Notes on Wasserstein distance and the JT gravitational ensemble

Ville Keranen* Claude (Anthropic)

July 2026 (revised November 2026)

hexagon:2611.00295v1

Abstract

This note completes the study of the Boltzmann–Wasserstein (BW) distance, an optimal-transport metric on quantum theories, begun in [1]. There the distance was evaluated between macroscopically different theories; here we take the opposite limit, two independent draws of the JT-gravity ensemble, which are identical at the classical level. The distance is computable and is controlled by the connected spectral form factor: the *ramp*, the double-trumpet wormhole, regulated by the *plateau*, the spectral discreteness. The mean is exponentially small, $\langle \mathcal{W}^2 \rangle \sim S_0\, \mathrm{e}^{-S_0}$, whereas uncorrelated (Poisson) levels would give an unsuppressed $O(1)$ distance: the self-distance is a sharp diagnostic of spectral chaos. As a metric, the BW distance renders the ensemble a point cloud: exponentially small, strongly anisotropic, its leading principal axis a coherent "breathing" deformation of the low-energy spectrum. Its effective dimension grows asymptotically as $D_{\text{eff}} \simeq 0.42\, S_0^2$, so the distance self-averages, its relative width shrinking as std/mean $\sim 1/S_0$. At low temperatures it localizes on the Airy edge, where a single Tracy–Widom mode dominates and self-averaging fails. We test the picture with exact-diagonalization GUE and SYK simulations: in GUE the effective dimension grows logarithmically at every temperature and is predicted at finite N, with no free parameters, by the same covariance; unfolded SYK exhibits the predicted anisotropy and mode structure of the cloud. Together with the companion result, this fixes the resolution of the metric: the microstate rigidity computed here is the noise floor beneath a macroscopic difference between theories.

Contents

1 Introduction **2**

2 Quick review of the gravitational ensemble and the spectral form factor **4**
2.1 The ensemble and the double trumpet 5
2.2 The spectral form factor from the double trumpet 5

3 The Boltzmann–Wasserstein distance and its cumulants **7**
3.1 The distance and the exact identity 7
3.2 The continuum representation 8
3.3 The covariance operator and the cumulant tower 9

4 The first two cumulants in gravity **10**
4.1 Number variance and the spectral form factor 10
4.2 The mean distance: $\langle \mathcal{W}^2 \rangle \sim S_0\, \mathrm{e}^{-S_0}$ 13
4.3 The variance: double trumpet squared 14

*vkeranen1@gmail.com

5 Geometry of the ensemble **15**
5.1 Flatness, typical separation and width 16
5.2 Principal modes: how the spectra of the ensemble members differ 16
5.3 Concentration, effective dimensionality and anisotropy 19

6 Low temperatures: the Airy edge and SYK **21**
6.1 SYK: the gravitational ensemble member by member 24

7 Discussion **26**

A Anatomy of the SYK ensemble **29**

B GUE tests of the bulk-regime predictions **30**

1 Introduction

The gravitational path integral appears to not compute the observables of a single quantum theory. Summing over bulk geometries with several asymptotic boundaries produces connected contributions, wormholes[1] so that products of partition functions fail to factorize, $\langle Z(\beta_1)Z(\beta_2)\rangle \neq \langle Z(\beta_1)\rangle\langle Z(\beta_2)\rangle$ [2]. The standard reading is that the path integral computes *ensemble averages*: it describes not one boundary theory but a distribution over them, the connected geometries supplying the ensemble's cumulants [2, 3]. The individual members are non-perturbative data the smooth bulk does not resolve; what it resolves are the *statistics* of the ensemble. This is at once the source of the factorization puzzle and an invitation: if gravity hands us an ensemble of theories, we may ask what that ensemble looks like as a *space*: how its members are spread, how far apart they sit, and what shape they form.

Such questions need a notion of distance between theories. The companion paper [1] introduced one: the Boltzmann–Wasserstein (BW) distance, a temperature-dependent optimal-transport metric on quantum systems,

$$\mathcal{W}^2(\beta) = \min_{U\in \mathrm{U}(\mathcal{D})} \mathrm{Tr}\big(\mathrm{e}^{-\beta H_1} - U\,\mathrm{e}^{-\beta H_2}\,U^\dagger\big)^2,$$

the least squared mismatch between the two unnormalized thermal operators $\mathrm{e}^{-\beta H_1}$ and $\mathrm{e}^{-\beta H_2}$, minimized over all unitary alignments U of their eigenbases. Its essential simplification is that the optimization can be carried out exactly: the optimal alignment is the monotone (rank) pairing, and $\mathcal{W}^2$ collapses to the Euclidean distance between the rank-ordered Boltzmann weights, $\mathcal{W}^2 = \sum_n (a_n - b_n)^2$ with $a_n = \mathrm{e}^{-\beta E_n}$ (Section 3.1). A genuine metric on the space of spectra thereby becomes something one computes one spectrum against another.

Why this distance and not another? The alternatives are instructive: some are ill-defined on discrete spectra, others are well-defined but hard to compute or compare the wrong thing, and these obstructions are the motivation. Distances that compare spectral densities pointwise, such as relative entropy or L^2, are ill-defined here before any gravity is attempted: two draws share a smooth density, so what distinguishes them is where the individual levels sit, and there the objects compared are mutually singular. They are sets of δ-functions at distinct energies, so the relative entropy between them is *infinite* however close the levels lie. Smearing makes such divergences finite but regulator-dependent. The Wasserstein distance is built instead on the

[1] A terminology clash with the companion paper [1] is worth flagging. There "wormhole" referred to TFD-type eternal black holes: geometries that are topologically discs in Euclidean signature, while containing a perfectly valid *Lorentzian* wormhole, the Einstein–Rosen bridge connecting the two exterior regions. Here we adopt the now-standard Euclidean usage: a wormhole is a space that is topologically a cylinder in Euclidean signature, connecting two disconnected asymptotic boundaries, the double trumpet throughout. The invariant distinction is the number of connected components of the Euclidean boundary: one for the SK/TFD geometry, two for the cylinder.

ground metric of energy: it measures how far the eigenvalues must *move* to align, staying finite between singular distributions, with no smearing scale, and small precisely when the spectra are close.

Quantum-information metrics between the Gibbs states themselves, such as fidelity or the relative entropy of the density matrices, are finite, but compare the wrong thing or at the wrong price. Without an optimization over the relative eigenbasis they are dominated by its Haar-random misalignment: two independent draws are nearly orthogonal however similar their spectra, and the answer depends on eigenvector statistics that the spectral ensemble does not even fix. With the optimization (the analogue of the $\min_U$ above, which is what converts a basis-noise meter into a spectral metric) they reduce to sorted-spectra functionals, but ones carrying normalizations $1/Z$ and fractional powers of $\mathrm{e}^{-\beta H}$, whose ensemble averages demand quenched ratios and replica continuations.

The BW distance avoids every one of these obstructions at once. The one-dimensional transport problem is solved exactly by the rank pairing, $\mathcal{W}^2 = \sum_n (a_n - b_n)^2$, the n-th level of one draw against the n-th of the other, so the distance is the Boltzmann-weighted displacement of one spectrum from the other (the rigidity), with the quadratic cost turning level displacements into the variances that Gaussian level statistics control. It is a true metric on theories identified when unitarily equivalent, that is, on the space of spectra, so the "geometry of the ensemble" below is geometry in the literal sense. (For an ensemble whose members carry distinguished operators beyond the Hamiltonian, such as SYK, this identification is a projection; we return to it in Section 7.) And it is *quadratic in the unnormalized Boltzmann operator*, the very currency of the gravitational path integral: one connected two-boundary amplitude for the mean, a fixed finite number of boundaries for every cumulant, no replicas anywhere.

The companion paper used this metric to compare two *distinct* theories (two BTZ black holes at different cosmological constant, or a $T\bar{T}$ deformation), where the leading separation is macroscopic: the difference of horizon areas $\frac{1}{8}(\delta A/4G)^2$, computed by a classical Schwinger–Keldysh wormhole saddle, an area comparator, or time-averaged correlators, depending on the regime. The present paper asks whether the metric survives at the opposite, *finest* scale. Two i.i.d. draws from the ensemble are identical at the classical level: they share a smooth spectrum, so the macroscopic term vanishes identically and each of those methods degenerates. What remains is microstate *rigidity*: the fluctuation of the discrete levels about the common smooth density. The questions become metric questions about the ensemble itself: how far apart two draws sit, how the distance fluctuates from pair to pair, and what size, dimension and principal axes the resulting cloud of theories has. (The mixed case, theories differing both macroscopically and in their microstates, superposes the two regimes; we return to it in Section 7.)

The ensemble-averaged BW distance is controlled by the connected *spectral form factor*, read through a weighted integral over time: built from the *ramp*, the double-trumpet wormhole, and cut off by the *plateau*, the spectral discreteness (23). A priori one might have guessed that a distance between two discrete spectra knows only about the discreteness itself. The surprise is the opposite: the magnitude is set by the ramp, the dynamical wormhole correlation, with the discreteness entering only as its regulator, and the mean comes out exponentially small, $\langle \mathcal{W}^2 \rangle \sim S\,\mathrm{e}^{-S}$.[2] Two draws of the same $O(1)$-temperature theory are thus $\mathrm{e}^{-S/2}$-close. The scale to hold this against: each thermal operator has squared norm $\mathrm{Tr}\,\mathrm{e}^{-2\beta H} = Z(2\beta) \sim \mathrm{e}^{S}$, so the *normalized* separation is $\langle \mathcal{W}^2 \rangle / Z(2\beta) \sim S\,\mathrm{e}^{-2S}$, against e^{-S} for uncorrelated levels; unless stated otherwise, we quote distances in the unnormalized units of the definition above, as the resolution-floor discussion of Section 7 switches to the normalized form. The linear entropy factor is itself a diagnostic of chaos: an uncorrelated (integrable) spectrum, all plateau and no ramp, gives an unsuppressed $O(1)$ distance instead (Section 4.2). The same two-level correlator

[2] Throughout, S_0 denotes the extremal entropy, the logarithm of the e^{S_0} near-extremal ground-state degeneracy, and $S = S(E) = \ln \bar{n}(E)$ the microcanonical entropy; at the $O(1)$ temperatures where our headline scalings are quoted the two agree up to subleading terms and are used interchangeably, the relation $S = S_0 - \frac{3}{2}\ln\beta + O(1)$ being made precise in Section 6.

fixes the variance (two double trumpets), and with it the *shape* of the cloud: an exponentially tight, strongly anisotropic shell whose effective dimension *grows* with the entropy, asymptotically as $D_{\text{eff}} \simeq 0.42\, S^2$, so the distance self-averages. We make the picture concrete in JT gravity and corroborate it with exact-diagonalization SYK and GUE, finding the cloud's principal axes to be the typical spectral-density fluctuations of its members.

Scope. The explicit results (the rigidity $\Sigma^2 \simeq \frac{1}{2\pi^2} \ln \bar{n}$, the distance $\langle \mathcal{W}^2 \rangle \sim S_0\, \mathrm{e}^{-S_0}$, and the self-averaging $D_{\text{eff}} \sim S^2$) are specific to black holes with a *large degeneracy at the bottom of the chaotic band*: the near-extremal regime, whose throat is the Schwarzian/JT matrix model. The mechanism is the weighting: $\mathrm{e}^{-2\beta E}/\bar{\rho}(E)$ concentrates the distance at the bottom of the spectrum, and an extremal degeneracy keeps the density there exponentially large. The relevant statistics are then the universal bulk ramp with a controlled Airy edge, and the e^{S_0} floor states supply exponentially many independent fluctuations, enough to self-average. Without such a degeneracy (non-rotating BTZ, a generic large AdS black hole) the same weighting localizes onto the few lowest, sparsely spaced microstates and the distance is expected to be power-law rather than e^{-S}. We therefore present the results as statements about the near-extremal regime; the non-extremal case still *poses* a sharp question, the rigidity of the lowest black-hole microstates, but not one we answer here.

Outline and relation to prior work. Section 2 reviews standard JT material; what is carried forward is the energy-resolved spectral form factor and the identification of the late-time plateau with the coincident-energy cutoff. Readers familiar with these may skip it. Section 3 reviews the BW distance (kept brief, with details in [1]) and computes its ensemble cumulants: Wick contractions organize them into necklace diagrams whose propagators are number covariances, the mean being the Boltzmann-weighted number variance, $\langle \mathcal{W}^2 \rangle = 2\beta^2 \int \mathrm{d}E\, \mathrm{e}^{-2\beta E}\, \Sigma^2(E)/\bar{\rho}(E)$. Section 4 evaluates the mean and the variance in JT. The number-variance/form-factor relation is classical random-matrix theory; the version needed here differs in that only the window's upper edge contributes, which we show explicitly in JT; we have not found this precise statement in the literature, though we suspect it is not new. The mean comes out $\sim S_0\, \mathrm{e}^{-S_0}$ and the variance $\sim \mathrm{e}^{-2S_0}$; the factor of S_0, the logarithm of the plateau scale, marks the distance as non-perturbative in the genus expansion. Rigorous bounds on Wasserstein distances between random-matrix spectra exist in the mathematics literature; our additions are the Boltzmann-weighted, unnormalized setting, the gravitational identification, and the explicit variance. Section 5 contains the main new results: the geometry of the ensemble as seen by the BW metric: the principal spectral-density fluctuations by which typical members differ (a Karhunen–Loève analysis of the wormhole covariance) and the effective dimensionality of the cloud. Section 6 develops the low-temperature limit, where the distance becomes a Tracy–Widom/FZZT-brane observable and the effective dimension a second Rényi entropy of the brane ensemble (the ingredients standard, the Rényi identification new), and closes with exact-diagonalization SYK, kept deliberately modest in scale, whose purpose is to exhibit the microscopic ensemble as a point cloud realizing the geometric picture. Section 7 summarizes and collects future directions; the appendices gather the SYK and GUE numerics.

2 Quick review of the gravitational ensemble and the spectral form factor

This section collects the standard facts we use later: that the gravitational path integral computes ensemble averages, that the leading connected two-boundary amplitude is the double trumpet, and that the spectral form factor can be written as a Fourier transform in the *energy difference*. Readers fluent in [2] can skip to Section 3; the one object worth isolating is the local form factor (4), which Section 3 feeds directly into the number variance.

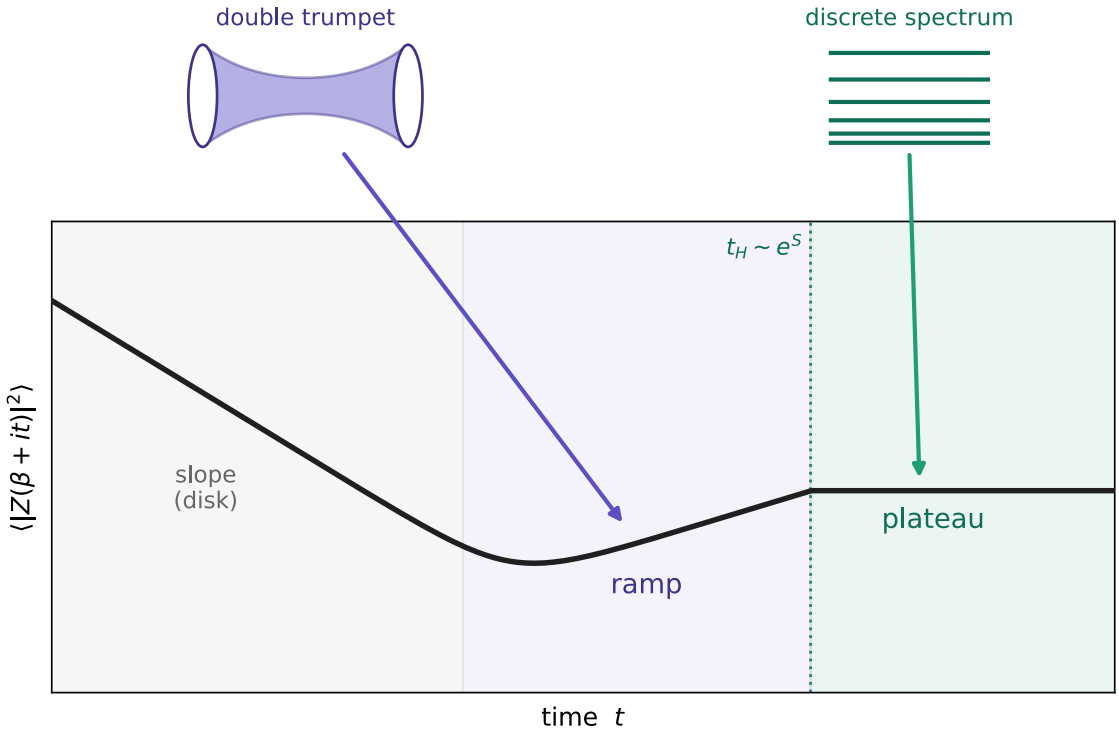


Figure 1: Anatomy of the spectral form factor (schematic). The early-time *slope* is the disconnected/disk contribution; the linearly rising *ramp* is produced by the connected double-trumpet (cylinder) two-level correlator; the late-time *plateau* reflects the discreteness of the spectrum and sets in at the Heisenberg time $t_H \sim \mathrm{e}^S$. The same two ingredients, ramp and plateau, organize the mean and variance of the BW distance below.

2.1 The ensemble and the double trumpet

A boundary partition function $Z(\beta) = \mathrm{Tr}\,\mathrm{e}^{-\beta H}$ is computed in the bulk by summing over geometries with one asymptotic boundary of renormalized length β. Products $Z(\beta_1)\cdots Z(\beta_k)$ additionally receive contributions from geometries with several boundaries, including *connected* ones (wormholes), so the products do not factorize, $\langle Z(\beta_1)Z(\beta_2)\rangle \neq \langle Z(\beta_1)\rangle\langle Z(\beta_2)\rangle$. The standard reading [2, 3] is that the path integral computes *averages* over an ensemble of boundary theories, the connected geometries supplying the ensemble cumulants. We write $\langle\cdot\rangle$ for this average and $\langle\cdot\rangle_{\mathrm{c}}$ for its connected part. In JT gravity the leading connected piece is the *double trumpet*, two trumpets (i.e. geometries ending on a hole cut into the bulk by hand, a geodesic of circumference b) glued along this hole and the $b\,\mathrm{d}b$ integral below then sums over the sizes of the hole

$$\langle Z(\beta_1)Z(\beta_2)\rangle_{\mathrm{c}} = \int_0^\infty b\,\mathrm{d}b\; Z_{\mathrm{tr}}(\beta_1, b)\, Z_{\mathrm{tr}}(\beta_2, b) = \frac{\sqrt{\beta_1\beta_2}}{2\pi(\beta_1+\beta_2)}, \tag{1}$$

with $Z_{\mathrm{tr}}(\beta, b) = \mathrm{e}^{-b^2/4\beta}/(2\sqrt{\pi\beta})$ the JT gravity Euclidean path integral in the trumpet geometry.

2.2 The spectral form factor from the double trumpet

The form factor the double trumpet gives. Continuing the two boundary lengths to complex-conjugate values $\beta_{1,2} = \beta \pm it$ turns (1) into the connected *spectral form factor*,

$$\big\langle Z(\beta+it)\, Z(\beta-it)\big\rangle_{\mathrm{c}} = \frac{\sqrt{\beta^2+t^2}}{4\pi\beta} \xrightarrow{t\gg\beta} \frac{|t|}{4\pi\beta}, \tag{2}$$

rising from $\langle Z(\beta)^2\rangle_{\rm c} = 1/4\pi$ at $t = 0$ to the linear *ramp* $|t|/4\pi\beta$.

