familiar rising and falling profile is a typical-state benchmark for suitable evaporation models [71, 72]; unitarity alone does not fix its entire shape or guarantee monotonic branches.

Page time The crossing of competing entropy branches in a specified evaporation model. In the flat-weight dimension scan it occurs at $k = L$, or $\gamma = 1$. Broadening, differing branch slopes, and finite-size corrections can shift the maximum of the entropy away from the crossing of the branch means.

Quantum extremal surface

A codimension-two surface that extremizes generalized entropy, including the area term and the renormalized bulk entropy. The QES prescription [27] includes quantum corrections to holographic entropy.

Quenched average

The average $\overline{\ln Z}$, where the logarithm is taken before averaging. It is still an ensemble average, not the value of a particular realization. For normalized density matrices the finite-dimensional quenched and annealed replica entropies share the limit $\mathbb{E}S(\rho)$ at replica index one. The entropy of the averaged matrix is a different quantity.

R -transform The free-probability analogue of the logarithm of the characteristic function. Additive under free convolution. See Appendix E.

Rényi entropy

$S_n = \frac{1}{1-n} \ln \text{tr}(\rho^n)$: a one-parameter family of entropies that reduces to the von Neumann entropy in the limit $n \rightarrow 1$.

Replica wormhole

A saddle-point geometry in the n -replica gravitational path integral that connects different replica copies. Replica wormholes are responsible for the transition from the growing to the decreasing phase of the Page curve [76].

Scrambling time

The time over which initially localized information becomes distributed among many degrees of freedom. Fast-scrambling black-hole models motivate $t_{\text{scr}} \sim \beta \log S_{\text{BH}}$. Recovery from radiation also depends on the channel, prior entanglement, and how much radiation is collected.

Schwarzian derivative

$\{f, \tau\} = f'''/f' - \frac{3}{2}(f''/f')^2$: the differential expression that enters the boundary action in JT gravity; see Appendix C.

Stieltjes transform

$G(z) = \int d\mu(\lambda)/(z - \lambda)$: the analytic transform that encodes a spectral distribution. See Appendix E.

Tracy–Widom distribution

The universal distribution governing fluctuations of the largest eigenvalue in suitable random-matrix ensembles at a regular soft edge; it is not a general law for fluctuations of a Page transition.

Wishart matrix

A random matrix of the form $W = XX^\dagger$, with X a $k \times L$ matrix having i.i.d. mean-zero entries of variance $1/L$. Models the (rescaled) reduced density matrix of a random pure state.

