Information Scrambling over Bipartitions:
Equilibration, Entropy Production, and Typicality

Georgios Styliaris,[1, 2, 3] Namit Anand,[3] and Paolo Zanardi[3]

[1]*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, 85748 Garching, Germany*
[2]*Munich Center for Quantum Science and Technology, Schellingstraße 4, 80799 München, Germany*
[3]*Department of Physics and Astronomy, and Center for Quantum Information Science and Technology,
University of Southern California, Los Angeles, California 90089, USA*
(Dated: January 21, 2021)

In recent years, the out-of-time-order correlator (OTOC) has emerged as a diagnostic tool for information scrambling in quantum many-body systems. Here, we present exact analytical results for the OTOC for a typical pair of random local operators supported over two regions of a bipartition. Quite remarkably, we show that this "bipartite OTOC" is equal to the operator entanglement of the evolution and we determine its interplay with entangling power. Furthermore, we compute long-time averages of the OTOC and reveal their connection with eigenstate entanglement. For Hamiltonian systems, we uncover a hierarchy of constraints over the structure of the spectrum and elucidate how this affects the equilibration value of the OTOC. Finally, we provide operational significance to this bipartite OTOC by unraveling intimate connections with average entropy production and scrambling of information at the level of quantum channels.

Introduction.— A characteristic feature of certain quantum many-body systems is their ability to quickly spread "localized" information over subsystems, thereby making it inaccessible to local observables. Although unitary evolution retains all information, this local inaccessibility manifests itself as equilibration in closed systems, and has been termed "information scrambling" [1–5].

For Hamiltonian quantum dynamics, scrambling can be probed by examining the overlap of a time-evolved local operator $V(t) := U_t^\dagger V U_t$ with a second static operator W. This overlap is commonly quantified via the strength of the commutator[1]

$$C_{V,W}(t) := \frac{1}{2} \operatorname{Tr} \left([V(t), W]^\dagger [V(t), W] \rho_\beta \right) \quad (1)$$

where ρ_β denotes the thermal state at inverse temperature β. From the perspective of information spreading, $C_{V,W}(t)$ is a natural quantity to consider since it constitutes a state-dependent variant of the Lieb-Robinson scheme; the latter enforces a fundamental restriction on the speed of correlations spreading in nonrelativistic quantum systems [6–9]. In Eq. (1), it is convenient to consider pairs of operators V, W which at $t = 0$ act nontrivially on different subsystems, thus, commute; we follow this convention here.

The commutator $C_{V,W}(t)$ is intimately linked to the out-of-time-order correlator (OTOC) [10, 11] which is a four-point function with an unconventional time-ordering

$$F_{V,W}(t) := \operatorname{Tr} \left(V^\dagger(t) W^\dagger V(t) W \rho_\beta \right). \quad (2)$$

The connection between the two arises when V, W are unitary; Eq. (1) then immediately reduces to $C_{V,W}(t) =$

$1 - \operatorname{Re}[F_{V,W}(t)]$. In this Letter we focus on the infinite-temperature, $\beta = 0$ case.

Through the years, several key signatures of quantum chaos [12–15] have been introduced. The initial exponential growth of the OTOC was proposed as a diagnostic of quantum chaos [16–23]. However, a careful analysis has revealed that information scrambling does not always necessitate chaos [24–29].

Per se, the OTOC's ability to probe dynamical features clearly depends on the choice of operators V, W. However, it is desirable to be able to capture these features as independently as possible from the specific choice of operators. This insensitivity can be achieved by averaging over a set of operators, a strategy also considered in Refs. [22, 30–35]. It is crucial to remark that, for the averaged OTOC to faithfully capture information spreading, the averaging process must *preserve the initial locality of the system*, i.e., which subsystems V, W initially act upon — an observation that was quintessential in revealing the correct behavior of the OTOC and its connection with Loschmidt echo [35].

Given a bipartition of a finite-dimensional Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \cong \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$, we will henceforth focus on averaging $C_{V_A, W_B}(t)$ over the (independent) unitary operators V_A and W_B, whose support is over subsystems A and B, respectively. The resulting quantity

$$G(t) := 1 - \frac{1}{d} \operatorname{Re} \int dV \, dW \, \operatorname{Tr} \left(V_A^\dagger(t) W_B^\dagger V_A(t) W_B \right), \quad (3)$$

depends only on the dynamics and the Hilbert space cut, where we denote $V_A = V \otimes I_B$, $W_B = I_A \otimes W$ and the averaging is performed according to the Haar measure [36]. We will refer to $G(t)$ for brevity as the *bipartite OTOC*, and analyzing its properties will be the focus of the present Letter.

It was recently shown in Ref. [35], where $G(t)$ was first introduced, under the assumptions of (i) weak coupling between A and B, and (ii) Markovianity, that

---

[1] In fact, $C_{V,W}(t) = \frac{1}{2} \| [V(t), W] \|^2$ for the norm associated with the inner product $\langle X, Y \rangle_\beta = \operatorname{Tr}(X^\dagger Y \rho_\beta)$, $\beta < \infty$.

$G(t)$ exhibits a close connection with the Loschmidt echo [37, 38]; the latter has been widely employed to characterize chaos [39, 40]. Here, we first show, without any of the previous assumptions, that $G(t)$ is, in fact, amenable to exact analytical treatment, and we uncover its direct relation with entropy production, information spreading, and entanglement. We also rigorously prove that the average case is also the typical one, hence justifying the averaging process. Our main results are stated in the theorems that follow. All proofs of the claims appearing in the text can be found in Appendix A.

The bipartite OTOC.— We begin by bringing $G(t)$ in a more explicit form which will be the starting point for a sequence of results. This can be achieved by working on the doubled space $\mathcal{H} \otimes \mathcal{H}'$, where $\mathcal{H}' = \mathcal{H}_{A'} \otimes \mathcal{H}_{B'}$ is a replica of the original Hilbert space.

Theorem 1. *Let $S_{AA'}$ be the operator over $\mathcal{H} \otimes \mathcal{H}'$ that swaps A with its replica A' and $d = \dim(\mathcal{H})$. Then*

$$G(t) = 1 - \frac{1}{d^2} \operatorname{Tr}\left(S_{AA'} U_t^{\otimes 2} S_{AA'} U_t^{\dagger \otimes 2} \right). \quad (4)$$

The analogous expression for BB' also holds.

The above formula immediately exposes a connection between the bipartite OTOC and the *operator entanglement* of the evolution $E_{\mathrm{op}}(U_t)$, as defined in Ref. [41] (see also Appendix A for the relevant definitions). The two quantities, remarkably, coincide exactly. This observation also allows one to express the *entangling power* [42] $e_{\mathrm{P}}(U_t)$ as a function of the bipartite OTOC for the symmetric case $d_A = d_B$. The former quantifies the average entanglement produced by the evolution and has been established as an indicator of global chaos in few-body systems [43–46].

Theorem 2. *Let G_U denote the bipartite OTOC for the evolution U. Then, (i) $E_{\mathrm{op}}(U_t) = G_{U_t}$, and (ii) for a symmetric bipartition $d_A = d_B$,*

$$e_{\mathrm{P}}(U_t) = \frac{d}{(\sqrt{d}+1)^2} \left(G_{U_t} + G_{U_t S_{AB}} - G_{S_{AB}} \right). \quad (5)$$

For the finite-temperature case, Eq. (4) admits a straightforward generalization which we report in Appendix A. However, a direct connection with operator entanglement and entangling power may not be so simple.

How informative is the average $G(t)$?— Usually, one is interested in behavior of the OTOC for a typical choice of random unitary operators. Because of measure concentration [47], we prove that the two essentially coincide; i.e., the probability that a random instance deviates significantly from the mean is exponentially suppressed as the dimension of either of the subsystems A and B grows large.

Proposition 3. *Let $P(\epsilon)$ be the probability that a random instance of $C_{V_A, W_B}(t)$ deviates from its Haar aver-*

age $G(t)$ more than ϵ. Then,

$$P(\epsilon) \leq 2 \exp\left(-\frac{\epsilon^2 d_{\max}}{64} \right), \quad (6)$$

where $d_{\max} = \max\{d_A, d_B\}$.

In the definition of the bipartite OTOC and to obtain the replica formula Eq. (4), we have so far considered averaging over the uniform (Haar) ensemble which continuously extends over the whole unitary group. Although natural from a mathematical viewpoint, this choice can turn out to be rather complicated on physical and numerical grounds [48]. Nonetheless, we show in Appendix B that Haar averaging can be replaced by any unitary ensemble that forms a 1-design [49–52] without altering $G(t)$. Such ensembles mimic the Haar randomness only up to the first moment, which is the depth of randomness that the OTOC can probe [22]. The latter assumption is thus much weaker than Haar randomness. For instance, consider the case of a spin-1/2 many-body system split into two parts, A and B. Instead of averaging over Haar random unitaries V_A and W_B, that typically do not factor, the 1-design (equivalent) picture prescribes to instead consider only fully factorized unitaries with support over A and B, e.g., products of local Pauli matrices.

Time-averaging the bipartite OTOC.— In finite-dimensional quantum systems, nontrivial quantum expectation values or quantities such as $C_{V,W}(t)$ do not converge to a limit for $t \to \infty$. Instead, after a long time they typically oscillate around an equilibrium value [53–58] which can be extracted by time-averaging $\overline{X(t)} := \lim_{T \to \infty} \frac{1}{T} \int_0^T dt\, X(t)$. We now turn to examine this long-time behavior $\overline{G(t)}$ of the bipartite OTOC as a function of the Hamiltonian and the Hilbert space cut.

Let us begin with the case of a chaotic dynamics, which entails level repulsion statistics [15] and an "incommensurable" relation among the energy levels. As such, chaotic Hamiltonians satisfy (either exactly or to very good approximation) the no-resonance condition (NRC): The energy levels and energy gaps feature nondegeneracy. This has important implications for the long-time behavior of their bipartite OTOC, as we will see soon.

Let us spectrally decompose $H = \sum_k E_k |\phi_k\rangle\langle\phi_k|$ and use $\rho_k^{(\chi)} := \operatorname{Tr}_{\overline{\chi}}(|\phi_k\rangle\langle\phi_k|)$ to denote the reduced density operator over $\chi = A, B$ corresponding to the kth Hamiltonian eigenstate ($\overline{\chi}$ corresponds to the complement). Below, $\langle X, Y \rangle := \operatorname{Tr}(X^\dagger Y)$ denotes the Hilbert-Schmidt inner product [59], which gives rise to the operator 2-norm $\|X\|_2 := \sqrt{\langle X, X \rangle}$.

Proposition 4. *Consider a Hamiltonian satisfying the NRC. Then*

$$\overline{G(t)}^{\mathrm{NRC}} = 1 - \frac{1}{d^2} \sum_{\chi \in \{A, B\}} \left(\|R^{(\chi)}\|_2^2 - \frac{1}{2} \|R_D^{(\chi)}\|_2^2 \right) \quad (7)$$

where $R^{(\chi)}$ is the Gram matrix of the reduced Hamilto-

nian eigenstates $\{\rho_k^{(\chi)}\}_{k=1}^d$, i.e.,

$$R_{kl}^{(\chi)} := \langle \rho_k^{(\chi)}, \rho_l^{(\chi)} \rangle \tag{8}$$

while $\left(R_D^{(\chi)}\right)_{kl} := R_{kl}^{(\chi)} \delta_{kl}$.

Let us first point out some basic, yet important properties of the above formula. The matrix $R^{(\chi)}$ is real and symmetric, while $R_D^{(\chi)}$ is positive semidefinite and diagonal. Moreover, the completeness of the Hamiltonian eigenvectors imposes $\sum_k \rho_k^{(\chi)} = d_{\overline{\chi}} I$; thus the rescaled $\tilde{R}^{(\chi)} := R^{(\chi)}/d_{\overline{\chi}}$ are doubly stochastic, i.e., $\sum_i \tilde{R}_{ij}^{(\chi)} = \sum_i \tilde{R}_{ji}^{(\chi)} = 1 \; \forall j$. As $\tilde{R}^{(\chi)}$ is a (rescaled) Gram matrix, its eigenvalues are non-negative, upper bounded by 1, and at most d_χ^2 of them are nonzero [59]. This last property follows from the fact that $\mathrm{Rank}\,\tilde{R}^{(\chi)} = \dim \mathrm{Span}\{\rho_k^{(\chi)}\}_k \leq d_\chi^2$. Observe also that $\left\|R_D^{(A)}\right\|_2^2 = \left\|R_D^{(B)}\right\|_2^2$ as two states $\rho_k^{(A)}$ and $\rho_k^{(B)}$ always have the same spectrum (up to irrelevant zeroes).