Resolving in energy. Fixed-energy observables need the form factor before the thermal smearing. The double inverse Laplace transform of (1) is the connected density–density correlator,

$$\big\langle \rho(E_1)\rho(E_2)\big\rangle_{\rm c} = -\frac{1}{4\pi^2}\,\frac{E_1+E_2}{\sqrt{E_1E_2}\,(E_1-E_2)^2} \xrightarrow{E_1\to E_2} -\frac{1}{2\pi^2(E_1-E_2)^2}, \tag{3}$$

where we record the near coincident energy limit for future use. Holding the mean energy $E = \frac{1}{2}(E_1+E_2)$ fixed and Fourier transforming only in $\omega = E_1 - E_2$ defines the *local* (microcanonical) form factor [4, 5],

$$\mathcal{K}(E,t) \equiv \int \mathrm{d}\omega\, \mathrm{e}^{i\omega t}\,\big\langle \rho(E+\tfrac{\omega}{2})\,\rho(E-\tfrac{\omega}{2})\big\rangle_{\rm c}, \tag{4}$$

the energy-resolved object the rest of the paper feeds into the number variance. It is related to the β-space version (2) through the Boltzmann smearing[3]

$$\langle Z(\beta+it)Z(\beta-it)\rangle_{\rm c} = \frac{1}{4\pi} + \int \mathrm{d}E\, \mathrm{e}^{-2\beta E}\,\mathcal{K}(E,t), \tag{5}$$

the constant being the spectral-edge $\frac{1}{4\pi}\,\delta(E)$ term written out, so that the integral carries only the bulk kernel. Carried out on (3), the transform (4) closes in Bessel functions,

$$\mathcal{K}(E,t) = \frac{|t|}{2\pi}\Big[\int_0^{2E|t|} J_0(s)\,\mathrm{d}s - J_1\big(2E|t|\big)\Big], \tag{6}$$

that is, $|t|$ times a function of $\lambda \equiv 2E|t|$ alone.

The ramp and a Thouless time. At late times (6) is the linear ramp,

$$\mathcal{K}(E,t) \xrightarrow{\lambda\gg 1} \frac{|t|}{2\pi} \tag{7}$$

The long time limit is related to small $\omega = E_1 - E_2$ by the Fourier transform (4) as ω and t are Fourier conjugates. Thus, the ramp is generated by the nearly coincident energy region of the density correlator $\propto 1/\omega^2$. On the other hand at early times $\mathcal{K} \to E\,t^2/2\pi$. Their crossover at $\lambda \sim 1$ is a *Thouless time*

$$t_{\rm dip}(E) \sim \frac{1}{E}, \tag{8}$$

the time to resolve the bandwidth (the band runs from the edge 0 up to E).

The plateau is non-perturbative. There is a maximum time up to which the ramp can continue. This can be seen from the energy space density correlator, as the smooth part creating the ramp $\propto 1/(E_1-E_2)^2$ is cut off in a system with discrete energy levels, at the level spacing $\omega = E_1 - E_2 \sim 1/\bar{\rho}(E)$. Again by the Fourier conjugacy a minimum in ω relates to a maximum time $t_H \sim 1/\omega_{\rm min} \sim \bar{\rho}(E)$ where t_H denotes the Heisenberg time defined as $t_H = 2\pi\bar{\rho}(E)$. So the true form factor saturates as,

$$\mathcal{K}(E,t) = \begin{cases} |t|/2\pi, & t_{\rm dip} \lesssim |t| < t_H \quad \text{(ramp)},\\ \bar{\rho}(E), & |t| > t_H \quad \text{(plateau)}. \end{cases} \tag{9}$$

[3]The δ-term bookkeeping: Boltzmann-averaging the bulk kernel (6) alone gives $\int_0^\infty \mathrm{d}E\,\mathrm{e}^{-2\beta E}\mathcal{K}(E,t) = \sqrt{\beta^2+t^2}/4\pi\beta - \frac{1}{4\pi}$: the constant deficit is the static value $\langle Z(\beta)^2\rangle_{\rm c} = \mathrm{Var}Z(\beta) = \frac{1}{4\pi}$, β-independent. Being the large-β asymptote of (2), it inverse-Laplace-transforms ($2\beta \to E$) to $\frac{1}{4\pi}\,\delta(E)$: in the energy-resolved bookkeeping the variance sits entirely at the spectral edge, while the Hadamard-regulated bulk $\mathcal{K}$, with $\mathcal{K}(E,0) = 0$, carries none of it. Every observable below uses the bulk $\mathcal{K}$ with the edge treated separately (Section 6), so the two are never double-counted.

The *plateau* comes not from any smooth geometry but from the spectral *discreteness*: it is the Fourier image of a contact term $\bar{\rho}(E)\,\delta(\omega)$, absent from the perturbative correlator (3) and non-perturbative in e^{-S} (the spectrum is a sum of e^{S} sharp δ-functions, the individual microstates; in the matrix model the eigenvalue-instanton / brane completion [6, 7, 10]). It is what cancels the ramp's continued growth at t_H. An effective-theory derivation of the plateau, formulated without reference to an ensemble, is given in [11, 12].

3 The Boltzmann–Wasserstein distance and its cumulants

For two *independent* draws from any common ensemble, the BW distance reduces to the trace of a single covariance operator C, and when the count fluctuations in the ensemble are Gaussian all of its cumulants form a closed tower $\kappa_p \propto \mathrm{Tr}C^p$. The gravitational evaluation of the first two cumulants, where the double trumpet enters, is deferred to Section 4.

3.1 The distance and the exact identity

The BW distance between two theories with Hamiltonians H_1, H_2 on a common $\mathcal{D}$-dimensional Hilbert space, at inverse temperature β, is [1]

$$\mathcal{W}^2(\beta) = \min_{U\in \mathrm{U}(\mathcal{D})} \mathrm{Tr}\big(\mathrm{e}^{-\beta H_1} - U\,\mathrm{e}^{-\beta H_2}\,U^\dagger\big)^2 : \tag{10}$$

the least-squares mismatch between the two unnormalized thermal operators, evaluated after the best possible unitary alignment of their eigenbases. The minimization is what makes this a sensible distance between *theories*: without it, the trace would depend on the arbitrary relative orientation of the two eigenbases; with it, that orientation is optimized away and only the spectra can contribute.

Expanding the square in (10),

$$\mathcal{W}^2(\beta) = Z_1(2\beta) + Z_2(2\beta) - 2\max_U \mathrm{Tr}\big(\mathrm{e}^{-\beta H_1}\,U\,\mathrm{e}^{-\beta H_2}\,U^\dagger\big),$$

with $Z_i(2\beta) = \mathrm{Tr}\,\mathrm{e}^{-2\beta H_i}$: the first two terms are U-independent, so minimizing the distance means *maximizing* the overlap. The maximum is a classical fact: by the von Neumann rearrangement inequality, $\mathrm{Tr}(AUBU^\dagger)$ is largest when the eigenvalues of A and B are paired in matching order: largest weight with largest weight, i.e. lowest energy with lowest energy. Denoting the ordered eigenvalues of each theory by $E_n^{(i)}$ (with $E_1^{(i)} \le E_2^{(i)} \le E_3^{(i)} \le ...$) and defining $a_n = \mathrm{e}^{-\beta E_n^{(1)}}$, $b_n = \mathrm{e}^{-\beta E_n^{(2)}}$ gives the maximum $\max_U \mathrm{Tr}(AUBU^\dagger) = \sum_n a_n b_n \equiv C_{\max}(\beta)$. This gives the perfect-square form

$$\mathcal{W}^2(\beta) = Z_1(2\beta) + Z_2(2\beta) - 2\,C_{\max}(\beta) = \sum_n \big(a_n - b_n\big)^2, \tag{11}$$

the statement that for one-dimensional optimal transport the monotone coupling is optimal [13] (for genuinely quantum formulations of optimal transport see [14, 15]; a different use of Wasserstein geometry in holography appears in [16, 17]: transport between phase-space representations of individual states *within* one system, the resulting space conjectured to play the role of the emergent bulk). In words: the optimal U compares the two spectra *level by level*, the n-th lowest of one draw against the n-th lowest of the other, and the distance is the Boltzmann-weighted sum of the squared level mismatches. The fact that the BW distance is a genuine metric in the space of spectra is immediate from the vector ℓ^2 norm form of (11).

Now let H_1, H_2 be *independent, identically distributed* draws from a common ensemble $\langle\,\cdot\,\rangle$ (a disorder average, or the gravitational ensemble). The average of (11) is then immediate:

independence factorizes the cross term, $\langle a_n b_n\rangle = \langle a_n\rangle\langle b_n\rangle$, and identical distribution equates the marginals, $\langle a_n^k\rangle = \langle b_n^k\rangle$, so

$$\langle \mathcal{W}^2\rangle = \sum_n \big\langle (a_n - b_n)^2\big\rangle = 2\sum_n \mathrm{Var}\big(\mathrm{e}^{-\beta E_n}\big). \tag{12}$$

This is exact, with no perturbative input, only the assumption that the draws are i.i.d.: the mean self-distance is twice the summed variance of the rank-ordered Boltzmann weights. The per-level statistics entering here — the means and variances of the individual low-lying levels — are exactly what Johnson's Fredholm-determinant technology computes for a non-perturbative completion of JT [8, 9], there in service of the quenched free energy; (12) reads the same data as a distance.

Where the second boundary comes from. The observable (10) is a single trace, and its semiclassical avatar in the companion paper is correspondingly a single-boundary geometry, the Schwinger–Keldysh disk computing $C_{\max}$ [1]; yet this paper computes it from the double trumpet, a two-boundary amplitude. There is no contradiction: the number of traces is tied to the topology of the spacetimes computing an observable only *perturbatively*, in the genus expansion, and the BW self-distance, as a functional of the smooth spectral data, vanishes to all orders in that expansion. Smooth corrections shift the two draws' spectra identically, the comonotone map between identical smooth spectra degenerates to the identity, and $\langle \mathcal{W}^2\rangle = \langle Z(2\beta)\rangle + \langle Z(2\beta)\rangle - 2\langle Z(2\beta)\rangle = 0$. An observable invisible to the genus expansion is under no obligation to obey its boundary counting, and the first non-vanishing contribution indeed violates it: the plateau-regulated cylinder of Section 4, which renders $\langle \mathcal{W}^2\rangle$ a non-analytic function of the genus expansion parameter e^{-S_0}. It takes two boundaries to know a variance.[4]

3.2 The continuum representation

Equation (12) is a rank sum of the variance of $\mathrm{e}^{-\beta E_n}$; two short steps turn it into a continuum representation. Both spectra fluctuate about the common mean staircase $\{\bar{E}_n\}$, defined by $\bar{n}(\bar{E}_n) = n$, or through the spectral density as $\bar{n}(\bar{E}_n) = \int_{E_0}^{\bar{E}_n} dE'\bar{\rho}(E')$. Writing $E_n^{(i)} = \bar{E}_n + \delta E_n^{(i)}$, $\mu_n \equiv \mathrm{e}^{-\beta\bar{E}_n}$, and linearizing the weight in the small displacement $\delta E \sim 1/\bar{\rho}$,

$$a_n - b_n \simeq -\beta\, e^{-\beta\bar{E}_n}\,(\delta E_n^{(1)} - \delta E_n^{(2)}) \implies \mathcal{W}^2 \simeq \beta^2 \sum_n e^{-2\beta\bar{E}_n}\,(\delta E_n^{(1)} - \delta E_n^{(2)})^2. \tag{13}$$

When the energy levels fluctuate from one ensemble member to another, the count function must also fluctuate to keep the rank of the eigenvalue $E_n = \bar{E}_n + \delta E_n$ at n

$$n = \bar{n}(\bar{E}_n + \delta E_n) + \delta n(\bar{E}_n + \delta E_n) \approx \bar{n}(\bar{E}_n) + \bar{\rho}(\bar{E}_n)\delta E_n + \delta n(\bar{E}_n), \tag{14}$$

and using $\bar{n}(\bar{E}_n) = n$ leads to $\delta E_n \approx -\delta n(\bar{E}_n)/\bar{\rho}(\bar{E}_n)$. Substituting to (13) leads to

$$\mathcal{W}^2 \simeq \beta^2 \sum_n e^{-2\beta\bar{E}_n}\,\frac{(n_1(\bar{E}_n) - n_2(\bar{E}_n))^2}{\bar{\rho}(\bar{E}_n)^2}, \tag{15}$$

where we used the fact that $\delta n_i(\bar{E}_n)$ share the same mean staircase $\bar{n}$ that cancels from their difference $\delta n_1 - \delta n_2 = n_1 - n_2$. Trading the rank sum for an energy integral, $\sum_n = \int \bar{\rho}\,\mathrm{d}E$ (one level per unit $\bar{n}$), turns the random variable $\mathcal{W}^2$ itself into a quadratic functional of a single fluctuation field,

$$\mathcal{W}^2(\beta) \;\simeq\; \beta^2 \int \mathrm{d}E\,\frac{\mathrm{e}^{-2\beta E}}{\bar{\rho}(E)}\,\Delta n(E)^2, \qquad \Delta n(E) \equiv n_1(E) - n_2(E), \tag{16}$$

[4] The eigenbrane construction of [18] makes the fluctuation content tangible: fixing the individual levels with eigenbrane pairs would set $\mathcal{W}^2 = 0$ draw by draw: the distance measures precisely the fluctuation the branes would remove.

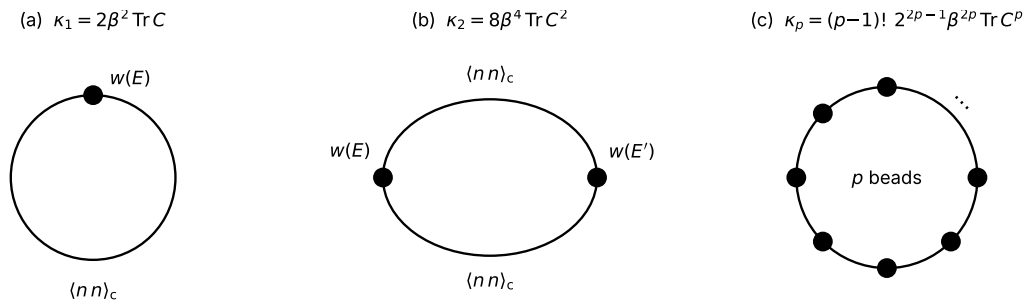


Figure 2: The cumulants of $\mathcal{W}^2$ as traces of the covariance operator C. Each vertex is a factor of $\mathcal{W}^2$ (an energy integral against the thermal weight $\mathrm{e}^{-2\beta E}/\bar{\rho}$), and each link is a connected number-fluctuation correlator $\langle n(E)\, n(E')\rangle_{\mathrm{c}}$, the Wick contraction of two Δn fields. (a) The mean $\kappa_1 = 2\beta^2 \mathrm{Tr}C$ is a single self-contracted kernel. (b) The variance $\kappa_2 = 8\beta^4 \mathrm{Tr}C^2$ is two kernels joined by two correlators. (c) The p-th cumulant $\kappa_p \propto \mathrm{Tr}C^p$ is a closed ring of p kernels linked by p correlators.

the difference of the two draws' counting functions ($n_i(E)$ = number of states below E). Using the fact that n_1 and n_2 are independent draws from the ensemble, $\langle n_1 n_2\rangle = \langle n_1\rangle\langle n_2\rangle = \langle n\rangle^2$ and $\langle n_i n_i\rangle = \langle nn\rangle$ leads to

$$\big\langle \Delta n(E)\, \Delta n(E')\big\rangle = \langle (n_1(E) - n_2(E))(n_1(E') - n_2(E'))\rangle = 2\,\big\langle n(E)\, n(E')\big\rangle_{\mathrm{c}}, \tag{17}$$

which we will use later.

3.3 The covariance operator and the cumulant tower

Taking the ensemble expectation value of (16) and identifying $\langle n(E)^2\rangle_c = \mathrm{Var}\, n(E) = \Sigma^2(E)$ as the *number variance*,

$$\langle \mathcal{W}^2\rangle = 2\beta^2 \int \mathrm{d}E\, \frac{\mathrm{e}^{-2\beta E}}{\bar{\rho}(E)}\, \Sigma^2(E) \;=\; 2\beta^2\, \mathrm{Tr}C, \tag{18}$$

where the second form names the object that controls every cumulant: the covariance operator C on $L^2(\mathrm{d}E)$, with symmetric kernel

$$C(E, E') = \frac{\mathrm{e}^{-\beta(E+E')}}{\sqrt{\bar{\rho}(E)\,\bar{\rho}(E')}}\, \big\langle n(E)\, n(E')\big\rangle_{\mathrm{c}}. \tag{19}$$

It is the count correlator $\langle n\, n\rangle_{\mathrm{c}}$ of (17) dressed by the Boltzmann weight that selects the thermal shell; the $1/\sqrt{\bar{\rho}}$ factors are positional, turning count fluctuations into the level displacements the optimal-transport (comonotone) map moves. Its diagonal is $C(E, E) = (\mathrm{e}^{-2\beta E}/\bar{\rho})\, \Sigma^2(E)$, so the integral is $2\beta^2 \mathrm{Tr}C$; its eigenmodes are the principal axes of the theory cloud (Section 5).

The counting function, and thus Δn of (16), is an asymptotically Gaussian field. This follows from the central limit theorem for determinantal linear statistics [19, 20, 22], together with the joint CLT for eigenvalue positions [21]. Since $\mathcal{W}^2$ is a quadratic form in the Gaussian field, Wick's theorem makes *every* cumulant a single trace of C,

$$\kappa_p(\mathcal{W}^2) \approx (p-1)!\, 2^{2p-1}\, \beta^{2p}\, \mathrm{Tr}\big(C^p\big)\ . \tag{20}$$

Diagrammatically $\mathrm{Tr}C^p$ is a closed *necklace of p kernels*: p thermal beads $\mathrm{e}^{-2\beta E}/\bar{\rho}$ on a ring, linked by p count correlators $\langle n\, n\rangle_{\mathrm{c}}$ (Fig. 2).

Validity. Three approximations enter (13)–(16), all controlled at the dominant energies: the weight linearization drops $O(\beta/\bar{\rho})$; the rank-sum-to-integral step is accurate to $O(1/\bar{n})$; and the smooth density $\bar{\rho}$ is frozen at its mean. One correction in this class can be identified: the distance is built from the *level-position* variance $\bar{\rho}^2\,\mathrm{Var}(E_n)$, which the continuum representation replaces by the *count* variance Σ^2. The two agree at leading logarithm but differ by a constant, $\bar{\rho}^2\,\mathrm{Var}(E_n) = \Sigma^2 - \frac{1}{12}$: heuristically, the 1/12 is what the fractional part of an integer-valued count, uniform and decoupled from the smooth part up to corrections expected to be exponentially small in Σ^2, contributes to Σ^2; we have not derived it, but it is what GUE measures (0.085 ± 0.004, App. B). It is immaterial wherever only the leading logarithm is kept, and it matters exactly where the additive constants do: the finite-N comparison of App. B and the constant quoted in Section 5.3. Because the distance is set by the *positional* fluctuations of individual levels, the coherent breathing of $\bar{\rho}$ is a parametrically separate effect: reinstating it corrects the level-position variance by a connected three-level term (a pair of pants $\langle\rho\rho\rho\rangle_{\mathrm{c}}$) and a two-cylinder term; a power counting using the exact $Z_{0,3}$ amplitude shows both to be exponentially suppressed relative to the leading rigidity, with the weakest suppression coming from the spectral threshold, far below the accuracy at which we work. (That estimate is a statement about the matrix-model ensemble; a *microscopic* disorder ensemble also carries smooth-density fluctuations that are only power-law small in the number of couplings (parametrically larger than any of these), and these must be unfolded away before comparing with the wormhole predictions, Section 6.1.) Finally, the Gaussian (Wick) step is underwritten by the central limit theorems cited above; its accuracy is logarithmic rather than exponential (the higher cumulants of δn saturate at finite Fisher–Hartwig constants, quantified in Section 4.3), and it degrades at the soft spectral edge, where the lowest levels obey Tracy–Widom [23] rather than Gaussian statistics, the operative caveat of the low-temperature regime (Section 6).