References

- [1] Almheiri, A., T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, 2020, Replica Wormholes and the Entropy of Hawking Radiation, *Journal of High Energy Physics* 2020, 013.
- [2] Almheiri, A., T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, 2021, The Entropy of Hawking Radiation, *Reviews of Modern Physics* 93, 035002.
- [3] Akers, C., and G. Penington, 2021, Leading Order Corrections to the Quantum Extremal Surface Prescription, *Journal of High Energy Physics* 2021, 062.
- [4] Almheiri, A., N. Engelhardt, D. Marolf, and H. Maxfield, 2019, The Entropy of Bulk Quantum Fields and the Entanglement Wedge of an Evaporating Black Hole, *Journal of High Energy Physics* 2019, 063.
- [5] Almheiri, A., R. Mahajan, J. Maldacena, and Y. Zhao, 2020, The Page Curve of Hawking Radiation from Semiclassical Geometry, *Journal of High Energy Physics* 2020, 149.
- [6] Almheiri, A., D. Marolf, J. Polchinski, and J. Sully, 2013, Black Holes: Complementarity or Firewalls?, *Journal of High Energy Physics* 2013, 062.
- [7] Araki, H., and E. H. Lieb, 1970, Entropy Inequalities, *Communications in Mathematical Physics* 18, 160–170.
- [8] Bachelier, L., 1900, Théorie de la Spéculation, *Annales Scientifiques de l'École Normale Supérieure* 17, 21–86.
- [9] Balasubramanian, V., A. Lawrence, J. M. Magán, and M. Sasieta, 2024, Microscopic Origin of the Entropy of Black Holes in General Relativity, *Physical Review X* 14, 011024.
- [10] Banks, T., L. Susskind, and M. E. Peskin, 1984, Difficulties for the Evolution of Pure States into Mixed States, *Nuclear Physics B* 244, 125–134.
- [11] Bardeen, J. M., B. Carter, and S. W. Hawking, 1973, The Four Laws of Black Hole Mechanics, *Communications in Mathematical Physics* 31, 161–170.
- [12] Bekenstein, J. D., 1973, Black Holes and Entropy, *Physical Review D* 7, 2333–2346.
- [13] Bell, J. S., 1964, On the Einstein Podolsky Rosen Paradox, *Physics Physique Fizika* 1, 195–200.
- [14] Bena, I., and N. P. Warner, 2008, Black Holes, Black Rings, and their Microstates, *Lecture Notes in Physics* 755, 1–92.
- [15] Bousso, R., and M. Miyaji, 2024, Fluctuations in the Entropy of Hawking Radiation, *Physical Review D* 109, 026006, arXiv:2307.13920.
- [16] Brandão, F. G. S. L., A. W. Harrow, and M. Horodecki, 2016, Local Random Quantum Circuits Are Approximate Polynomial-Designs, *Communications in Mathematical Physics* 346, 397–434.
- [17] Calabrese, P., and J. Cardy, 2004, Entanglement Entropy and Quantum Field Theory, *Journal of Statistical Mechanics: Theory and Experiment* 2004, P06002.
- [18] Carlson, F., 1914, Sur une Classe de Séries de Taylor, Ph.D. thesis, Uppsala University.
- [19] Carter, B., 1971, Axisymmetric Black Hole Has Only Two Degrees of Freedom, *Physical Review Letters* 26, 331–333.

- [20] A. Chamblin, R. Emparan, C. V. Johnson and R. C. Myers, *Charged AdS black holes and catastrophic holography*, Phys. Rev. D **60** (1999) 064018 [[hep-th/9902170](#)].
- [21] Cheng, N., Lancien, C., Penington, G., Walter, M., and Witteveen, F., 2024, Random Tensor Networks with Non-trivial Links, *Annales Henri Poincaré* 25(4), 2107–2212.
- [22] 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.
- [23] Coffman, V., J. Kundu, and W. K. Wootters, 2000, Distributed Entanglement, *Physical Review A* 61, 052306.
- [24] Colin-Ellerin, S., X. Dong, D. Marolf, M. Rangamani, and Z. Wang, 2021, Real-Time Gravitational Replicas: Formalism and a Variational Principle, *Journal of High Energy Physics* 2021, 117.
- [25] Cori, R., and G. Hetyei, 2014, Counting genus one partitions and permutations, *Séminaire Lotharingien de Combinatoire* 70, B70e, arXiv:1306.4628.
- [26] X. Dong and H. Wang, *Enhanced corrections near holographic entanglement transitions: a chaotic case study*, JHEP **11** (2020) 007 [[2006.10051](#)].
- [27] Engelhardt, N., and A. C. Wall, 2015, Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime, *Journal of High Energy Physics* 2015, 073.
- [28] Faulkner, T., A. Lewkowycz, and J. Maldacena, 2013, Quantum Corrections to Holographic Entanglement Entropy, *Journal of High Energy Physics* 2013, 074.
- [29] Feynman, R. P., 1948, Space-Time Approach to Non-Relativistic Quantum Mechanics, *Reviews of Modern Physics* 20, 367–387.
- [30] Gao, P., D. L. Jafferis, and A. C. Wall, 2017, Traversable Wormholes via a Double Trace Deformation, *Journal of High Energy Physics* 2017, 151.
- [31] Gibbons, G. W., and S. W. Hawking, 1977, Action Integrals and Partition Functions in Quantum Gravity, *Physical Review D* 15, 2752–2756.
- [32] Giddings, S. B., 1992, Black Holes and Massive Remnants, *Physical Review D* 46, 1347–1352.
- [33] 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.
- [34] Goupil, A., and G. Schaeffer, 1998, Factoring n -cycles and counting maps of given genus, *European Journal of Combinatorics* 19, 819–834.
- [35] Harer, J., and D. Zagier, 1986, The Euler characteristic of the moduli space of curves, *Inventiones Mathematicae* 85, 457–485.
- [36] Harlow, D., 2016, Jerusalem Lectures on Black Holes and Quantum Information, *Reviews of Modern Physics* 88, 015002.
- [37] Hawking, S. W., 1971, Gravitational Radiation from Colliding Black Holes, *Physical Review Letters* 26, 1344–1346.
- [38] Hawking, S. W., 1975, Particle Creation by Black Holes, *Communications in Mathematical Physics* 43, 199–220.

