

## Semiclassical Wormholes toward Typical Entangled States

Javier M. Magán<sup>1</sup>, Martin Sasieta<sup>2</sup>, and Brian Swingle<sup>2</sup>

<sup>1</sup>*Instituto Balseiro, Centro Atómico Bariloche, 8400-S.C. de Bariloche, Río Negro, Argentina*

<sup>2</sup>*Martin Fisher School of Physics, Brandeis University, Waltham, Massachusetts 02453, USA*



(Received 22 April 2025; accepted 17 September 2025; published 15 October 2025)

What do the typical entangled states of two black holes look like? Do they contain semiclassical interiors? We approach these questions constructively, providing ensembles of states that densely explore the black hole Hilbert space. The states contain very long *Einstein-Rosen caterpillars*: semiclassical wormholes with large numbers of matter inhomogeneities. Distinguishing these ensembles from the typical entangled states of the black holes is hard. We quantify this by deriving the correspondence between a microscopic notion of quantum randomness and the geometric length of the wormhole. This formalizes a “complexity = geometry” relation.

DOI: 10.1103/btw6-44ry

**Introduction**—ER = EPR postulates the remarkable equivalence between spatial wormholes in general relativity (Einstein-Rosen bridges, ER) and quantum entanglement (Einstein-Podolsky-Rosen, EPR) [1]. In essence, when two systems are quantum-mechanically entangled, a “wormhole” is conjectured to be connecting them.

The proposal is fundamentally rooted in holographic entanglement in the anti-de Sitter space-conformal field theory correspondence (AdS-CFT) [2]. There, the observation is that an honest geometric wormhole can emerge between two black holes when the microscopic quantum correlations are large enough [3]. The classic example is the thermofield double state (TFD) of two copies of the holographic CFT on  $\mathbf{R} \times \mathbf{S}^{d-1}$ ,

$$|\text{TFD}\rangle = \frac{1}{\sqrt{Z(\beta)}} \sum_i e^{-\frac{\beta}{2} E_i} |E_i^*\rangle_L |E_i\rangle_R, \quad (1)$$

where  $|E_i\rangle_{L,R}$  are energy eigenstates of the left ( $L$ ) and right ( $R$ ) CFT Hamiltonians  $H_{L,R}$ . The star in  $|E_i^*\rangle$  represents the action under CRT. On the one hand, this is an entangled state between the CFTs, which canonically purifies the thermal state of a single CFT. On the other hand, the TFD is dual to a connected spatial wormhole, corresponding to the Hartle-Hawking state of an eternal black hole in AdS space [4].

The TFD example is very illuminating but, as often emphasized [1,4–7], it is rather atypical, given that the TFD is a very special entangled state. For one thing, it lies in the

diagonal subspace, annihilated by  $H_L - H_R$ . Additionally, its wave function has large correlations between the two systems, captured geometrically by a short wormhole with a bifurcate Killing horizon. Such fine-tuned correlations are dynamically destroyed by small perturbations injected at a scrambling time in the past [5].

If ER = EPR is to represent a general principle, one might ask: is there a wormhole for a typical entangled state of two black holes?

**Typical EPR**—To motivate a definition of “typical EPR” that makes sense for CFTs in thermal equilibrium, we first consider a toy example consisting of two collections of  $N$  qubits, prepared in the maximally entangled state

$$|I\rangle = \frac{1}{\sqrt{D}} \sum_{i=1}^D |i^*\rangle_L |i\rangle_R, \quad (2)$$

where the  $|i\rangle \equiv |i_1\rangle \otimes \dots \otimes |i_N\rangle$  form an orthonormal basis and the dimension of each Hilbert space is  $D = 2^N$ . The state (2) is the infinite-temperature TFD, in which the systems share  $N$  Bell pairs.

We will define the ensemble of *typical entangled states*,  $\mathcal{E}_{\text{EPR}}$ , between these systems as the collection of states obtained by applying a random single-sided unitary to the TFD,  $|\Psi_{\text{EPR}}\rangle = U_R |I\rangle$ , or, more explicitly,

$$|\Psi_{\text{EPR}}\rangle = \frac{1}{\sqrt{D}} \sum_{i,j=1}^D U_{ij} |i^*\rangle_L |j\rangle_R, \quad (3)$$

where  $U$  is a Haar random unitary in  $\mathbf{U}(D)$ . The state (3) is a random purification of the maximally mixed state  $\rho_L = \rho_R = (1/D)\mathbf{1}$ .

Now, to define the typical entangled states of two black holes, we will proceed by analogy and generalize the notion above to finite temperature. In AdS-CFT, an equilibrium

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.

black hole has an associated equilibrium density matrix  $\rho$  of the CFT. For a neutral black hole, the state  $\rho$  is ensemble equivalent to the canonical Gibbs state  $\rho_\beta = e^{-\beta H}/Z(\beta)$  in the large- $N$  limit of the CFT.

We define the ensemble of *typical equilibrium entangled states*,  $\mathcal{E}_{\text{EPR}}^\rho$ , as the collection of states of the form  $(\sqrt{\rho} \otimes \sqrt{\rho})|\Psi_{\text{EPR}}\rangle$ , or explicitly,

$$|\Psi_{\text{EPR}}^\rho\rangle \propto \sum_{i,j} \sqrt{\rho_i \rho_j} X_{ij} |E_i^*\rangle_L |E_j\rangle_R. \quad (4)$$

Here,  $\rho_i$  is the eigenvalue of  $\rho$  in the state  $|E_i\rangle$ , and we have introduced  $X = \sqrt{D}U$  for the Haar random  $U \in \mathbf{U}(D)$ . With this normalization, we can safely take the infinite Hilbert space dimension limit  $D \rightarrow \infty$ . In this limit, the coefficients  $X_{ij}$  converge to i.i.d. Gaussian random variables with zero mean and unit variance. In Appendix A, we outline some of the properties of the ensemble  $\mathcal{E}_{\text{EPR}}^\rho$ . For example, on average, the reduced state is the equilibrium density matrix,  $\mathbb{E}[\rho_{L,R}] = \rho$ .

Our question becomes, do such typical equilibrium entangled states have wormholes associated with them? Our approach in this Letter is to explicitly construct ensembles of states with wormholes that better and better approximate such typical EPRs.

*A dense exploration of Hilbert space*—Random states and unitaries are ubiquitous in quantum physics, but implementing them in practice is computationally hard. In the present context, the objective is to, even in principle, construct a typical-looking entangled state of two black holes starting from the TFD.