4 The first two cumulants in gravity

We now evaluate the first two cumulants of the BW distance for the gravitational ensemble. The only input is the connected density correlator which fixes the count correlator $\langle n\,n\rangle_{\mathrm{c}}$, hence the operator C of (19): its trace is the mean (Section 4.2), the trace of its square the variance (Section 4.3). We begin with the number variance $\Sigma^2 = \langle n(E)\,n(E)\rangle_{\mathrm{c}}$, the diagonal of C.

The logic, summarized in Fig. 3, runs in three steps. First, the number variance is a windowed time integral of the form factor, so the double trumpet fixes it up to the coincident-energy region, where the plateau takes over: $\Sigma^2 \simeq \frac{1}{2\pi^2}\ln\bar{n}$ (Section 4.1). Second, feeding this into (18) gives the mean, $\langle\mathcal{W}^2\rangle \sim S_0\,\mathrm{e}^{-S_0}$. Third, the variance is the same correlator squared, and needs no plateau cutoff at leading order, so it carries no factor of S_0 (Section 4.3); the mismatch between the two is what later drives the growth of the effective dimension (Section 5).

4.1 Number variance and the spectral form factor

The diagonal of C is the number variance $\Sigma^2(\bar{E}_n) = \mathrm{Var}\,n(\bar{E}_n)$. Taking the window from the bottom of the spectrum up to E, it is the double integral of the connected density–density correlator,

$$\Sigma^2(E) = \int_0^E\!\!\int_0^E \mathrm{d}E_1\,\mathrm{d}E_2\,\big\langle\rho(E_1)\rho(E_2)\big\rangle_{\mathrm{c}}. \tag{21}$$

This is where gravity enters: the connected correlator $\langle\rho\rho\rangle_{\mathrm{c}}$ is exactly what the two-boundary path integral computes: the double trumpet (1), equivalently (3). At the level of the perturbative wormhole, the BW distance is then nothing but a double energy integral of the double trumpet.

The coincident-energy singularity. Taken at face value, however, (21) on the double trumpet diverges: the ramp $-1/2\pi^2(E_1 - E_2)^2$ in (3) is non-integrable as $E_1 \to E_2$. The divergence

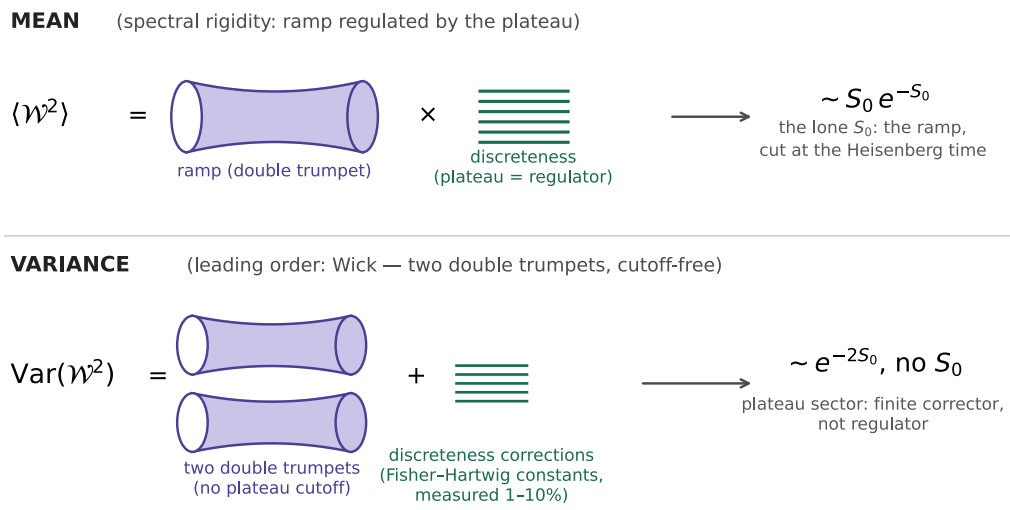


Figure 3: Mean versus variance of the BW distance (schematic). The mean $\langle \mathcal{W}^2 \rangle \sim S\,\mathrm{e}^{-S}$ is *one* power of the connected correlator (one double trumpet, i.e. the ramp) regulated by the spectral discreteness (the plateau). The variance $\mathrm{Var}(\mathcal{W}^2) \sim \mathrm{e}^{-2S}$ is its entrywise *square* (two double trumpets), whose leading (Wick) term is built from the ramp alone and needs no plateau cutoff (Section 4.3 for the finite discreteness corrections). Their ratio sets the effective dimension, $D_{\mathrm{eff}} \sim S^2$.

is spurious: it is cut off by the spectral discreteness, the *plateau* of the form factor, which is non-perturbative in e^{-S} and invisible in the smooth double trumpet on its own. So the double trumpet hands us the correlator immediately, but treating it *precisely* requires supplying the plateau as a coincident-energy regulator. There is no tension with the all-orders vanishing of Section 3.1: an observable that vanishes identically order by order can surface in the expansion only through an expression the expansion cannot evaluate on its own, and the non-integrable coincident-energy singularity is precisely that signal, resolved by the non-perturbative plateau.

The form factor, exactly. An illustrative way to handle the singularity is to pass to the time domain, where the discreteness is just the statement that the form factor stops growing beyond the Heisenberg time. Write the connected correlator as the inverse transform of the local form factor (4),

$$\Sigma^2 = \int_0^E d\bar{E} \int_{-a(\bar{E})}^{a(\bar{E})} \mathrm{d}\omega \langle \rho(\bar{E}+\omega/2)\rho(\bar{E}-\omega/2)\rangle_{\mathrm{c}} = \int_0^E d\bar{E} \int_{-a(\bar{E})}^{a(\bar{E})} \mathrm{d}\omega \int \frac{\mathrm{d}t}{2\pi}\, \mathrm{e}^{-i\omega t}\, \mathcal{K}(\bar{E},t) \quad (22)$$

with $\bar{E} = \frac{1}{2}(E_1+E_2)$, $\omega = E_1 - E_2$. In the mean/difference variables the square integration region $E_{1,2} \in [0, \bar{E}_n]$ becomes a diamond: $\bar{E}$ runs across the window while $|\omega| \le a(\bar{E}) \equiv 2\min(\bar{E},\, \bar{E}_n - \bar{E})$; here $a(\bar{E})$ is simply twice the distance from $\bar{E}$ to the nearer window edge. The ω-integral is then elementary, $\int_{-a}^{a} \mathrm{d}\omega\, \mathrm{e}^{-i\omega t} = 2\sin(at)/t$, leaving a single mean-energy integral across the window, with both window edges intact:

$$\Sigma^2(E) = \int_{-\infty}^{\infty} \frac{\mathrm{d}t}{2\pi} \int_0^E \mathrm{d}\bar{E}\, \frac{2\sin\big(a(\bar{E})\,t\big)}{t}\, \mathcal{K}(\bar{E},t), \qquad a(\bar{E}) = 2\min(\bar{E},\, E-\bar{E}). \quad (23)$$

This is the general relation between the number variance and the spectral form factor.[5]

[5]The relation between the number variance and the form factor is itself classical in random-matrix theory: the rigidity as a windowed time integral of the form factor goes back to Berry's semiclassical analysis [24] and the Dyson–Mehta statistics [25], and its classic dynamical application is the number variance of disordered metals computed from the diffusive form factor [26]; (23) is its finite-window form with both window edges kept.

The window splits at its midpoint. The window enters (23) only through $a(\bar{E})$, twice the distance to the *nearer* edge, so the integral splits exactly at the midpoint into two halves. The two halves see very different physics: the upper half sits deep in the spectrum, while the lower half rides on the spectral threshold. We take them in turn.

Upper half: freezing the form factor. In the upper half $a = 2(E - \bar{E})$, and the only obstacle to doing the $\bar{E}$-integral outright is the $\bar{E}$-dependence of $\mathcal{K}$. We can trade it for a derivative: integrating by parts with the antiderivative anchored at the midpoint, $U(\bar{E}, t) = \int_{E/2}^{\bar{E}} \mathrm{d}\bar{E}'\, 2\sin\big(2(E-\bar{E}')t\big)/t = \big[\cos\big(2(E-\bar{E})t\big) - \cos(Et)\big]/t^2$ with $U(E/2, t) = 0$,

$$\Sigma^2_{\text{upper}} = \int \frac{\mathrm{d}t}{2\pi}\, \frac{1-\cos(Et)}{t^2}\, \mathcal{K}(E,t) \;-\; \int \frac{\mathrm{d}t}{2\pi} \int_{E/2}^{E} \mathrm{d}\bar{E}\, U(\bar{E},t)\, \partial_{\bar{E}}\mathcal{K}(\bar{E},t). \tag{24}$$

The first term is a time integral of the spectral form factor weighted by the *Fejér kernel* $(1 - \cos Et)/t^2$. The interesting physics lives in the first term while the second term from integration by parts results in non-universal order one corrections.[6]

The ramp diverges; the plateau is the cutoff. Evaluating the boundary term on the ramp alone exposes the physics:

$$\int \frac{\mathrm{d}t}{2\pi}\, \frac{1-\cos(Et)}{t^2}\, \frac{|t|}{2\pi} = \frac{1}{2\pi^2} \int_0^\infty \frac{1-\cos(Et)}{t}\, \mathrm{d}t$$

while the oscillatory cosine piece decays at late times, the non-oscillating $\int \mathrm{d}t/t$ diverges logarithmically. The $1/t$ tail is the announcement that the smooth ramp cannot by itself answer a counting question about discrete levels; the divergence is cut off precisely by the plateau at the Heisenberg time $t_H = 2\pi\bar{\rho}(E)$, the coincident-energy regulator of (9) doing its job. With the ramp below t_H and the plateau above, the two regions are additive and elementary,

$$\Sigma^2_{\text{ramp}} = \frac{1}{2\pi^2} \int_0^{t_H} \frac{1-\cos(Et)}{t}\, \mathrm{d}t = \frac{1}{2\pi^2}\big[\ln(2\pi\bar{n}) + \gamma - \mathrm{Ci}(2\pi\bar{n})\big], \tag{25}$$

$$\Sigma^2_{\text{plateau}} = \frac{\bar{\rho}}{\pi} \int_{t_H}^\infty \frac{1-\cos(Et)}{t^2}\, \mathrm{d}t = \frac{1-\cos(2\pi\bar{n})}{2\pi^2} + \frac{\bar{n}}{\pi}\Big(\frac{\pi}{2} - \mathrm{Si}(2\pi\bar{n})\Big), \tag{26}$$

both at $E\, t_H = 2\pi\bar{n}$, with $\bar{n} \equiv E\,\bar{\rho}(E)$ the frozen-density count (equal to the true $\bar{n}(E)$ up to an $O(1)$ factor, absorbed in the additive constant below), the logarithm and its argument coming from the ramp, the additive constant from the plateau. For what follows the $\ln\bar{n} \propto S_0$ piece in (25) is the main interesting one as it is much larger than any order one corrections.

Lower half: the threshold is silent. In the upper half the ramp/plateau parts are enabled by the upper limit of the $\bar{E}$ integral at $\bar{E} = E$ which, together with the ramp, gives rise to the non-oscillating $1/t$ term. In the lower half, the analogous non-oscillating piece would come from the opposite edge $\bar{E} = 0$, but there the ramp would need times larger than the Thouless time $\sim 1/\bar{E} \to \infty$: the same boundary mechanism as in the upper half cannot give rise to a log term.

There is a second possible mechanism, distinct from the boundary term: the t-integral at fixed $\bar{E}$ could produce a contribution proportional to $1/\bar{E}$, whose $\bar{E}$-integral would then accumulate a log divergence at the lower limit $\bar{E} = 0$, cut off only by the discreteness of the lowest levels. To rule this out we calculate the t integral explicitly in JT. Using the exact form

[6] Quantitatively: on the plateau $\partial_{\bar{E}}\mathcal{K} = \bar{\rho}'(\bar{E})$ and $|U| \le 2/t^2$, bounding the drift piece by $\frac{1}{2\pi^2}\ln[\bar{\rho}(E)/\bar{\rho}(E/2)]$, an $O(1)$ constant with no factor of S, both densities carrying the same e^{S_0}. The Thouless-scale piece, $\partial_{\bar{E}}\mathcal{K} = |t|J_1(2\bar{E}|t|)/2\pi\bar{E}$ from (6), converges absolutely against U and is likewise $O(1)$.

factor (6), $\mathcal{K}(\bar{E},t) = \frac{|t|}{2\pi}F(2\bar{E}|t|)$ with $F(\lambda) = \int_0^\lambda J_0(s)\,\mathrm{d}s - J_1(\lambda)$, and substituting $u = 2\bar{E}t$, the $\bar{E}$ integrand in the lower window becomes

$$\left.\frac{\mathrm{d}\,\Sigma^2}{\mathrm{d}\bar{E}}\right|_{\text{lower}} = \int_{-\infty}^{\infty}\frac{\mathrm{d}t}{2\pi}\,\frac{2\sin(2\bar{E}t)}{t}\,\mathcal{K}(\bar{E},t) = \frac{1}{2\pi^2\bar{E}}\int_0^\infty \mathrm{d}u\ \sin u\ F(u)$$
$$= \frac{1}{2\pi^2\bar{E}}\int_0^\infty \mathrm{d}u\ \frac{\cos u\, J_1(u)}{u} = 0, \tag{27}$$

the second step integrated by parts ($F(0) = 0$, $F'(u) = J_1(u)/u$; the boundary term oscillates about zero since $F \to 1$). The last step is a classical discontinuous Bessel integral, $\int_0^\infty J_1(u)\cos(bu)\,\mathrm{d}u/u = \sqrt{1-b^2}$ for $b \le 1$ [27], evaluated at $b = 1$.[7] A non-zero value here would have produced precisely such a log; as we have explicitly seen, this does not happen in JT.

The number variance. The number variance is therefore carried by the window's upper edge alone: summing (25) and (26),

$$\Sigma^2(\bar{E}_n) = \frac{1}{2\pi^2}\big[\ln(2\pi\bar{n}) + \gamma + 1\big] + \ldots \ \simeq\ \frac{1}{2\pi^2}\ln\bar{n}(\bar{E}_n), \tag{28}$$

the dots collecting the additive constants above (in any case not universal, since at the threshold the statistics are Airy rather than sine-kernel), and only the leading logarithm is used in what follows.[8] One bookkeeping remark for readers who know the random-matrix literature: a window with *both* edges in the interior of the spectrum receives the same logarithm from each of its two fluctuating edges (the classical number variance of Dyson and Mehta [25]), while (28) is the single-edge (index) variance [21], because our window's lower edge is nailed to the threshold.

4.2 The mean distance: $\langle \mathcal{W}^2\rangle \sim S_0\,\mathrm{e}^{-S_0}$

With the number variance (28) in hand the integrals can be done. Inserting (28) and using that the logarithm is dominated by its leading term, $\ln\bar{n}(E) \simeq S(E)$ (the microcanonical entropy),

$$\langle \mathcal{W}^2\rangle(\beta) = 2\beta^2\int_0^\infty \mathrm{d}E\ \frac{\mathrm{e}^{-2\beta E}}{\bar{\rho}(E)}\,\Sigma^2(E) \simeq \frac{\beta^2}{\pi^2}\int_0^\infty \mathrm{d}E\ \frac{S(E)}{\bar{\rho}(E)}\,\mathrm{e}^{-2\beta E} \tag{29}$$

With the Schwarzian disk density $\bar{\rho}(E) = \frac{\mathrm{e}^{S_0}}{2\pi^2}\sinh(2\pi\sqrt{E})$[9] and $S(E) \simeq S_0$

$$\langle \mathcal{W}^2\rangle(\beta) \simeq 2\beta^2\, J(\beta) S_0 \mathrm{e}^{-S_0}, \qquad J(\beta) = \int_0^\infty \mathrm{d}E\ \frac{\mathrm{e}^{-2\beta E}}{\sinh(2\pi\sqrt{E})}, \tag{30}$$

with J a convergent $O(1)$ integral. Crucially, the integrand carries the *reciprocal* density $1/\bar{\rho} \propto 1/\sinh(2\pi\sqrt{E})$, the mean level spacing, so it is dominated not by the thermal energy but by the *bottom* of the spectrum: the square-root edge makes the integrand $\sim 1/(2\pi\sqrt{E})$ near

[7] Two finite-t_H details, neither disturbing the conclusion. First, the physical form factor plateaus at the local Heisenberg time, truncating the u-integral at $u_H = 2\bar{E}t_H$: the truncation leaves $-\cos(u_H)/2\pi^2\bar{E}$, while the plateau region contributes $\frac{2\bar{\rho}}{\pi}\big[\frac{\pi}{2} - \mathrm{Si}(u_H)\big] = +\cos(u_H)/2\pi^2\bar{E}$ up to $O(\mathrm{e}^{-S})$: the same oscillation, cancelling. Second, the integration by parts is justified by damping the integrand with $\mathrm{e}^{-\epsilon u}$: the boundary term then vanishes exactly for every $\epsilon > 0$, and the surviving integrand is $O(u^{-3/2})$ (the $u^{-1/2}$ tails inside $J_1(u)/u = \frac{1}{2}[J_0 + J_2]$ cancel), hence absolutely convergent uniformly in ϵ, and continuous in b through $b = 1$.

[8] The sharp ramp/plateau split fixes the additive constant only approximately. Cutting the coincident-energy logarithm of (37) at $\delta E = 1/2\pi\bar{\rho}$ gives the logarithm $\ln(8\pi\bar{n})$; the additive $\gamma + 1$ is the separate sine-kernel counting constant. The GUE measurement of App. B lands on the combination ($c_0 = 3.7 \pm 0.2$ measured against $\gamma + 1 + 3\ln 2 \simeq 3.66$ in the variables used there).

[9] A convention note: SSS [2] normalize $\bar{\rho} = \mathrm{e}^{S_0}\sinh(2\pi\sqrt{E})/4\pi^2$, so our S_0 is theirs minus $\ln 2$. Nothing depends on the choice: the cylinder (1) is S_0-independent, and S_0 enters only through $\bar{\rho}$.

threshold (integrable), and the weight piles up at low-lying order one energies $E \sim 1/(2\pi)^2$. The temperature sets only a soft upper cutoff. Only when β grows past the exponentially large scale $\mathrm{e}^{2S_0/3}$ does the Boltzmann factor squeeze the weight down to the threshold, where the smooth double trumpet (3) must be completed by the Airy/Tracy–Widom edge and the distance crosses over to a ground-state-dominated regime which we study separately in section 6.

