



CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

Evolution of Entanglement Spectra under Generic Quantum Dynamics

Po-Yao Chang, Xiao Chen, Sarang Gopalakrishnan, and J. H. Pixley

Phys. Rev. Lett. **123**, 190602 — Published 6 November 2019

DOI: [10.1103/PhysRevLett.123.190602](https://doi.org/10.1103/PhysRevLett.123.190602)

Evolution of entanglement spectra under generic quantum dynamics

Po-Yao Chang,^{1,2,3} Xiao Chen,⁴ Sarang Gopalakrishnan,⁵ and J. H. Pixley³

¹Department of Physics, National Tsing Hua University, Hsinchu 30013, Taiwan

²Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Strasse 38, D-01187 Dresden, Germany

³Department of Physics and Astronomy, Center for Materials Theory, Rutgers University, Piscataway, NJ 08854 USA

⁴Kavli Institute for Theoretical Physics, University of California at Santa Barbara, Santa Barbara, CA 93106, USA

⁵Department of Physics and Astronomy, CUNY College of Staten Island, Staten Island, NY 10314, USA, and Physics Program and Initiative for the Theoretical Sciences, CUNY Graduate Center, New York, NY 10016, USA

We characterize the early stages of the approach to equilibrium in isolated quantum systems through the evolution of the entanglement spectrum. We find that the entanglement spectrum of a subsystem evolves with three distinct timescales. First, on an $o(1)$ timescale, independent of system or subsystem size and the details of the dynamics, the entanglement spectrum develops nearest-neighbor level repulsion. The second timescale sets in when the light-cone has traversed the subsystem. Between these two times, the density of states of the reduced density matrix takes a universal, scale-free $1/f$ form; thus, random-matrix theory captures the local statistics of the entanglement spectrum but not its global structure. The third time scale is that on which the entanglement saturates; this occurs well after the light-cone traverses the subsystem. Between the second and third times, the entanglement spectrum compresses to its thermal Marchenko-Pastur form. These features hold for chaotic Hamiltonian and Floquet dynamics as well as a range of quantum circuit models.

Understanding how an isolated quantum system reaches thermal equilibrium is a central problem in quantum statistical physics. Substantial progress has been made on the late-time aspects of thermalization, based on the eigenstate thermalization hypothesis [1–5], which implies that small enough subsystems are well described by thermal density matrices if one waits long enough for information to have traversed the entire system. Much numerical [3, 5, 6] and experimental [7] evidence now exists for eigenstate thermalization. However, the mechanism by which a local density matrix goes from being disentangled to being fully thermal is still poorly understood. Some coarse grained features of the thermalization process have recently been characterized numerically through the study of random unitary circuits (RUCs) [8–16]. In special limits of RUCs and fine-tuned models such as the self-dual kicked Ising model [17], exact solutions are available for entanglement growth and the scrambling of local operators. However, these solvable cases are non-generic and miss important aspects of the generic thermalization process.

The present work addresses the dynamics of entanglement and thermalization at early times in generic systems (i.e. non-integrable models with a low-dimensional on-site Hilbert space): here, entanglement spectra [i.e., eigenvalues of the reduced density matrix] [18–23] evolve in a highly nontrivial way: behavior that is absent in the aforementioned solvable limits. The picture that emerges is independent of how the dynamics is generated, holding for Hamiltonian, Floquet, and temporally random dynamics; for systems with and without conservation laws; and for chaotic as well as many-body localized systems. Here, we focus on Hamiltonian dynamics and RUCs; for other cases see [24].