A natural possibility is to consider the time evolution of the TFD state  $|\text{TFD}_t\rangle = e^{-itH_R}|\text{TFD}\rangle$ , or explicitly,

$$|\text{TFD}_t\rangle = \frac{1}{\sqrt{Z(\beta)}} \sum_i e^{-\left(\frac{\beta}{2} + it\right)E_i} |E_i^*\rangle_L |E_i\rangle_R. \quad (5)$$

As we elaborate in Appendix B,  $|\text{TFD}_t\rangle$  will never become a random EPR, even if the Hamiltonian is completely structureless. The main reason is the conservation of energy along the trajectory, which prevents the state vector from exploring most of Hilbert space.

This example nevertheless provides a more general instance of ER = EPR. From the spacetime perspective, the wormhole grows in time [1,8]. The geometric volume of the wormhole is conjectured to encode a suitable notion of “quantum complexity” of the time evolution [9].

The way to continuously reach random-looking entangled states is to instead consider time-dependent Hamiltonians driving the evolution. For the sake of clarity, we shall first describe this in the qubit system. We will consider a general time-dependent Hamiltonian of the form

$$H(t) = \sum_{\alpha=1}^K g_\alpha(t) \mathcal{O}_\alpha, \quad (6)$$

where the  $\mathcal{O}_\alpha$  form a collection of  $K$  Hermitian operators, normalized so that  $\text{Tr}(\mathcal{O}_\alpha^2) = D$ , with associated time-dependent couplings  $g_\alpha(t)$ . The time-evolution operator now becomes

$$U(t) = \mathbf{T} \left\{ \exp \left( -i \int_0^t ds H(s) \right) \right\}. \quad (7)$$

Applying it to the TFD generates states  $|\Psi_t\rangle = U(t)_R |I\rangle$ , or, explicitly,

$$|\Psi_t\rangle = \frac{1}{\sqrt{D}} \sum_{i,j=1}^D U(t)_{ij} |i^*\rangle_L |j\rangle_R. \quad (8)$$

In this Letter, we will select the couplings from a collection of white-noise correlated Gaussian random couplings,

$$\mathbb{E}[g_\alpha(t)] = 0, \quad \mathbb{E}[g_\alpha(t)g_{\alpha'}(t')] = J\delta_{\alpha\alpha'}\delta(t-t'). \quad (9)$$

Such couplings define an independent random Hamiltonian at each time step. Upon exponentiation,  $U(t)$  follows a multiplicative Brownian motion in  $\mathbf{U}(D)$ , which corresponds to the continuous-time analog of a random quantum circuit. In particular, at each time step, the time evolution implements an infinitesimal “random gate”  $\exp[-i\delta t H(t)]$ . At the level of the states (8), this Markov process defines an ensemble of states  $\mathcal{E}_t$ .

Under general assumptions for the drive operators, and in the absence of symmetries, it is possible to show that the measure of the ensemble of states  $\mathcal{E}_t$  converges weakly to the uniform measure at infinite times [10,11],

$$\lim_{t \rightarrow \infty} \mathcal{E}_t = \mathcal{E}_{\text{EPR}}, \quad (10)$$

or, equivalently, that  $U(\infty)$  is Haar random.

*Gradual cooling*—For black holes, we will consider a holographic random quantum circuit selecting a collection of perturbations  $\mathcal{O}_\alpha$  of the CFT associated with low-energy matter fields in the bulk. An obvious problem that arises here is that applying  $U(t)$  to a finite-temperature TFD will eventually heat up the corresponding black hole, and this process will not generate a typical equilibrium EPR of the form (5). To stabilize the energy of the black hole, we will instead gradually cool down the random circuit applying the operator [12],

$$U(\Gamma_t) = e^{-\delta\beta H_0} U(n, \delta t) e^{-\delta\beta H_0} \dots e^{-\delta\beta H_0} U(1, \delta t), \quad (11)$$

where  $n = t/\delta t$ . Formally, this defines an operator

$$U(\Gamma_t) = \mathbf{P} \left\{ \exp \left( -i \int_{\Gamma_t} ds H_c(s) \right) \right\} \quad (12)$$

by path ordering on a contour  $\Gamma_t$ , shown in Fig. 1. We follow an infinitesimal cooling-and-evolving process, with

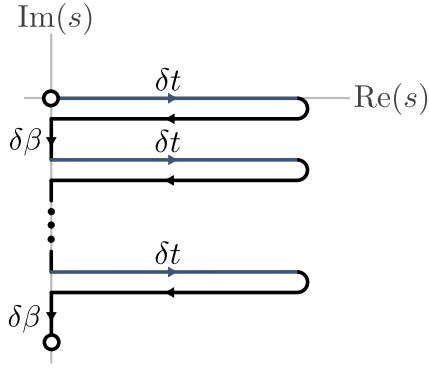


FIG. 1. Complex time contour  $\Gamma_t$  defining the operator  $U(\Gamma_t)$ . The Hamiltonian  $H_c(s)$  is taken to be  $H_0$  on the black parts of the contour and  $H_0 + H(s)$  on the blue parts. The time folds effectively implement the interaction Hamiltonian  $H(s)$ .

$\delta t = \delta\beta \rightarrow 0$ , where the relative magnitude between  $\delta t$  and  $\delta\beta$  has been absorbed into the definition of  $J$  in (9).

We will utilize the gradually cooled random circuit to define the ensemble of states  $\mathcal{E}_{\Gamma_t}$  at fixed time,

$$|\Psi_{\Gamma_t}\rangle \propto \sum_{i,j} e^{-\frac{\beta}{4}(E_i + E_j)} U(\Gamma_t)_{ij} |E_i^*\rangle_L |E_j\rangle_R. \quad (13)$$

At  $t = 0$ , the state (13) corresponds to the TFD (1). As we will later explain, when  $t \rightarrow \infty$ ,  $U(\Gamma_\infty)$  converges weakly to a random matrix  $\sqrt{\rho_0} X \sqrt{\rho_0}$  for Gaussian random  $X$  and a density matrix  $\rho_0$ . The state  $\rho_0$  is determined by the steady state of the random quantum circuit [12]. Accordingly, the ensemble  $\mathcal{E}_{\Gamma_t}$  converges to the ensemble of typical entangled states of the two black holes,

$$\lim_{t \rightarrow \infty} \mathcal{E}_{\Gamma_t} = \mathcal{E}_{\text{EPR}}^{\rho}, \quad (14)$$

for the equilibrium state

$$\rho = e^{-\frac{\beta}{4}H_0} \rho_0 e^{-\frac{\beta}{4}H_0}. \quad (15)$$

Thus, the random circuit allows us to continuously connect the TFD and a typical EPR of two black holes.