Uncorrelated levels: the integrable comparison. It is worth pausing on what the same integral gives when the levels are *uncorrelated*, the Poisson statistics of generic integrable spectra [28]. The energy states behave as if they were independent draws from a common distribution. The probability for a state at $[E, E+\Delta E]$ to be occupied is $\bar{\rho}(E)\,\Delta E$ independent of other states. Thus, dividing the window into M cells of width ΔE, one can model the number of states with

$$n(E) = \sum_{i=1}^{M} X_i \tag{31}$$

where X_i are independent Bernoulli random variables taking values 0 (unoccupied) with probability $1 - \bar{\rho}(E_i)\,\Delta E$ or 1 (occupied) with probability $\bar{\rho}(E_i)\,\Delta E$. The number expectation value is then

$$\langle n(E)\rangle = \sum_{i=1}^{M} \langle X_i\rangle = \sum_{i=1}^{M} \bar{\rho}(E_i)\,\Delta E \;\to\; \bar{n}(E), \tag{32}$$

where we used $\langle X_i\rangle = \mathrm{Prob}(\mathrm{X_i} = 1)$. The number variance then becomes

$$\langle n(E)^2\rangle_c = \sum_{i,j}^{M} (\langle X_i X_j\rangle - \langle X_i\rangle\langle X_j\rangle) = \sum_{i}^{M} (\langle X_i\rangle - \langle X_i\rangle^2) \to \bar{n}(E), \tag{33}$$

where in the last step we took the continuum limit (and neglected the ΔE^2 term). Feeding the Poisson value to the BW distance (18) gives an unsuppressed

$$\langle \mathcal{W}^2\rangle = 2\beta^2 \int_0^\infty \mathrm{d}E\; \frac{\mathrm{e}^{-2\beta E}}{\bar{\rho}(E)}\; \bar{n}(E) \;\sim\; \mathcal{O}(1), \tag{34}$$

in place of $S_0\,\mathrm{e}^{-S_0}$: the self-distance separates chaotic from level-uncorrelated ensembles.

4.3 The variance: double trumpet squared

The variance is the second cumulant in (20) and thus, given by

$$\mathrm{Var}(\mathcal{W}^2) = 8\beta^4 \mathrm{Tr} C^2 = 8\beta^4 \iint \mathrm{d}E\,\mathrm{d}E'\; \frac{\mathrm{e}^{-2\beta E}\,\mathrm{e}^{-2\beta E'}}{\bar{\rho}(E)\,\bar{\rho}(E')} \left[\langle n(E)\,n(E')\rangle_{\mathrm{c}}\right]^2. \tag{35}$$

The covariance $\langle n(E)\,n(E')\rangle_{\mathrm{c}}$ follows by integrating the energy-resolved cylinder (3) over both counting windows, $\langle n(E)\,n(E')\rangle_{\mathrm{c}} = \int_0^E \mathrm{d}E_1 \int_0^{E'} \mathrm{d}E_2\, \langle\rho(E_1)\rho(E_2)\rangle_{\mathrm{c}}$. In the variables $k_i = \sqrt{E_i}$ the measure combines with the correlator into a partial-fraction form,

$$\mathrm{d}E_1\,\mathrm{d}E_2\, \big\langle\rho(E_1)\rho(E_2)\big\rangle_{\mathrm{c}} = -\frac{\mathrm{d}k_1\,\mathrm{d}k_2}{2\pi^2} \left[\frac{1}{(k_1-k_2)^2} + \frac{1}{(k_1+k_2)^2}\right], \tag{36}$$

and both terms integrate elementarily. The coincident $(k_1 - k_2)^{-2}$ singularity is handled by taking the finite part (in the Hadamard sense), whose justification is the plateau, leading to

$$\big\langle n(E)\,n(E')\big\rangle_{\mathrm{c}} = \frac{1}{2\pi^2}\, \ln \frac{\sqrt{E}+\sqrt{E'}}{|\sqrt{E}-\sqrt{E'}|}\,. \tag{37}$$

As a nice consistency check, the coincident divergence reproduces the number variance Σ^2 as the log is cut at the mean spacing $\Delta E \sim \mathrm{e}^{-S_0}$ and is therefore extensive in the entropy, $\Sigma^2(E) = \frac{1}{2\pi^2}\ln(E\bar{\rho}) \sim S_0/2\pi^2$. The *off-diagonal*, for a typical pair across the thermally active band, is $\frac{1}{2\pi^2}$ times the $O(1)$ logarithm $\ln\frac{\sqrt{E}+\sqrt{E'}}{|\sqrt{E}-\sqrt{E'}|}$: cutoff-free, carrying *no* factor of S_0.

Inserting the Schwarzian density $\bar{\rho} = (\mathrm{e}^{S_0}/2\pi^2)\sinh(2\pi\sqrt{E})$ and the double-trumpet correlator (37) into (35), the two factors of $1/\bar{\rho}$ give e^{-2S_0} and the $1/2\pi^2$'s cancel; in $k = \sqrt{E}$,

$$\mathrm{Var}(\mathcal{W}^2) \simeq 8\beta^4\, I(\beta)\, \mathrm{e}^{-2S_0}, \qquad I(\beta) = \int_0^\infty\!\!\int_0^\infty \mathrm{d}k\, \mathrm{d}k'\, \frac{2k\,\mathrm{e}^{-2\beta k^2}}{\sinh 2\pi k}\, \frac{2k'\,\mathrm{e}^{-2\beta k'^2}}{\sinh 2\pi k'}\, \ln^2\frac{k+k'}{|k-k'|}. \tag{38}$$

Beyond Wick: the discreteness constants. As we used the leading Wick contractions to calculate the variance, it is natural to ask how large are the corrections to the Wick tower, and where do they come from? Two statements above, Wick factorization and the variance needing no plateau, hold at leading order only. The counting field (and thus Δn) is Gaussian only in a weak sense: its variance grows, $\Sigma^2 \sim S_0/2\pi^2$, while its higher cumulants saturate at finite *universal* constants, fixed by the Fisher–Hartwig asymptotics of counting determinants [29, 30]. For a window with bulk (sine-kernel) endpoints the generating function is Gaussian times Barnes G-functions [30], so the tower of constants is odd zeta values; the first is

$$\kappa_4\big(n(E)\big) \xrightarrow{\bar{n}\to\infty} -\frac{24\,\zeta(3)}{(2\pi)^4} \simeq -0.0185, \tag{39}$$

which we have confirmed in GUE counting statistics. These constants originate in the oscillatory part of the kernel, the same discreteness sector as the plateau. The plateau therefore does not decouple from the higher cumulants after all: it stops being a *regulator* (the variance is finite without it) and becomes a finite *corrector*, contributing to $\mathrm{Var}(\mathcal{W}^2)$ at the same e^{-2S_0} order as the two-cylinder term, suppressed only by the smallness of the constants themselves: standardized against the growing number variance, the field's non-Gaussianity falls off as $1/S_0^2$, a power, not an exponential. In the cases measured in GUE in Appendix B the corrections shift the variance of $\mathcal{W}^2$ at the few-percent level, a numerically fairly small effect (we refer the reader to the Appendix for details). While understanding these effects in gravity would be very interesting, we will neglect them in what follows, as they do not change our qualitative results.

Relation to the mathematical transport literature. With both cumulants in hand it is worth recording what was already known. The scaling of the *mean* Wasserstein distance is classical in the probability literature: for Wigner/GUE matrices the expected W_2^2 distance between the empirical spectral measure and the semicircle law is bounded by $C\ln N/N^2$ (the rigidity logarithm over the squared density), built from exactly the Gustavsson-type eigenvalue variances used here [31] (proven as an upper bound, which those variances saturate); and for ensembles with concentrated entries the distance concentrates around its mean at the finer scale $\sim 1/N$ [32], so the mean is the whole story at leading order. Equation (30) is the Boltzmann-weighted, unnormalized analogue of that statement, with e^{S_0} in place of N. What is added in this paper is the gravitational identification (the correlator is the double trumpet, the logarithm's cutoff the plateau), together with the variance, the S_0-mismatch between the two cumulants, and the ensemble geometry built on it in Section 5.

5 Geometry of the ensemble

Sections 3–4 reduced the cumulants of $\mathcal{W}^2$ to traces of the transport covariance operator C of (19), with the mean and variance $\langle\mathcal{W}^2\rangle = 2\beta^2\mathrm{Tr}C$, $\mathrm{Var} = 8\beta^4\mathrm{Tr}C^2$ evaluated for the double trumpet. This section reads off the *geometry* the same C encodes: the size, principal axes,

and concentration of the ensemble seen as a point cloud, including the relative spread (effective dimension) that sets its self-averaging.

5.1 Flatness, typical separation and width

Flatness. Each ensemble member is a point in $\mathbb{R}^L$ (L being the dimensionality of the Hilbert space) with coordinate $e^{-\beta E_n}$. The BW metric is the flat ℓ^2 metric on this space.

Typical separation. $\sqrt{\langle \mathcal{W}^2 \rangle} \sim \sqrt{S_0}\, \mathrm{e}^{-S_0/2}$: an exponentially tight cluster against the points' own norm, $|\mathbf{a}|^2 = Z(2\beta) \sim \mathrm{e}^{S_0}$.

Width and self-averaging. The width of the cloud (averaged over all directions in $\mathbb{R}^L$) is simply the standard deviation

$$\Delta \mathcal{W}^2 \equiv \sqrt{\mathrm{Var}(\mathcal{W}^2)} = \sqrt{8\beta^4 \mathrm{Tr} C^2} \sim \mathrm{e}^{-S_0}, \tag{40}$$

where we used the scaling in (38). When drawing a typical theory pair from the ensemble, their squared distance typically lies in the range $\langle \mathcal{W}^2 \rangle \pm \Delta \mathcal{W}^2 = S_0\, \mathrm{e}^{-S_0}\, (1 \pm S_0^{-1})$. So in other words, as $S_0 \to \infty$ the typical value of the distance is close to the mean in units of the mean, i.e. it is self averaging. This is in contrast with the spectral form factor from which the BW distance is built, which is known to not self average [33]. So the typical ensemble members do not converge to the same spectral form factor as $S_0 \to \infty$, while they do converge to the same BW distance.

5.2 Principal modes: how the spectra of the ensemble members differ

Take two random ensemble members, points in $\mathbb{R}^L$

$$\mathbf{a} = e^{-\beta \mathbf{E}^{(1)}}, \quad \mathbf{b} = e^{-\beta \mathbf{E}^{(2)}}. \tag{41}$$

These can differ in any of the L directions. For example they might only differ in one or a few directions (eigenvalues). Or they might differ by some smooth fluctuation shared by many energies. Next, we will find what the shape of the typical differences is. Rather than working with the discrete eigenvalues, the same information can be organized in the continuum by saying that the two theories have different densities of states $\rho_1(E)$ and $\rho_2(E)$ differing by $\Delta\rho(E) = \rho_1(E) - \rho_2(E)$. So a single energy difference would be a local peak in $\Delta\rho(E)$ around the differing energy, while a smooth collective difference in many states would be a smooth fluctuation. As we have seen, the natural variable for the BW distance is not necessarily the density but the count variable $n(E) = \int_0^E dE' \rho(E')$; we therefore characterize the difference between the theories through the number difference

$$\Delta n(E) = \int_0^E dE' (\rho_1(E') - \rho_2(E')). \tag{42}$$

The number difference can be expanded in terms of modes as

$$\Delta n(E) = \sum_m c_m \varphi_m(E), \tag{43}$$

where c_m are random numbers (approximately Gaussian [21]) that depend on the theory pair we have drawn from the ensemble, while $\varphi_m(E)$ is a basis of modes, unspecified at this point. What is the basis of modes in which the geometry of the cloud of ensemble members is most clear? The answer is provided by principal component (Karhunen–Loève) analysis [34, 35]. If we choose the

modes $\varphi_m(E) = \chi_m(E)/\sqrt{w(E)}$ where χ_m diagonalize the operator C and $w(E) = e^{-2\beta E}/\bar{\rho}(E)$, something rather nice happens as

$$\frac{\mathcal{W}^2}{\beta^2} = \int dE\, w(E)\Delta n(E)^2 = \sum_{m,m'} c_m c_{m'} \int dE\, w(E)\varphi_m(E)\varphi_{m'}(E) = \sum_m c_m^2, \tag{44}$$

where $\mathcal{W}^2$ is still the random squared distance between ensemble members and we used orthonormality of the modes $\int dE\, w(E)\varphi_m(E)\varphi_{m'}(E) = \delta_{mm'}$ (which follows from orthonormality of the eigenmodes χ_m). So the BW distance is simply the squared length of the $\mathbf{c}$ vector in flat space. The statistics of the coefficients follow directly: they are projections, $c_m = \int \mathrm{d}E\, w(E)\, \varphi_m(E)\, \Delta n(E)$, so their covariance is inherited from the count correlator. Using $w\,\varphi_m = \sqrt{w}\,\chi_m$, the correlator $\langle \Delta n(E)\Delta n(E')\rangle = 2\langle n(E)n(E')\rangle_{\mathrm{c}}$ of (17), and the eigenvalue equation of C,

$$\begin{aligned}\langle c_m c_{m'}\rangle &= \int \mathrm{d}E\, \mathrm{d}E'\, \sqrt{w(E)}\, \chi_m(E)\, \big\langle \Delta n(E)\, \Delta n(E')\big\rangle\, \sqrt{w(E')}\, \chi_{m'}(E') \\ &= 2\int \mathrm{d}E\, \mathrm{d}E'\, \chi_m(E)\, C(E,E')\, \chi_{m'}(E') = 2\lambda_m\, \delta_{mm'}.\end{aligned} \tag{45}$$

The coefficients of different modes are uncorrelated (this is what makes the principal axes principal), and they are Gaussian by the central limit theorems of Section 3.3, hence independent. Taking the mean over the ensemble reproduces $\langle \mathcal{W}^2\rangle = \beta^2 \sum_m \langle c_m^2\rangle = 2\beta^2 \mathrm{Tr}C$, matching (18). To recap the geometric meaning, (45) is saying that the typical squared spread of theory pairs in the m'th direction is $2\lambda_m$, set by the eigenvalue λ_m of C, and the corresponding density/number fluctuation is $\varphi_m(E)$. This becomes even more obvious when introducing the unit variance zero mean independent Gaussian random variables $z_m = c_m/\sqrt{2\lambda_m}$: we obtain the form

$$\mathcal{W}^2 = 2\beta^2 \sum_m \lambda_m z_m^2. \tag{46}$$

This tells us that the BW distance follows a generalized χ^2 distribution[10]. We will come back to plot the distribution, once we calculate the eigenvalues explicitly.

While φ_m modes relate most intuitively to the spectral density fluctuations, the χ_m modes are more natural for the BW distance because they carry the suppression of high energy fluctuations, precisely suppressing the region to which the BW distance is blind. The resulting picture of the modes is shown in Fig. 4.

To construct that eigensystem we write C explicitly. Inserting the Schwarzian density $\bar{\rho}(E) = (\mathrm{e}^{S_0}/2\pi^2)\sinh(2\pi\sqrt{E})$ and the double-trumpet correlator (37) into the kernel (19) gives the covariance operator of JT:

$$C(E,E') = \frac{1}{2\pi^2}\sqrt{w(E)\,w(E')}\,\ln\frac{\sqrt{E}+\sqrt{E'}}{|\sqrt{E}-\sqrt{E'}|}, \qquad w(E) = \frac{\mathrm{e}^{-2\beta E}}{\bar{\rho}(E)} = \frac{2\pi^2\,\mathrm{e}^{-S_0}\,\mathrm{e}^{-2\beta E}}{\sinh(2\pi\sqrt{E})}, \tag{47}$$

a symmetric kernel on $L^2(\mathrm{d}E)$: the smooth Boltzmann envelope $w(E)$, peaked at small $\mathcal{O}(1)$ energies, times the scale-free logarithmic kernel.

The logarithmic kernel on its own is not compact but the Boltzmann envelope $w(E)$ discretizes it: w decays like a Gaussian in $k = \sqrt{E}$, so C is Hilbert–Schmidt, $\mathrm{Tr}C^2 < \infty$, hence compact, with a discrete ladder of positive eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \to 0$ and orthonormal eigenfunctions $\chi_m(E)$. The discreteness is only marginal: $\sum_m \lambda_m^2 = \mathrm{Tr}C^2$ converges while $\sum_m \lambda_m = \mathrm{Tr}C$ diverges logarithmically, finite only once the plateau caps the ladder at $m \sim \mathrm{e}^{S_0}$. This is the same cutoff-free-versus-cutoff-dependent split behind the mean and the variance.

[10]The χ^2 distribution is simply the distribution of a sum of squares of independent unit Gaussians. It is positively skewed and has longer tails than a Gaussian distribution.

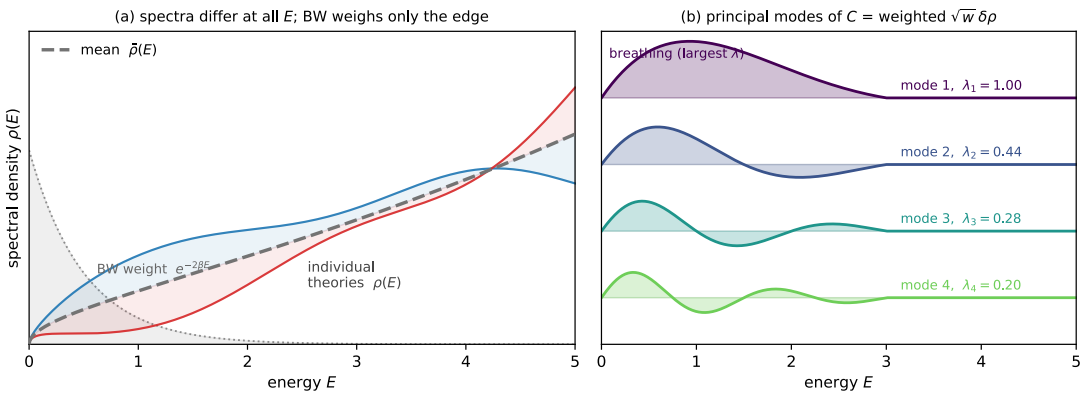


Figure 4: What the eigenmodes of C mean (schematic). (a) Different theories have spectral densities $\rho(E)$ that scatter around the mean $\bar{\rho}(E)$ at *all* energies: the bare fluctuation does not decay at large E. The BW distance, however, weights the spectrum by the Boltzmann factor $w(E) = \mathrm{e}^{-2\beta E}/\bar{\rho}$ (shaded window), so it only registers the low-energy edge. (b) Through that weight, the principal modes the distance sees (the eigenmodes $\chi_m = \sqrt{w}\,\delta\rho$ of C, ranked by eigenvalue) are edge-localized: the softest "breathing" mode (largest λ) is a coherent low-energy reweighting, higher modes resolve finer structure with smaller amplitude. The localization is the weight, not the fluctuations; the bare patterns $\varphi_m = \chi_m/\sqrt{w}$ extend to high E. The numerically diagonalized JT modes of Fig. 5 are the precise version of this cartoon. *Amplitudes are exaggerated for visibility*: relative to the mean density (taken $O(1)$ here) the true fluctuations are $O(\mathrm{e}^{-S})$ (the microscopic rigidity is exponentially small), whereas the figure draws them at $O(1)$.

We diagonalize C in the natural variable $k = \sqrt{E}$, where the kernel depends on $k \pm k'$ and $\mathrm{d}E = 2k\,\mathrm{d}k$; folding the Jacobian into the weight, C is unitarily equivalent to the symmetric operator on $L^2(\mathrm{d}k)$ with kernel

$$A(k,k') = u(k)\,u(k')\,\ln\frac{k+k'}{|k-k'|}, \qquad u(k) = \sqrt{\frac{k}{\sinh 2\pi k}}\;\mathrm{e}^{-\beta k^2} \tag{48}$$

(up to the overall scale e^{-S_0} that only multiplies each eigenvalue). Discretized on a uniform k-grid by the Nyström rule (off-diagonal entries $A(k_a,k_b)\,\Delta k$, the integrable diagonal log replaced by its exact cell average $\ln(4k_a/\Delta k)+1$), this real symmetric matrix is diagonalized directly, with no discreteness put in by hand; the ladder is a property of the smooth operator (47).