We show that the process of thermalization takes place in three stages; our main new results are that the entanglement spectra behave universally even at relatively early times as demonstrated in Fig. 1, although its early- and late-time properties belong to different universality classes. To explain these

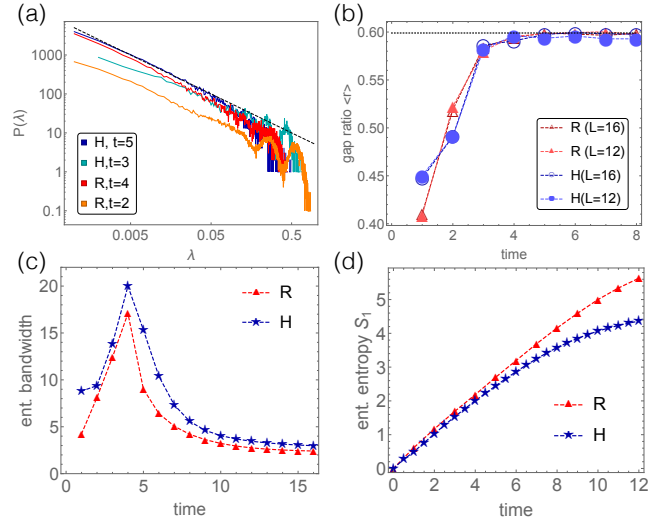


FIG. 1. *Spectral properties of the reduced density matrix under generic time evolution.* (a) Spectral density of the reduced density matrix for random unitary (R) and chaotic Hamiltonian (H) dynamics at early times; this follows a $1/f$ distribution (dashed line). (b) Adjacent gap ratio of the entanglement spectrum, as a function of time, comparing R and H dynamics. Colors denote the model (red for R and blue for H); for each color, empty symbols are for system size $L = 12$ and subsystem size $l_A = 6$, whereas filled symbols are for $L = 16, l_A = 8$. The random-matrix prediction [$\langle r \rangle \approx 0.599$] is marked with a dashed black line. (c, d) Evolution of entanglement bandwidth w and von Neumann entanglement entropy S_1 , for R and H dynamics with $L = 16, l_A = 8$. For both R and H dynamics the entanglement bandwidth grows until $t = l_A/2$, then shrinks, whereas the entanglement entropy keeps growing.

regimes, we introduce multiple characteristic timescales in the entanglement evolution: (i) the timescale on which the entanglement spectrum develops nearest-neighbor level repulsion; (ii) the timescale on which the rank of the density matrix

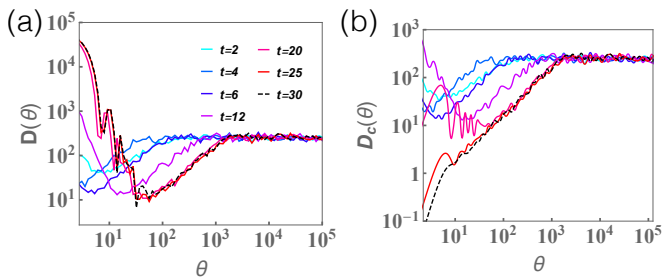


FIG. 2. *Pure random unitary circuits*: (a) The disconnected ESFF $D(\theta)$ and (b) the connected ESFF $D_c(\theta)$ for various different times. The legend in (a) is shared across these figures.

[i.e., the Rényi entropy S_0 in Eq. (3)] saturates; and (iii) the timescale on which the RDM saturates to its late-time behavior. One of our main results is that timescale (i) is independent of system and subsystem size and insensitive to the nature of the dynamics. A second main result is that the spectrum of the reduced density matrix between timescales (i) and (ii) exhibits a universal, scale-invariant $1/f$ density of states. This distribution spreads over increasingly many decades as time passes, until we hit timescale (ii). Once again, this behavior is present in all the models we have considered, but is absent in the exactly solvable limits. Finally, between timescales (ii) and (iii) the range of the $1/f$ distribution shrinks, and narrows toward the late-time Marchenko-Pastur form [19]; during this entire process the entanglement entropy is still growing. For quantum circuits, which have a strict light cone, there is a sharp transition between these regimes, set by the subsystem size. For Hamiltonian dynamics this is rounded into a crossover (due to the exponential tails in the Lieb-Robinson bound [25]) but the two temporal regimes are still clearly distinguished in practice [Fig. 1]. Our main findings are absent in exactly solvable limits, where the entanglement density of states is a delta function at all times, and consequently the nearest-neighbor level spacing is not defined.