*ER caterpillars*—We will now elucidate the geometric properties of the semiclassical duals to  $|\Psi_{\Gamma_t}\rangle$ , using the gravitational path integral to prepare them. While the precise geometry of individual states will be difficult to characterize, we will argue that they include two black holes connected by an Einstein-Rosen (ER) caterpillar [13]: a spatial wormhole with a large number of matter inhomogeneities and geometric features qualitatively similar to those illustrated in Fig. 2.

Unlike the multiple-shock spacetimes of [6], the caterpillars are not spherically symmetric and are constructed by introducing matter from the Euclidean section directly into

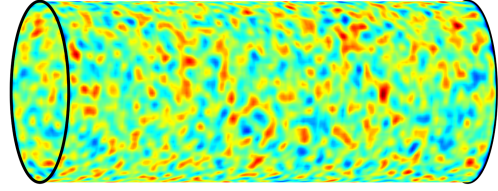


FIG. 2. The ER caterpillar is a long, bumpy wormhole supported by an inhomogeneous matter distribution, with correlation scale set by  $\ell_\Delta$  and average length set by  $\ell(t)$ .

the black hole interior, thereby preserving equilibrium and allowing the wormhole to become much longer.

In order to construct them, we start from the norm of the right-hand side of (13), which is proportional to the correlation function  $Z_\Psi(t) = Z(\beta)^{-1} \text{Tr}[e^{-(\beta/2)H_0} U(\Gamma_t) e^{-(\beta/2)H_0} U(\Gamma_t)^\dagger]$ . This quantity can be alternatively expressed as the survival amplitude [12],

$$Z_\Psi(t) = \langle \text{TFD} | U(\Gamma_t)_1 \otimes U(\Gamma_t)_1^* | \text{TFD} \rangle. \quad (16)$$

Here, we have introduced the TFD state between forward (1) and backward ( $\bar{1}$ ) contours at inverse temperature  $\beta$ . In this form,  $Z_\Psi(t)$  is computed by the CFT path integral along a complex time contour  $\mathcal{C}_t$ , shown in Fig. 3. The time orientation of the backward contour is reversed, as  $U(\Gamma_t)_1^*$  is applied on this contour. The un-normalized state (13) is prepared by the lower half of this contour.

Using the AdS-CFT dictionary,  $Z_\Psi(t)$  can be evaluated semiclassically by the gravitational path integral with sources for the bulk fields  $g_\alpha(t)$  given by the time-dependent couplings of  $U(\Gamma_t)$ . For single instances of the circuit, evaluating the gravitational path integral in a saddle point approximation is highly nontrivial. A trick that we will use is to average over the time-dependent couplings and evaluate the average norm,

$$Z(t) \equiv \mathbb{E}[Z_\Psi(t)]. \quad (17)$$

Given that the perturbations that we have chosen are semiclassical for single realizations of the couplings,  $Z(t)$

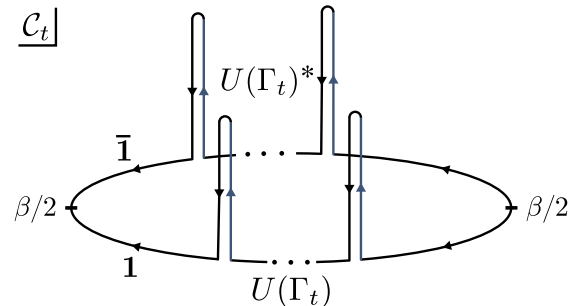


FIG. 3. Complex time contour  $\mathcal{C}_t$  composed of forward (1) and backward ( $\bar{1}$ ) contours.

will give us information about the average geometry over the ensemble of states. This is justified because the norm of the states—and thus, the semiclassical geometry that computes it—is self-averaging over  $\mathcal{E}_{\Gamma_t}$  [12]. At late times,  $\text{Std}[Z_{\Psi}(t)]/Z(t) \sim e^{-S_2(\rho)}$  where  $S_2(\rho) = -\log \text{Tr}(\rho^2)$  is the second Rényi entropy of the equilibrium state (15).

The average over the Brownian couplings can be performed exactly, using the identity

$$\mathbb{E}[U(\Gamma_t)_1 \otimes U(\Gamma_t)_1^*] = \exp(-tH_{\text{eff},1}) \quad (18)$$

for the effective time-independent Hamiltonian

$$H_{\text{eff},1} = H_0^1 + H_0^{\bar{1}} + \frac{J}{2} \sum_{\alpha=1}^K (\mathcal{O}_\alpha^1 - \mathcal{O}_\alpha^{\bar{1}*})^2. \quad (19)$$

That is, the disorder average over the Brownian couplings effectively produces local-in-time interactions between both contours in the form of a time-independent Hamiltonian  $H_{\text{eff},1}$ . The quadratic interaction in (19) arises from the Gaussianity of the random couplings.

Inserting the identity (18) into (17), we find that the average norm corresponds to the matrix element

$$Z(t) = \langle \text{TFD} | \exp(-tH_{\text{eff},1}) | \text{TFD} \rangle. \quad (20)$$

This is considerably simpler because (20) is a purely Euclidean object and so is the gravitational saddle point manifold  $M_t$  that evaluates it semiclassically,

$$Z(t) \sim Z_{\text{grav}}[M_t] \quad \text{as } N \rightarrow \infty, \quad (21)$$

where  $Z_{\text{grav}}[M_t] = e^{-I_{\text{grav}}[M_t]} Z_{1\text{-loop}}$ , with  $I_{\text{grav}}[M_t]$  the Euclidean gravitational action and  $Z_{1\text{-loop}}$  the one-loop partition function for the bulk fields on  $M_t$ .

The relevant properties of  $M_t$  follow from an assumption for the effective Hamiltonian of the random circuit:

At large  $N$ , the effective Hamiltonian  $H_{\text{eff},1}$  (19) is gapped, and it contains a unique ground state  $|\text{GS}\rangle$  with a semiclassical description as a connected spatial wormhole.