These modes are shown in Fig. 5. Mode m carries $m-1$ nodes and is localized on the thermally active edge $E \lesssim \beta^{-1}$; the oscillations bunch toward $E = 0$ because the natural variable is $\sqrt{E}$, in which they are evenly spaced. Ranked by eigenvalue, they are the independent patterns of spectral deviation $\delta\rho(E)$ along which ensemble members differ.

The largest eigenvalue in Fig. 5, and thus the most typical fluctuation through which ensemble members differ, corresponds to the smoothest eigenmode: a "breathing" deformation, almost a parallel shift of the spectrum. Higher, more rapidly fluctuating modes are subleading. This is the same structure as one finds when diagonalizing a log correlation in position space by a Fourier transform into momentum space, in which the correlator is simply $1/k$; for GUE the exact statement is Johansson's diagonalization of the eigenvalue fluctuations in Chebyshev/Fourier modes with the $1/m$ ladder [22], the arc-variable form of which appears in App. B.

The ladder reproduces the moments of Section 4 as a check: $\sum_m \lambda_m^2 < \infty$ gives the finite variance $\mathrm{Tr}C^2$, while $\sum_m \lambda_m$ diverges logarithmically and, cut by the plateau at $m \sim \mathrm{e}^{S_0}$ ($\sum_m^{\sim \mathrm{e}^{S_0}} 1/m \sim S_0$), returns the lone S_0 of $\mathrm{Tr}C$ found directly there. Geometrically the cloud is therefore a thin, highly anisotropic shell: members differ *most* in the coarse, low-energy weight of

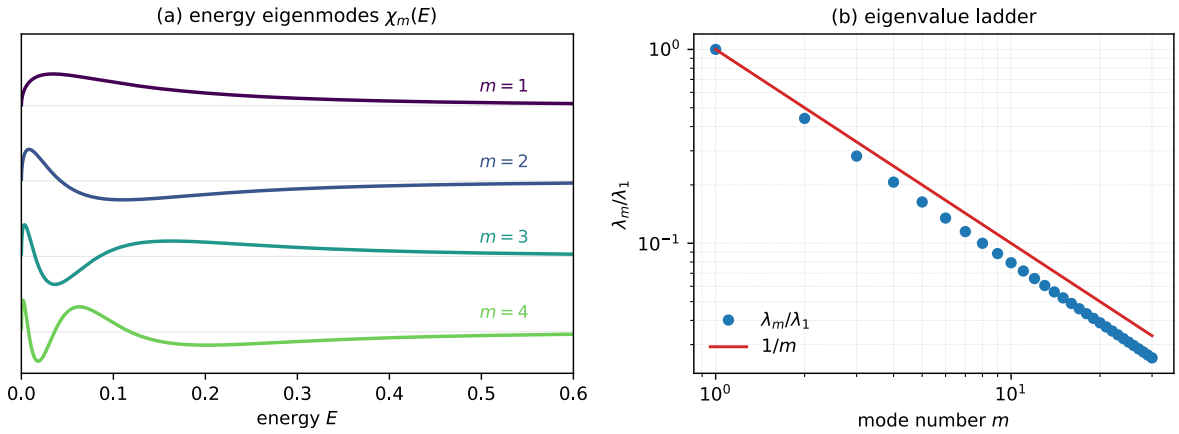


Figure 5: Principal modes of the theory cloud, from a fine-grid continuum diagonalization of the genus-zero BW covariance (37) at $\beta = 1$ (no discreteness imposed). (a) The lowest energy eigenmodes $\chi_m(E)$: mode m has $m-1$ nodes and sits on the thermal edge, the oscillations bunching toward $E = 0$ because the natural variable is $\sqrt{E}$. (b) The eigenvalue ladder λ_m/λ_1 (log–log) against the $1/m$ guide: the canonical Hilbert–Schmidt-but-not-trace-class spectrum $\lambda_m \propto 1/m$, falling slightly faster than $1/m$ at the edge (so $\sum_m \lambda_m^2 < \infty$ but $\sum_m \lambda_m$ diverges, requiring the plateau cutoff $m \lesssim e^{S_0}$).

their spectra and *least* in fine, level-by-level detail. One distinction is worth keeping sharp, since the two statements "dominated by a single mode" and "$D_{\text{eff}} \sim S_0^2$" would otherwise collide: the breathing mode dominates the *fluctuations* of the distance, carrying $\lambda_1^2/\mathrm{Tr}C^2 \simeq 0.69$ of $\mathrm{Var}(\mathcal{W}^2)$ at $\beta = 1$, while of the *mean* distance $\mathrm{Tr}C$, whose ladder must be summed to the plateau cutoff, it carries only $\lambda_1/\mathrm{Tr}C \simeq 1.3/S_0$. Both are true at once because the variance is cutoff-free and top-heavy while the mean is spread democratically over the $\sim e^{S_0}$-mode tail; it is the latter that makes the effective dimension large.

The full distribution of $\mathcal{W}^2$, the generalized χ^2 form of (46), is shown in Fig. 6 together with a Gaussian comparison. The distribution is clearly non-Gaussian with positive skewness (right tail longer than left tail) and longer tailed (excess kurtosis). All dependence on S_0 is in the mean and the variance, which disappear in the relative units used in the plot. The generalized χ^2 (46) is confirmed directly in GUE, where it reproduces the measured distribution in both peak and tail (Fig. 13, App. B).

5.3 Concentration, effective dimensionality and anisotropy

In $\mathbf{c}$ space the distribution of pairs of theories (recalling that a single vector $\mathbf{c}$ relates to the difference between two randomly sampled theories) has the measure

$$d\mu = d^L c\, e^{-\sum_m \frac{c_m^2}{4\lambda_m}}. \tag{49}$$

Since the dimensionality L is high the measure increases rapidly with $|c|$ as $d^L c \propto |c|^{L-1} d|c|$, while the exponential factor decays with $|c|$. This has the effect that the space of typical $\mathbf{c}$ values gets squeezed to a thin shell where the hypersphere volume and the Gaussian suppression get balanced. The shell structure explains the at first contradictory sounding statement that the mean distance between the theory pairs is much larger than the standard deviation of the distance: the typical separation of a pair is the shell's radius, while the fluctuation in separations is only the shell's thickness. Along the directions where λ_m is very small the theories barely move at all, so those directions are effectively frozen. We can introduce an effective dimensionality

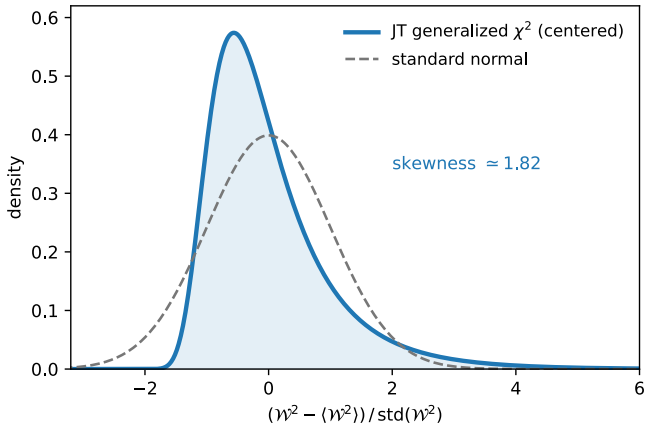


Figure 6: The JT BW-distance distribution, centered and standardized (hence S_0-independent): the generalized χ^2 (46) built from the eigenvalues λ_m of the numerically diagonalized covariance operator C (at $\beta = 1$; the standardized shape moves by $< 2\%$ across $\beta \in [\frac{1}{4}, 4]$). It is strongly right-skewed (skewness $\simeq 1.82$) and never approaches the Gaussian (dashed).

that measures how many directions are participating in the BW distance. If there were K eigenvalues of roughly equal size $\bar{\lambda}$ and the others were much smaller, then

$$\mathrm{Tr}C \approx K\bar{\lambda}, \quad \mathrm{Tr}C^2 \approx K\bar{\lambda}^2. \tag{50}$$

Solving for K leads to the effective dimension

$$D_{\mathrm{eff}} = \frac{(\mathrm{Tr}C)^2}{\mathrm{Tr}C^2} = \frac{2\,\langle\mathcal{W}^2\rangle^2}{\mathrm{Var}(\mathcal{W}^2)}. \tag{51}$$

While in our case the split between large and small eigenvalues is not clear cut, we will use this as a definition of the effective dimensionality of the theory cloud. From the established results (30) and (38) we see that

$$D_{\mathrm{eff}} = \frac{J(\beta)^2}{I(\beta)} S_0^2, \tag{52}$$

at leading logarithm.[11] The coefficient $J(\beta)^2/I(\beta)$ is plotted in Fig. 7: it is constant to a few percent across three decades of temperature, equal to 0.42 at $\beta = 1$ and flowing onto the universal pure-edge value $4/\pi^2$ as $\beta \to \infty$.

The growth with S_0^2 is tested directly in random-matrix numerics in App. B, with the number of thermally active levels, $n_{\mathrm{part}} \equiv (\sum_n a_n)^2/\sum_n a_n^2$ for the Boltzmann weights $a_n = \mathrm{e}^{-\beta E_n}$, playing the role of e^{S_0}: across GUE ensembles of size $N = 200$–600 the measured effective dimension grows logarithmically at every temperature, climbing from 3 to 25 with no sign of saturation. At these sizes the *coefficient* of the growth is, however, not yet the universal one: resolved in temperature, the measured slopes $[\partial\sqrt{D_{\mathrm{eff}}}/\partial \ln N]^2$ drift from 0.23 at $\beta = 4$ to 0.49 at $\beta = \frac{1}{8}$, tracking a genuine β-dependence of the GUE target (the finite semicircle band interpolates between the single-edge $4/\pi^2$ and a both-edges-active $6/\pi^2$, App. B) on top of a finite-size suppression that decays only logarithmically. The numerics therefore test the logarithmic form and the predicted drift of its coefficient rather than the asymptotic value itself — and, with the two $O(1)$ covariance constants supplied (the sine-kernel bracket and the level-versus-count shift of Section 3.3), the same kernel predicts the measured finite-N values with

[11] Keeping the constants: with the sine-kernel bracket (footnote under (28)) and the level-versus-count shift of Section 3.3, $\mathrm{Tr}C \propto S_0 + c(\beta)$ with $c(\beta) = \langle\ln(8\pi E\bar{\rho})\rangle_w + \gamma + 1 - \frac{\pi^2}{6} - S_0 \simeq -3.5$ at $\beta = 1$ ($w = \mathrm{e}^{-2\beta E}/\bar{\rho}$), so the finite-$S_0$ form is $D_{\mathrm{eff}} \simeq 0.42\,(S_0 - 3.5)^2$: at $S_0 = 20$ a 33% suppression. These are the same constants that put the finite-N kernel prediction on the GUE data in App. B.

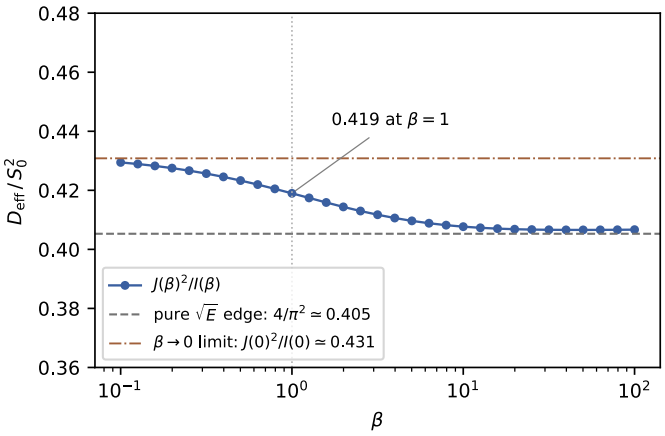


Figure 7: The coefficient of the effective dimension, $D_{\rm eff}/S_0^2 = J(\beta)^2/I(\beta)$, across three decades of temperature: constant to a few percent, drifting monotonically onto the universal pure-$\sqrt{E}$-edge value $4/\pi^2 \simeq 0.405$ (dashed) as the thermal weight concentrates on the spectral edge. At $\beta = 1$ the value is 0.419. The $\beta \to 0$ limit is finite as well, $J(0)^2/I(0) \simeq 0.431$ (dash-dotted) with $J(0) = 1/8$ exactly: at high temperature the integrals are cut off by the sinh growth of the density rather than by the Boltzmann factor, so nothing diverges as $\beta \to 0$.

no free parameters, to a few percent across the grid (App. B). The single-edge rigidity behind the lone logarithm — the silent threshold of Section 4.1 — is confirmed separately and directly: the measured variance of the count in the threshold-anchored window carries the index-variance prefactor $1/2\pi^2$, half the two-edge Dyson–Mehta value (App. B).

For the purpose of distances the cloud behaves like points on a sphere of *concentration dimension* $D_{\rm eff} \sim S_0^2$. The shell is, however, far from round. As the eigenvalues have the approximate $\lambda_m \sim 1/m$ form, the first eigenmode is the dominant direction in which the cloud is most elongated, while smaller eigenvalues (larger m) correspond to directions in which the cloud is less and less spread. So it could be characterized as a thin shell ellipsoid with approximately $\mathrm{e}^{-S_0/2}/\sqrt{m}$ spread in the m'th direction.

6 Low temperatures: the Airy edge and SYK

The mean of Section 4.2, and the geometry built on it, used the *smooth* Schwarzian density, valid while the Boltzmann weight $\mathrm{e}^{-2\beta E}$ is spread over many levels. As the temperature drops the weight is squeezed toward the bottom of the spectrum, and once β reaches the exponentially large scale $\mathrm{e}^{2S_0/3}$ it resolves only the lowest few levels, where the spectrum is discrete and the smooth density fails. The physics there is the universal *soft edge* of a chaotic spectrum [36]: the lowest eigenvalues form the *Airy point process* [37] and the ground state E_1 fluctuates with the *Tracy–Widom* law. This is the regime SYK at accessible N (Section 6.1) occupies, so we set up the BW distance there.

Edge scaling. Near threshold the Schwarzian density vanishes as $\bar\rho(E) \simeq c\sqrt{E}$, $c = \mathrm{e}^{S_0}/\pi$, so the mean count $\bar n(E) = \frac{2}{3}cE^{3/2}$ becomes of order unity (the first level) at the edge scale $E \sim (\pi c)^{-2/3} = \mathrm{e}^{-2S_0/3}$. This fixes the crossover: the ramp results hold for $\beta \ll \mathrm{e}^{2S_0/3}$, the edge regime is $\beta \gtrsim \mathrm{e}^{2S_0/3}$, and the whole edge physics is a universal function of the single dimensionless temperature $\tau = \beta\,\mathrm{e}^{-2S_0/3}$. In Airy units, $E_k = \mathrm{e}^{-2S_0/3}\,b_k$: the levels near threshold form the Airy point process $\{b_k\}$, with b_1 the (reflected) Tracy–Widom variable, and the Boltzmann weights become $\mathrm{e}^{-\beta E_k} = \mathrm{e}^{-\tau b_k}$.

High temperature: matching the bulk. For $\tau \ll 1$ many edge levels are thermally active, and the BW distance returns to the continuum form (18) with the edge density $\bar{\rho} = c\sqrt{E}$ and the number variance $\Sigma^2 \simeq S(E)/2\pi^2$ of (28)[12], with the difference that the energy dependence from the entropy cannot be dropped

$$\langle \mathcal{W}^2 \rangle \simeq \frac{\beta^2}{\pi^2} \int_0^\infty \mathrm{d}E\, \mathrm{e}^{-2\beta E}\, \frac{S(E)}{c\sqrt{E}} \simeq \frac{S\,\beta^2}{\pi^2 c} \sqrt{\frac{\pi}{2\beta}} = \frac{1}{\sqrt{2\pi}}\, S\, \beta^{3/2}\, \mathrm{e}^{-S_0}, \tag{53}$$

with $S = \ln \bar{n}(1/\beta) = S_0 - \frac{3}{2}\ln\beta + O(1)$ the entropy at the thermally dominant energy, in agreement with the large-β limit of the bulk result (30) ($J(\beta) \to 1/(2\sqrt{2\pi\beta})$), with *no free constant*. In the dimensionless variable, $\bar{n}(1/\beta) = \frac{2}{3\pi}\tau^{-3/2}$ gives $S = \frac{3}{2}\ln(1/\tau) + O(1)$, so

$$\langle \mathcal{W}^2 \rangle \equiv F(\tau) \simeq \frac{3}{2\sqrt{2\pi}}\, \tau^{3/2} \ln(1/\tau), \qquad \tau \ll 1, \tag{54}$$

where we defined the scaling function $F(\tau)$. The $\langle \mathcal{W}^2 \rangle \sim S\,\mathrm{e}^{-S}$ of Section 4.2 is the $\beta \sim 1$ value of a curve that in fact *grows* as $\beta^{3/2}$ on cooling, reaching $O(1)$ at the crossover. The variance matches in parallel: from (38), $\mathrm{Var}(\mathcal{W}^2) \propto \tau^3$ at large β while $\langle \mathcal{W}^2 \rangle^2 \propto \tau^3 \ln^2(1/\tau)$, so $D_{\mathrm{eff}} \propto \ln^2(1/\tau)$ falls smoothly through $O(1)$ at the same crossover (Fig. 8b).

The very low temperature BW distance. For very low temperatures the linearization in β times the level fluctuations, leading to (16), breaks down and we need to step back to the exact formula (12), which in Airy units gives,

$$\langle \mathcal{W}^2 \rangle(\beta) = 2 \sum_k \mathrm{Var}\big(\mathrm{e}^{-\beta E_k}\big) = 2 \sum_k \mathrm{Var}_{\mathrm{Airy}}\big(\mathrm{e}^{-\tau b_k}\big), \tag{55}$$

now summed over the Airy levels; in a non-perturbative completion of JT, the level-by-level inputs to this sum are available from the Fredholm-determinant construction of [8, 9].

Single Tracy–Widom mode, and the FZZT brane. For $\tau \gtrsim 1$ only the lowest level survives the Boltzmann filter[13],

$$F(\tau) \;\longrightarrow\; 2\mathrm{Var}\big(\mathrm{e}^{-\tau b_1}\big) = 2\Big[\big\langle \mathrm{e}^{-2\tau b_1}\big\rangle - \big\langle \mathrm{e}^{-\tau b_1}\big\rangle^2\Big], \tag{56}$$

and its statistics are set by a single eigenvalue-instanton [40, 41], an FZZT brane [42, 43]. In the matrix model, the lowest eigenvalue dipping below zero is one charge of the Coulomb gas pulled out of the droplet into the classically forbidden region $E < 0$, and its action is the work done against the mean force of the remaining gas. That force is encoded in the same function as the density: the resolvent $R(E) = \int \mathrm{d}E'\, \bar{\rho}(E')/(E - E')$ jumps by $2\pi i\bar{\rho}$ across the support, and the saddle point condition *on* the support fixes its continuation *off* the support to $\mathcal{F}(E) = 2\pi|\bar{\rho}(E)|$: the analytically continued density *is* the tunneling force. The action is therefore the area under the continued density,

$$I(E) = 2\pi \int_E^0 |\bar{\rho}(E')|\, \mathrm{d}E' = \tfrac{4\pi c}{3}\, |E|^{3/2} = \tfrac{4}{3}\, \mathrm{e}^{S_0} |E|^{3/2}, \tag{57}$$

with $|\bar{\rho}| = c\sqrt{|E|}$ the near-threshold continuation. This is the JT FZZT-brane action [2, 6, 10] $\propto \mathrm{e}^{S_0}|E|^{3/2}$; in Airy units it reads $\frac{4}{3}|s|^{3/2}$, the Tracy–Widom tail exponent. The ground state

[12] the Airy-process counting variance carries the same single-edge logarithm, $\frac{3}{4\pi^2}\ln T = \frac{1}{2\pi^2}\ln\bar{n}$ in the edge variable [20]

[13] Doubly so: the next level sits a typical gap $\langle b_2 - b_1 \rangle \approx 1.9$ Airy units higher, and its deep-tail events are two-brane configurations of action $2I$, giving $\langle \mathrm{e}^{-2\tau b_2}\rangle \sim \mathrm{e}^{\tau^3/6} \ll \mathrm{e}^{2\tau^3/3}$.