We capture level statistics beyond nearest-neighbor using an appropriate *entanglement* spectral form factor. At short times the spectral form factor of the entanglement spectrum has a “ramp” feature characteristic of level repulsion, but does not quantitatively behave as random-matrix theory would predict. Further, the spectral form factor drifts with time until very late times when the entanglement has saturated; only then does it take on its universal shape dictated by random matrix theory. Thus our results clarify the sense in which such systems are “locally thermal”: although the coarse structure of the reduced density matrix is far from that of a thermal state, its “short-distance” level statistics look thermal.

Models.—The main results outlined above were checked for a variety of models, both under discrete-time evolution (i.e., quantum circuits) and continuous-time Hamiltonian evolution. The quantum circuits considered here all involve time-evolution operators of the form $U(t) = U(t, t-1)U(t-1, t-$

$2) \cdots U(1, 0)$, where

$$U(t', t'-1) = \bigotimes_{i \in 2\mathbb{Z}} U_{i,i+1}(t', t'-1) \bigotimes_{i \in 2\mathbb{Z}+1} U_{i,i+1}(t', t'-1), \quad (1)$$

with i being the site index and $U_{i,i+1}$ being unitary matrices. When written as a matrix in the many-body Hilbert space the gates are very sparse, and therefore we simulate them exactly using sparse matrix-vector multiplication. In the main text we present results for circuits in which these unitaries are randomly chosen at each point in space and time; we draw them either completely randomly (with Haar measure) or from an ensemble of random matrices with a single conservation law [12]. We have also simulated the Floquet versions of these circuits, but find no noticeable differences in entanglement spectra between the temporally random and Floquet cases. One other case—a Floquet model that is many-body localized [26] rather than chaotic—is shown in [24]. Although the evolution of S_1 is very different in this case, the entanglement spectrum still shows level repulsion and a $1/f$ distribution in its bulk: chaos matters only for the largest few Schmidt coefficients.

To study Hamiltonian evolution we consider the Ising model with both transverse and longitudinal fields:

$$H = \sum_i J \sigma_i^z \sigma_{i+1}^z + h_x \sigma_i^x + h_z \sigma_i^z \quad (2)$$

where σ_i^α are spin-1/2 Pauli operators. For our simulations we choose the parameters $(h_x/J, h_z/J) = (0.9045, 0.809)$, corresponding to a nonintegrable regime in which thermalization is known to be fast [27, 28]. We use a Krylov-space method to efficiently time-evolve the state [29].

Measured quantities.—The RDM of any subsystem has non-negative real eigenvalues $\{\lambda_n\}$. Since broad distributions are present, it is helpful to work with the entanglement spectrum, which has eigenvalues $\{E_n\} = \{-\log \lambda_n\}$. The entanglement density of states is given by $\rho_S(E) = D^{-1} \sum_n \delta(E - E_n)$ where D is Hilbert space dimension of subsystem A , and the entanglement bandwidth is the width of this probability distribution [30]. The Rényi entropies are moments of the $\{\lambda_n\}$:

$$S_\alpha \equiv \frac{1}{1-\alpha} \log \left(\sum_n \lambda_n^\alpha \right). \quad (3)$$

We quantify level statistics via the adjacent gap ratio r [31],

$$r_m \equiv \frac{\min(\delta_m, \delta_{m+1})}{\max(\delta_m, \delta_{m+1})}, \quad (4)$$

where $\delta_m = E_m - E_{m-1}$ and the E_m are arranged in ascending order. The average adjacent gap ratio takes the value $\langle r \rangle \approx 0.599$ for the Gaussian unitary ensemble (GUE); its probability distribution also approaches a universal form [31].

The adjacent gap ratio is only sensitive to the level repulsion of neighboring eigenvalues. To quantify “longer-range”

level repulsion we study the spectral form factor of the entanglement spectrum, which is the Fourier transform of the two-point function of the spectral density in the entanglement spectrum. We term this the “entanglement spectral form factor” (ESFF). The ESFF characterizes the *global* level statistics of the entanglement spectrum, and is expressed as:

$$D(\theta) \equiv \left\langle \sum_{n,m} e^{i\theta(E_n - E_m)} \right\rangle. \quad (5)$$