We will later motivate specific choices of the drive operators  $\mathcal{O}_\alpha$ , which lead to an effective Hamiltonian satisfying this assumption, but we keep the argument general here. Under the assumption above, the operator  $\exp(-sH_{\text{eff},1})$  is approximately proportional to the projector  $|\text{GS}\rangle\langle\text{GS}|$  for  $s \geq t_*$ . What this means for  $Z(t)$  is that there is a Euclidean time-translation symmetry of the CFT path integral  $Z(t) \approx \langle \text{TFD} | \exp(-sH_{\text{eff},1}) | \text{GS}\rangle\langle\text{GS} | \exp[-(t-s)H_{\text{eff},1}] | \text{TFD} \rangle$ , and the bulk saddle point  $M_t$  must inherit this isometry. The onset time of this symmetry is  $t_* \sim E_{\text{gap}}^{-1} \log N_*$ , where  $E_{\text{gap}}$  is the gap and  $N_*$  is the number of first excited states of  $H_{\text{eff},1}$ . As illustrated in Fig. 4, this symmetry deforms the geometry of the disk and makes it very long. Cutting the path integral open at a

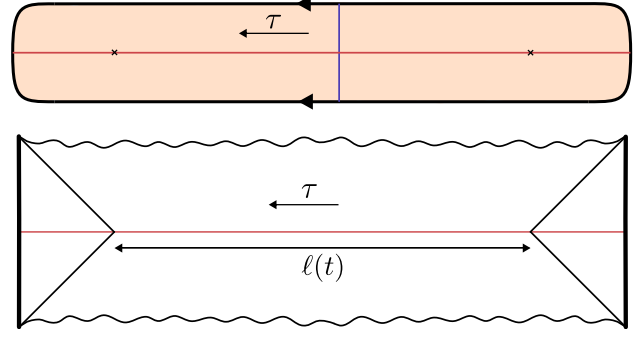


FIG. 4. On top, the Euclidean saddle point  $M_t$  of topology  $D \times \mathbf{S}^{d-1}$ , where  $D$  is a disk, has an approximate translation isometry in the  $\tau$  direction. At each constant- $\tau$  (blue) slice, the path integral prepares the semiclassical dual to  $|\text{GS}\rangle$ . On the bottom, the Lorentzian continuation across the red slice is preparing a two-sided black hole with an approximate translation symmetry in the interior. The length of the wormhole  $\ell(t)$  is proportional to the circuit time  $t - t_*$ .

constant Euclidean time prepares the un-normalized ground state, represented in the bulk as a connected spatial wormhole (blue slice). For this reason, the metric on  $M_t$  is that of a Euclidean eternal traversable wormhole [12]. The relevant slice that determines the average geometry of the ensemble of ER caterpillars is the red reflection-symmetric slice on  $M_t$ . Because of the isometry, this slice contains an approximately cylindrical spatial wormhole of length scaling linearly with  $t - t_*$ .

Since the effective Hamiltonian is gapped, the lightest particle excitations on this wormhole will correspond to the first excited states, with energy  $E_{\text{gap}}$ . These excitations serve as a bulk “clock” to measure the wormhole length  $\ell(t)$  and relate it to the asymptotic circuit time. The precise relation is [12]

$$\frac{\ell(t)}{\ell_\Delta} = E_{\text{gap}}(t - t_*), \quad (22)$$

where  $\ell_\Delta$  is the correlation scale of the lightest excitations on the wormhole. The quantities  $E_{\text{gap}}$  and  $\ell_\Delta$  are determined by quantizing the bulk fields on  $M_t$  with suitable boundary conditions and reading their energy and the decay of the bulk two-point function.

For explicit constructions of the ER caterpillars, we could specify the drive operators  $\mathcal{O}_\alpha$  to be local single-trace relevant conformal primaries, in which case  $\alpha$  labels spatial points. In this case, the interaction (19) corresponds to an eternal version of the Gao-Jafferis-Wall double-trace perturbation [16]. Coincidentally, the effective Hamiltonian (19) was analyzed in detail in the Sachdev-Ye-Kitaev (SYK) model by Maldacena and Qi, with an interaction term formed by a large number  $K \sim N$  of relevant Majorana bilinears [17]. In that case, the lesson is that the low-temperature phase of the coupled systems is gapped, with a

unique ground state that corresponds to a connected spatial wormhole, satisfying our assumptions. The ground state is similar to the TFD at some coupling-dependent temperature  $\beta(J)$  (in units of the SYK couplings). Moreover, it has  $O(1)$  overlap with it for large- $q$  SYK and for  $1/N \ll J \ll 1$  in the regular  $q = 4$  SYK (in units where the SYK couplings  $\mathcal{J} = 1$ ). In [12], we provided a totally explicit construction of the ER caterpillars of the SYK model based on this construction. A similar picture is expected to hold in higher dimensions, provided that the Hamiltonian  $H_0$  is sufficiently chaotic [18].

*Wormhole length = randomness*—We will now derive a quantitative relation between the average length of the wormhole  $\ell(t)$  and a microscopic feature of the ensemble of states  $\mathcal{E}_{\Gamma}$ . Let us explain this notion in the qubit system first. The idea is that, at finite but large circuit time  $t$ , the states (8) are not fully typical EPRs because  $\mathcal{E}_t$  and  $\mathcal{E}_{\text{EPR}}$  are statistically distinguishable. The approach to typical is quantified by the notion of a *quantum state  $k$  design*: an ensemble of states that reproduces the first  $k$  moments of  $\mathcal{E}_{\text{EPR}}$ . More precisely, the  $k$ th moment superstate of a general ensemble of states  $\mathcal{E}_{\Psi}$  is defined as

$$\rho_k(\mathcal{E}_{\Psi}) = \mathbb{E}_{\mathcal{E}_{\Psi}}[(|\Psi\rangle\langle\Psi|)^{\otimes k}]. \quad (23)$$

The state  $\rho_k(\mathcal{E}_{\Psi})$  is defined on  $2k$  replicas of the original Hilbert space. It is generally a mixed state, given the correlations between replicas that arise from the average over  $\mathcal{E}_{\Psi}$ . Given this, a quantum state  $k$  design is defined by the condition  $\rho_k(\mathcal{E}_{\Psi}) = \rho_k(\mathcal{E}_{\text{EPR}})$ . For  $\mathcal{E}_{\Psi} = \mathcal{E}_t$ , this is equivalent to the condition that the random circuit  $U(t)$  forms a *unitary  $k$  design*: an ensemble of unitaries that reproduces the first  $k$  moments of the Haar ensemble.