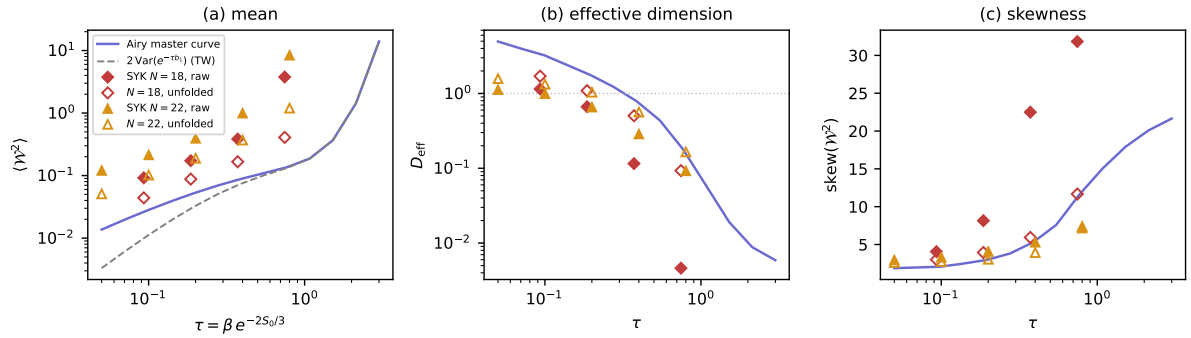


Figure 8: The BW distance in the Airy edge regime as a universal function of the dimensionless temperature $\tau = \beta\, \mathrm{e}^{-2S_0/3}$ (edge of the Dumitriu–Edelman tridiagonal ensemble at Dyson index two [38], whose soft-edge limit is the Airy point process [39]; 1050 realizations of the 100 lowest edge levels b_k, with weights $\mathrm{e}^{-\tau b_k}$). (a) The mean; at low temperature it approaches the lowest-level Tracy–Widom fluctuation $2\mathrm{Var}(\mathrm{e}^{-\tau b_1})$ (dashed). (b) The effective dimension falls through unity toward zero. (c) The skewness grows. Diamonds: SYK N=18 (2100 realizations); triangles: N=22 (590 realizations), with $\tau = \beta\sigma$ and the edge scale σ fitted to the ensemble-mean edge levels; filled symbols raw spectra, open symbols width-unfolded. The systematic offset from the master curve, and its origin in the collective coupling fluctuations, are analyzed in Section 6.1 and App. A.

then has the large-deviation tail $\propto \mathrm{e}^{-I(E_1)}$, and the moments of its weight are forbidden-region integrals evaluated by saddle point,

$$\left\langle \mathrm{e}^{-n\beta E_1} \right\rangle \;\sim\; \int_{-\infty}^{0} \mathrm{d}E\; \mathrm{e}^{\,n\beta|E|-I(E)}, \qquad n\beta = 2\,\mathrm{e}^{S_0}\sqrt{|E_1^*|} \;\;\Rightarrow\;\; |E_1^*| = \tfrac{1}{4}(n\beta)^2\mathrm{e}^{-2S_0}, \tag{58}$$

$$\left\langle \mathrm{e}^{-n\beta E_1} \right\rangle \;\sim\; \exp\left[\tfrac{1}{12}(n\beta)^3\mathrm{e}^{-2S_0}\right] = \mathrm{e}^{(n\tau)^3/12}\,: \tag{59}$$

the cubic exponent is the Legendre transform of the $\frac{4}{3}|s|^{3/2}$ brane action (the Tracy–Widom right-tail exponent [44]), whence the known cubic growth of the Tracy–Widom moment generating function, the brane action accumulating as cooling pushes the saddle $|E_1^*|$ deeper. The Gaussian fluctuation about the saddle restores the prefactor: the moments are large-argument Airy partition functions, $Z_1(n\tau) \equiv \langle \mathrm{e}^{-n\tau b_1}\rangle \simeq \mathrm{e}^{(n\tau)^3/12}/\left[2\sqrt{\pi}\,(n\tau)^{3/2}\right]$ [10]. Hence $F(\tau) \simeq 2Z_1(2\tau) = \mathrm{e}^{2\tau^3/3}/\left[2\sqrt{2\pi}\,\tau^{3/2}\right]$ at large τ, dominated by the second moment. The effective dimension follows from the same moments, with no new input: the tail-dominated limit of $D_{\mathrm{eff}} = 2\langle \mathcal{W}^2\rangle^2/\mathrm{Var}(\mathcal{W}^2)$ is

$$D_{\mathrm{eff}} \;\simeq\; \frac{4\,Z_1(2\tau)^2}{Z_1(4\tau)} \;=\; 4\,\mathrm{e}^{S_2} \;\simeq\; \frac{2}{\sqrt{\pi}\,\tau^{3/2}}\,\mathrm{e}^{-4\tau^3}, \qquad S_2 = 2\ln Z_1(2\tau) - \ln Z_1(4\tau)\,: \tag{60}$$

S_2 is precisely the second Rényi entropy of the thermal ensemble of brane positions at inverse temperature 2τ: the effective dimension counts how many brane configurations the thermal weight effectively occupies, in the same way that $(\mathrm{Tr}C)^2/\mathrm{Tr}C^2$ counts the effectively excited covariance modes in the bulk: a counting that is literal while $S_2 > 0$, and which past $D_{\mathrm{eff}} \sim 1$ turns into a measure of how completely the rare deep-tail configurations dominate. It *collapses* ($D_{\mathrm{eff}} \to 0$) as they take over, and the distribution of $\mathcal{W}^2$ develops a tail heavier than the bulk generalized-χ^2 of Section 5.2.

The collapse of $D_{\mathrm{eff}} \to 0$ can be seen as a massive failure of self averaging: typical values not only fluctuate by a large amount compared to the mean, but even the mean diverges from typical draws, because the BW distance is dominated by rare tail events: branes tunneling into the negative-energy region. The analogue of the breathing mode of Section 5 is here literally E_1.

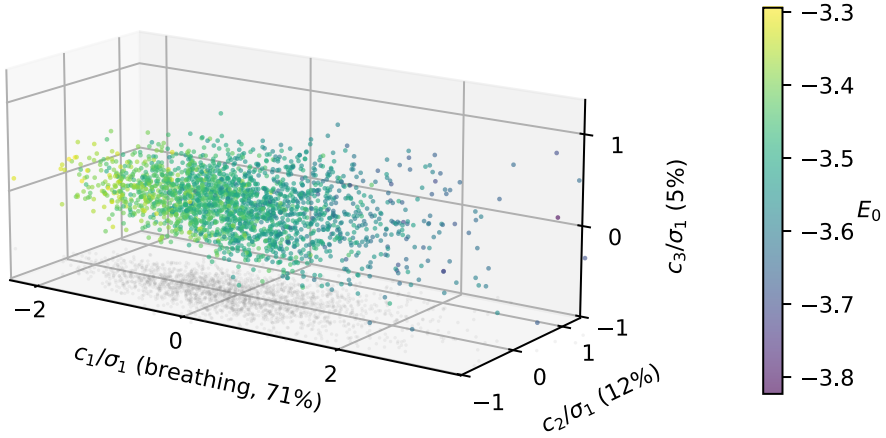


Figure 9: The unfolded SYK theory cloud (2100 realizations, N=18, β=2) in its three leading covariance eigendirections, coloured by ground-state energy E_0 and drawn at true relative scale: the box proportions equal the data ranges (variance shares in parentheses), and the c_3 axis carries the same range as c_2 to make the flattening visible. The cloud is a thin, strongly anisotropic shell dominated by the breathing direction, with the ground-state energy grading along it; the grey floor points are the projection onto the (c_1, c_2) plane. Coordinates are in units of the breathing standard deviation σ_1.

6.1 SYK: the gravitational ensemble member by member

We test the picture against the Sachdev–Ye–Kitaev model, whose disorder ensemble is a microscopic realization of the near-extremal gravitational ensemble, and, crucially, one whose members can be held in the hand: each disorder realization is an actual theory, a point of the cloud, so the geometry of Section 5 can be reconstructed directly. We exactly diagonalize Majorana q=4 SYK [45, 46, 47] in a single fermion-parity sector, chosen so that the sector statistics lie in the GUE class, at N=18 (2100 realizations) and N=22 (590 realizations), at $\beta = 2$; the model conventions, the symmetry bookkeeping and the statistics checks are collected in App. A. The thermal window then holds $n_{\rm part} \approx 27$ levels at N=18 (77 at N=22), the soft-edge regime of this section. The asymptotic ramp-regime scalings ($D_{\rm eff} \sim S^2$ foremost) live at large $n_{\rm part}$ (the ramp itself is visible in SYK exact diagonalization [48], but the slow $\ln \bar{n}$ and S^2 growths require windows of 10^2–10^3 levels at high statistics, out of reach there) and are tested with GUE in App. B.

The SYK disorder ensemble is not, as such, the ensemble dual to a single gravitational theory: its dominant fluctuations move each member's smooth spectral data (the overall energy scale first among them), whereas the double-scaled matrix ensemble of JT holds the spectral curve fixed, with only the microstates fluctuating around it. Studying the raw ensemble therefore means studying a family of theories with slightly different coarse data (to leading order the same theory in different units, $H \to \lambda H$), while the question posed in this paper is the opposite one: the shape of the ensemble at *fixed* coarse data. Reducing the former to the latter is the standard unfolding procedure of spectral statistics [25, 48, 49]: it quotients out the coarse directions member by member, leaving the microstate fluctuations that the wormhole computes. The details of the unfolding process we implement are in App. A.

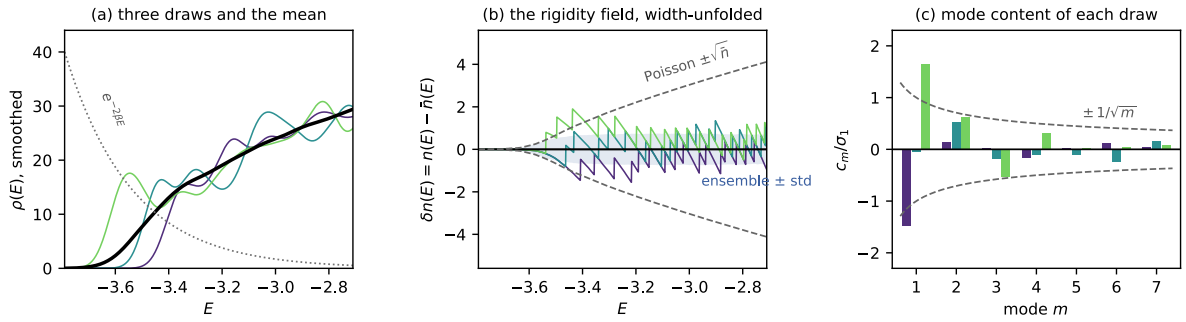


Figure 10: The SYK ensemble microscopically (N=18, β=2). (a) Gaussian-smoothed spectral densities (width 0.05, about one edge gap) of three draws spanning the breathing distribution (the 6%, 50% and 94% quantiles of c_1), against the ensemble mean (black); the dotted line indicates the Boltzmann window $\mathrm{e}^{-2\beta E}$ through which the BW distance reads the spectrum. (b) The counting fluctuation $\delta n(E) = n(E) - \bar{n}(E)$ of the same three draws after width unfolding: it stays of order one (the shaded band is the ensemble standard deviation), while uncorrelated levels would fill the growing Poisson envelope $\pm\sqrt{\bar{n}}$ (dashed). This bounded wander is the spectral rigidity the double trumpet computes. (c) The same fluctuations decomposed onto the principal modes of the cloud covariance: bars are each draw's coefficients c_m/σ_1; the dashed envelope is the $\pm 1/\sqrt{m}$ scaling found to be a reasonable approximation of the mode sizes at order-one temperatures (Section 5.2). The fluctuation is dominated by the leading "breathing" mode.

On the master curves, and the large-N trend. On the Airy master curves of Fig. 8 the SYK points probe the high-temperature side of the edge crossover ($\tau \lesssim 1$), the Tracy–Widom-dominated tail being charted by the Airy ensemble itself. The SYK means sit systematically above the master curve, and the excess grows along it: the raw means by a factor 3–4 at $\tau \approx 0.1$, rising to 30–60 by $\tau \approx 0.75$, reduced by width-unfolding to roughly 2–9. The unfolded *skewness*, by contrast, lies on the master curve at every τ, the best agreement in Fig. 8, while the raw skewness scatters far off it. The offset in the means arises from the same collective fluctuations unfolding attempts to remove, a removal that gets more and more difficult at large N.[14]

The ensemble cloud. With the unfolding at hand, Figure 9 shows the resulting theory cloud: the strongly anisotropic ellipsoid of Section 5, dominated by a single "breathing" axis that tracks the ground-state energy. One possible source of confusion is worth clearing up. In Section 5.2 we described the cloud as a thin shell, yet the displayed point cloud looks filled in rather than hollow. There is no tension: the shell lives in the *radial* coordinate of the full high-dimensional cloud, the squared distance to the ensemble-mean theory (equal in expectation to half the BW distance between two draws). Any low-dimensional projection of a high-dimensional shell looks filled, just as the marginals of a high-dimensional Gaussian do. The shell is thin when many modes contribute, std/mean $\sim 1/\sqrt{D_{\mathrm{eff}}}$: sharply so in the bulk regime, where asymptotically $D_{\mathrm{eff}} \simeq 0.42\, S_0^2$, while in the edge regime shown here $D_{\mathrm{eff}} = O(1)$ (App. A) and the shell is only incipient. What is fully developed at accessible sizes is the anisotropy.

[14] Width-unfolding removes roughly half of the offset at both sizes (the coupling-norm share), and the remainder is the residual wander of the spectral edge, which *grows* in Airy units with N. The fluctuations that move us from one gravitational theory to another are relatively small, of order $1/N^2$; but the edge sits at extensive energy, so they move the density of states at the absolute scale $\Delta E \sim 1/N$. The Airy edge, by contrast, lives in an exponentially narrow window, $\Delta E_{\mathrm{Airy}} \sim \mathrm{e}^{-2S_0/3}$ with $S_0 \propto N$. Zooming to the edge at larger N, these fluctuations therefore overwhelm the universal edge physics, and unfolding cannot remove them all (App. A).

The ensemble microscopically. Figure 10 shows the same content at the level of the spectra themselves: the smoothed densities of individual draws breathe coherently around the ensemble mean (a); after unfolding, their counting fluctuations stay of order one inside the growing Poisson envelope (b), the spectral rigidity the double trumpet computes, and decompose onto the principal modes of the cloud covariance with the breathing mode dominant (c). The mode structure itself survives the split intact (the actual mode functions and the eigenvalues are shown in Fig. 11, App. A): raw or unfolded, the eigenvectors of the cloud covariance carry $m-1$ nodes and are localized on the thermal edge, the structure predicted by the continuum diagonalization of Section 5.2, while the eigenvalue ladder falls distinctly faster than the asymptotic $1/m$ in this edge regime: a factor $\sim$3 below it already at $m = 2$ for the unfolded cloud ($\sim$6 raw) in Fig. 11(b), with the GUE control, at $\sim$1.5, sitting closest to $1/m$. The per-draw mode amplitudes in panel (c) of Figure 10 show the same steepening against the dashed $1/\sqrt{m}$ envelope. The robust microscopic statement of this section is geometric: the gravitational ensemble is visible as an anisotropic, breathing-dominated point cloud whose universal core carries the wormhole's mode structure.

7 Discussion

A tight, self-averaging ensemble. The BW self-distance is a direct *metric* measure of the width of the cloud of theories the gravitational path integral averages over, and the cloud it reveals is unusually simple: exponentially small, of typical separation $\sqrt{\langle \mathcal{W}^2 \rangle} \sim \sqrt{S}\, \mathrm{e}^{-S/2}$ (Section 5), and self-averaging (because its effective dimension grows, asymptotically as $D_{\text{eff}} \simeq 0.42\, S^2$), a single typical pair of draws realizing the ensemble-mean separation to relative accuracy $\sim 1/S$ (Section 5.3). This sharpens, at the metric level, the sense in which gravity computes an ensemble rather than a theory: the ensemble is genuine (the separation is nonzero, set by the wormhole and regulated by the plateau), but it is narrow and self-averaging, so for distance-based observables a single typical member nearly stands in for the whole. The averaging behind the factorization puzzle is, in this metric, a parametrically small effect. This bears on a live interpretational question: Johnson has argued, from the same non-perturbative matrix-model data, that a definite discrete microstate spectrum underlies JT gravity, extracted level by level via Fredholm determinants [8, 9], with the ensemble a calculational device and the factorization puzzle thereby dissolved. The cloud computed here is the complementary, ensemble reading of the same level distributions: the per-level distributions, which in [9] blur a single physical spectrum, are, read as an ensemble, precisely the cloud's extent. The BW distance quantifies how much hangs on the choice between the readings: a normalized self-distance $\sim S\, \mathrm{e}^{-2S}$, so at order-one temperatures the ensemble, if it is one, is operationally almost a point (at the Airy-edge temperatures of Section 6, where a single level dominates, the readings genuinely part ways). The BW distance is a functional of the spectra alone (the minimization over U discards the relative eigenbasis), so the cloud characterized here is the ensemble's *spectral* projection. In the matrix model, where a member is its spectrum, the distinction is empty; for a microscopic ensemble like SYK it is not.

Distinct theories, and a resolution floor. Everything above concerns two draws from *one* ensemble, where the smooth spectra coincide and the entire distance is the variance term of $\langle \Delta n^2 \rangle = \langle \Delta n \rangle^2 + \mathrm{Var}(\Delta n)$. The companion paper [1] occupies the opposite end: two macroscopically distinct theories (BTZ at different cosmological constant, a $T\bar{T}$ deformation), where the mean term dominates and becomes the horizon-area comparator $\frac{1}{8}(\delta A/4G)^2$. When both terms are present, the rigidity computed here is the *noise floor* under the comparator: in the normalised distance of [1], $\langle \tilde{\mathcal{W}}^2 \rangle \simeq \frac{1}{8}(\delta A/4G)^2 + O(S\, \mathrm{e}^{-2S})$, so a difference in horizon area is resolvable only down to $\delta A/4G \sim \sqrt{S}\, \mathrm{e}^{-S}$. Below that scale two geometries are, for the BW metric, the same point of the cloud, an irreducible graininess of theory space at the scale of

the level spacing itself, as it should be: with e^S states to tell apart, no spectral metric can resolve theories more finely than their levels are spaced. (Within JT itself the area reading of the smooth term is unavailable, as the controlled deformations act on the density essentially linearly, adding states rather than moving a horizon [50, 51]; the mixed regime therefore properly lives in higher-dimensional near-extremal settings.)