Here θ denotes an auxiliary “time” variable conjugate to the entanglement “energy.” For a GUE random matrix, the spectral form factor has a linear growth in θ , called the *ramp*, followed by a sudden saturation, reaching its *plateau* value [32]. A precise *ramp-plateau* structure can be obtained by subtracting out the disconnected parts $|\langle \sum_n \exp(i\theta E_n) \rangle|^2$, which defines the connected ESFF $D_c(\theta) = D(\theta) - |\langle \sum_n \exp(i\theta E_n) \rangle|^2$. These form factors have the advantage of capturing gap correlations beyond nearest neighbor, but the disadvantage of being sensitive to the overall entanglement density of states (DOS) [Fig. 1 (c)]. Note that the ESFF is not the unique spectral form factor one can construct for the reduced density matrix; we could instead have constructed a spectral form factor from the eigenvalues of the reduced density matrix [24]. However, the ESFF has the crucial advantage that its asymptotic large- θ behavior is set by the large Schmidt coefficients, and is therefore sensitive to the late stages of the thermalization process.

Under Hamiltonian dynamics, the eigenstate thermalization hypothesis implies that at late times the reduced density matrix takes the form $\rho_A = \exp(-H_A/T)$, where T is the temperature set by the global energy density [5]. Thus, the ESFF matches the spectral form factor of the Hamiltonian (projected into the subsystem), up to rescaling. On the other hand, under random unitary dynamics, even when there is a conservation law, the conserved quantity is not the generator of the dynamics. Hence, the ESFF acts as a measure of how random the state is, and its late-time structure is what one would predict from a random pure state [33]. We find that both spectral form factors settle down to a time-independent function that is consistent with the shape predicted from random matrix theory, once the entanglement entropy has completely saturated (see Fig. 2 and [24]).

Purely random circuits.—We first discuss our results for the purely random case. In this case each gate is picked Haar-randomly at each space and time point. The distribution of RDM eigenvalues becomes broad at short times ($t < l_A/2$, where l_A is the size of the subsystem) following a universal scale free $1/f$ distribution [Fig. 1 (a)]. The entanglement level statistics rapidly approaches its random-matrix value on an $o(1)$ timescale [Fig. 1 (b)], that is independent of the system and subsystem size [24]. The entanglement bandwidth initially grows linearly in time, out to a time $t = l_A/2$ when the light cone hits the edge of the subsystem and then decays algebraically to a small steady state value [Fig. 1 (c)]. During this short time dynamical process the entanglement entropy continues to grow until it saturates at time scale set by the

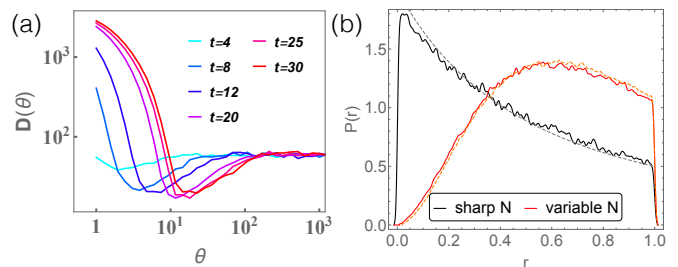


FIG. 3. *Random unitary circuits with a single conservation law.* (a) Evolution of the ESFF in the conserving case for an initial state with definite particle number, at $L = 16, l_A = 6$, averaged over 600 samples. Note the appearance of ramp-plateau structure despite the Poisson level statistics in (b). (b) Level statistics parameter r for the conserving circuit with fixed- and variable-number initial states, which respectively approach Poisson and random-matrix behavior (dashed lines show the exact distributions [34] of the r ratio for Poisson and GUE distributions respectively).

system size [Fig. 1 (d)]. In Fig. 2 we show the behavior of the ESFF in this model, for $L = 20, l_A = 8$. The ESFF develops a ramp-plateau structure at early times, corresponding to the short timescale on which level repulsion sets in among the entanglement “energy levels”. However, the overall shape of the ESFF drifts over time, until the entanglement bandwidth and entanglement entropy have saturated.