Note that these notions are physically very strong, since a  $k$  design is indistinguishable from  $\mathcal{E}_{\text{EPR}}$  by any statistical  $k$ -copy measure. In practice, only approximate notions of  $k$  designs make sense at finite time. For these, one typically bounds some quantum distance between moment superstates.

We define a ( $\rho$  equilibrium) quantum state  $k$  design by the condition  $\rho_k(\mathcal{E}_{\Psi}) = \rho_k(\mathcal{E}_{\text{EPR}}^{\rho})$ . For  $\mathcal{E}_{\Psi} = \mathcal{E}_{\Gamma}$ , we will compute the distance to  $k$  design

$$\Delta_k^{\rho}(t) \equiv \frac{\|\rho_k(\mathcal{E}_{\Gamma}) - \rho_k(\mathcal{E}_{\text{EPR}}^{\rho})\|_2}{\|\rho_1(\mathcal{E}_{\text{EPR}}^{\rho})\|_2^k}, \quad (24)$$

and the corresponding time to  $kt_k = \min\{t: \Delta_k^{\rho}(t) \leq \varepsilon\}$ . We are now ready to state the main result of this Letter.

*Length-randomness correspondence*: The ensemble of ER caterpillars of average length  $\ell$  and matter correlation scale  $\ell_{\Delta}$  forms an  $\varepsilon$ -approximate quantum state  $k$  design of the black holes for

$$k \sim \frac{\ell - \ell_{\varepsilon}}{\ell_{\Delta}}, \quad (25)$$

where  $\ell_{\varepsilon} = \ell_{\Delta} \log \varepsilon^{-1}$  and  $k \ll O(e^{S_2(\rho)})$ .

The correspondence follows from evaluating the distance  $\Delta_k^{\rho}(t)$  using the gravitational path integral. The calculation boils down to gravitationally computing the  $k$ th moment *two-point function* of the ensemble,

$$G_k(t, t') \equiv \text{Tr}(\rho_k(\mathcal{E}_{\Gamma_t})\rho_k(\mathcal{E}_{\Gamma_{t'}})) = \mathbb{E}[|\langle\Phi_{\Gamma_t}|\Psi_{\Gamma_{t'}}\rangle|^{2k}]. \quad (26)$$

This quantity corresponds to the  $2k$ th moment of the overlap between two independent draws of the ensemble of states. In the CFT, it is evaluated as a disordered path integral on  $2k$  replicas of a closed time contour, with two independent instances of the random circuit on each contour, generalizing Fig. 3. The moment two-point function is relevant because it controls the distance to design

$$\Delta_k^{\rho}(t)^2 = \frac{F_k(t) - 2G_k(t, \infty) + F_k(\infty)}{F_1(\infty)^k}, \quad (27)$$

where we have defined the purity of the moment state  $F_k(t) = G_k(t, t)$ , also dubbed the  $k$ th frame potential.

Now, the average over Brownian couplings in (26) can be performed exactly. This produces time-independent interactions between the  $2k$  contours, and the moment two-point function corresponds to a thermal-looking partition function,

$$G_k(t, t') \propto \text{Tr}\left(e^{-\frac{\ell}{2}H_{0,k}} e^{-tH_{\text{eff},k}} e^{-\frac{\ell}{2}H_{0,k}} e^{-t'H_{\text{eff},k}}\right), \quad (28)$$

for the time-independent  $2k$  replica Hamiltonian

$$H_{\text{eff},k} = H_{0,k} + \frac{J}{2} \sum_{\alpha=1}^K \left( \sum_{r=1}^k \mathcal{O}_{\alpha}^r - \mathcal{O}_{\alpha}^{r*} \right)^2 \quad (29)$$

and the bare  $2k$  replica Hamiltonian

$$H_{0,k} \equiv \sum_{r=1}^k (H_0^r + H_0^{r*}). \quad (30)$$

For large  $t, t'$ , the dominant contribution to this partition function comes from the ground space and first excited states of  $H_{\text{eff},k}$ . The assumption, consistent with semiclassical and infinite-temperature (or large enough coupling  $J$ ) considerations, is that the ground states break replica symmetry of  $H_{\text{eff},k}$  and factorize into  $k$  copies of the ground state of  $H_{\text{eff},1}$  connecting two replicas. The two-boundary Euclidean wormhole that reproduces the ground state contribution is shown in Fig. 5. The first excited states correspond to single-particle excitations on each replica of this wormhole, with energy given by the two-replica gap

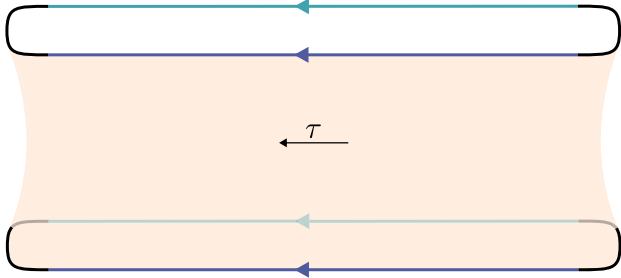


FIG. 5. Euclidean two-boundary wormhole providing the late-time “plateau” value of  $G_1(t, t')$ . The saddle-point topology is  $\mathbf{S}^1 \times \mathbf{R} \times \mathbf{S}^{d-1}$ . The wormhole geometry includes two long regions with a Euclidean time  $\tau$  translation isometry geometrially equivalent to the one for the disk in Fig. 4.

$E_{\text{gap}}$ . For the caterpillars of the SYK model, this wormhole is a stabilized double trumpet [12].

This produces a late-time behavior

$$G_k(t, t') \approx k! F_1(\infty)^k \left(1 + kN_* e^{-E_{\text{gap}}(t+t')}\right), \quad (31)$$

where  $k!$  comes from the number of ground states, and  $kN_*$  is the number of first excited states, corresponding to the possible single-particle excitations on each of the  $k$  two-boundary wormholes. The late-time value is expressed in terms of the  $k = 1$  frame potential  $F_1(\infty) = e^{-S_2(\rho)}$  for  $\rho$  defined in (15). Using (27), this leads to a late-time distance to  $k$  design of

$$\Delta_k^\rho(t)^2 \approx N_* k! k e^{-2tE_{\text{gap}}}, \quad (32)$$

which for large  $k$  produces a time to  $k$  design of

$$t_k \approx \frac{1}{2} E_{\text{gap}}^{-1} (k \log k - k + \log k + \log N_* + 2 \log \epsilon^{-1}). \quad (33)$$

Up to subleading logarithmic factors, the growth is linear in  $k$ , with the slope  $(2E_{\text{gap}})^{-1}$ . The semiclassical analysis requires  $k \ll e^{S_2(\rho)}$  to be able to neglect other nonperturbative effects. The length-randomness relation (25) follows directly from (33) together with (22).