- [39] Hawking, S. W., 1976, Breakdown of Predictability in Gravitational Collapse, *Physical Review D* 14, 2460–2473.
- [40] Hawking, S. W., and D. N. Page, 1983, Thermodynamics of Black Holes in Anti-de Sitter Space, *Communications in Mathematical Physics* 87, 577–588.
- [41] Hayden, P., and J. Preskill, 2007, Black Holes as Mirrors: Quantum Information in Random Subsystems, *Journal of High Energy Physics* 2007, 120.
- [42] Holevo, A. S., 1973, Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel, *Problems of Information Transmission* 9, 177–183.
- [43] Hubeny, V. E., M. Rangamani, and T. Takayanagi, 2007, A Covariant Holographic Entanglement Entropy Proposal, *Journal of High Energy Physics* 2007, 062.
- [44] Islam, R., R. Ma, P. M. Preiss, M. E. Tai, A. Lukin, M. Rispoli, and M. Greiner, 2015, Measuring Entanglement Entropy in a Quantum Many-Body System, *Nature* 528, 77–83.
- [45] Israel, W., 1967, Event Horizons in Static Vacuum Space-Times, *Physical Review* 164, 1776–1779.
- [46] Israel, W., 1968, Event Horizons in Static Electrovac Space-Times, *Communications in Mathematical Physics* 8, 245–260.
- [47] Jackiw, R., 1985, Lower Dimensional Gravity, *Nuclear Physics B* 252, 343–356.
- [48] Kreweras, G., 1972, Sur les Partitions Non Croisées d’un Cycle, *Discrete Mathematics* 1, 333–350.
- [49] D. Kubizňák and R. B. Mann, *P–V criticality of charged AdS black holes*, JHEP **07** (2012) 033 [[1205.0559](#)].
- [50] Lewkowycz, A., and J. Maldacena, 2013, Generalized Gravitational Entropy, *Journal of High Energy Physics* 2013, 090.
- [51] Lieb, E. H., and M. B. Ruskai, 1973, Proof of the Strong Subadditivity of Quantum-Mechanical Entropy, *Journal of Mathematical Physics* 14, 1938–1941.
- [52] Lloyd, S., and H. Pagels, 1988, Complexity as Thermodynamic Depth, *Annals of Physics* 188, 186–213.
- [53] Lubkin, E., 1978, Entropy of an n-System from Its Correlation with a k-Reservoir, *Journal of Mathematical Physics* 19, 1028–1031.
- [54] Maldacena, J., 1998, The Large N Limit of Superconformal Field Theories and Supergravity, *Advances in Theoretical and Mathematical Physics* 2, 231–252.
- [55] Maldacena, J., 2003, Eternal Black Holes in Anti-de Sitter, *Journal of High Energy Physics* 2003, 021.
- [56] Maldacena, J., D. Stanford, and Z. Yang, 2016, Conformal Symmetry and its Breaking in Two Dimensional Nearly Anti-de-Sitter Space, *Progress of Theoretical and Experimental Physics* 2016, 12C104.
- [57] Maldacena, J., and L. Susskind, 2013, Cool Horizons for Entangled Black Holes, *Fortschritte der Physik* 61, 781–811.
- [58] Marchenko, V. A., and L. A. Pastur, 1967, Distribution of Eigenvalues for Some Sets of Random Matrices, *Matematicheskii Sbornik* 72, 507–536.