Bipartite OTOC and entanglement.— Proposition 4 makes it possible to bridge the long-time behavior of the bipartite OTOC with the entanglement structure of the Hamiltonian eigenstates. Let us begin with the symmetric case where $d_A = d_B$ and all $|\phi_k\rangle$ are maximally entangled with respect to the A-B Hilbert space cut. This limit uniquely determines the time average for the NRC case, regardless of the exact Hamiltonian eigenbasis. In general, however, knowledge of the entanglement is not enough to uniquely determine the equilibration value; the inner products $R_{kl}^{(\chi)}$ go beyond probing just the spectrum of the reduced states. A simple substitution in Eq. (7) gives for the maximally entangled case $\overline{G_{\mathrm{ME}}(t)}^{\mathrm{NRC}} = (1 - 1/d)^2$. We will later show the upper bound $G(t) \leq 1 - 1/d_{\min}^2$; therefore the equilibrium value for the bipartite OTOC in this case is nearly maximal, as expected for highly entangled models (e.g., [60, 61]).

How robust is this conclusion for chaotic Hamiltonians with a possibly asymmetric bipartition? Typical eigenstates of chaotic Hamiltonians, as also predicted by the eigenstate thermalization hypothesis [62–64], are believed to obey a volume law for the entanglement entropy. Moreover, their entanglement properties in the bulk resemble those of Haar random pure states [65–67]. We will now show that high entanglement for the Hamiltonian eigenstates necessarily implies that the deviation of the actual equilibration value from $\overline{G_{\mathrm{ME}}(t)}^{\mathrm{NRC}}$ is small.

It is convenient for this purpose to quantify the amount of entanglement via the linear entropy [68, 69] of the reduced state $E(|\psi_{AB}\rangle) := S_{\mathrm{lin}}\left(\mathrm{Tr}_\chi |\psi_{AB}\rangle\langle\psi_{AB}|\right)$, where $S_{\mathrm{lin}}(\rho) := 1 - \mathrm{Tr}(\rho^2)$. The latter will also emerge naturally later when we express the bipartite OTOC in terms of entropy production. Notice that $E \leq 1 - 1/d_{\max} := E_{\max}$, which is achievable only for $d_A = d_B$.

Proposition 5. *If $E_{\max} - E(|\phi_k\rangle) \leq \epsilon$ holds for at least a fraction α of the Hamiltonian eigenstates, then*

$$\left|\overline{G_{\mathrm{ME}}(t)}^{\mathrm{NRC}} - \overline{G(t)}^{\mathrm{NRC}}\right| \leq \alpha J + (1 - \alpha)K, \text{ where}$$

$$J := \frac{6\epsilon}{d_{\min}} + \frac{5\epsilon^2}{2} + 2\frac{\lambda^2 - 1}{d_{\max}^2} \tag{9a}$$

$$K := \left(1 + \frac{2}{d_{\min}}\right)(1 - \alpha) + \frac{2}{d} + 4(\epsilon + \sqrt{\epsilon}) \tag{9b}$$

and $\lambda = d_{\max}/d_{\min}$.

The above bound provides a sufficient condition such that the bipartite OTOC equilibrates around $\overline{G_{\mathrm{ME}}(t)}^{\mathrm{NRC}}$. It is expressed in terms of the fraction α of the highly entangled eigenstates, their entanglement and the asymmetry of the A-B bipartition. Notice that the bound simplifies considerably for the case $\alpha = 1$ and $d_{\min} = d_{\max} = \sqrt{d}$, that is, $\left|\overline{G_{\mathrm{ME}}(t)}^{\mathrm{NRC}} - \overline{G(t)}^{\mathrm{NRC}}\right| \leq \epsilon(6/\sqrt{d} + 5\epsilon/2)$ which should hold to a good approximation for Hamiltonians with high entanglement in the bulk of the energies. Applied to chaotic Hamiltonians[2], the bound of Proposition 5 indicates that the bipartite OTOC will equilibrate near $\overline{G_{\mathrm{ME}}(t)}^{\mathrm{NRC}}$, with deviations up to $O(1/d_{\min}^2)$. For a fixed ratio λ and as d grows, $\overline{G(t)}^{\mathrm{NRC}}$ hence converges to $\overline{G_{\mathrm{ME}}(t)}^{\mathrm{NRC}}$ for all chaotic systems. Since $G(t) \leq 1 - 1/d_{\min}^2$, fluctuations around the time average are necessarily insignificant, justifying the term equilibration.

Beyond chaotic Hamiltonians.— We now relax the "strong" level repulsion, i.e., NRC, criterion and uncover how a hierarchy of constraints, each implying a different strength of chaos, is reflected in the equilibration value of the bipartite OTOC.

Integrable models, which possess a structured spectrum, are expected to violate the NRC. Nevertheless, notice that Eq. (7), although derived under the NRC, can still be evaluated for an (arbitrary) choice of orthonormal eigenvectors of the Hamiltonian. We will refer to the resulting value as the *NRC estimate* of the time average and we will shortly show that this estimate always constitutes an upper bound of the actual equilibration value (and coincides with it for chaotic Hamiltonians). This is of both conceptual and practical importance, as evaluating the NRC estimate is considerably less intensive than calculating the exact value.

In fact, one can make a broader claim. For that, we first sketch three types of averaging processes over G, increasingly shifting away from the strong chaoticity limit. Each of them gives rise to a corresponding estimate for the (exact) equilibration time-average value $\overline{G(t)}$. (i) $\overline{G}^{\mathrm{Haar}}$: Averaging over (global) Haar random

_____

[2] Here chaoticity concretely means that the Hamiltonian spectrum satisfies the NRC and that the entanglement of the typical eigenvectors in the bulk, which determine the equilibration value, resembles that of Haar random vectors [70, 71], i.e., $\mathrm{Tr}\left(\rho_\chi^2\right) \approx (d_A + d_B)/(d+1)$ thus $\epsilon = O(1/d_{\min})$ and $\alpha \approx 1$.
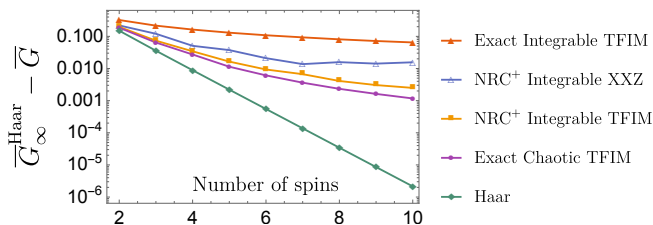
FIG. 1. Logarithmic plot of various $\overline{G}$ estimates, along with the exact time average, for fixed $d_A = 2$ as a function of the total number of spins n. $\overline{G}^{\text{Haar}}_\infty = 3/4$ corresponds to the Haar estimate for $n \to \infty$. For the chaotic phase of the TFIM ($g = -1.05$, $h = 0.5$), the NRC constitutes a satisfactory, though imperfect, approximation. The chaotic and integrable phases ($h = 0$) can be clearly distinguished through the equilibration behavior of the bipartite OTOC. For the integrable XXZ model (we set $J = 0.4$, $\Delta = 2.5$), the NRC$^+$ estimate coincides (up to numerical error) with the exact time average. Inequality (11) holds valid in all cases.

unitary operators $U \in U(d)$ in place of the time evolution. This averaging process is "beyond chaos", in the sense that it does not conserve energy, in contrast with time averaging over any Hamiltonian evolutions. Its estimate (only a function of the dimension) is given later in Eq. (10). (ii) $\overline{G(t)}^{\text{NRC}}$: Time-average, assuming the Hamiltonian has nondegenerate energy levels and nondegenerate energy gaps. The corresponding estimate is Eq. (7). (iii) $\overline{G(t)}^{\text{NRC}^+}$: As before, but assuming the Hamiltonian may have degenerate spectrum, but the energy gaps (between the different levels) are nondegenerate. Its estimate depends only on the eigenprojectors of the Hamiltonian and can be found in Appendix (A).

The value of the Haar average can be performed exactly, with the result

$$\overline{G}^{\text{Haar}} = \frac{(d_A^2 - 1)(d_B^2 - 1)}{d^2 - 1}. \tag{10}$$

The following ordering holds.

Theorem 6. *For any given Hamiltonian, the corresponding estimates are related with the exact time average $\overline{G(t)}$ as*

$$\overline{G}^{\text{Haar}} \geq \overline{G(t)}^{\text{NRC}} \geq \overline{G(t)}^{\text{NRC}^+} \geq \overline{G(t)}. \tag{11}$$

The above constitutes a proof that coincidences in the spectrum of a Hamiltonian up to the "gaps of gaps" (i.e., degeneracy over the energy levels and their gaps) always *reduce* the equilibration value of the bipartite OTOC.

Let us now numerically compare each of the estimates for two models of spin-1/2 chains with open-boundary conditions: (i) transverse-field Ising model (TFIM) with nearest neighbour interaction, $H_{\text{I}} = -\sum_i(\sigma_i^z \sigma_{i+1}^z + g\sigma_i^x + h\sigma_i^z)$ (ii) nearest-neighbor XXZ interaction $H_{\text{XXZ}} = -J\sum_i(\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + \Delta\sigma_i^z \sigma_{i+1}^z)$. Recall that H_{I} for $h = 0$ is integrable in terms of free-fermions, while

H_{XXZ} by Bethe Ansatz techniques. The two types of solutions yield qualitatively different spectra; free fermion solutions necessarily violate nondegeneracy of the gaps. This is reflected in the accuracy of the estimates (see Figure 1). Although the NRC estimate provides essentially the exact equilibration values for the chaotic phase of the TFIM, it overestimates them in the integrable phase. On the other hand, NRC$^+$ is essentially exact for the integrable case of the H_{XXZ} due to the lack of coincidences in the gaps. The results obtained here corroborate existing studies in the literature, where the (short- and) long-time behavior of the OTOC was studied for various many-body systems; see Refs. [72–74].

Bipartite OTOC and subsystem evolution.— We have so far focused on examining the behavior of the bipartite OTOC from the perspective of closed systems, i.e., over the full bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. One can instead express $G(t)$ as a function of the reduced time dynamics over only either $\mathcal{H}_A$ or $\mathcal{H}_B$ (and the corresponding duplicate), at the expense of giving up unitarity. This can be easily realized by formally performing a partial trace in Eq. (4), which immediately results in the following equivalent expression for the bipartite OTOC.

Proposition 7. *Let $\Lambda_t^{(A)}(\rho_A) := \text{Tr}_B\left[U_t\left(\rho_A \otimes \frac{I_B}{d_B}\right)U_t^\dagger\right]$ be the reduced dynamics over A when the environment B is initialized in a maximally mixed state. Then,*

$$G(t) = 1 - \frac{1}{d_A^2}\text{Tr}\left[S_{AA'}\left(\Lambda_t^{(A)}\right)^{\otimes 2}(S_{AA'})\right]. \tag{12}$$

The analogous expression for BB' also holds.

The quantum map $\Lambda_t^{(\chi)}$ is unital; i.e., the maximally mixed state is a fixed point. As such, the transformation $\rho_\chi \mapsto \Lambda_t^{(\chi)}(\rho_\chi)$ results always in an output state whose spectrum is more disordered than the input one [75]. As a result, when ρ_χ is pure, the effect of the reduced time dynamics is to scramble and, hence, produce entropy. Let us now turn to examine this connection more closely.

Bipartite OTOC as entropy production.— We now show that the bipartite OTOC $G(t)$ is nothing but a measure of the average entropy production over pure states, with the latter quantified by linear entropy S_{lin}.

Theorem 8.

$$G(t) = \frac{d_\chi + 1}{d_\chi}\int dU\, S_{\text{lin}}\left[\Lambda_t^{(\chi)}(|\psi_U\rangle\langle\psi_U|)\right] \tag{13}$$

where $\chi = A, B$ and $|\psi_U\rangle := U|\psi_0\rangle$ corresponds to Haar random pure states over $\mathcal{H}_\chi$.

In this manner, the bipartite OTOC can be fully characterized by linear entropy measurements over any of the A, B subsystems. To obtain a satisfactory estimate of the mean in the rhs of Eq. (13), one does not, in practice, need to sample over the full Haar ensemble. An adequate estimate can be obtained with a rapidly decreasing number of necessary samples, as the dimension

d_χ grows. More precisely, let $\tilde{P}(\epsilon)$ be the probability of the entropy $S_{\lin}\big[\Lambda_t^{(\chi)}\big(\,|\psi\rangle\langle\psi|\,\big)\big]$ deviating from $\frac{d_\chi}{d_\chi+1}G(t)$ more than ϵ for an instance of a random state. We show in Appendix A that

$$\tilde{P}(\epsilon) \le \exp\left(-\frac{d_\chi\epsilon^2}{64}\right). \tag{14}$$

The linear entropy, although, per se, a nonlinear functional, can be turned into an ordinary expectation value if two (uncorrelated) copies of the quantum state are simultaneously available, $1 - S_{\lin} = \Tr\left(S\rho^{\otimes 2}\right)$ for $S = S_{AA'}S_{BB'}$. This fact can be exploited to simplify its experimental accessibility [76–80]. More recently, protocols based on correlating measurements over random bases have also been developed to measure entropies [81–84], as well as OTOCs [85, 86]. As a result, Theorem 8 and the typicality result Eq. (14) suggest that the bipartite OTOC is, in turn, tractable via linear entropy measurements. We provide more details in Appendix C.