*Discussion*—The ensembles of caterpillars constructed in this Letter provide a window into the generic structure of the black hole Hilbert space in any theory of gravity with low-energy matter. The construction and main result of this Letter support a vastly more general form of ER = EPR and seem to be in some tension with arguments against semiclassicality of typical interiors [7, 19–22]. We believe that resolving this tension could be relevant for the typical-state firewall paradox. We think caterpillars can be useful for this purpose because, from the randomness perspective, we see no reason why they should have their properties dramatically altered before exponential times, even though we cannot show this with semiclassical methods. In future

work, this extrapolation of the semiclassical analysis could be addressed in explicit UV complete models at finite  $N$  using numerical or even experimental methods. Another open question is whether caterpillars offer any insight into weaker notions of firewalls in typical states, such as highly boosted matter shocks near the horizon [6, 23–27]. Naively, since the matter in the caterpillar begins at rest, the states do not seem to exhibit this type of firewall. Past-evolved caterpillars, evolved over a few scrambling times, do, however, because the matter becomes highly boosted in the frame of the infaller. At the level of ensembles of microscopic quantum states, both the caterpillars and the past-evolved caterpillars form approximate quantum state  $k$  designs under the 2-norm definition used in this Letter. While, at the technical level, this points to the need for a stricter 1-norm definition of an approximate  $k$  design, it conceptually suggests that distinguishing states with firewalls from those without is extremely difficult. In fact, if we could extrapolate our result to exponentially large values of  $k$ , the indistinguishability would be so strong that the very question of whether a state has a firewall might become meaningless. It would be interesting to know whether this limitation is what the overlap analysis of [24–26] is indicating. It would also be worth exploring whether the state-dependent interior reconstruction of [28] plays any role here, given that our conclusion is largely state independent.

Finally, in the spirit of the “quantum gravity in the lab” program [29], it would be interesting to construct and study these states in the lab as a way of directly probing the black hole interior. Physically instantiating these states is not trivial since they involve nonunitary elements due to the gradual cooling, but such evolutions can be implemented inefficiently using postselection or a variety of other methods, e.g., the recently discussed double-bracket flow approach [30]. Alternatively, by constructing an isometry that embeds the  $e^S$  low-energy states of the black hole into the microscopic Hilbert space (e.g., as was effectively done in [31] for a noninteracting fermion system), one might be able to combine this isometry with a conventional random circuit in the low-energy Hilbert space to produce random low-energy states embedded in the microscopic Hilbert space.

*Acknowledgments*—We thank Stefano Antonini, José Barbón, Adam Bouland, Horacio Casini, Matt Headrick, Don Marolf, Douglas Stanford, Lenny Susskind, and Mark Van Raamsdonk for discussions. M. S. and B. S. acknowledge support from the U.S. Department of Energy through DE-SC0009986, QuantISED DE-SC0020360, and GeoFlow DE-SC0019380. The work of J. M. is supported by CONICET, Argentina. This preprint is assigned the code BR-TH-6724.

*Data availability*—No data were created or analyzed in this study.

- [1] J. Maldacena and L. Susskind, Cool horizons for entangled black holes, *Fortschr. Phys.* **61**, 781 (2013).
- [2] S. Ryu and T. Takayanagi, Holographic derivation of entanglement entropy from AdS/CFT, *Phys. Rev. Lett.* **96**, 181602 (2006).
- [3] M. Van Raamsdonk, Building up spacetime with quantum entanglement, *Gen. Relativ. Gravit.* **42**, 2323 (2010).
- [4] J. M. Maldacena, Eternal black holes in anti-de Sitter, *J. High Energy Phys.* **04** (2003) 021.
- [5] S. H. Shenker and D. Stanford, Black holes and the butterfly effect, *J. High Energy Phys.* **03** (2014) 067.
- [6] S. H. Shenker and D. Stanford, Multiple shocks, *J. High Energy Phys.* **12** (2014) 046.
- [7] D. Marolf and J. Polchinski, Gauge/gravity duality and the black hole interior, *Phys. Rev. Lett.* **111**, 171301 (2013).
- [8] T. Hartman and J. Maldacena, Time evolution of entanglement entropy from black hole interiors, *J. High Energy Phys.* **05** (2013) 014.
- [9] L. Susskind, Computational complexity and black hole horizons, *Fortschr. Phys.* **64**, 24 (2016); **64**, 44(E) (2016).
- [10] S.-K. Jian, G. Bentsen, and B. Swingle, Linear growth of circuit complexity from Brownian dynamics, *J. High Energy Phys.* **08** (2023) 190.
- [11] S. Guo, M. Sasieta, and B. Swingle, Complexity is not enough for randomness, *SciPost Phys.* **17**, 151 (2024).
- [12] J. M. Magan, M. Sasieta, and B. Swingle, Random circuits in the black hole interior, *SciPost Phys.* **19**, 007 (2025).
- [13] To our knowledge, related sorts of caterpillars were first discussed by H. Lin and L. Susskind in the different context of exponentially long wormholes arising from a fixed Hamiltonian evolution [13,14].
- [14] L. Susskind, Inside black holes at exp-time, in *Talk at Complexity from Quantum Information to Black Holes Workshop* (2020), <https://www.youtube.com/watch?v=OVrMphTSkrE>.
- [15] L. Susskind, Black holes at exp-time, [arXiv:2006.01280](https://arxiv.org/abs/2006.01280).
- [16] P. Gao, D. L. Jafferis, and A. C. Wall, Traversable wormholes via a double trace deformation, *J. High Energy Phys.* **12** (2017) 151.
- [17] J. Maldacena and X.-L. Qi, Eternal traversable wormhole, [arXiv:1804.00491](https://arxiv.org/abs/1804.00491).
- [18] W. Cottrell, B. Freivogel, D. M. Hofman, and S. F. Lokhande, How to build the thermofield double state, *J. High Energy Phys.* **02** (2019) 058.
- [19] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully, Black holes: Complementarity or firewalls?, *J. High Energy Phys.* **02** (2013) 062.
- [20] A. Almheiri, D. Marolf, J. Polchinski, D. Stanford, and J. Sully, An apologia for firewalls, *J. High Energy Phys.* **09** (2013) 018.
- [21] R. Bousso, Complementarity is not enough, *Phys. Rev. D* **87**, 124023 (2013).
- [22] M. Van Raamsdonk, Evaporating firewalls, *J. High Energy Phys.* **11** (2014) 038.
- [23] L. Susskind, The typical-state paradox: Diagnosing horizons with complexity, *Fortschr. Phys.* **64**, 84 (2016).
- [24] D. Stanford and Z. Yang, Firewalls from wormholes, [arXiv:2208.01625](https://arxiv.org/abs/2208.01625).
- [25] A. Blommaert, C.-H. Chen, and Y. Nomura, Firewalls at exponentially late times, *J. High Energy Phys.* **10** (2024) 131.
- [26] L. V. Iliesiu, A. Levine, H. W. Lin, H. Maxfield, and M. Mezei, On the non-perturbative bulk Hilbert space of JT gravity, *J. High Energy Phys.* **10** (2024) 220.
- [27] R. Bousso, Firewalls from general covariance, *Phys. Rev. Lett.* **135**, 021501 (2025).
- [28] K. Papadodimas and S. Raju, An infalling observer in AdS/CFT, *J. High Energy Phys.* **10** (2013) 212.
- [29] A. R. Brown, H. Gharibyan, S. Leichenauer, H. W. Lin, S. Nezami, G. Salton, L. Susskind, B. Swingle, and M. Walter, Quantum gravity in the lab. I. Teleportation by size and traversable wormholes, *PRX Quantum* **4**, 010320 (2023).
- [30] M. Gluza, J. Son, B. H. Tiang, Y. Suzuki, Z. Holmes, and N. H. Y. Ng, Double-bracket quantum algorithms for quantum imaginary-time evolution, [arXiv:2412.04554](https://arxiv.org/abs/2412.04554).
- [31] T. J. Sewell, C. D. White, and B. Swingle, Thermal multi-scale entanglement renormalization ansatz for variational Gibbs state preparation, [arXiv:2210.16419](https://arxiv.org/abs/2210.16419).
- [32] M. L. Mehta, *Random Matrices* (Elsevier, New York, 2004).
- [33] J. S. Cotler, G. Gur-Ari, M. Hanada, J. Polchinski, P. Saad, S. H. Shenker, D. Stanford, A. Streicher, and M. Tezuka, Black holes and random matrices, *J. High Energy Phys.* **05** (2017) 118; **09** (2018) 002(E).
- [34] J. Cotler, N. Hunter-Jones, J. Liu, and B. Yoshida, Chaos, complexity, and random matrices, *J. High Energy Phys.* **11** (2017) 048.
- [35] P. Saad, S. H. Shenker, and D. Stanford, A semiclassical ramp in SYK and in gravity, [arXiv:1806.06840](https://arxiv.org/abs/1806.06840).