What the wormhole computes. The double trumpet does not bridge two members of the ensemble; it is the *covariance* of the ensemble: the shape of the cloud, not a tube between theories. (This statistics reading of wormholes has been developed for boundary data beyond the spectrum [52, 53, 54]; within a single draw the wormhole saddle persists, factorization being restored by coupling-sensitive "half-wormhole" contributions [55].) Indeed, to the order computed here the BW distance is a functional of the connected spectral form factor: (23) reads $\langle Z(\beta+it)Z(\beta-it)\rangle_c$ over the ramp, from the window scale to the Heisenberg time (the rigidity–form-factor link being itself classical [24, 25]), and the cumulant tower is built from the same two-boundary data (Section 3.3): no gravitational amplitude enters that the moments of Z would not already supply. Read deflationarily, the calculation is a repackaging. But the question ran the other way: the BW distance is *given* (an optimal-transport metric defined with no reference to gravity), and the result is that this metric has a sharp gravitational identity: it *is* the spectral rigidity, the ramp regulated by the plateau, explicit in JT. What survives either reading is structural. The metric organizes the two-boundary data into a geometry (size, principal modes, concentration) and into a self-averaging pairwise diagnostic of chaos: $\langle \mathcal{W}^2\rangle \sim S\, e^{-S}$ for a rigid spectrum against an unsuppressed $O(1)$ for uncorrelated levels (Section 4.2), sharp already from a single pair of draws, where the form factor at fixed time is not self-averaging at all [33].

Scope: the near-extremal band. The relation (23) is itself dimension-independent; what is not automatic is the clean output. The weight $e^{-2\beta E}/\bar{\rho}$ drives the distance to the bottom of the spectrum, so everything turns on whether the bottom is a macroscopic chaotic band or the lowest few levels. A near-extremal throat [56] makes it the former: the e^{S_0}-enhanced density puts $\sim e^{S_0}$ microstates inside the thermal window, the rigidity logarithm builds to $\Sigma^2 \sim S_0/2\pi^2$, and the JT results follow verbatim, for charged or rotating black holes and any setting with a Schwarzian throat. Absent the enhancement, the distance localizes on the first handful of levels: black-hole microstate statistics *without* a near-extremal limit fall outside the clean $S\, e^{-S}$ form. The fine-grained statistics of these lowest states in pure 3d gravity is a developing subject that has seen a great deal of interesting recent progress [57, 51, 58, 59, 60].

Open directions. The main structural gap is precisely that *non-extremal* case. Away from a near-extremal throat the Boltzmann weight still drives the distance to the bottom of the spectrum, but the bottom now holds only the lowest few chaotic levels: the BW distance becomes a near-edge observable: the machinery of Section 6, the master function and its brane tail, anchored to the edge of a generic black-hole band rather than an e^{S_0}-enhanced one. Developing this regime would extend the analysis to black holes without a near-extremal limit, in any dimension. BTZ would be the natural first target.

In one line: the gravitational path integral equips the space of theories it averages over with a metric; in JT that metric is the spectral form factor read as geometry: an exponentially small, stiff cloud, its fluctuations led by a single breathing mode, sharp enough that a single pair of draws already measures it.

Note on the preparation of this work. This paper was written in a back-and-forth collaboration between Claude (Anthropic; models Opus 4.8 and Fable 5) and VK, over a time frame of a few weeks around June 2026, outside VK's main professional activity (quantitative finance, currently machine learning). Claude did all the calculations first, and VK checked them by hand

or by independent simulation — except for the SYK appendix, which is Claude's throughout. Writing was likewise done first by Claude; VK reviewed it and rewrote some sections. The GUE appendix (App. B) was reworked in November 2026, after VK judged the original pooled fit too optimistic: the β-resolved reanalysis and the regenerated ensembles there are Claude's (Fable 5), cross-checked against the anchors of the original run. As internal quality control before any human review, the manuscript was examined by several referee agents prompted as experts in their subfields (JT gravity, random matrix theory, and quantum information/black holes), whose reports led to revisions. A near-final version was then reviewed, in several passes, by Claude Fable 5.2 in a separate session that had taken no part in the writing. It read the manuscript and figures, recomputed the quoted JT constants, and ran its own GUE simulations. Its reports led to the sharpened scope statements in the abstract and Sections 5 and 6, the corrected bookkeeping in (5), and a number of smaller fixes. Two quantitative statements originate in that review: the level-versus-count constant of Section 3.3, which was found numerically and is not derived, and the finite-N prediction it enables in App. B; both were then reproduced independently by the authors. This was not independent peer review: the reviewer is a model of the same family as the co-author, and it re-examined revisions made in response to its own comments. It did not rerun the SYK numerics, and of the references it checked only the Johnson entries. Responsibility for all content rests with VK.

A Anatomy of the SYK ensemble

This appendix records the model bookkeeping behind Section 6.1 and the quantitative decomposition of the SYK theory cloud into its collective and universal parts.

Model and data sets. We diagonalize $H = \sum_{i<j<k<l} J_{ijkl}\,\gamma_i\gamma_j\gamma_k\gamma_l$ with $\langle J_{ijkl}^2\rangle = 3!/N^3$ and $\gamma_i^2 = 1$, in a single fermion-parity sector: dimension 256 at N=18 and 1024 at N=22. For $N \equiv 2, 6 \pmod 8$ each parity sector lies in the GUE class [61, 62, 48], matching the double-scaled/JT ramp [4], and the two parity sectors are exactly degenerate, so fixing one is mandatory for level statistics; the measured mean gap-ratio statistic [63] is $\langle r\rangle = 0.60$, the GUE value [64]. Width unfolding rescales each member's spectrum by the one overall factor that matches its second moment to the ensemble mean, $E \to E\sqrt{\langle m_2\rangle/m_2}$; no shift arises, since $\mathrm{Tr}H = 0$ identically in a parity sector. This is the one-parameter version of the standard unfolding by the smooth counting function used for GUE in App. B; asymptotically in N every low sample moment fluctuates at relative order $1/N^2$, so ever more collective modes cross the exponentially shrinking universal scales and a full per-sample unfolding by each member's smooth density becomes necessary, the standard practice in the SFF literature [48, 49].

The collective fluctuations, measured. The sample second moment $\mathrm{Tr}H^2/\mathcal{D} = \sum_a J_a^2$ fluctuates between draws by $\sqrt{2/M} = 2.6\%$ with $M = \binom{N}{4} = 3060$ at N=18, and the measured value agrees (2.56%; since $\mathrm{Tr}H^2/\mathcal{D} = \sum_a J_a^2$ is an identity, this is a consistency check of the pipeline rather than independent physics); the cubic and quartic sample moments fluctuate at similar order [65]. These collective fluctuations are power-law small in $1/N$, parametrically *larger* than the $\sqrt{S_0}\,\mathrm{e}^{-S_0}$ amplitude at which the smooth density breathes in the matrix model (Section 3.3), and at N=18 they dominate the raw cloud: the three lowest sample moments explain 72% of $\mathrm{Var}(E_0)$, against 22% in a GUE control at matched n_{part}, which has no collective modes; and the matched GUE cloud has breathing share (the leading eigenvalue's fraction $\lambda_1/\sum_m \lambda_m$ of the cloud covariance) 52%, not the raw SYK 84% (with a comparable E_0-alignment $|\mathrm{corr}(c_1, E_0)| = 0.90$, where the alignment is generic to the edge regime; the anomalous *magnitude* is not).

The split. The two contributions separate cleanly. Unfolding each member by its own spectral width drops the breathing share from 84% to 71%; removing all three low moments (regressing the weighted-level vectors on the sample (m_2, m_3, m_4) and subtracting the fitted part) lands the cloud on the GUE baseline (share 0.48 against 0.46 for the GUE control processed through the same moment removal, with the residual E_0 fluctuation 0.42 mean gaps vs. GUE's 0.47; the slight undershoot of the Tracy–Widom value reflects the regression absorbing a little genuine edge fluctuation, which is exactly why the baseline is processed identically). The raw SYK cloud is therefore a *mixture*: roughly two thirds collective smooth-density fluctuation (each pair of draws differs slightly in its coarse density, the microscopic finite-N analogue of comparing two theories with different smooth spectra, the regime of the companion paper [1]), riding on a universal rigidity cloud that quantitatively matches GUE. Unfolding performs, member by member, precisely the smooth-versus-rigidity split that the master relation (23) formalizes: the wormhole cloud is the unfolded one. In the JT matrix model there is nothing to remove: the coarse density carries no independent fluctuation at all (its "breathing" is the long-wavelength part of the double trumpet itself, of relative size $\sqrt{S_0}\,\mathrm{e}^{-S_0}$, already included), whereas the SYK coupling scale wanders at $\sim\sqrt{2/M} \sim 1/N^2$, which with $S_0 \propto N$ dwarfs every e^{-S_0} effect in this paper at any N, not merely at accessible ones. Unfolding is thus not a discarding of data but a scheme choice: measuring each member's temperature in its own coupling: fixing β in units of the sample's bandwidth rather than absolutely.

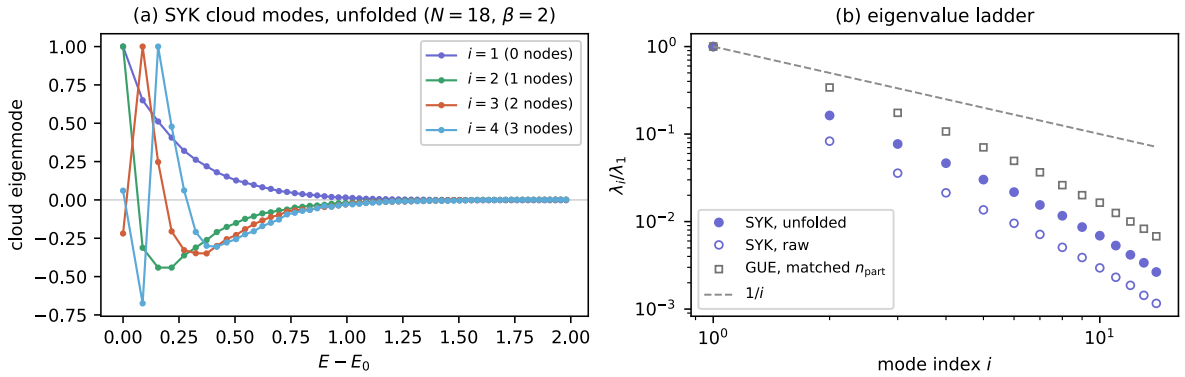


Figure 11: Cloud principal modes from SYK (N=18, β=2, 2100 realizations, unfolded spectra). (a) Eigenfunctions in energy: mode m has $m-1$ nodes, edge-localized, matching the double-trumpet prediction of Fig. 5; the structure is unchanged between raw and unfolded clouds. (b) Eigenvalue ladders: steeper than the $1/m$ ramp asymptote in the edge regime; the collective mode inflates λ_1 of the raw ladder, and the GUE control (no collective modes) sits closest to $1/m$. The panels' mode index i is the text's m.

Edge-regime numbers and the size trend. At accessible N the effective dimension is $O(1)$ ($D_{\text{eff}} = 1.4$ raw, 1.9 unfolded, against 3.1 for the matched GUE control), the variance is carried by essentially the single breathing mode, and $P(\mathcal{W}^2)$ is strongly right-skewed. The N=22 set confirms the size trend of Section 6.1 directly: the coupling-norm fluctuation is 1.6% against the predicted $\sqrt{2/\binom{22}{4}} = 1.65\%$, the three low moments now explain 76% of Var(E_0) (against 72% at N=18), and the raw breathing share rises from 84% to 90%, a trend that could be mistaken for deepening edge domination, and is in fact the collective mode outgrowing the exponentially shrinking universal scale: in Airy units the residual edge wander grows as $\sim e^{2S_0/3}/N^2$ up to factors polynomial in N (for SYK $S_0 = s_0 N$ with $s_0 \approx 0.232$ at q=4 [47]), so the exponential wins. On the master curves of Fig. 8 (SYK points placed by fitting the edge scale σ, $\tau = \beta\sigma$, to the ensemble-mean edge levels) the growth is seen directly: the width-unfolded std(b_1) is 1.15 at N=18 and 1.56 at N=22 (raw: 1.48 and 2.04), against the Tracy–Widom 0.90. The σ-fit protocol itself introduces no such offset: a GUE control (N=200, 7500 realizations) placed on the master curve by the identical fit lands on it to 1–2% across $\tau = 0.04$–0.34, with std(b_1) = 0.89: the SYK offset is collective, not a density-shape artifact of the protocol.

B GUE tests of the bulk-regime predictions

The SYK runs of Section 6.1 realize the gravitational ensemble *member by member*, but at accessible N the Boltzmann weight sits on the soft spectral edge (the Airy/Tracy–Widom regime), where $D_{\text{eff}} = O(1)$ and the variance is carried by a single breathing mode, itself dominated by the collective coupling fluctuations identified there. The bulk, ramp-regime predictions of Sections 3–5 (the logarithmic rigidity, the growth $D_{\text{eff}} \sim S^2$, the generalized-χ^2 law, and the mode tower) only become visible once the thermal window holds many levels, $n_{\text{part}} \sim 10^2$–$10^3$, which is out of reach for SYK exact diagonalization. The Gaussian unitary ensemble is the natural testbed: it is the universal bulk limit whose two-level correlator is exactly the sine-kernel ramp (the bulk double trumpet), its *lower* soft edge plays the role of the JT threshold (same Airy class, same $\sqrt{E}$ density exponent, same logarithmic count covariance (37) near it), and n_{part} is tuned freely through the matrix size N and the temperature. We diagonalize GUE in the semicircle normalization (eigenvalues on $[-\sqrt{2}, \sqrt{2}]$), with 7500/3600/2400 realizations at $N = 200/400/600$ on a grid $\beta \in [0.125, 4]$, and form the BW distance by the exact comonotone

(rank) pairing of (11) on the absolute weights $a_n = \mathrm{e}^{-\beta E_n}$ (without subtracting each draw's ground state), pairing distinct realizations. One methodological note: pair-based estimates of the moments of $\mathcal{W}^2$ reuse each realization in many pairs, so naive errors are invalid; every error bar below is a bootstrap over *realizations*, re-pairing within each resample.

Rigidity and the linearization. With the spectra unfolded by the exact semicircle counting function, the number variance in a bulk window of L mean levels grows logarithmically with the predicted slope: measured 0.103 against $1/\pi^2 = 0.101$, the *two*-endpoint (Dyson–Mehta) slope, twice the single-edge $1/2\pi^2$ of (28), as appropriate for a window with both edges in the bulk; the additive constant matches the two-edge $(\ln 2\pi L + \gamma + 1)/\pi^2$ at the percent level. With the window instead *anchored at the spectral threshold*, the configuration the BW distance uses, the variance follows the *single*-edge (index) form: in the arc variable $E = -\sqrt{2}\cos\varphi$ the measured $\mathrm{Var}\, n(E) = \frac{1}{2\pi^2}\big[\ln N + 3\ln\sin\varphi + c_0\big]$ across the band ($c_0 \simeq 3.7$; deviations only inside the Airy zones), and its growth from $N = 200$ to 600 is 0.047 ± 0.012 against the single-edge $\frac{1}{2\pi^2}\ln 3 = 0.056$, the two-edge 0.111 excluded: the direct form of the silent threshold of Section 4.1. (Polynomial unfolding inflates the slope by $\sim 20\%$; a residual quoted in an earlier version of these numerics was that artifact.) Independently, the exact distance $\langle \mathcal{W}^2 \rangle = 2\sum_n \mathrm{Var}(a_n)$ agrees with the linearized rigidity form $2\beta^2 \sum_n \mu_n^2\, \mathrm{Var}(E_n)$ of (13) to 0.1% for $\beta \le 2$ and 1% at $\beta = 6$, confirming the small-fluctuation linearization of Section 3.3. Neither warrants a figure.

Counting statistics and the Wick baseline. The discreteness constants of Section 4.3 are measured on the same unfolded spectra ($N = 600$). The fourth cumulant of the count in a window of L mean spacings, pooled over disjoint windows in the central third of the spectrum, carries a unit-period oscillation in L (the subleading Fisher–Hartwig branches [30]) about the asymptotic constant; phase-averaged over $L \in [4, 5)$ it gives $\kappa_4 = -0.017(2)$ against the predicted $-24\zeta(3)/(2\pi)^4 = -0.0185$ of (39). The Wick baseline for the moments of $\mathcal{W}^2$ is built from the measured level-weight covariance $\hat{S}_{nm} = \mathrm{Cov}(a_n, a_m)$: comparing the measured mean, variance and skewness against $2\mathrm{Tr}\hat{S}$, $8\mathrm{Tr}\hat{S}^2$ and $64\mathrm{Tr}\hat{S}^3/(8\mathrm{Tr}\hat{S}^2)^{3/2}$ gives the excesses quoted in Section 4.3 (1% on the variance deep in the bulk, growing to $\sim 10\%$ as the weight moves onto the edge), and the joint fourth cumulant of the counts in adjacent windows, standardized by their variances, is bounded, $|\kappa_{22}| \lesssim 0.01$.

Self-averaging: D_{eff} grows, with a drifting coefficient (Fig. 12). The headline of the GUE suite: the effective dimension grows logarithmically at *every* temperature, from 3.2 at $n_{\mathrm{part}} = 30$ to 25 at $n_{\mathrm{part}} \simeq 600$, with no sign of the saturation an equicorrelated cloud would force. An earlier version of these numerics summarized the growth by a single pooled fit $\sqrt{D_{\mathrm{eff}}} = a \ln n_{\mathrm{part}} + b$ over the whole (N, β) grid, finding $a^2 = 0.42 \pm 0.07$, numerically the universal edge coefficient of Fig. 7; that agreement was too optimistic, and we retract it in favor of the β-resolved analysis below. The pooled form assumes D_{eff} is a function of n_{part} alone, but $\ln n_{\mathrm{part}} = \ln N + \ln g(\beta)$ conflates the two variables, and for GUE the true coefficient depends on β: in the arc variable $E = -\sqrt{2}\cos\varphi$ the smoothed count covariance of the semicircle is

$$\big\langle n(E)\, n(E') \big\rangle_{\mathrm{c}} = \frac{1}{2\pi^2} \ln \frac{\sin\frac{\varphi+\varphi'}{2}}{\big|\sin\frac{\varphi-\varphi'}{2}\big|} = \frac{1}{2\pi^2} \sum_{m\ge 1} \frac{2}{m} \sin m\varphi \, \sin m\varphi' \tag{61}$$

(the Fourier diagonalization of GUE fluctuations that underlies Johansson's CLT [22]; verified on the data with no free parameters, rms deviation below the statistical error; its lower-edge limit is (37)), and the trace ratio built on it exactly as in Section 5.3 gives a β-dependent target $a_{\mathrm{sc}}^2(\beta) = J_{\mathrm{sc}}^2/I_{\mathrm{sc}}$, with $J_{\mathrm{sc}} = \int_0^\pi \mathrm{d}\varphi\, \mathrm{e}^{-2\beta E(\varphi)}$ and $I_{\mathrm{sc}} = \iint_0^\pi \mathrm{d}\varphi\, \mathrm{d}\varphi'\, \mathrm{e}^{-2\beta(E+E')} \ln^2[\,\cdot\,]$ the squared-log double integral. It interpolates between the single-edge $4/\pi^2 \simeq 0.41$ as $\beta \to \infty$ and, because

the Fourier modes in (61) carry the $1/m$ ladder of Section 5.2, the exact value $6/\pi^2 \simeq 0.61$ at $\beta \to 0$, where both band edges are active: across the measured grid the target itself moves by 50%.