From Eq. (13) one can also infer the upper bound $G(t) \le 1 - 1/d_\chi^2 \coloneqq G_{\max}^{(\chi)}$ announced earlier that follows from the range of the linear entropy function. The bound is thus achievable only when $\Lambda_t^{(\chi)}$ is equal to the completely depolarizing map $\mathcal{T}^{(\chi)}(\cdot) \coloneqq \Tr(\cdot)\frac{I_\chi}{d_\chi}$.

Finally, we remark that linear entropy occurs rather naturally in relation with the bipartite OTOC, as demonstrated by Theorem 2 (where it lies implicitly in the definition of operator entanglement and entangling power) and Theorem 8. This fact has its roots in the definition of the OTOC, which is intimately related to the Frobenius norm. Relevant relations for the linear entropy have been also reported in [31]. Starting from the inequality $S_{\lin}(\rho) \le S(\rho)$ between the linear and von Neumann entropies $(S(\rho) \coloneqq -\Tr[\rho\log(\rho)])$, one can also obtain the corresponding estimates for the latter.

Bipartite OTOC and information spreading.— The bipartite OTOC measures the average ability of the reduced time evolution to erase information, as captured by the entropy production over a random pure state. This naturally raises the question as to whether $G(t)$ can also be understood as a measure of distance between $\Lambda_t^{(\chi)}$ and the depolarizing map $\mathcal{T}^{(\chi)}$, that is, in the space of quantum channels (i.e., Completely Positive and Trace Preserving (CPTP) maps [87]).

A straightforward answer can be obtained by resorting to the duality between quantum states and operations [87]. Let $\rho_{\mathcal{E}} \coloneqq \mathcal{E}\otimes\mathcal{I}(|\phi^+\rangle\langle\phi^+|)$ denote the (Choi) state corresponding to the CPTP map $\mathcal{E}$, where $|\phi^+\rangle \coloneqq d^{-1/2}\sum_{i=1}^d |ii\rangle$ is a maximally entangled state.

Proposition 9. *The bipartite OTOC is a measure of the distance between the reduced time evolution and the depolarizing map:*

$$G(t) = G_{\max}^{(\chi)} - \left\|\rho_{\Lambda_t^{(\chi)}} - \rho_{\mathcal{T}^{(\chi)}}\right\|_2^2. \tag{15}$$

As an application, the proposition above can be utilized to bound the distance $\left\|\Lambda_t^{(\chi)} - \mathcal{T}^{(\chi)}\right\|_\Diamond$ given by the diamond norm [88, 89]; the latter is a well-established measure of distance between quantum channels[3] since it admits an operational interpretation in terms of discrimination on the level of quantum processes [90]. The distinguishability of the two operations satisfies $\left\|\Lambda_t^{(\chi)} - \mathcal{T}^{(\chi)}\right\|_\Diamond \le d_\chi^{3/2}\sqrt{G_{\max}^{(\chi)} - G(t)}$ (see Appendix A); therefore if $G_{\max}^{(\chi)} - G(t)$ decays faster than d_χ^{-3}, then asymptotically the two channels are essentially indistinguishable.

Summary.— We showed that the bipartite OTOC is amenable to exact analytical treatment and, quite remarkably, is equal to the operator entanglement of the dynamics. This identity allows one to establish a rigorous quantitative connection between the OTOC and the notion of entangling power, a well-established quantifier of few-body chaos. This may provide insights into recent work involving "dual-unitaries" and many-body chaos [91–94]; the latter maximize operator entanglement [94, 95]. We then turned to late-time averages of the bipartite OTOC and provided a hierarchy of estimates for systems that violate the conditions of a "generic spectrum". Finally, we unraveled the operational significance of the OTOC by establishing intimate connections with entropy production and information scrambling at the level of quantum channels. Possible future directions include applying further these theoretical tools to concrete many-body systems and uncovering relations with thermalization, localization, and other many-body phenomena.

---

[3] Bounding the difference in terms of the quantum processes

[1] D. N. Page, Average entropy of a subsystem, Physical Review Letters **71**, 1291 (1993).

[2] P. Hayden and J. Preskill, Black holes as mirrors: quantum information in random subsystems, Journal of High Energy Physics **2007**, 120 (2007).

[3] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, Chaos in quantum channels, Journal of High Energy Physics **2016**, 4 (2016).

[4] C. Von Keyserlingk, T. Rakovszky, F. Pollmann, and S. L. Sondhi, Operator hydrodynamics, OTOCs, and entanglement growth in systems without conservation laws, Physical Review X **8**, 021013 (2018).

[5] S. Moudgalya, T. Devakul, C. Von Keyserlingk, and S. Sondhi, Operator spreading in quantum maps, Physical Review B **99**, 094312 (2019).

[6] E. H. Lieb and D. W. Robinson, The finite group velocity of quantum spin systems, in *Statistical mechanics* (Springer, 1972) pp. 425–431.

[7] M. B. Hastings, Lieb-Schultz-Mattis in higher dimensions, Physical Review B **69**, 104431 (2004).

[8] D. A. Roberts and B. Swingle, Lieb-Robinson bound and the butterfly effect in quantum field theories, Physical Review Letters **117**, 091602 (2016).

[9] N. Lashkari, D. Stanford, M. Hastings, T. Osborne, and P. Hayden, Towards the fast scrambling conjecture, Journal of High Energy Physics **2013**, 22 (2013).

[10] A. Larkin and Y. N. Ovchinnikov, Quasiclassical method in the theory of superconductivity, Sov Phys JETP **28**, 1200 (1969).

[11] A. Kitaev, A simple model of quantum holography, in *Proceedings of the KITP Program: Entanglement in Strongly-Correlated Quantum Matter*, Vol. 7 (2015).

[12] S. Fishman, D. Grempel, and R. Prange, Chaos, quantum recurrences, and anderson localization, Physical Review Letters **49**, 509 (1982).

[13] S. Adachi, M. Toda, and K. Ikeda, Quantum-classical correspondence in many-dimensional quantum chaos, Physical Review Letters **61**, 659 (1988).

[14] M. C. Gutzwiller, *Chaos in classical and quantum mechanics* (Springer, 1990).

[15] F. Haake, *Quantum Signatures of Chaos* (Springer, 2013).

[16] J. Maldacena, S. H. Shenker, and D. Stanford, A bound on chaos, Journal of High Energy Physics **2016**, 106 (2016).

[17] D. A. Roberts and D. Stanford, Diagnosing chaos using four-point functions in two-dimensional conformal field theory, Physical Review Letters **115**, 131603 (2015).

[18] J. Polchinski and V. Rosenhaus, The spectrum in the Sachdev-Ye-Kitaev model, Journal of High Energy Physics **2016**, 1 (2016).

[19] M. Mezei and D. Stanford, On entanglement spreading in chaotic systems, Journal of High Energy Physics **2017**, 65 (2017).

[20] Y. Huang, Y.-L. Zhang, and X. Chen, Out-of-time-ordered correlators in many-body localized systems, Annalen der Physik **529**, 1600318 (2017).

[21] Y.-L. Zhang, Y. Huang, X. Chen, *et al.*, Information scrambling in chaotic systems with dissipation, Physical Review B **99**, 014303 (2019).

[22] D. A. Roberts and B. Yoshida, Chaos and complexity by design, Journal of High Energy Physics **2017**, 121 (2017).

[23] R. Prakash and A. Lakshminarayan, Scrambling in strongly chaotic weakly coupled bipartite systems: Universality beyond the Ehrenfest timescale, Physical Review B **101**, 121108 (2020).

[24] S. Pappalardi, A. Russomanno, B. Žunkovič, F. Iemini, A. Silva, and R. Fazio, Scrambling and entanglement spreading in long-range spin chains, Physical Review B **98**, 134303 (2018).

[25] Q. Hummel, B. Geiger, J. D. Urbina, and K. Richter, Reversible quantum information spreading in many-body systems near criticality, Physical Review Letters **123**, 160401 (2019).

[26] D. J. Luitz and Y. B. Lev, Information propagation in isolated quantum systems, Physical Review B **96**, 020406 (2017).

[27] S. Pilatowsky-Cameo, J. Chávez-Carlos, M. A. Bastarrachea-Magnani, P. Stránský, S. Lerma-Hernández, L. F. Santos, and J. G. Hirsch, Positive quantum Lyapunov exponents in experimental systems with a regular classical limit, Physical Review E **101**, 010202 (2020).

[28] T. Xu, T. Scaffidi, and X. Cao, Does scrambling equal chaos?, Physical Review Letters **124**, 140602 (2020).

[29] K. Hashimoto, K.-B. Huh, K.-Y. Kim, and R. Watanabe, Exponential growth of out-of-time-order correlator without chaos: inverted harmonic oscillator, arXiv:2007.04746 (2020).

[30] J. Cotler, N. Hunter-Jones, J. Liu, and B. Yoshida, Chaos, complexity, and random matrices, Journal of High Energy Physics **2017**, 48 (2017).

[31] R. Fan, P. Zhang, H. Shen, and H. Zhai, Out-of-time-order correlation for many-body localization, Science bulletin **62**, 707 (2017).

[32] R. de Mello Koch, J.-H. Huang, C.-T. Ma, and H. J. Van Zyl, Spectral form factor as an OTOC averaged over the Heisenberg group, Physics Letters B **795**, 183 (2019).

[33] C.-T. Ma, Early-time and late-time quantum chaos, International Journal of Modern Physics A , 2050082 (2020).

[34] A. Touil and S. Deffner, Quantum scrambling and the growth of mutual information, Quantum Science and Technology **5**, 035005 (2020).

[35] B. Yan, L. Cincio, and W. H. Zurek, Information scrambling and Loschmidt echo, Physical Review Letters **124**, 160603 (2020).

[36] J. Watrous, *The theory of quantum information* (Cambridge University Press, 2018).

[37] A. Peres, Stability of quantum motion in chaotic and regular systems, Physical Review A **30**, 1610 (1984).

[38] R. A. Jalabert and H. M. Pastawski, Environment-independent decoherence rate in classically chaotic sys-

also constraints the distinguishability in terms of states: $\|\mathcal{E}_1(\rho) - \mathcal{E}_2(\rho)\|_1 \leq \|\mathcal{E}_1 - \mathcal{E}_2\|_\diamond$ for all states and quantum processes.

tems, Physical Review Letters **86**, 2490 (2001).

[39] T. Gorin, T. Prosen, T. H. Seligman, and M. Žnidarič, Dynamics of Loschmidt echoes and fidelity decay, Physics Reports **435**, 33 (2006).

[40] A. Goussev, R. A. Jalabert, H. M. Pastawski, and D. A. Wisniacki, Loschmidt echo, Scholarpedia **7**, 11687 (2012).

[41] P. Zanardi, Entanglement of quantum evolutions, Physical Review A **63**, 040304 (2001).

[42] P. Zanardi, C. Zalka, and L. Faoro, Entangling power of quantum evolutions, Physical Review A **62**, 030301 (2000).

[43] X. Wang, S. Ghose, B. C. Sanders, and B. Hu, Entanglement as a signature of quantum chaos, Physical Review E **70**, 016217 (2004).

[44] A. Lakshminarayan, Entangling power of quantized chaotic systems, Physical Review E **64**, 036207 (2001).

[45] A. J. Scott and C. M. Caves, Entangling power of the quantum baker's map, Journal of Physics A: Mathematical and General **36**, 9553 (2003).

[46] R. Pal and A. Lakshminarayan, Entangling power of time-evolution operators in integrable and nonintegrable many-body systems, Physical Review B **98**, 174304 (2018).

[47] M. Ledoux, *The concentration of measure phenomenon*, Mathematical Surveys and Monographs, Vol. 89 (American Mathematical Society, 2001).

[48] J. Emerson, Y. S. Weinstein, M. Saraceno, S. Lloyd, and D. G. Cory, Pseudo-random unitary operators for quantum information processing, Science **302**, 2098 (2003).

[49] D. P. DiVincenzo, D. W. Leung, and B. M. Terhal, Quantum data hiding, IEEE Transactions on Information Theory **48**, 580 (2002).

[50] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, Symmetric informationally complete quantum measurements, Journal of Mathematical Physics **45**, 2171 (2004).

[51] A. J. Scott, Tight informationally complete quantum measurements, Journal of Physics A: Mathematical and General **39**, 13507 (2006).

[52] D. Gross, K. Audenaert, and J. Eisert, Evenly distributed unitaries: On the structure of unitary designs, Journal of Mathematical Physics **48**, 052104 (2007).

[53] P. Reimann, Foundation of statistical mechanics under experimentally realistic conditions, Physical Review Letters **101**, 190403 (2008).

[54] N. Linden, S. Popescu, A. J. Short, and A. Winter, Quantum mechanical evolution towards thermal equilibrium, Physical Review E **79**, 061103 (2009).

[55] L. C. Venuti, N. T. Jacobson, S. Santra, and P. Zanardi, Exact infinite-time statistics of the Loschmidt echo for a quantum quench, Physical Review Letters **107**, 010403 (2011).

[56] A. Nahum, S. Vijay, and J. Haah, Operator spreading in random unitary circuits, Physical Review X **8**, 021014 (2018).