## End Matter

*Appendix A: Properties of typical equilibrium EPRs*—In this appendix, we study the ensemble of random purifications  $\mathcal{E}_{\text{EPR}}^\rho$  of a generic equilibrium density matrix  $\rho$ . The states have the form

$$|\Psi_{\text{EPR}}^\rho\rangle = \frac{1}{\sqrt{Z_X}} \sum_{i,j} \sqrt{\rho_i \rho_j} X_{ij} |E_i^*\rangle_L |E_j\rangle_R, \quad (\text{A1})$$

where  $|E_i\rangle$  is an eigenstate of the density matrix  $\rho$  with eigenvalue  $\rho_i$  and  $X_{ij}$  are Gaussian random coefficients

with zero mean and unit variance. The quantity  $Z_X = \text{Tr}(\rho X \rho X^\dagger)$  is a normalization factor. On average over the ensemble,

$$\mathbb{E}[Z_X] = \int dX \text{Tr}(\rho X \rho X^\dagger) = 1. \quad (\text{A2})$$

The variance, on the other hand, is

$$\text{Var}[Z_X] = \int dX \text{Tr}(\rho X \rho X^\dagger)^2 - 1 = e^{-2S_2(\rho)}, \quad (\text{A3})$$

where  $S_2(\rho) = -\log \text{Tr}(\rho^2)$  is the second Rényi entropy. Since in our cases of interest,  $S_2(\rho) = O(N^2)$ , we can forget about the normalization and consider “annealed averages” for the states normalized only on average. The average reduced density matrices are given by the equilibrium state

$$\mathbb{E}[\rho_{L,R}] = \rho. \quad (\text{A4})$$

Moreover, for individual instances, the 2-norm distance to  $\rho$  is exponentially suppressed,

$$\text{Tr}(\rho_{L,R}^2) \approx \int dX \text{Tr}(\rho X \rho X^\dagger \rho X \rho X^\dagger)^2 = 2e^{-2S_2(\rho)}.$$

From here, it directly follows that the 2-norm distance is exponentially small,

$$\mathbb{E}[\|\rho_{L,R} - \rho\|_2] \leq \mathbb{E}[\|\rho_{L,R} - \rho\|_2^2]^{1/2} \approx e^{-S_2(\rho)}. \quad (\text{A5})$$

Note, however, that the average 1-norm distance to  $\rho$  can be  $O(1)$ , so individual instances are distinguishable from the equilibrium state  $\rho$ .

*Appendix B: Exploration via time-independent vs time-dependent Hamiltonians*—Let us focus on the  $N$  qubit system. Consider time-evolving the TFD  $|I_t\rangle = e^{-itH_R}|I\rangle$ , or explicitly,

$$|I_t\rangle = \frac{1}{\sqrt{D}} \sum_{i=1}^D e^{-itE_i} |E_i^*\rangle_L |E_i\rangle_R. \quad (\text{B1})$$

In order to compare this to the exploration via the random quantum circuit, we will diagonalize  $U(t)$  and write the state (8) as

$$|\Psi_t\rangle = \frac{1}{\sqrt{D}} \sum_{i=1}^D e^{-i\varphi_i(t)} |\varphi_i(t)^*\rangle_L |\varphi_i(t)\rangle_R, \quad (\text{B2})$$

where  $|\varphi(t)_i\rangle$  is an eigenvector of  $U(t)$  with eigenvalue  $\exp[-i\varphi_i(t)]$  and  $-\pi \leq \varphi_i(t) < \pi$ . Finally, we will do the same thing for the typical EPR (3) and express it as

$$|\Psi_{\text{EPR}}\rangle = \frac{1}{\sqrt{D}} \sum_{i=1}^D e^{-i\theta_i} |\psi_i^*\rangle_L |\psi_i\rangle_R, \quad (\text{B3})$$

where  $|\psi_i\rangle$  is an eigenvector of  $U_{\text{Haar}}$  with eigenvalue  $\exp(-i\theta_i)$  and  $-\pi \leq \theta_i < \pi$ . In this form, the structural similarity between (B1)–(B3) is clear.