Resolved in temperature, the measured slopes $a^2_{\rm meas}(\beta) = [\partial\sqrt{D_{\rm eff}}/\partial \ln N]^2$ (three sizes per β) drift from 0.23 ± 0.03 at $\beta = 4$ to 0.49 ± 0.11 at $\beta = \frac{1}{8}$ — the predicted direction and a comparable magnitude of drift — while sitting systematically below the asymptotic curve (Fig. 12b). The gap is not Gaussianity: Wick traces of the *measured* weight covariance reproduce the pair-measured $D_{\rm eff}$ to 1–4% across the whole grid. It is two $O(1)$ constants in the covariance, and both can be identified. First, the diagonal of (61) regulated at the universal local cutoff carries the bracket constant $\gamma+1+3\ln 2 \simeq 3.66$, measured here as 3.7 ± 0.2: the 8π form of the footnote under (28). Second, the distance is built from the *level-position* variance, which sits below the count variance by the integer-count constant of Section 3.3: $\bar\rho^2\,{\rm Var}(E_n) = \Sigma^2 - \frac{1}{12}$, measured (bootstrap over realizations, pooled over $N = 200$–600 with the residual inter-size scatter folded into the error) as 0.085 ± 0.004 against $\frac{1}{12} = 0.083$. With both constants in place, the kernel (61) evaluated at the ensemble-mean level positions predicts, with *no free parameters*, every measured $D_{\rm eff}$ on the grid to 1–3%, degrading to 6% only at the most edge-dominated point (N=200, β=4), where Airy statistics take over; the open squares of Fig. 12b are the slopes of this prediction. The asymptotic coefficient itself is still approached only logarithmically, the constants being $O(1)$ against $\ln N \simeq 5.3$–6.4, but the approach is now modeled rather than merely bounded. The pooled fit, regressing this drifting, constant-shifted slope on the conflated variable, averaged it into a number that happened to land on the universal value (its $\chi^2/{\rm dof} \approx 13$ was the warning). What the data establish is therefore: logarithmic growth of $D_{\rm eff}$ at every β (the self-averaging), the predicted drift of its coefficient, quantitative agreement with the finite-N covariance model, and the single-edge rigidity confirmed directly above; the two-edge coefficient ($a^2 \simeq 0.8$ in the pooled language) is excluded by every measured slope. The mechanism behind the growth is the decay of the displacement correlation, measured at N=600 in the lower-mid spectrum: ${\rm corr}(E_n, E_m) = 0.77, 0.70, 0.62, 0.52, 0.44, 0.37, 0.28$ at separations $|n-m| = 1, 2, 4, \ldots, 64$: the cutoff-free off-diagonal of (37).

The distribution: a split-sample test (Fig. 13). Diagonalizing the sample covariance of the weighted levels on *one half* of the realizations gives eigenvalues ν_m, and (46) predicts $\mathcal{W}^2 = \sum_m 2\nu_m z_m^2$; the histogram is measured on the held-out half ($N = 400$, $\beta = 0.25$), so the comparison has no free parameters and no reuse of data. The law is reproduced in peak and tail; the moments give std/mean = 0.304 measured versus 0.316 predicted and skewness 1.39 versus 1.62. The residuals have three identified sources: finite-sample noise in the covariance inflates the predicted ${\rm Tr}S^p$ (a relative bias $\sim D_{\rm eff}/R$ with R the realizations behind the covariance), the leading-Wick corrections of Section 4.3 enter at the few-percent level, and the sample skewness of a right-skewed law is itself biased downward at $O(10^3)$ effective samples.

Principal modes. The eigenmodes of the sample covariance reproduce the tower of Section 5.2 exactly as in SYK: the leading modes carry $m-1$ nodes, edge-localized, with the leading "breathing" mode holding 25% of the variance at $\beta = 1$ followed by the slow tail $(12\%, 7\%, 5\%, 4\%, \ldots)$ of the anisotropic shell, and the eigenvalue ladder falls close to (though slightly steeper than) the $1/m$ ramp asymptote, as expected at these modest, edge-regime $n_{\rm part}$; the GUE ladder is included directly in Fig. 11(b), so no separate figure is needed here.

Together with Section 6.1, GUE and SYK bracket the picture: SYK realizes the gravitational ensemble member by member (its universal core, once the collective coupling fluctuations are unfolded away, matching the GUE edge), while GUE, reaching the large $n_{\rm part}$ where the universal ramp controls everything, tests the bulk-regime predictions the edge cannot expose: the logarithmic growth of $D_{\rm eff}$ (the content of $D_{\rm eff} \sim S^2$) is confirmed at every temperature, with a coefficient that drifts as the finite band predicts and approaches its asymptote only

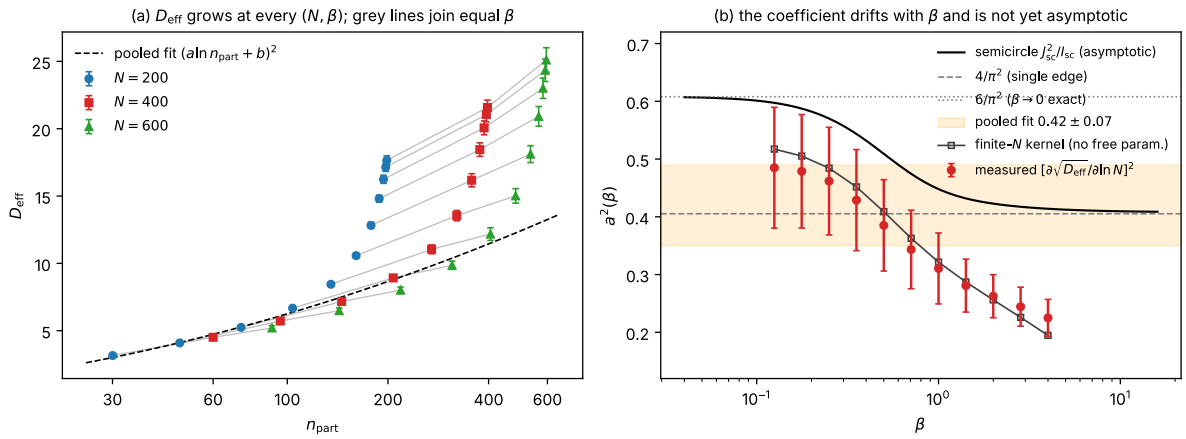


Figure 12: The effective dimension in GUE (comonotone BW distance on absolute weights; bootstrap-over-realizations errors). (a) D_{eff} against n_{part}: growth at every temperature; grey lines join equal-β points across $N = 200, 400, 600$ and fan away from the retired pooled fit (dashed), which mixes the two variables. (b) The coefficient, resolved in temperature: measured slopes $[\partial\sqrt{D_{\text{eff}}}/\partial \ln N]^2$ against the asymptotic semicircle target $J_{\text{sc}}^2/I_{\text{sc}}$ of (61), which interpolates between the single-edge $4/\pi^2$ and the both-edges $6/\pi^2$. The slopes drift as predicted; the open squares are the slopes of the finite-N kernel prediction with both $O(1)$ constants included (no free parameters), which tracks the data within errors. The pooled 0.42 ± 0.07 band is shown for reference.

logarithmically.

References

[1] V. Keranen, *Notes on Wasserstein distance and wormholes*, arXiv:2606.01296.

[2] P. Saad, S. Shenker, D. Stanford, *JT gravity as a matrix integral*, arXiv:1903.11115.

[3] D. Marolf, H. Maxfield, *Transcending the ensemble: baby universes, spacetime wormholes, and the order and disorder of black hole information*, JHEP **08** (2020) 044, arXiv:2002.08950.

[4] P. Saad, S.H. Shenker, D. Stanford, *A semiclassical ramp in SYK and in gravity*, arXiv:1806.06840.

[5] P. Saad, *Late time correlation functions, baby universes, and ETH in JT gravity*, arXiv:1910.10311.

[6] C.V. Johnson, *Nonperturbative Jackiw–Teitelboim gravity*, Phys. Rev. D **101** (2020) 106023, arXiv:1912.03637.

[7] C.V. Johnson, *Explorations of nonperturbative Jackiw–Teitelboim gravity and supergravity*, arXiv:2006.10959.

[8] C.V. Johnson, *Quantum gravity microstates from Fredholm determinants*, Phys. Rev. Lett. **127** (2021) 181602, arXiv:2106.09048.

[9] C.V. Johnson, *The microstate physics of JT gravity and supergravity*, arXiv:2201.11942.

[10] K. Okuyama, K. Sakai, *JT gravity, KdV equations and macroscopic loop operators*, JHEP **01** (2020) 156, arXiv:1911.01659.

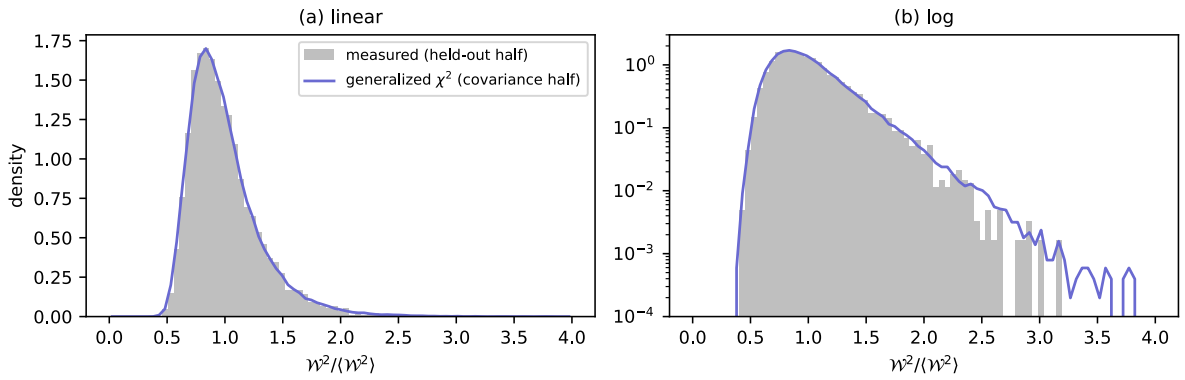


Figure 13: Split-sample test of the generalized χ^2 law ($N = 400$, $\beta = 0.25$, normalized by the mean): the prediction (46) is built from the covariance eigenvalues of one half of the realizations, the histogram from the held-out half: no free parameters, no data reuse. (a) Linear; (b) log scale.

[11] A. Altland, J. Sonner, *Late time physics of holographic quantum chaos*, SciPost Phys. **11** (2021) 034, arXiv:2008.02271.

[12] A. Altland, B. Post, J. Sonner, J. van der Heijden, E.P. Verlinde, *Quantum chaos in 2D gravity*, SciPost Phys. **15** (2023) 064, arXiv:2204.07583.

[13] C. Villani, *Optimal Transport: Old and New*, Springer (2009).

[14] G. De Palma, D. Trevisan, *Quantum optimal transport with quantum channels*, Ann. Henri Poincaré **22** (2021) 3199, arXiv:1911.00803.

[15] G. De Palma, D. Trevisan, *Quantum optimal transport: quantum channels and qubits*, in *Optimal Transport on Quantum Structures*, Springer (2024), arXiv:2307.16268.

[16] K. Hashimoto, N. Tanahashi, *Holography and optimal transport: emergent Wasserstein spacetime in harmonic oscillator, SYK and Krylov complexity*, arXiv:2604.17649.

[17] K. Hashimoto, N. Tanahashi, K. Yoshida, *Wasserstein space of quantum chaos*, arXiv:2605.20995.

[18] A. Blommaert, T.G. Mertens, H. Verschelde, *Eigenbranes in Jackiw–Teitelboim gravity*, JHEP **02** (2021) 168, arXiv:1911.11603.

[19] O. Costin, J.L. Lebowitz, *Gaussian fluctuation in random matrices*, Phys. Rev. Lett. **75** (1995) 69.

[20] A. Soshnikov, *Gaussian fluctuation for the number of particles in Airy, Bessel, sine, and other determinantal random point fields*, J. Stat. Phys. **100** (2000) 491.

[21] J. Gustavsson, *Gaussian fluctuations of eigenvalues in the GUE*, Ann. Inst. H. Poincaré **41** (2005) 151.

[22] K. Johansson, *On fluctuations of eigenvalues of random Hermitian matrices*, Duke Math. J. **91** (1998) 151.

[23] C.A. Tracy, H. Widom, *Level-spacing distributions and the Airy kernel*, Commun. Math. Phys. **159** (1994) 151.

[24] M.V. Berry, *Semiclassical theory of spectral rigidity*, Proc. R. Soc. Lond. A **400** (1985) 229.

[25] M.L. Mehta, *Random Matrices*, 3rd ed., Elsevier (2004).

[26] B.L. Altshuler, B.I. Shklovskii, *Repulsion of energy levels and conductivity of small metal samples*, Sov. Phys. JETP **64** (1986) 127.

[27] I.S. Gradshteyn, I.M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed., Academic Press (2007), Eq. 6.693.2.

[28] M.V. Berry, M. Tabor, *Level clustering in the regular spectrum*, Proc. R. Soc. Lond. A **356** (1977) 375.

[29] P. Deift, A. Its, I. Krasovsky, *Asymptotics of Toeplitz, Hankel, and Toeplitz+Hankel determinants with Fisher–Hartwig singularities*, Ann. of Math. **174** (2011) 1243, arXiv:0905.0443.

[30] A.G. Abanov, D.A. Ivanov, Y. Qian, *Quantum fluctuations of one-dimensional free fermions and Fisher–Hartwig formula for Toeplitz determinants*, J. Phys. A **44** (2011) 485001, arXiv:1108.1355.

[31] S. Dallaporta, *Eigenvalue variance bounds for Wigner and covariance random matrices*, Random Matrices Theory Appl. **1** (2012) 1250007, arXiv:1203.1597.

[32] E.S. Meckes, M.W. Meckes, *Concentration and convergence rates for spectral measures of random matrices*, Probab. Theory Relat. Fields **156** (2013) 145, arXiv:1109.5997.

[33] R.E. Prange, *The spectral form factor is not self-averaging*, Phys. Rev. Lett. **78** (1997) 2280.

[34] K. Karhunen, *Über lineare Methoden in der Wahrscheinlichkeitsrechnung*, Ann. Acad. Sci. Fenn. Ser. A. I. Math.-Phys. **37** (1947) 1.

[35] M. Loève, *Probability Theory II*, 4th ed., Graduate Texts in Mathematics **46**, Springer (1978), Ch. X.

[36] O. Bohigas, M.J. Giannoni, C. Schmit, *Characterization of chaotic quantum spectra and universality of level fluctuation laws*, Phys. Rev. Lett. **52** (1984) 1.

[37] P.J. Forrester, *The spectrum edge of random matrix ensembles*, Nucl. Phys. B **402** (1993) 709.

[38] I. Dumitriu, A. Edelman, *Matrix models for beta ensembles*, J. Math. Phys. **43** (2002) 5830, arXiv:math-ph/0206043.

[39] J.A. Ramírez, B. Rider, B. Virág, *Beta ensembles, stochastic Airy spectrum, and a diffusion*, J. Amer. Math. Soc. **24** (2011) 919, arXiv:math/0607331.

[40] F. David, *Phases of the large-N matrix model and non-perturbative effects in 2d gravity*, Nucl. Phys. B **348** (1991) 507.

[41] S.H. Shenker, *The strength of nonperturbative effects in string theory*, in *Random Surfaces and Quantum Gravity* (Cargèse 1990), Plenum (1991).

[42] V. Fateev, A.B. Zamolodchikov, Al.B. Zamolodchikov, *Boundary Liouville field theory I. Boundary state and boundary two-point function*, arXiv:hep-th/0001012.

[43] J. Teschner, *Remarks on Liouville theory with boundary*, arXiv:hep-th/0009138.

[44] S.N. Majumdar, G. Schehr, *Top eigenvalue of a random matrix: large deviations and third-order phase transition*, J. Stat. Mech. (2014) P01012, arXiv:1311.0580.

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

[46] A. Kitaev, *A simple model of quantum holography*, KITP seminars, April and May 2015.

[47] J. Maldacena, D. Stanford, *Remarks on the Sachdev–Ye–Kitaev model*, Phys. Rev. D **94** (2016) 106002, arXiv:1604.07818.

[48] J.S. Cotler, G. Gur-Ari, M. Hanada, J. Polchinski, P. Saad, S.H. Shenker, D. Stanford, A. Streicher, M. Tezuka, *Black holes and random matrices*, arXiv:1611.04650.

[49] H. Gharibyan, M. Hanada, S.H. Shenker, M. Tezuka, *Onset of random matrix behavior in scrambling systems*, arXiv:1803.08050.

[50] E. Witten, *Matrix models and deformations of JT gravity*, Proc. R. Soc. Lond. A **476** (2020) 20200582, arXiv:2006.13414.

[51] H. Maxfield, G.J. Turiaci, *The path integral of 3D gravity near extremality; or, JT gravity with defects as a matrix integral*, JHEP **01** (2021) 118, arXiv:2006.11317.

[52] J. Pollack, M. Rozali, J. Sully, D. Wakeham, *Eigenstate thermalization and disorder averaging in gravity*, Phys. Rev. Lett. **125** (2020) 021601, arXiv:2002.02971.

[53] A. Belin, J. de Boer, *Random statistics of OPE coefficients and Euclidean wormholes*, Class. Quantum Grav. **38** (2021) 164001, arXiv:2006.05499.

[54] J. Chandra, S. Collier, T. Hartman, A. Maloney, *Semiclassical 3D gravity as an average of large-c CFTs*, JHEP **12** (2022) 069, arXiv:2203.06511.

[55] P. Saad, S.H. Shenker, D. Stanford, S. Yao, *Wormholes without averaging*, JHEP **09** (2024) 133, arXiv:2103.16754.

[56] L.V. Iliesiu, G.J. Turiaci, *The statistical mechanics of near-extremal black holes*, arXiv:2003.02860.

[57] N. Benjamin, H. Ooguri, S.-H. Shao, Y. Wang, *Light-cone modular bootstrap and pure gravity*, Phys. Rev. D **100** (2019) 066029, arXiv:1906.04184.

[58] S. Collier, L. Eberhardt, M. Zhang, *Solving 3d gravity with Virasoro TQFT*, SciPost Phys. **15** (2023) 151, arXiv:2304.13650; S. Collier, L. Eberhardt, B. Mühlmann, V.A. Rodriguez, *The Virasoro minimal string*, SciPost Phys. **16** (2024) 057, arXiv:2309.10846.

[59] G. Di Ubaldo, E. Perlmutter, *AdS_3/RMT_2 duality*, arXiv:2307.03707; J. Boruch, G. Di Ubaldo, F.M. Haehl, E. Perlmutter, M. Rozali, *Modular-invariant random matrix theory and AdS_3 wormholes*, Phys. Rev. Lett. **135** (2025) 121602, arXiv:2503.00101.

[60] D.L. Jafferis, L. Rozenberg, D. Sarkar, D. Wang, *On random matrix statistics of 3d gravity*, arXiv:2512.05045.

[61] Y.-Z. You, A.W.W. Ludwig, C. Xu, *Sachdev–Ye–Kitaev model and thermalization on the boundary of many-body localized fermionic symmetry-protected topological states*, Phys. Rev. B **95** (2017) 115150, arXiv:1602.06964.

[62] A.M. García-García, J.J.M. Verbaarschot, *Spectral and thermodynamic properties of the Sachdev–Ye–Kitaev model*, Phys. Rev. D **94** (2016) 126010, arXiv:1610.03816.

[63] V. Oganesyan, D.A. Huse, *Localization of interacting fermions at high temperature*, Phys. Rev. B **75** (2007) 155111, arXiv:cond-mat/0610854.

[64] Y.Y. Atas, E. Bogomolny, O. Giraud, G. Roux, *Distribution of the ratio of consecutive level spacings in random matrix ensembles*, Phys. Rev. Lett. **110** (2013) 084101, arXiv:1212.5611.

[65] Y. Jia, J.J.M. Verbaarschot, *Spectral fluctuations in the Sachdev–Ye–Kitaev model*, arXiv:1912.11923.