(B) *Random circuits with a conservation law.*—To test whether these results are restricted to fully random circuits we turn to the case with a conserved quantity, which we take to be the z -component of the spin. For spin-1/2 degrees of freedom the most general conserving two-spin gate acts as a random phase on the states $|\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$, and a random 2×2 matrix on the space spanned by $|\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle$. The conserved quantity is $N \equiv \sum_i \sigma_i^z$. We consider two separate classes of initial product states: (i) random eigenstates of N (i.e., random binary strings) and (ii) random product states that are superpositions of different N sectors. The results are shown in Fig. 3.

For (i) states that are initially random binary strings, the Schmidt decomposition is block-diagonal. Each partition of N into N_A “particles” in the sub-interval has $N - N_A$ particles in the complement, so ρ_A has no coherence between states of different N_A . Different- N_A blocks do not repel each other, so the global level statistics is Poisson [Fig. 3(b)]. Nevertheless, level repulsion persists within each *individual* block, and yields a ramp-plateau structure in the ESFF [Fig. 3(a)]. The ESFF is sensitive to level repulsion effects beyond nearest-neighbor levels, and is therefore able to detect intra-block structure, unlike the adjacent gap ratio.

For random product states (ii), the behavior is qualitatively similar to that of random circuits, although there are quantitative differences in entanglement growth rates [24]. Again, GUE level statistics emerges on a fixed size-independent timescale when the bond dimension of ρ_A is still growing [24]. The entanglement DOS behaves qualitatively as in

the Haar random unitary circuit model although its bandwidth grows even wider for conserving dynamics. One might have expected level repulsion in the entanglement spectrum to signal chaos in the underlying dynamics; from this perspective, the irrelevance of the conservation law is unexpected. We observe this feature persists even in dynamics that is not chaotic but localized [24]. To summarize, for random product states, the presence of a conservation law has no qualitative effect on the evolution of the entanglement spectrum. Only when the initial states are also eigenstates of the conserved charge does one see qualitatively different evolution in the entanglement spectrum.

(C) *Ising model with transverse and longitudinal fields.*—To test the generality of our results we now turn to Hamiltonian dynamics. We consider the nonintegrable Ising Hamiltonian [Eq. (2)] and time-evolve a random product state. We consider the total system $L = 16$ with the subsystem size $l_A = 8$. We observe the same scale-free $1/f$ probability distribution of the eigenvalues of the reduced density matrix [24] [Fig. 1 (a)] and find that the adjacent gap ratio [Fig. 1 (b)] approaches the GUE value on a $o(1)$ time scale, independent of subsystem size [24]. In addition we find the entanglement bandwidth grows for times $t < l_A/2$ and then shrinks at late times. Distinct from RUCs, the entanglement bandwidth starts from a non-zero initial value because the RDM is full rank for Hamiltonian dynamics (since the light-cone set by the Lieb-Robinson bounds is not strict but has exponential tails). Lastly, the entanglement bandwidth shrinks well before the entropy saturates [Fig. 1(d)]. This case evidently behaves like the previous ones.

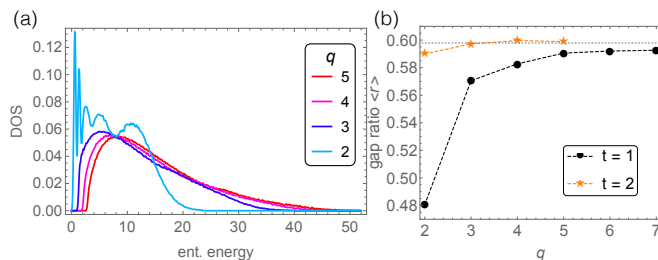


FIG. 4. *Dependence on the local Hilbert space q .* (a) Entanglement DOS at a fixed time, $t = 2$, as a function of local Hilbert space dimension q . The shape of the DOS does not seem to change much with q , though the average entanglement “energy” goes down as one might expect. (b) Adjacent gap ratio r vs. q at a fixed time $t = 1$ and $t = 2$, these are the only times at which there are appreciable deviations from GUE level statistics for $q > 2$ (the dashed line marks the exact GUE value $r \approx 0.599$).