[57] C. B. Dağ, K. Sun, and L.-M. Duan, Detection of quantum phases via out-of-time-order correlators, Physical Review Letters **123**, 140602 (2019).

[58] Á. M. Alhambra, J. Riddell, and L. P. García-Pintos, Time evolution of correlation functions in quantum many-body systems, Physical Review Letters **124**, 110605 (2020).

[59] R. Bhatia, *Matrix analysis*, Vol. 169 (Springer-Verlag, 2013).

[60] Y. Huang, F. G. Brandao, Y.-L. Zhang, *et al.*, Finite-size scaling of out-of-time-ordered correlators at late times, Physical Review Letters **123**, 010601 (2019).

[61] A. W. Harrow, L. Kong, Z.-W. Liu, S. Mehraban, and P. W. Shor, A separation of out-of-time-ordered correlator and entanglement, arXiv:1906.02219 (2019).

[62] J. M. Deutsch, Quantum statistical mechanics in a closed system, Physical Review A **43**, 2046 (1991).

[63] M. Srednicki, Chaos and quantum thermalization, Physical Review E **50**, 888 (1994).

[64] M. Rigol, V. Dunjko, and M. Olshanii, Thermalization and its mechanism for generic isolated quantum systems, Nature **452**, 854 (2008).

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

[66] Y. Huang, Universal eigenstate entanglement of chaotic local hamiltonians, Nuclear Physics B **938**, 594 (2019).

[67] T.-C. Lu and T. Grover, Renyi entropy of chaotic eigenstates, Physical Review E **99**, 032111 (2019).

[68] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Review of Modern Physics **81**, 865 (2009).

[69] S. Bose and V. Vedral, Mixedness and teleportation, Physical Review A **61**, 040101 (2000).

[70] E. Lubkin, Entropy of an n-system from its correlation with a k-reservoir, Journal of Mathematical Physics **19**, 1028 (1978).

[71] A. Hamma, S. Santra, and P. Zanardi, Quantum entanglement in random physical states, Physical Review Letters **109**, 040502 (2012).

[72] E. M. Fortes, I. García-Mata, R. A. Jalabert, and D. A. Wisniacki, Gauging classical and quantum integrability through out-of-time-ordered correlators, Physical Review E **100**, 042201 (2019).

[73] I. García-Mata, M. Saraceno, R. A. Jalabert, A. J. Roncaglia, and D. A. Wisniacki, Chaos signatures in the short and long time behavior of the out-of-time ordered correlator, Physical Review Letters **121**, 210601 (2018).

[74] J. Rammensee, J. D. Urbina, and K. Richter, Many-body quantum interference and the saturation of out-of-time-order correlators, Physiscal Review Letters **121**, 124101 (2018).

[75] I. Bengtsson and K. Życzkowski, *Geometry of quantum states: An introduction to quantum entanglement* (Cambridge University Press, 2017).

[76] A. K. Ekert, C. M. Alves, D. K. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, Direct estimations of linear and nonlinear functionals of a quantum state, Physical Review Letters **88**, 217901 (2002).

[77] F. A. Bovino, G. Castagnoli, A. Ekert, P. Horodecki, C. M. Alves, and A. V. Sergienko, Direct measurement of nonlinear properties of bipartite quantum states, Physical Review Letters **95**, 240407 (2005).

[78] C. Moura Alves and D. Jaksch, Multipartite entanglement detection in bosons, Physical Review Letters **93**, 110501 (2004).

[79] A. Daley, H. Pichler, J. Schachenmayer, and P. Zoller, Measuring entanglement growth in quench dynamics of

bosons in an optical lattice, Physical Review Letters **109**, 020505 (2012).

[80] R. Islam, R. Ma, P. M. Preiss, M. E. Tai, A. Lukin, M. Rispoli, and M. Greiner, Measuring entanglement entropy in a quantum many-body system, Nature **528**, 77 (2015).

[81] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, Probing Rényi entanglement entropy via randomized measurements, Science **364**, 260 (2019).

[82] A. Elben, B. Vermersch, C. F. Roos, and P. Zoller, Statistical correlations between locally randomized measurements: A toolbox for probing entanglement in many-body quantum states, Physical Review A **99**, 052323 (2019).

[83] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, Nature Physics **16**, 1050–1057 (2020).

[84] A. Elben, R. Kueng, H.-Y. R. Huang, R. van Bijnen, C. Kokail, M. Dalmonte, P. Calabrese, B. Kraus, J. Preskill, P. Zoller, and B. Vermersch, Mixed-state entanglement from local randomized measurements, Physical Review Letters **125**, 200501 (2020).

[85] B. Vermersch, A. Elben, L. M. Sieberer, N. Y. Yao, and P. Zoller, Probing scrambling using statistical correlations between randomized measurements, Physical Review X **9**, 021061 (2019).

[86] M. K. Joshi, A. Elben, B. Vermersch, T. Brydges, C. Maier, P. Zoller, R. Blatt, and C. F. Roos, Quantum information scrambling in a trapped-ion quantum simulator with tunable range interactions, Physical Review Letters **124**, 240505 (2020).

[87] M. A. Nielsen and I. Chuang, Quantum computation and quantum information (Cambridge University Press, 2000).

[88] A. Y. Kitaev, Quantum computations: algorithms and error correction, Russian Mathematical Surveys **52**, 1191 (1997).

[89] A. Y. Kitaev, A. Shen, and M. N. Vyalyi, Classical and quantum computation (American Mathematical Society, 2002).

[90] M. M. Wilde, Quantum information theory (Cambridge University Press, 2013).

[91] M. Akila, D. Waltner, B. Gutkin, and T. Guhr, Particle-time duality in the kicked Ising spin chain, Journal of Physics A: Mathematical and Theoretical **49**, 375101 (2016).

[92] B. Bertini, P. Kos, and T. Prosen, Exact correlation functions for dual-unitary lattice models in $1+1$ dimensions, Physiscal Review Letters **123**, 210601 (2019).

[93] L. Piroli, B. Bertini, J. I. Cirac, and T. Prosen, Exact dynamics in dual-unitary quantum circuits, Physiscal Review B **101**, 094304 (2020).

[94] B. Bertini, P. Kos, and T. Prosen, Operator Entanglement in Local Quantum Circuits I: Chaotic Dual-Unitary Circuits, SciPost Phys. **8**, 67 (2020).

[95] S. A. Rather, S. Aravinda, and A. Lakshminarayan, Creating ensembles of dual unitary and maximally entangling quantum evolutions, Physical Review Letters **125**, 070501 (2020).

[96] G. W. Anderson, A. Guionnet, and O. Zeitouni, An introduction to random matrices (Cambridge university press, 2010).

[97] J. Eisert, M. Cramer, and M. B. Plenio, Colloquium: Area laws for the entanglement entropy, Reviews of Modern Physics **82**, 277 (2010).

[98] R. Goodman and N. R. Wallach, Symmetry, representations, and invariants (Springer, 2009).

[99] J. Watrous, Is there any connection between the diamond norm and the distance of the associated states?, Theoretical Computer Science Stack Exchange (2011).

[100] Z. Webb, The Clifford group forms a unitary 3-design, Quantum Inf. Comput. **16**, 1379 (2016).

APPENDICES

Appendix A: Proofs

Here we restate the Theorems and Propositions, as well as other mathematical claims appearing in the main text, and give their proof.

Theorem 1

Theorem 1. *Let $S_{AA'}$ be the operator over $\mathcal{H} \otimes \mathcal{H}'$ that swaps A with its replica A' and $d = \dim(\mathcal{H})$. Then*

$$G(t) = 1 - \frac{1}{d^2} \operatorname{Tr}\left(S_{AA'} U_t^{\otimes 2} S_{AA'} U_t^{\dagger \otimes 2}\right). \tag{4}$$

The analogous expression for BB' also holds.

Proof. Let S be the operator over $\mathcal{H} \otimes \mathcal{H}'$ that swaps $\mathcal{H}$ with its replica $\mathcal{H}'$. Then for any operators X, Y acting over $\mathcal{H}$ it holds that

$$\operatorname{Tr}(XY) = \operatorname{Tr}\left[S(X \otimes Y)\right], \tag{A1}$$

as it can be easily verified by expressing both sides in a basis. Notice that in our case, where $\mathcal{H}$ carries a bipartition, one can further decompose $S = S_{AA'} S_{BB'}$.

Using the above identity the OTOC averaging in Eq. (3) can be written as

$$
\begin{aligned}
G(t) &= 1 - \frac{1}{d} \operatorname{Re} \int dV \, dW \operatorname{Tr}\left(S \, V_A^\dagger(t) W_B^\dagger \otimes V_A(t) W_B\right) \\
&= 1 - \frac{1}{d} \operatorname{Re} \int dV \, dW \operatorname{Tr}\left(S U_t^{\dagger \otimes 2}(V_A^\dagger \otimes V_A) U_t^{\otimes 2}(W_B^\dagger \otimes W_B)\right) \\
&= 1 - \frac{1}{d} \operatorname{Re} \operatorname{Tr}\left[S U_t^{\dagger \otimes 2}\left(\int dV V_A^\dagger \otimes V_A\right) U_t^{\otimes 2}\left(\int dW W_B^\dagger \otimes W_B\right)\right].
\end{aligned}
$$

Now the two independent averages can be easily performed since for unitary operators over $\mathcal{H} \cong \mathbb{C}^d$ the corresponding Haar integrals evaluate to

$$\int dU \, U \otimes U^\dagger = \frac{S}{d} \tag{A2}$$

where S is again the swap operator over the doubled space.

A quick way to prove the well-known identity (A2) is by using Eq. (A1) to write

$$U X U^\dagger = \operatorname{Tr}_{\mathcal{H}'}\left[(U \otimes U^\dagger)(X \otimes I)S\right]$$

and then using the fact that

$$\int dU \, U X U^\dagger = \frac{\operatorname{Tr}(X)}{d} \tag{A3}$$

which follows directly from the left/right invariance of the Haar measure [36].

Using Eq. (A2) twice, we get

$$
\begin{aligned}
G(t) &= 1 - \frac{1}{d} \operatorname{Re} \operatorname{Tr}\left(S U_t^{\dagger \otimes 2} \frac{S_{AA'}}{d_A} U_t^{\otimes 2} \frac{S_{BB'}}{d_B}\right) \\
&= 1 - \frac{1}{d^2} \operatorname{Tr}\left(S_{AA'} U_t^{\otimes 2} S_{AA'} U_t^{\dagger \otimes 2}\right).
\end{aligned}
$$

Since $\left[S, X^{\otimes 2}\right] = 0$ for all operators X, the analogous expression for BB' holds, i.e.,

$$G(t) = 1 - \frac{1}{d^2} \operatorname{Tr}\left(S_{BB'} U_t^{\otimes 2} S_{BB'} U_t^{\dagger \otimes 2}\right). \tag{A4}$$

∎

Notice that the symmetry of the Haar measure forces the bipartite OTOC to be time reversal invariant, i.e., $G(t) = G(-t)$.

Finally, we also note that that there is a straightforward generalization of Theorem 1 to any finite temperature thermal state. Following similar steps as above, one gets for for the thermal version of the bipartite OTOC

$$G(t) = 1 - \frac{1}{d} \operatorname{Re} \operatorname{Tr} \left((\rho_\beta \otimes I_{A'B'}) U_t^{\dagger \otimes 2} S_{AA'} U_t^{\otimes 2} S_{AA'} \right). \tag{A5}$$

Theorem 2

Theorem 2. *Let G_U denote the bipartite OTOC for the evolution U. Then, (i) $E_{\mathrm{op}}(U_t) = G_{U_t}$, and (ii) for a symmetric bipartition $d_A = d_B$,*

$$e_{\mathrm{P}}(U_t) = \frac{d}{(\sqrt{d}+1)^2} \left(G_{U_t} + G_{U_t S_{AB}} - G_{S_{AB}} \right). \tag{5}$$

Before giving the proof, let us first recall the definitions of operator entanglement [41] and entangling power [42].

The main idea behind operator entanglement is to first express the unitary evolution U (over the bipartite Hilbert space $\mathcal{H}_{AB}$) as a state in the doubled space $\mathcal{H}_{AB} \otimes \mathcal{H}_{A'B'}$ via

$$|U\rangle = U \otimes I_{A'B'} |\phi^+\rangle \tag{A6}$$

for the maximally entangled state $|\phi^+\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |i\rangle_{AB} |i\rangle_{A'B'}$ and then evaluate the linear entropy of the state $\sigma_U = \operatorname{Tr}_{BB'} (|U\rangle\langle U|)$, i.e.,

$$E_{\mathrm{op}}(U) \coloneqq S_{\mathrm{lin}}(\sigma_U) = 1 - \operatorname{Tr}(\sigma_U^2). \tag{A7}$$

The entangling power [42] of a quantum evolution U over a bipartite quantum system $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ is defined as the average entanglement that the evolution generates when acting on random separable pure states. More specifically,

$$e_{\mathrm{P}}(U) \coloneqq \int dV dW E \left[U \left(|\psi_V\rangle_A |\psi_W\rangle_B \right) \right], \tag{A8}$$

where $|\psi_V\rangle_A = V |\psi_0\rangle_A$ corresponds to Haar random pure states over A ($|\psi_0\rangle_A$ is an irrelevant reference state), and similarly for B, while $E(|\psi_{AB}\rangle) \coloneqq S_{\mathrm{lin}} (\operatorname{Tr}_B |\psi_{AB}\rangle\langle\psi_{AB}|)$ is the entanglement of the resulting state, as measured by the linear entropy.