The eigenvectors  $|\psi_i\rangle$  in (B3) are i.i.d. random vectors. The eigenphases  $\theta_i$  are distributed according to the circular unitary ensemble (CUE) [32],

$$p_{\text{CUE}}(\theta_1, \dots, \theta_D) = \frac{1}{D!(2\pi)^D} \prod_{i < j} |e^{-i\theta_i} - e^{-i\theta_j}|^2. \quad (\text{B4})$$

In the CUE, the eigenphases are uniformly distributed on average with a strong repulsion given by

$$\mathbb{E}_{\text{CUE}}[\delta\rho(\theta)\delta\rho(\theta')] \supset -\frac{\sin\left(D\frac{\theta-\theta'}{2}\right)^2}{\left[2\pi\sin\left(\frac{\theta-\theta'}{2}\right)\right]^2}, \quad (\text{B5})$$

where  $\rho(\theta) = \sum_{i=1}^D \delta(\theta - \theta_i)$  is the eigenphase density, and  $\delta\rho(\theta) = \rho(\theta) - (2\pi)^{-1}$  is the fluctuation from the mean.

The question is whether the exploration via time-independent and time-dependent Hamiltonians is able to reproduce these statistical distributions of eigenphases and eigenvectors at sufficiently late times.

For time-independent Hamiltonians, there are two natural ensembles of states to be considered. One is to fix the Hamiltonian and consider the time evolution trajectory  $\mathcal{E}_{\text{erg}} = \{|I_t\rangle : t \in \mathbf{R}\}$  with the uniform long-time measure  $\lim_{T \rightarrow \infty} (1/T) \int_{-T/2}^{T/2} dt$ . Any draw of  $\mathcal{E}_{\text{erg}}$  lives on a  $D$ -dimensional submanifold of  $\mathcal{H}_{\text{EPR}}$ , of torus topology, characterized by the phases in the energy eigenbasis (B1) (we are neglecting that the global phase is unphysical). Assuming that the spectrum of the Hamiltonian lacks rational relations, the ensemble  $\mathcal{E}_{\text{erg}}$  covers this set ergodically, and the long-time distribution of phases is the uniform measure on this torus,

$$p_{\text{erg}}(\theta_1, \dots, \theta_D) = \frac{1}{(2\pi)^D}, \quad (\text{B6})$$

are Poisson distributed.

The second possibility is to define an ensemble at fixed time,  $\mathcal{E}_t^H$ , by introducing disorder in the time-independent Hamiltonian  $H$ . Generally, the draws of the ensemble  $\mathcal{E}_t^H$  will explore the full  $\mathcal{H}_{\text{EPR}}$  but will never become typical EPRs in  $\mathcal{E}_{\text{EPR}}$ . In order to see this, consider a general spectral distribution for the eigenvalues of the Hamiltonian  $E_i$ , which may universally include eigenvalue repulsion. Essentially, without an extreme fine-tuning, at late enough times, the distribution of phases  $\theta_i = tE_i \pmod{2\pi}$  of the entangled state (B1) follows (B6). The reason is that, at timescales comparable with the mean level spacing of the Hamiltonian, the eigenphases  $\exp(-itE_i)$  will have wound around the unit circle so many times that all correlations between the  $E_i$ 's will be erased. Thus, no matter which of the two ensembles we consider, the exploration via time-independent Hamiltonians does not generate typical EPRs.

On the other hand, the ensemble  $\mathcal{E}_t$  generated with the random quantum circuit will have a statistical distribution of phases  $\theta_i = \varphi_i(t)$ , distributed according to  $p_t(\theta_1, \dots, \theta_D)$  which, at infinite times, converges weakly to the CUE [11],

$$p_t(\theta_1, \dots, \theta_D) \xrightarrow{t \rightarrow \infty} p_{\text{CUE}}(\theta_1, \dots, \theta_D). \quad (\text{B7})$$

Likewise,  $\mathcal{E}_t$  will define a distribution of eigenvectors that converges weakly to the random state distribution.

All of these differences are quantified by the frame potential  $F_k(\mathcal{E}_\Psi) = \text{Tr}[\rho_k(\mathcal{E}_\Psi)^2]$ . For the ensemble of typical EPRs [10–12],

$$F_k(\mathcal{E}_{\text{EPR}}) = \frac{k!}{D^{2k}} \quad \text{for } k \leq D. \quad (\text{B8})$$

The  $k$ th frame potential is controlled by the  $2k$ th moment of the CUE,  $\mathbb{E}_{\text{CUE}}[\rho(\theta_1) \dots \rho(\theta_{2k})]$ .

For the ensemble  $\mathcal{E}_{\text{erg}}$ , the frame potential is a purely spectral quantity. It is controlled by the long-time average of the  $k$ th spectral form factor [33], i.e., its plateau value, which for  $k \leq D$  is

$$F_k(\mathcal{E}_{\text{erg}}) = \frac{1}{D^{2k}} \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} dt |\text{Tr}(e^{-itH})|^{2k} = \frac{k!}{D^k}, \quad (\text{B9})$$

where in the last equality we have used the spectral ergodicity of  $H$  to evaluate the long-time average. Already for the second moment ( $k = 1$ ), the frame potential differs from (B8) by factors of the dimension, which signals that the distribution of phases in  $\mathcal{E}_{\text{erg}}$  fails to incorporate the eigenphase repulsion (B5). A similar conclusion follows for  $\mathcal{E}_t^H$  at late enough times [34].

For  $\mathcal{E}_t$ , the frame potential  $F_k(t) \equiv F_k(\mathcal{E}_t)$  is also a purely spectral quantity. In this case, it corresponds to the  $2k$ th spectral form factor of the random quantum circuit, which at late times converges to the CUE value [10–12],

$$F_k(t) = \frac{1}{D^{2k}} \mathbb{E} |\text{Tr} U(t)|^{2k} \xrightarrow{t \rightarrow \infty} F_k(\mathcal{E}_{\text{EPR}}). \quad (\text{B10})$$

The quantity  $F_1(t)$  was, in fact, analyzed in [35] as a generalization of the spectral form factor to a periodically driven Brownian SYK model.