Dependence on local Hilbert space.— We next compare our results for different dimensions of local Hilbert space $q > 2$, focusing on purely random circuits. Surprisingly, the entanglement DOS stays broad for all the q we have considered [Fig. 4 (a)]; despite the expectation that this quantity narrows as $q \rightarrow \infty$ [8, 23], we see no clear sign of narrowing. Thus, the approach to the known $q \rightarrow \infty$ behavior is slow and

possibly singular. Turning to the gap ratio r we find [Fig. 4 (b)] that for $q \geq 6$ one has GUE statistics in the entanglement spectrum for $t = 1$. Thus, at large q , the onset of level repulsion in the entanglement spectrum is essentially instantaneous.

Discussion.—Our results can be qualitatively understood [35] invoking operator spreading [9, 36–38], as follows: one can expand the reduced density matrix in a basis of strings of Pauli matrices, and study the evolution of these strings in the Heisenberg picture. Strings initially localized on either side of the cut spread out, under time evolution, to more complicated operators that straddle the cut. Under the partial trace, most such operators vanish. While the unitary evolution of strings is rank-preserving, the partial trace “dephases” components of the reduced density matrix and thereby increases its rank. Heuristically, operators with a given amplitude, when traced out, generate entries of that amplitude in the reduced density matrix. At early times the density matrix is low-rank, so adding a new entry of some size almost always creates a new eigenvalue of the same size. This picture qualitatively captures the entanglement DOS and level statistics. In RUCs, the speed of the strict causal light-cone ($v_{LC} = 2$) exceeds the butterfly velocity v_B at which generic operators spread. Thus, terms that extend beyond the operator front but within the causal light-cone get generated with small amplitude; those closest to the light-cone are generated at time t with amplitude $\exp\{-[t(v_{LC} - v_B)]^2/(Dt)\}$ [9, 10, 39], where D is the rate at which the front broadens. These exponentially small-amplitude operators generate correspondingly small eigenvalues in the reduced density matrix, leading to entanglement energies that grow linearly in t and thus accounting for the observed linear bandwidth expansion. Once the light-cone hits the edge of the subsystem, the density matrix is full rank, and tracing out further operators cannot create new eigenvalues, but instead redistributes weight among existing eigenvalues, causing the spectrum to narrow. The entanglement level statistics can be understood in similar terms: operators that contribute nonzero Schmidt coefficients are those that have crossed the entanglement cut; by virtue of this property they all have overlapping support and are in causal contact. Therefore it is natural for the corresponding eigenvalues to have the statistics described by the random matrix theory [40].

Although we presented this argument for RUCs, it can straightforwardly be adapted to Hamiltonian dynamics. The density of states and level statistics of the entanglement spectrum behave qualitatively the same as with RUCs. The main difference is that the reduced density matrix is always full-rank so S_0 is not meaningful. However, if one “regularizes” S_0 to include only eigenvalues above a certain threshold (that is well above numerical precision), the resulting evolution is qualitatively the same as in RUCs.

These results imply that the approach to full local equilibrium involves multiple stages, which the entanglement spectrum can distinguish. In principle this can be experimentally tested by measuring Rényi entropies [41]; however, since the local RDM is nonthermal, there exist local correlation func-

tions that also diagnose this structure. Detecting scalable signatures of multiple-stage relaxation, and of the nonthermal, universal entanglement structure in this intermediate regime, is an important task for future work.

P.-Y. C. thanks M. Tezuka and C.-T. Ma for discussions. S.G. thanks A. Lamacraft, S. Parameswaran, F. Pollmann, and T. Rakovszky for discussions and collaborations on related topics. P.-Y.C. was supported by the Young Scholar Fellowship Program by Ministry of Science and Technology (MOST) in Taiwan, under MOST Grant for the Einstein Program MOST 108-2636-M-007-004, and the Rutgers Center for Materials Theory postdoctoral grant. X.C. was supported by postdoctoral fellowships from the Gordon and Betty Moore Foundation, under the EPIQS initiative, Grant GBMF4304, at the Kavli Institute for Theoretical Physics. S.G. acknowledges support from NSF Grant No. DMR-1653271. S.G. and J.H.P. performed part of this work at the Aspen Center for Physics, which is supported by NSF Grant No. PHY-1607611, and at the Kavli Institute for Theoretical Physics, which is supported by NSF Grant No. PHY-1748958. The authors acknowledge the Beowulf cluster at the Department of Physics and Astronomy of Rutgers University and the Office of Advanced Research Computing (OARC) at Rutgers, The State University of New Jersey (<http://oarc.rutgers.edu>) for providing access to the Amarel cluster and associated research computing resources that have contributed to the results reported here.