- [59] Markov, A. A., 1958, Insolubility of the Problem of Homeomorphy, in *Proceedings of the International Congress of Mathematicians 1958* (Cambridge University Press), 300–306.
- [60] Marolf, D., and H. Maxfield, 2020, Transcending the Ensemble: Baby Universes, Spacetime Wormholes, and the Order and Disorder of Black Hole Information, *Journal of High Energy Physics* 2020, 044.
- [61] 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.
- [62] Mathur, S. D., 2005, The Fuzzball Proposal for Black Holes: an Elementary Review, *Fortschritte der Physik* 53, 793–827.
- [63] Mathur, S. D., 2009, The information paradox: A pedagogical introduction, *Classical and Quantum Gravity* 26, 224001, arXiv:0909.1038.
- [64] Mertens, T. G., and G. J. Turiaci, 2023, Solvable Models of Quantum Black Holes: A Review on Jackiw–Teitelboim Gravity, *Living Reviews in Relativity* 26, 4.
- [65] Mück, W., 2024, Black holes and Marchenko–Pastur distribution, *Physical Review D* 109, 126001.
- [66] Murthy, C. and Srednicki, M., 2019, Structure of Chaotic Eigenstates and Their Entanglement Entropy, *Physical Review E* 100, 022131.
- [67] Nielsen, M. A., and I. L. Chuang, 2000, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge).
- [68] Osterwalder, K., and R. Schrader, 1973, Axioms for Euclidean Green’s Functions, *Communications in Mathematical Physics* 31, 83–112.
- [69] Osterwalder, K., and R. Schrader, 1975, Axioms for Euclidean Green’s Functions II, *Communications in Mathematical Physics* 42, 281–305.
- [70] Page, D. N., 1976, Particle Emission Rates from a Black Hole: Massless Particles from an Uncharged, Nonrotating Hole, *Physical Review D* 13, 198–206.
- [71] Page, D. N., 1993, Average Entropy of a Subsystem, *Physical Review Letters* 71, 1291–1294.
- [72] Page, D. N., 1993, Information in Black Hole Radiation, *Physical Review Letters* 71, 3743–3746.
- [73] Papadodimas, K., and S. Raju, 2013, An Infalling Observer in AdS/CFT, *Journal of High Energy Physics* 2013, 212.
- [74] Penington, G., 2020, Entanglement Wedge Reconstruction and the Information Paradox, *Journal of High Energy Physics* 2020, 002.
- [75] Penrose, R., 1963, Asymptotic Properties of Fields and Space-Times, *Physical Review Letters* 10, 66–68.
- [76] Penington, G., S. H. Shenker, D. Stanford, and Z. Yang, 2022, Replica Wormholes and the Black Hole Interior, *Journal of High Energy Physics* 2022, 205.
- [77] Raju, S., 2022, Lessons from the Information Paradox, *Physics Reports* 943, 1–80.
- [78] Ryu, S., and T. Takayanagi, 2006, Holographic Derivation of Entanglement Entropy from the anti-de Sitter Space/Conformal Field Theory Correspondence, *Physical Review Letters* 96, 181602.
- [79] Saad, P., S. H. Shenker, and D. Stanford, 2019, JT Gravity as a Matrix Integral, arXiv:1903.11115.