Proof. **(i)** The key observation here is that the bipartite OTOC G_{U_t}, in the form of Eq. (4), coincides with the operator entanglement $E(U_t)$ as defined in Ref. [41] (see Eq. (6) therein). Evaluating the expression (A7), as in the proof of Theorem 1, one obtains exactly Eq. (4), hence $E_{\mathrm{op}}(U_t) = G_{U_t}$.

(ii) For the symmetric case $d_A = d_B$, the result follows by combining the first part of the current Theorem and Eq. (12) of Ref. [41].

Finally, we note that by direct substitution, one has $G_{S_{AB}} = 1 - 1/d$. ∎

Proposition 3

Proposition 3. *Let $P(\epsilon)$ be the probability that a random instance of $C_{V_A, W_B}(t)$ deviates from its Haar average $G(t)$ more than ϵ. Then,*

$$P(\epsilon) \leq 2 \exp \left(-\frac{\epsilon^2 d_{\max}}{64} \right), \tag{6}$$

where $d_{\max} = \max\{d_A, d_B\}$.

The proof relies on measure concentration and, in particular, Levy's lemma which we shall recall shortly (see, e.g., [96]). Below we are also going use various operator (Schatten) k-norms [59]; the latter are defined as $\|X\|_k \coloneqq \left(\sum_i s_i^k \right)^{1/k}$ where $\{s_i\}_i$ are the singular values of X. The case $\|X\|_\infty \coloneqq \max_i \{s_i\}_i$ corresponds to the usual operator norm. For $k \geq l$, one always has $\|X\|_k \leq \|X\|_l$.

We also remind the reader that a function $f : U(d) \to \mathbb{R}$ is said to be Lipschitz continuous with constant K if it satisfies

$$|f(V) - f(W)| \leq K \|V - W\|_2 \tag{A9}$$

for all $V, W \in U(d)$. For brevity, in this section we denote the Haar averages as $\langle(\cdot)\rangle_U$ and also occasionally drop the explicit time dependence.

Theorem (Levy's lemma). *Let $U \in U(d)$ be distributed according to the Haar measure and $f : U(d) \to \mathbb{R}$ be a Lipschitz continuous function. Then for any $\epsilon > 0$*

$$\text{Prob}\{|f(U) - \langle f(U)\rangle_U| \geq \epsilon\} \leq \exp\left(-\frac{d\epsilon^2}{4K^2}\right), \tag{A10}$$

where K is a Lipschitz constant.

During the course of the proof of Proposition 3, the following two continuity results will come in handy.

Lemma 1. *(i) The function $f_W(V) : U(d_A) \to \mathbb{R}$ with $f_W(V) := C_{V_A, W_B}(t)$ is Lipschitz continuous with constant $K_f = 2$ for all $t \in \mathbb{R}$ and $W \in U(d_B)$.*

(ii) The function $g(W) : U(d_B) \to \mathbb{R}$ with $g(W) := \langle C_{V_A, W_B}(t)\rangle_V$ is Lipschitz continuous with constant $K_g = 2/d_A$ for all $t \in \mathbb{R}$.

Proof of lemma. **(i)** Let $X, Y \in U(d_A)$. We need to show that

$$|f_W(X) - f_W(Y)| \leq K_f \|X - Y\|_2.$$

Following the proof of Theorem 1, we can express

$$f_W(V) = 1 - \frac{1}{d} \text{Re} \, \text{Tr}\left[S U_t^{\dagger \otimes 2} (V_A^\dagger \otimes V_A) U_t^{\otimes 2} (W_B^\dagger \otimes W_B) \right]$$

therefore

$$|f_W(X) - f_W(Y)| \leq \frac{1}{d} \left| \text{Tr}\left[U_t^{\otimes 2}(W_B^\dagger \otimes W_B) S U_t^{\dagger \otimes 2}(X_A^\dagger \otimes X_A - Y_A^\dagger \otimes Y_A) \right] \right|$$
$$\leq \frac{1}{d} \left\| X_A^\dagger \otimes X_A - Y_A^\dagger \otimes Y_A \right\|_1,$$

where in the last step we used the inequality $\|\text{Tr}(AB)\| \leq \|A\|_1 \|B\|_\infty$ and the fact that $\left\| U_t^{\otimes 2}(W_B^\dagger \otimes W_B) S U_t^{\dagger \otimes 2} \right\|_\infty = 1$ since the operator within the norm is unitary.

In order to express the last norm as a function of the difference $X_A - Y_A$, we first add and subtract $Y_A^\dagger \otimes X_A$ and then use the triangle inequality. This results in

$$\frac{1}{d} \left\| X_A^\dagger \otimes X_A - Y_A^\dagger \otimes Y_A \right\|_1 \leq \frac{1}{d} \left(\left\| (X_A^\dagger - Y_A^\dagger) \otimes X_A \right\|_1 + \left\| Y_A^\dagger \otimes (X_A - Y_A) \right\|_1 \right)$$
$$\leq \frac{1}{d} \left(\left\| X_A^\dagger - Y_A^\dagger \right\|_\infty \left\| I \otimes X_A \right\|_1 + \left\| X_A - Y_A \right\|_\infty \left\| Y_A^\dagger \otimes I \right\|_1 \right)$$

where for the last step we utilized the inequality $\|AB\|_1 \leq \|A\|_1 \|B\|_\infty$. Now notice that $\left\| I \otimes X_A \right\|_1 = d$ since X_A is unitary, and similarly for $\left\| Y_A^\dagger \otimes I \right\|_1$. Therefore we can bound

$$|f_W(X) - f_W(Y)| \leq \left\| X_A - Y_A \right\|_\infty + \left\| X_A^\dagger - Y_A^\dagger \right\|_\infty \leq 2 \left\| X_A - Y_A \right\|_\infty = 2 \left\| X - Y \right\|_\infty \leq 2 \left\| X - Y \right\|_2,$$

from which clearly one can take $K_f = 2$.

(ii) First notice that the Haar average over $V_A = V \otimes I_B$ can be performed, as was done in the proof of Theorem 1. The result is

$$g(W) = 1 - \frac{1}{d} \text{Re} \, \text{Tr}\left[S U_t^{\dagger \otimes 2} \frac{S_{AA'}}{d_A} U_t^{\otimes 2} W_B^\dagger \otimes W_B \right]$$
$$= 1 - \frac{1}{d} \text{Re} \, \text{Tr}\left[U_t^{\dagger \otimes 2} \frac{S_{BB'}}{d_A} U_t^{\otimes 2} W_B^\dagger \otimes W_B \right].$$

Considering the relevant difference, we can bound

$$|g(X) - g(Y)| \leq \frac{1}{d_A}\frac{1}{d}\left|\text{Tr}\left[U_t^{\dagger\otimes 2} S_{BB'} U_t^{\otimes 2}(X_B^\dagger \otimes X_B - Y_B^\dagger \otimes Y_B)\right]\right|$$

$$\leq \frac{1}{d_A}\frac{1}{d}\left\|X_B^\dagger \otimes X_B - Y_B^\dagger \otimes Y_B\right\|_1.$$

Now one can follow the exact same steps as in part (i); the result is identical except of the extra factor $1/d_A$ that carries through, which originates from the averaging. This results in

$$|g(X) - g(Y)| \leq \frac{2}{d_A}\left\|X - Y\right\|_2$$

from which one can take $K_g = 2/d_A$. ∎

Everything is now in place to give the proof of Proposition 3.

Proof. Let $\epsilon > 0$. We want to show that, for $V \in U(d_A)$ and $W \in U(d_B)$ distributed independently according to the Haar measure, it holds

$$\text{Prob}\,(\gamma \geq \epsilon) \leq \exp\left(-\frac{\epsilon^2 d_{\max}}{64}\right)$$

where $\gamma := |C_{V_A,W_B} - G|$ and by definition $G = \langle C_{V_A,W_B}\rangle_{V,W}$.

Let us consider any pair V_A, W_B that satisfies $\epsilon \leq \gamma$. Then, from the triangle inequality also

$$\epsilon \leq \alpha + \beta,$$

where we set $\alpha := \left|C_{V_A,W_B} - \langle C_{V_A,W_B}\rangle_V\right|$ and $\beta := \left|\langle C_{V_A,W_B}\rangle_V - G\right|$. Hence we have for the corresponding probabilities

$$\text{Prob}\,\{\gamma \geq \epsilon\} \leq \text{Prob}\,\{\alpha + \beta \geq \epsilon\}.$$

However, if $\alpha + \beta \geq \epsilon$ then necessarily $\alpha \geq \epsilon/2$ or $\beta \geq \epsilon/2$, therefore we also have

$$\text{Prob}\,\{\alpha + \beta \geq \epsilon\} \leq \text{Prob}\,(\{\alpha \geq \epsilon/2\} \cup \{\beta \geq \epsilon/2\}).$$

Using the standard union bound over the last expression results in

$$\text{Prob}\,\{\gamma \geq \epsilon\} \leq \text{Prob}\,\{\alpha \geq \epsilon/2\} + \text{Prob}\,\{\beta \geq \epsilon/2\}. \tag{A11}$$

The two Probabilities in Eq. (A11) can be bounded using Levy's lemma. For that, let us first define the auxiliary functions $f_W(V)$ and $g(W)$ as in Lemma 1. Combining the Lipschitz continuity result from there with Levy's lemma, one gets measure concentration bounds

$$\text{Prob}_V\{\left|C_{V_A,W_B} - \langle C_{V_A,W_B}\rangle_V\right| \geq \epsilon/2\} \leq \exp\left(-\frac{d_A\epsilon^2}{64}\right) \quad \forall W \tag{A12a}$$

$$\text{Prob}\{\langle C_{V_A,W_B}\rangle_V - G \geq \epsilon/2\} \leq \exp\left(-\frac{d_A^2 d_B \epsilon^2}{64}\right) \tag{A12b}$$

We are almost done; it suffices to notice that the bound (A12a) is uniform in W, hence it is also applicable to $\text{Prob}\,\{\alpha \geq \epsilon/2\}$. Therefore we arrive at

$$\text{Prob}\{|C_{V_A,W_B}(t) - G(t)| \geq \epsilon\} \leq \exp\left(-\frac{d_A\epsilon^2}{64}\right) + \exp\left(-\frac{d_A^2 d_B\epsilon^2}{64}\right) \leq 2\exp\left(-\frac{d_A\epsilon^2}{64}\right). \tag{A13}$$

Notice the resulting bound is independent of the dynamics, as long as the latter is unitary. Finally, one can obtain the analogous bound for $A \leftrightarrow B$ by inverting the roles of V and W in the proof. Therefore we obtain Eq. (6). ∎

Proposition 4

Proposition 4. *Consider a Hamiltonian satisfying the NRC. Then*

$$\overline{G(t)}^{\text{NRC}} = 1 - \frac{1}{d^2} \sum_{\chi \in \{A,B\}} \left(\left\| R^{(\chi)} \right\|_2^2 - \frac{1}{2} \left\| R_D^{(\chi)} \right\|_2^2 \right) \tag{7}$$

where $R^{(\chi)}$ is the Gram matrix of the reduced Hamiltonian eigenstates $\{\rho_k^{(\chi)}\}_{k=1}^d$, i.e.,

$$R_{kl}^{(\chi)} := \langle \rho_k^{(\chi)}, \rho_l^{(\chi)} \rangle \tag{8}$$

while $\left(R_D^{(\chi)} \right)_{kl} := R_{kl}^{(\chi)} \delta_{kl}$.

Here we give a straightforward proof assuming the NRC holds exactly. For a more detailed discussion, see also the section of the proof of Theorem 6.

Proof. Our starting point is Eq. (4), which we need to time average. Since the Hamiltonian is by assumption nondegenerate, we can spectrally decompose $H = \sum_{k=1}^d E_k P_k$, where $P_k := |\phi_k\rangle\langle\phi_k|$. We then have

$$\overline{G(t)}^{\text{NRC}} = 1 - \frac{1}{d^2} \sum_{klmn} \overline{\exp\left[i(E_k + E_l - E_m - E_n)t\right]} \text{Tr}\left[S_{AA'}(P_k \otimes P_l) S_{AA'}(P_m \otimes P_n)\right].$$

Time averaging the exponential results in

$$\overline{\exp\left[i(E_k + E_l - E_m - E_n)t\right]} = \delta_{E_k + E_l - E_m - E_n, 0} \stackrel{\text{NRC}}{=\!=} \delta_{k,m}\delta_{l,n} + \delta_{k,n}\delta_{l,m} - \delta_{k,l}\delta_{l,m}\delta_{m,n}$$

where in the last step we used the fact that energy gaps are nondegenerate. Thus

$$\overline{G(t)}^{\text{NRC}} = 1 - \frac{1}{d^2}\left(\sum_{kl} \text{Tr}\left[S_{AA'}(P_k \otimes P_l) S_{AA'}(P_k \otimes P_l)\right] + \sum_{kl} \text{Tr}\left[S_{AA'}(P_k \otimes P_l) S_{AA'}(P_l \otimes P_k)\right] \right.$$

$$\left. - \sum_k \text{Tr}\left[S_{AA'}(P_k \otimes P_k) S_{AA'}(P_k \otimes P_k)\right] \right)$$

$$= 1 - \frac{1}{d^2}\left(\sum_{kl} \left|\text{Tr}\left[(P_k \otimes P_l) S_{AA'}\right]\right|^2 + \sum_{kl} \left|\text{Tr}\left[(P_k \otimes P_l) S_{BB'}\right]\right|^2 - \sum_k \left|\text{Tr}\left[(P_k \otimes P_k) S_{AA'}\right]\right|^2 \right),$$

where for the second term we used that $P_l \otimes P_k = S(P_k \otimes P_l)S$ and $S = S_{AA'}S_{BB'}$.