- [80] Sakurai, J. J., and J. Napolitano, 2020, *Modern Quantum Mechanics*, 3rd ed. (Cambridge University Press, Cambridge).
- [81] Schumacher, B., 1996, Sending entanglement through noisy quantum channels, *Physical Review A* 54, 2614–2628, doi:10.1103/PhysRevA.54.2614, arXiv:quant-ph/9604023.
- [82] Schwarzschild, K., 1916, Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften* 189–196.
- [83] Sekino, Y., and L. Susskind, 2008, Fast Scramblers, *Journal of High Energy Physics* 2008, 065.
- [84] Sen, S., 1996, Average Entropy of a Quantum Subsystem, *Physical Review Letters* 77, 1–3, arXiv:hep-th/9601132.
- [85] Speicher, R., 1994, Multiplicative Functions on the Lattice of Non-Crossing Partitions and Free Convolution, *Mathematische Annalen* 298, 611–628.
- [86] Speicher, R., 1997, Free Probability Theory and Non-Crossing Partitions, *Séminaire Lotharingien de Combinatoire* 39, B39c.
- [87] Stanford, D., and L. Susskind, 2014, Complexity and Shock Wave Geometries, *Physical Review D* 90, 126007.
- [88] Stanford, D., and E. Witten, 2017, Fermionic Localization of the Schwarzian Theory, *Journal of High Energy Physics* 2017, 008.
- [89] Susskind, L., 1995, The World as a Hologram, *Journal of Mathematical Physics* 36, 6377–6396.
- [90] Teitelboim, C., 1983, Gravitation and Hamiltonian Structure in Two Spacetime Dimensions, *Physics Letters B* 126, 41–45.
- [91] 't Hooft, G., 1974, A Planar Diagram Theory for Strong Interactions, *Nuclear Physics B* 72, 461–473.
- [92] 't Hooft, G., 1993, Dimensional Reduction in Quantum Gravity, in *Salamfestschrift*, edited by A. Ali, J. Ellis, and S. Randjbar-Daemi, World Scientific, pp. 284–296.
- [93] Tracy, C. A., and H. Widom, 1994, Level-Spacing Distributions and the Airy Kernel, *Communications in Mathematical Physics* 159, 151–174.
- [94] Vidmar, L. and Rigol, M., 2017, Entanglement Entropy of Eigenstates of Quantum Chaotic Hamiltonians, *Physical Review Letters* 119, 220603.
- [95] Voiculescu, D., 1991, Limit Laws for Random Matrices and Free Products, *Inventiones Mathematicae* 104, 201–220.
- [96] Wang, J., 2023, Beyond Islands: A Free Probabilistic Approach, *Journal of High Energy Physics* 2023(10), 040.
- [97] E. Witten, *Anti-de Sitter space, thermal phase transition, and confinement in gauge theories*, Adv. Theor. Math. Phys. **2** (1998) 505–532 [[hep-th/9803131](#)].
- [98] Witten, E., 2022, Gravity and the Crossed Product, *Journal of High Energy Physics* 2022, 008.
- [99] Wuebben, B. J., 2026, The Page Curve Through the Transition: An Order-One Correction to the Diagonal Approximation, preprint.
- [100] Wuebben, B. J., 2026, The Page Curve in Higher Dimensions: Phase Transitions, Critical Phenomena, and Finite-Size Broadening, preprint.

- [101] Wuebben, B. J., 2026, The Page Curve Is Not Enough: Universality, Ensemble Dependence, and Microscopic Ambiguity in Black Hole Information Recovery, preprint.
- [102] York, J. W., Jr., 1972, Role of Conformal Three-Geometry in the Dynamics of Gravitation, *Physical Review Letters* 28, 1082–1085.
- [103] J. W. York, Jr., *Black-hole thermodynamics and the Euclidean Einstein action*, Phys. Rev. D **33** (1986) 2092, doi:[10.1103/PhysRevD.33.2092](https://doi.org/10.1103/PhysRevD.33.2092).