Now, notice that partial traces can be formally performed, giving

$$\text{Tr}_{AA'BB'}\left[(P_k \otimes P_l)S_{AA'}\right] = \text{Tr}_{AA'}\left[\text{Tr}_{BB'}(P_k \otimes P_l)S_{AA'}\right] = \text{Tr}_{AA'}\left[(\rho_k^{(A)} \otimes \rho_l^{(A')})S_{AA'}\right] = \text{Tr}\left(\rho_k^{(A)}\rho_l^{(A)}\right) = R_{kl}^{(A)},$$

and similarly

$$\text{Tr}_{AA'BB'}\left[(P_k \otimes P_l) S_{BB'}\right] = R_{kl}^{(B)}$$

$$\text{Tr}_{AA'BB'}\left[(P_k \otimes P_k) S_{AA'}\right] = \text{Tr}_{AA'BB'}\left[(P_k \otimes P_k) S_{BB'}\right] = R_{kk}^{(A)} = R_{kk}^{(B)}$$

where in the last line we used the fact that the spectra of $\rho_k^{(A)}$ and $\rho_k^{(B)}$ are equal, up to (irrelevant for the trace) zeroes. The result follows by expressing the matrix 2-norm as $\|X\|_2^2 = \sum_{ij}|X_{ij}|^2$. ∎

Proposition 5

Before proceeding with the proof, let us briefly comment on the need of including the parameter α, which corresponds to the fraction of the highly entangled eigenstates of the Hamiltonian. For certain Hamiltonian models (e.g., the class of gapped, local Hamiltonians over one-dimensional lattice systems) it is well known that the ground state follows an area law for the entanglement entropy [97]. Thus for larger system sizes ϵ cannot be chosen to be small for the ground state (and also possibly for the low lying excited states), even for the symmetric $d_A = d_B$ bipartition. Nevertheless, in the bulk of the spectrum, typical eigenstates are expected to obey instead a volume law, which is compatible with an ϵ that can be chosen to be suitably small. Therefore, we expect that, for certain physically relevant models, a large fraction α can be assumed to satisfy this condition.

Proposition 5. *If $E_{\max} - E(|\phi_k\rangle) \le \epsilon$ holds for at least a fraction α of the Hamiltonian eigenstates, then $\left|\overline{G_{\mathrm{ME}}(t)}^{\mathrm{NRC}} - \overline{G(t)}^{\mathrm{NRC}}\right| \le \alpha J + (1-\alpha)K$, where*

$$J := \frac{6\epsilon}{d_{\min}} + \frac{5\epsilon^2}{2} + 2\frac{\lambda^2 - 1}{d_{\max}^2} \tag{9a}$$

$$K := \left(1 + \frac{2}{d_{\min}}\right)(1-\alpha) + \frac{2}{d} + 4(\epsilon + \sqrt{\epsilon}) \tag{9b}$$

and $\lambda = d_{\max}/d_{\min}$.

Proof. To simplify the notation, we assume $d_A \le d_B$. Let us also define $I = \{k : E_{\max} - E(|\phi_k\rangle) \le \epsilon\}$, i.e., I is the index set of those Hamiltonian eigenstates that deviate at most by ϵ from $E_{\max}$, while we use $\bar{I}$ to label the rest of the eigenstates. By assumption, $|I| \ge \alpha d$.

First of all, notice that one can express the difference $E_{\max} - E(|\psi_{AB}\rangle)$ as the distance

$$E_{\max} - E(|\psi_{AB}\rangle) = \mathrm{Tr}(\rho_B^2) - 1/d_B = \left\|\rho_B - I/d_B\right\|_2^2 \ge \left\|\rho_A - I/d_A\right\|_2^2 = \mathrm{Tr}(\rho_A^2) - 1/d_A .$$

Setting for brevity $\Delta_k^{(\chi)} := \rho_k^{(\chi)} - I/d_\chi$ ($\chi = A, B$), we have for all $k \in I$ that $E_{\max} - E(|\phi_k\rangle) = \left\|\Delta_k^{(B)}\right\|_2^2 \le \epsilon$ and hence also $\left\|\Delta_k^{(A)}\right\|_2^2 = \left\|\rho_k^{(A)} - I/d_A\right\|_2^2 \le \epsilon$. It will be convenient for later to express

$$\left|\langle\rho_k^{(\chi)}, \rho_l^{(\chi)}\rangle\right|^2 = \left|\langle I/d_\chi + \Delta_k^{(\chi)}, I/d_\chi + \Delta_l^{(\chi)}\rangle\right|^2 = \left|\frac{1}{d_\chi} + \langle\Delta_k^{(\chi)}, \Delta_l^{(\chi)}\rangle\right|^2 = \frac{1}{d_\chi^2} + \frac{2}{d_\chi}\langle\Delta_k^{(\chi)}, \Delta_l^{(\chi)}\rangle + \langle\Delta_k^{(\chi)}, \Delta_l^{(\chi)}\rangle^2 . \tag{A14}$$

Moreover, by the Cauchy-Schwartz inequality,

$$\left|\langle\Delta_k^{(\chi)}, \Delta_l^{(\chi)}\rangle\right| \le \left\|\Delta_k^{(\chi)}\right\|_2 \left\|\Delta_l^{(\chi)}\right\|_2 \tag{A15a}$$

while

$$\left\|\Delta_k^{(\chi)}\right\|_2^2 \le \begin{cases} \epsilon & \text{if } k \in I, \\ 1 - \frac{1}{d_\chi} & \text{otherwise.} \end{cases} \tag{A15b}$$

Let's start from Eq. (7). Using the fact that $\left\|R_D^{(A)}\right\|_2^2 = \left\|R_D^{(B)}\right\|_2^2$ and recalling $\overline{G_{\mathrm{ME}}(t)}^{\mathrm{NRC}} = (1 - 1/d)^2$ we get by the triangle inequality

$$\left|\overline{G_{\mathrm{ME}}(t)}^{\mathrm{NRC}} - \overline{G(t)}^{\mathrm{NRC}}\right| \le \left|\frac{1}{d^2}\|R^{(A)}\|_2^2 - \frac{1}{d}\right| + \left|\frac{1}{d^2}\|R^{(B)}\|_2^2 - \frac{1}{d}\right| + \frac{1}{d^2}\left|\|R_D^{(A)}\|_2^2 - 1\right|. \tag{A16}$$

To bound the first term we write

$$\left|\frac{1}{d^2}\|R^{(A)}\|_2^2 - \frac{1}{d}\right| = \left|\frac{1}{d^2}\sum_{kl}\left|\langle\rho_k^{(A)}, \rho_l^{(A)}\rangle\right|^2 - \frac{1}{d}\right| \le \frac{1}{d_A^2} - \frac{1}{d} + \frac{1}{d^2}\sum_{kl}\left(\frac{2}{d_A}\left|\langle\Delta_k^{(A)}, \Delta_l^{(A)}\rangle\right| + \langle\Delta_k^{(A)}, \Delta_l^{(A)}\rangle^2\right)$$

where we used Eq. (A14). Splitting both of the sums as $\sum_k = \sum_{k \in I} + \sum_{k \notin I}$ and using Eqs. (A15) we have

$$\frac{1}{d^2}\sum_{kl}\left|\langle\Delta_k^{(A)}, \Delta_l^{(A)}\rangle\right| \le \epsilon\alpha^2 + 2\alpha(1-\alpha)\sqrt{\epsilon\left(1 - \frac{1}{d_A}\right)} + (1-\alpha)^2\left(1 - \frac{1}{d_A}\right)$$

and

$$\frac{1}{d^2}\sum_{kl}\langle\Delta_k^{(A)}, \Delta_l^{(A)}\rangle^2 \le \epsilon^2\alpha^2 + 2\alpha(1-\alpha)\epsilon\left(1 - \frac{1}{d_A}\right) + (1-\alpha)^2\left(1 - \frac{1}{d_A}\right)^2 .$$

Putting them together, and relaxing some inequalities for clarity, we obtain for the first term of Eq. (A16)

$$\left|\frac{1}{d^2}\|R^{(A)}\|_2^2 - \frac{1}{d}\right| \le \frac{1}{d_A^2} - \frac{1}{d} + \alpha\epsilon\left(\frac{2}{d_A} + \epsilon\right) + (1-\alpha)^2(1 + \frac{2}{d_A}) + 2(1-\alpha)(\epsilon + \sqrt{\epsilon}).$$

Analogously for the second term of Eq. (A16),

$$\left| \frac{1}{d^2} \|R^{(B)}\|_2^2 - \frac{1}{d} \right| \leq \frac{1}{d} - \frac{1}{d_B^2} + \alpha\epsilon \left(\frac{2}{d_A} + \epsilon \right) + (1-\alpha)^2(1 + \frac{2}{d_A}) + 2(1-\alpha)(\epsilon + \sqrt{\epsilon}).$$

For the third one, we have

$$\|R_D^{(A)}\|_2^2 = \sum_k |\langle \rho_k^{(A)}, \rho_k^{(A)} \rangle|^2 = \frac{d_B}{d_A} + \frac{2}{d_A} \sum_k \langle \Delta_k^{(A)}, \Delta_k^{(A)} \rangle + \sum_k \langle \Delta_k^{(A)}, \Delta_k^{(A)} \rangle^2.$$

Using similar manipulations as above, and under the convention $d_A \leq d_B$,

$$\frac{1}{d^2} |\|R_D^{(A)}\|_2^2 - 1| \leq \frac{1}{d^2} \left(\frac{d_B}{d_A} - 1 \right) + \frac{1}{d} \left[\alpha \left(\frac{2\epsilon}{d_A} + \epsilon^2 \right) + (1-\alpha) \left(\frac{2}{d_A} + 1 \right) \right]$$

Putting the inequalities together, we have

$$\left| \overline{G_{\mathrm{ME}}(t)}^{\mathrm{NRC}} - \overline{G(t)}^{\mathrm{NRC}} \right| \leq$$
$$\frac{\lambda - 1}{d^2} + \frac{\lambda^2 - 1}{d_B^2} + \alpha \left[2\epsilon \left(\frac{2}{d_A} + \frac{1}{d_A^2 d_B} \right) + \epsilon^2 \left(2 + \frac{1}{d} \right) \right] + (1-\alpha) \left[2(1-\alpha) \left(1 + \frac{2}{d_A} \right) + \frac{2}{d} + 4 \left(\epsilon + \sqrt{\epsilon} \right) \right] \quad \text{(A17)}$$

which can be relaxed to give the final result by using $\dfrac{\lambda^2 - 1}{d_B^2} \geq \dfrac{\lambda - 1}{d^2}$. ∎

Theorem 6

Theorem 6. *For any given Hamiltonian, the corresponding estimates are related with the exact time average $\overline{G(t)}$ as*

$$\overline{G}^{\mathrm{Haar}} \geq \overline{G(t)}^{\mathrm{NRC}} \geq \overline{G(t)}^{\mathrm{NRC}^+} \geq \overline{G(t)}. \tag{11}$$

Before giving the proof of the Theorem, we first briefly discuss some general facts regarding infinite time averages, their connection with the NRC and the NRC$^+$, and how they give rise to the corresponding estimates.

Let us consider unitary quantum dynamics $\mathcal{U}_t(\cdot) = U_t(\cdot)U_t^\dagger$ generated by a Hamiltonian $H = \sum_k \tilde{E}_k \Pi_k$, where Π_k denotes the projector onto the kth eigenspace. As a warm-up, let us calculate the time average of the superoperator $\mathcal{U}_t$. The latter can be easily performed by noticing that $\overline{\exp\left[-i(\tilde{E}_k - \tilde{E}_l)t \right]} = \delta_{kl}$. It results to

$$\mathcal{P}_H := \overline{\mathcal{U}_t} = \sum_k \Pi_k(\cdot)\Pi_k \tag{A18}$$

which is the (Hilbert-Schmidt orthogonal) projector onto the commutant of the algebra generated by $\{\Pi_k\}_k$, i.e., the projector whose range is the space of operators commuting with H.

The object of interest for us is, in fact, $\overline{\mathcal{U}_t^{\otimes 2}}$ since

$$\overline{G(t)} = 1 - \frac{1}{d^2} \langle S_{AA'}, \overline{\mathcal{U}_t^{\otimes 2}}(S_{AA'}) \rangle. \tag{A19}$$

Reasoning as above, it follows that the resulting superoperator is again a projector, whose range is the space of operators over the replicated Hilbert space $\mathcal{H}^{\otimes 2}$ that commute with $H^{(2)} := H \otimes I + I \otimes H$. The projector can be explicitly expressed as

$$\mathcal{P}_{H^{(2)}} := \overline{\mathcal{U}_t^{\otimes 2}} = \sum_{klmn} \delta_{\tilde{E}_k - \tilde{E}_m, \tilde{E}_l - \tilde{E}_n} \Pi_k \otimes \Pi_l(\cdot)\Pi_m \otimes \Pi_n \tag{A20}$$

To evaluate the above sum, let us for a moment examine what happens when the energy gaps $\{\tilde{E}_k - \tilde{E}_l\}_{kl}$ are nondegenerate. i.e.,

$$\mathrm{NRC}^+: \quad \tilde{E}_k + \tilde{E}_l = \tilde{E}_m + \tilde{E}_n \iff (k = m \ \wedge \ l = n) \vee (k = n \ \wedge \ l = m). \tag{A21}$$

We will refer to this condition over the spectrum as NRC^+, since it constitutes a relaxed version of the NRC. Without any assumption over the spectrum, one can always separate two contributions

$$\mathcal{P}_{\mathcal{H}^{(2)}} = \mathcal{P}_{\text{NRC}^+} + \mathcal{P}_{\overline{\text{NRC}^+}} \tag{A22}$$

where

$$\mathcal{P}_{\text{NRC}^+} := \sum_{kl} \Pi_k \otimes \Pi_l(\cdot)\Pi_k \otimes \Pi_l + \sum_{kl} \Pi_k \otimes \Pi_l(\cdot)\Pi_l \otimes \Pi_k - \sum_k \Pi_k \otimes \Pi_k(\cdot)\Pi_k \otimes \Pi_k \tag{A23}$$

and $\mathcal{P}_{\overline{\text{NRC}^+}}$ is any possibly remaining piece, which vanishes if and only if the Hamiltonian does indeed satisfy NRC^+.

Disregarding $\mathcal{P}_{\overline{\text{NRC}^+}}$, one gets the estimate

$$\overline{G(t)}^{\text{NRC}^+} := 1 - \frac{1}{d^2} \text{Tr} \left[S_{AA'} \mathcal{P}_{\text{NRC}^+} (S_{AA'}) \right] \tag{A24}$$

$$= 1 - \frac{1}{d^2} \Big(\sum_{kl} \text{Tr} \left[S_{AA'}(\Pi_k \otimes \Pi_l) S_{AA'} (\Pi_k \otimes \Pi_l) \right] + \sum_{kl} \text{Tr} \left[S_{AA'}(\Pi_k \otimes \Pi_l) S_{AA'} (\Pi_l \otimes \Pi_k) \right]$$

$$- \sum_k \text{Tr} \left[S_{AA'}(\Pi_k \otimes \Pi_k) S_{AA'} (\Pi_k \otimes \Pi_k) \right] \Big), \quad \text{(A25)}$$

where the second equation follows from the proof of Proposition 4. Clearly, if all projectors $\{\Pi_k\}$ are rank-1, then Eq. (A25) collapses to the corresponding one for NRC, Eq. (7). Notice that one can evaluate $\overline{G(t)}^{\text{NRC}^+}$ regardless of whether the Hamiltonian spectrum actually satisfies NRC^+, and obtain the NRC^+ estimate mentioned in the main text.

Evidently, one can also express the NRC time average, Eq. (7), in terms of the corresponding projector

$$\overline{G(t)}^{\text{NRC}} = 1 - \frac{1}{d^2} \text{Tr} \left[S_{AA'} \mathcal{P}_{\text{NRC}} (S_{AA'}) \right]. \tag{A26}$$

If the Hamiltonian does not satisfy NRC, performing a (possibly nonunique) decomposition $H = \sum_k E_k |\phi_k\rangle\langle\phi_k|$ and evaluating Eq. (7) gives rise to the corresponding NRC estimate.

Finally, for the case of Haar random unitaries, one has the corresponding projector $\overline{\mathcal{U}^{\otimes 2}}^{\text{Haar}} := \mathcal{P}_{\text{Haar}}$ whose range is given by the algebra generated by $\{I, S\}$ [98]. We evaluate its explicit expression in the next section.

We are now ready to give the proof of Theorem 6.

Proof. The key observation here is that, by construction, the range of each projector satisfies

$$\text{Ran}\left(\mathcal{P}_{H^{(2)}}\right) \supseteq \text{Ran}\left(\mathcal{P}_{\text{NRC}^+}\right) \supseteq \text{Ran}\left(\mathcal{P}_{\text{NRC}}\right) \supseteq \text{Ran}\left(\mathcal{P}_{\text{Haar}}\right). \tag{A27}$$

Since all of the above are Hilbert-Schmidt orthogonal projectors, it also follows that

$$\mathcal{P}_{H^{(2)}} \geq \mathcal{P}_{\text{NRC}^+} \geq \mathcal{P}_{\text{NRC}} \geq \mathcal{P}_{\text{Haar}}. \tag{A28}$$

As a result,

$$\langle S_{AA'}, \mathcal{P}_{H^{(2)}}(S_{AA'}) \rangle \geq \langle S_{AA'}, \mathcal{P}_{\text{NRC}^+}(S_{AA'}) \rangle \geq \langle S_{AA'}, P_{\text{NRC}}(S_{AA'}) \rangle \geq \langle S_{AA'}, \mathcal{P}_{\text{Haar}}(S_{AA'}) \rangle, \tag{A29}$$

from which Eq. (11) follows immediately. ∎

Proof of Eq. (10)

The Haar average

$$\overline{G}^{\text{Haar}} = \frac{(d_A^2 - 1)(d_B^2 - 1)}{d^2 - 1}$$

can be derived using fact that $\overline{\mathcal{U}^{\otimes 2}}^{\text{Haar}}$ is the CPTP orthogonal projector over the algebra generated by $\{I, S\}$ [98], i.e.,

$$\mathcal{P}_{\text{Haar}}(X) := \overline{\mathcal{U}^{\otimes 2}}^{\text{Haar}}(X) = \frac{1}{2} \sum_{\alpha = \pm 1} \frac{I + \alpha S}{d(d + \alpha)} \langle I + \alpha S, X \rangle, \tag{A30}$$

where S swaps $\mathcal{H}$ and its duplicate $\mathcal{H}'$, as usual. Plugging the above into Eq. (4), one gets

$$\overline{G}^{\text{Haar}} = 1 - \frac{1}{2d^2} \sum_{\alpha=\pm 1} \frac{|\langle I + \alpha S, S_{AA'}\rangle|^2}{d(d+\alpha)}$$

which, after some simple algebra, simplifies to the announced result.

<div align="center">

Theorem 8

</div>

Theorem 8.

$$G(t) = \frac{d_\chi + 1}{d_\chi} \int dU \, S_{\text{lin}} \left[\Lambda_t^{(\chi)}(|\psi_U\rangle\langle\psi_U|) \right] \tag{13}$$

where $\chi = A, B$ and $|\psi_U\rangle := U\,|\psi_0\rangle$ corresponds to Haar random pure states over $\mathcal{H}_\chi$.

Proof. Let us do the $\chi = A$ case. The result relies on the observation that one can express $S_{AA'}$ in Eq. (12) through the Haar average [98]

$$\int dU \, (|\psi_U\rangle\langle\psi_U|)^{\otimes 2} = \frac{1}{d_A(d_A+1)} \left(I_{AA'} + S_{AA'} \right). \tag{A31}$$

Performing the substitution results in

$$\begin{aligned}
G(t) &= 1 + \frac{1}{d_A^2}\text{Tr}\left(S_{AA'}\right) - \frac{d_A+1}{d_A}\int dU \, \text{Tr}\left(S_{AA'}\left[\Lambda_t^{(A)}(|\psi_U\rangle\langle\psi_U|)\right]^{\otimes 2}\right) \\
&= \frac{d_A+1}{d_A}\left(1 - \int dU \, \text{Tr}\left[\left(\Lambda_t^{(A)}(|\psi_U\rangle\langle\psi_U|)\right)^2\right]\right) \\
&= \frac{d_A+1}{d_A}\int dU \, S_{\text{lin}}\left[\Lambda_t^{(A)}(|\psi_U\rangle\langle\psi_U|)\right]
\end{aligned}$$

where we used the fact that $\Lambda_t^{(A)}(I) = I$ and the identity of Eq. (A1).

The $\chi = B$ case follows similarly. ∎

<div align="center">

Proof of Eq. (14)

</div>

We need to prove that

$$\text{Prob}\left\{\left|S_{\text{lin}}\left[\Lambda_t^{(\chi)}\left(|\psi\rangle\langle\psi|\right)\right] - \frac{d_\chi}{d_\chi+1}G(t)\right| \geq \epsilon\right\} \leq \exp\left(-\frac{d_\chi\epsilon^2}{64}\right) \tag{A32}$$

where $|\psi\rangle$ is a Haar random pure state. We will make use of the concentration of measure machinery, briefly presented before the proof of Proposition 3.

The result follows by the use of Levy's lemma and Theorem 8, if one shows that the function $f : U(d_\chi) \to \mathbb{R}$ with $f(V) := S_{\text{lin}}\left[\Lambda_t^{(\chi)}(|\psi_V\rangle\langle\psi_V|)\right]$ is Lipschitz continuous with $K = 4$. As before, we denote $|\psi_V\rangle := V\,|\psi_0\rangle$ for some (irrelevant) reference state $|\psi_0\rangle$.

Indeed, let us show the Lipschitz continuity. We have

$$\begin{aligned}
\left|f(V) - f(W)\right| &= \left|\left\|\Lambda_t^{(\chi)}(|\psi_V\rangle\langle\psi_V|)\right\|_2^2 - \left\|\Lambda_t^{(\chi)}(|\psi_W\rangle\langle\psi_W|)\right\|_2^2\right| \\
&= \left(\left\|\Lambda_t^{(\chi)}(|\psi_V\rangle\langle\psi_V|)\right\|_2 + \left\|\Lambda_t^{(\chi)}(|\psi_W\rangle\langle\psi_W|)\right\|_2\right)\left|\left\|\Lambda_t^{(\chi)}(|\psi_V\rangle\langle\psi_V|)\right\|_2 - \left\|\Lambda_t^{(\chi)}(|\psi_W\rangle\langle\psi_W|)\right\|_2\right| \\
&\leq 2\left\|\Lambda_t^{(\chi)}(|\psi_V\rangle\langle\psi_V|) - \Lambda_t^{(\chi)}(|\psi_W\rangle\langle\psi_W|)\right\|_1 \\
&\leq 2\left\|\mathcal{U}_t\left(|\psi_V\rangle\langle\psi_V| \otimes \frac{I_{d_{\overline{\chi}}}}{d_{\overline{\chi}}}\right) - \mathcal{U}_t\left(|\psi_W\rangle\langle\psi_W| \otimes \frac{I_{d_{\overline{\chi}}}}{d_{\overline{\chi}}}\right)\right\|_1 \\
&\leq 2\left\|\left(|\psi_V\rangle\langle\psi_V| - |\psi_W\rangle\langle\psi_W|\right) \otimes \frac{I_{d_{\overline{\chi}}}}{d_{\overline{\chi}}}\right\|_1 = 2\left\||\psi_V\rangle\langle\psi_V| - |\psi_W\rangle\langle\psi_W|\right\|_1,
\end{aligned}$$

where in the second to last line we used the monotonicity of the 1-norm under the partial trace and in the last line that it is unitarily invariant. Utilizing the inequality $\|X\|_1 \leq \sqrt{\operatorname{Rank}(X)}\,\|X\|_2$, we have

$$
\begin{aligned}
\left| f(V) - f(W) \right| &\leq 2\sqrt{2} \big\| |\psi_V\rangle\langle\psi_V| - |\psi_W\rangle\langle\psi_W| \big\|_2 = 4\sqrt{1 - |\langle\psi_V|\psi_W\rangle|^2} \\
&\leq 4\sqrt{2(1 - |\langle\psi_V|\psi_W\rangle|)} \leq 4\sqrt{2(1 - \operatorname{Re}\langle\psi_V|\psi_W\rangle)} \\
&\leq 4\big\| |\psi_V\rangle - |\psi_W\rangle \big\| \leq 4\|V - W\|_\infty \\
&\leq 4\|V - W\|_2
\end{aligned}
$$

hence one can take $K = 4$.

Proposition 9

Proposition 9. *The bipartite OTOC is a measure of the distance between the reduced time evolution and the depolarizing map:*

$$
G(t) = G_{\max}^{(\chi)} - \big\| \rho_{\Lambda_t^{(\chi)}} - \rho_{\mathcal{T}^{(\chi)}} \big\|_2^2. \tag{15}
$$

Proof. Let us first express the Choi states explicitly as

$$
\rho_{\Lambda_t^{(\chi)}} = \big(\Lambda_t^{(\chi)} \otimes \mathcal{I}\big) |\phi^+\rangle\langle\phi^+| = \frac{1}{d_\chi} \sum_{ij} \Lambda_t^{(\chi)}\big(|i\rangle\langle j|\big) \otimes |i\rangle\langle j|
$$

$$
\rho_{\mathcal{T}^{(\chi)}} = \big(\mathcal{T}^{(\chi)} \otimes \mathcal{I}\big) |\phi^+\rangle\langle\phi^+| = \left(\frac{I_\chi}{d_\chi}\right)^{\otimes 2}.
$$

Writing $S_{\chi\chi'} = \sum_{i,j=1}^{d_\chi} |i\rangle\langle j| \otimes |j\rangle\langle i|$ one also has from Eq. (12)

$$
G(t) = 1 - \frac{1}{d_\chi^2} \sum_{ij} \big\| \Lambda_t^{(\chi)}\big(|i\rangle\langle j|\big) \big\|_2^2.
$$

Thus, expanding the Choi state distance,

$$
\begin{aligned}
\big\| \rho_{\Lambda_t^{(\chi)}} - \rho_{\mathcal{T}^{(\chi)}} \big\|_2^2 &= \langle \rho_{\Lambda_t^{(\chi)}} - \rho_{\mathcal{T}^{(\chi)}}, \rho_{\Lambda_t^{(\chi)}} - \rho_{\mathcal{T}^{(\chi)}} \rangle = \langle \rho_{\Lambda_t^{(\chi)}}, \rho_{\Lambda_t^{(\chi)}} \rangle - 2\langle \rho_{\Lambda_t^{(\chi)}}, \rho_{\mathcal{T}^{(\chi)}} \rangle + \langle \rho_{\mathcal{T}^{(\chi)}}, \rho_{\mathcal{T}^{(\chi)}} \rangle \\
&= \big\| \rho_{\Lambda_t^{(\chi)}} \big\|_2^2 - \frac{1}{d_\chi^2} = \frac{1}{d_\chi^2} \sum_{ij} \big\| \Lambda_t^{(\chi)}\big(|i\rangle\langle j|\big) \big\|_2^2 - \frac{1}{d_\chi^2} \\
&= 1 - G(t) - \frac{1}{d_\chi^2}
\end{aligned}
$$

which is what we wanted. ∎

Proof of $\big\| \Lambda_t^{(\chi)} - \mathcal{T}^{(\chi)} \big\|_\diamond \leq d_\chi^{3/2} \sqrt{G_{\max}^{(\chi)} - G(t)}$ and an application on information spreading

We first remind the reader that the diamond norm can be defined as $\|\mathcal{X}\|_\diamond \coloneqq \|\mathcal{X} \otimes \mathcal{I}_d\|_{1,1}$ where $\mathcal{I}_d$ denotes the identity quantum channel over $\mathcal{H} \cong \mathbb{C}^d$ and $\|\mathcal{X}\|_{1,1} \coloneqq \sup_{\|A\|_1=1} \|\mathcal{X}(A)\|_1$. One of the reasons for this definition is the property that $\|\mathcal{X} \otimes \mathcal{Y}\|_\diamond = \|\mathcal{X}\|_\diamond \|\mathcal{Y}\|_\diamond$, which in general fails for the $\|(\cdot)\|_{1,1}$ norm (see, e.g., [89]).

Let us now prove that

$$
\sqrt{G_{\max}^{(\chi)} - G(t)} \leq \big\| \Lambda_t^{(\chi)} - \mathcal{T}^{(\chi)} \big\|_\diamond \leq d_\chi^{3/2} \sqrt{G_{\max}^{(\chi)} - G(t)}.
$$

Proof. The result follows easily by utilizing the inequalities

$$
\big\| \rho_{\mathcal{E}_1} - \rho_{\mathcal{E}_2} \big\|_1 \leq \big\| \mathcal{E}_1 - \mathcal{E}_2 \big\|_\diamond \leq d \big\| \rho_{\mathcal{E}_1} - \rho_{\mathcal{E}_2} \big\|_1 \tag{A33}
$$

that hold for any pair of CPTP maps. The inequality was reported by John Watrous in [99]. The result follows by use of the inequality $\|X\|_1 \leq \sqrt{d}\,\|X\|_2$ and Proposition 9. ∎

As an additional application of Eq. (A33), we can utilize it to bound from above the fraction of time such that $\left\|\Lambda_t^{(\chi)} - \mathcal{T}^{(\chi)}\right\|_\diamond \geq \epsilon$ holds true. This can be done by combining Eq. (A33) with our earlier time averages. The result

$$\text{Prob}\left\{t \mid \left\|\Lambda_t^{(\chi)} - \mathcal{T}^{(\chi)}\right\|_\diamond \geq \epsilon\right\} \leq \frac{2d_\chi^{3/2}}{\epsilon d_{\overline{\chi}}}\kappa\,, \tag{A34}$$

where $\kappa \coloneqq \sqrt{1 + \frac{d_{\overline{\chi}}^2}{2}(\overline{G}^{\text{Haar}} - \overline{G(t)})}$, demonstrates in yet another way that if $d_{\overline{\chi}} \gg d_\chi$ and $\kappa = O(1)$ (i.e., the equilibration is sufficiently close to the Haar estimate), then the reduced evolution is necessarily close to the maximally mixing one for a large fraction of time.

Proof. Our starting point will be inequality (A33), $\left\|\Lambda_t^{(\chi)} - \mathcal{T}^{(\chi)}\right\|_\diamond \leq d_\chi^{3/2}\sqrt{G_{\max}^{(\chi)} - G(t)}$. By taking the time average of both sides, and then using the concavity of the square root, we obtain

$$\overline{\left\|\Lambda_t^{(\chi)} - \mathcal{T}^{(\chi)}\right\|_\diamond} \leq d_\chi^{3/2}\sqrt{G_{\max}^{(\chi)} - \overline{G(t)}} \leq d_\chi^{3/2}\sqrt{\left(G_{\max}^{(\chi)} - \overline{G}^{\text{Haar}}\right) + \left(\overline{G}^{\text{Haar}} - \overline{G(t)}\right)} \leq 2\frac{d_\chi^{3/2}}{d_{\overline{\chi}}}\kappa\,,$$

where we approximated the difference

$$G_{\max}^{(\chi)} - \overline{G(t)}^{\text{Haar}} = \frac{(d_\chi^2 - 1)^2}{d_\chi^2(d^2 - 1)} \leq \frac{2}{d_{\overline{\chi}}^2}\,.$$

Finally, Eq. (A34) follows by the use of Markov's inequality. ∎

Appendix B: Haar measure, unitary k-designs and the bipartite OTOC

Here we discuss in more details how the Haar measure in the definition of the bipartite OTOC, Eq. (3), can be replaced by other possible averaging choices, in a way that Eq. (4) (and everything that stems from it) remains valid.

Let us first recall the definition of a (unitary) k-design [22, 49–52]. Consider an ensemble of unitary operators $\Lambda = \{(p_i, U_i)\}_i$ and define the family of CPTP maps

$$\mathcal{E}_\Lambda^{(k)} \coloneqq \sum_i p_i U_i^{\otimes k}(\cdot)U_i^{\dagger\otimes k} \tag{B1}$$

$$\mathcal{E}_{\text{Haar}}^{(k)} \coloneqq \int dU\, U^{\otimes k}(\cdot)U^{\dagger\otimes k} \tag{B2}$$

for $k \in \mathbb{N}$. The ensemble Λ forms a k-design if $\mathcal{E}_\Lambda^{(k)} = \mathcal{E}_{\text{Haar}}^{(k)}$. In words, a k-design emulates Haar averaging up to (at least) the kth moment.

Now, let us investigate what is the freedom over the possible probability measures of V_A and W_B in Eq. (3), such that Eq. (4) holds true without modification. It is easy to see, by the proof of Theorem 1, that we are in fact looking for a unitary ensemble Λ retaining the validity of Eq. (A2). In turn, the latter is just a vectorized form of the 1-design condition $\mathcal{E}_\Lambda^{(1)} = \mathcal{E}_{\text{Haar}}^{(1)}$. One can therefore substitute the Haar measure over $U(d_A)$ and $U(d_B)$ with 1-designs over the corresponding spaces; the full Haar randomness is not probed by the OTOC [22].

Moreover, 1-designs factorize, i.e., if $\Lambda_1 = \{(p_i^{(1)}, U_i^{(1)})\}_i$ and $\Lambda_2 = \{(p_j^{(2)}, U_j^{(2)})\}_j$ are 1-designs over $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively, then $\Lambda_1 \otimes \Lambda_2 \coloneqq \{(p_i^{(1)}p_j^{(2)}, U_i^{(1)} \otimes U_j^{(2)})\}_{ij}$ is a 1-design over $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. This follows just by the 1-design condition in the form of Eq. (A2) and the fact that the swap operator over the duplicated space $\mathcal{H} \otimes \mathcal{H}'$ factorizes $S_{AB;A'B'} = S_{AA'}S_{BB'}$.

This last fact has an important implication for the physically relevant case of many-body systems. Consider the case where $\mathcal{H}_\chi = \bigotimes_i \mathcal{H}_\chi^{(i)}$ for $\chi = A, B$, i.e., when A and B are made up of (not necessarily identical) individual subsystems. Then the OTOC of Eq. (3) remains unchanged if the averages $\int dV_A$ and $\int dW_b$ are replaced by the unitary ensemble $\bigotimes_i \Lambda_\chi^{(i)}$, where each $\Lambda_\chi^{(i)}$ is a 1-design on $\mathcal{H}_\chi^{(i)}$. In other words, it is always enough to average over unitary operators that factorize completely. For instance, in the case of a spin-1/2 many-body system $\mathcal{H}_\chi^{(i)} \cong \mathbb{C}^2$ such an example is given by the Pauli 1-design $\Lambda_{\chi,\text{Pauli}}^{(i)} \coloneqq \{1/4, \sigma_k\}_{k=0}^3$ [100].

Appendix C: Estimating the bipartite OTOC via linear entropy measurements of random pure states

Here we present a basic protocol, stemming directly from Theorem 8, for the estimation of the bipartite OTOC via repeated measurements of a single expectation value.
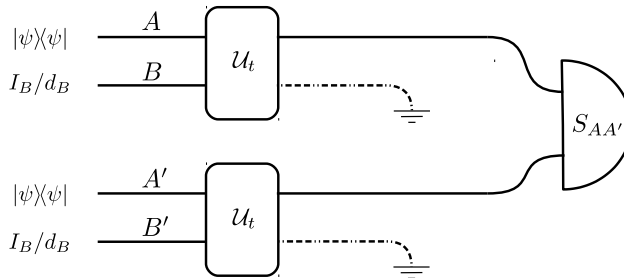

FIG. 2. Protocol to for the estimation of the purity $1 - S_{\lin}\left[\Lambda_t^{(A)}(|\psi\rangle\langle\psi|)\right]$ according to Eq. (13). The resulting purity constitutes also an estimate of the bipartite OTOC, up to a simple proportionality factor. The final measurement of the swap operator can be realized, for instance, by measuring the expectation value of A and A' over any preferred product basis $\{|i\rangle \otimes |j\rangle\}_{i,j=1}^{d_A}$, without the need for coherences.

As pointed out in the main text, the linear entropy of a state can be expressed as an expectation value, $1 - S_{\lin}(\rho) = \mathrm{Tr}\left(S\rho^{\otimes 2}\right)$ at the expense of requiring two copies of the state ρ, though uncorrelated. Combining Theorem 8 with the above observation, one can realize a simple protocol for estimating the bipartite OTOC via measuring the expectation value of the swap operator over pairs of randomly generated states $|\psi\rangle \in \mathcal{H}_A$. We schematically draw the protocol in Figure 2.

Averaging the resulting expectation value over Haar random pure states $|\psi\rangle$ converges to the exact value of the bipartite OTOC. In light of Eq. (14), the expected number of sample for this convergence to a given accuracy drops fast as d_A increases. Clearly, the corresponding protocol with the roles of A and B interchanged is formally equivalent.

Along conceptually similar lines, there have been a number of proposals for probing the linear entropy of a state in an experimentally accessible way. For example, in a recent experiment [80] quantum purity (which is directly related to the second-order Rényi entanglement entropy) was measured by interfering two uncorrelated but identical copies of a many-body quantum state; similar ideas have also been considered previously [76–79]. In particular, this scheme neither requires full quantum state tomography nor the use of entanglement witnesses to estimate entanglement of a quantum state.

Furthermore, there have been recent proposals for protocols based on measurements over random local bases that can probe entanglement given just a single copy of the quantum state, and, in this sense, go beyond traditional quantum state tomography. The main idea consists of directly expressing the linear entropy [81, 82], as well as other functions of the state [83], as an ensemble average of measurements over random bases. Related ideas have also been adapted to probe OTOCs [85, 86] and mixed state entanglement [84].