-
- [1] J. M. Deutsch, *Phys. Rev. A* **43**, 2046 (1991).
 [2] M. Srednicki, *Phys. Rev. E* **50**, 888 (1994).
 [3] M. Rigol, V. Dunjko, and M. Olshanii, *Nature* **452**, 854 (2008).
 [4] J. Cardy, *Phys. Rev. Lett.* **112**, 220401 (2014).
 [5] J. R. Garrison and T. Grover, *Phys. Rev. X* **8**, 021026 (2018).
 [6] M. Rigol and M. Srednicki, *Phys. Rev. Lett.* **108**, 110601 (2012).
 [7] A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, and M. Greiner, *Science* **353**, 794 (2016), <http://science.sciencemag.org/content/353/6301/794.full.pdf>.
 [8] A. Nahum, J. Ruhman, S. Vijay, and J. Haah, *Phys. Rev. X* **7**, 031016 (2017).
 [9] C. W. von Keyserlingk, T. Rakovszky, F. Pollmann, and S. L. Sondhi, *Phys. Rev. X* **8**, 021013 (2018).
 [10] A. Nahum, S. Vijay, and J. Haah, *Phys. Rev. X* **8**, 021014 (2018).
 [11] A. Nahum, J. Ruhman, and D. A. Huse, arXiv preprint arXiv:1705.10364 (2017).
 [12] V. Khemani, A. Vishwanath, and D. A. Huse, *Phys. Rev. X* **8**, 031057 (2018).
 [13] T. Rakovszky, F. Pollmann, and C. W. von Keyserlingk, *Phys. Rev. X* **8**, 031058 (2018).
 [14] S. Pai, M. Pretko, and R. M. Nandkishore, arXiv preprint arXiv:1807.09776 (2018).
 [15] A. Chan, A. De Luca, and J. T. Chalker, ArXiv e-prints (2017), arXiv:1712.06836 [cond-mat.stat-mech].
 [16] X. Chen and T. Zhou, arXiv preprint arXiv:1808.09812 (2018).
 [17] B. Bertini, P. Kos, and T. c. v. Prosen, *Phys. Rev. X* **9**, 021033 (2019).
 [18] M. Žnidarič et al., *Journal of Physics A: Mathematical and Theoretical* **45**, 125204 (2012).
 [19] Z.-C. Yang, C. Chamon, A. Hamma, and E. R. Mucciolo, *Phys. Rev. Lett.* **115**, 267206 (2015).
 [20] C. Chamon, A. Hamma, and E. R. Mucciolo, *Phys. Rev. Lett.* **112**, 240501 (2014).
 [21] Z.-C. Yang, A. Hamma, S. M. Giampaolo, E. R. Mucciolo, and C. Chamon, *Phys. Rev. B* **96**, 020408 (2017).
 [22] M. Mierzejewski, T. Prosen, D. Crivelli, and P. Prelovšek, *Phys. Rev. Lett.* **110**, 200602 (2013).
 [23] T. Zhou and A. Nahum, arXiv preprint arXiv:1804.09737 (2018).
 [24] See online supplemental material for details.
 [25] E. H. Lieb and D. W. Robinson, in *Statistical mechanics* (Springer, 1972) pp. 425–431.
 [26] A. Chan, A. De Luca, and J. T. Chalker, *Phys. Rev. Lett.* **121**, 060601 (2018).
 [27] H. Kim and D. A. Huse, *Phys. Rev. Lett.* **111**, 127205 (2013).
 [28] H. Kim, M. C. Bañuls, J. I. Cirac, M. B. Hastings, and D. A. Huse, *Phys. Rev. E* **92**, 012128 (2015).
 [29] M. Brenes, V. K. Varma, A. Scardicchio, and I. Girotto, *Computer Physics Communications* **235**, 477 (2019).
 [30] Because very small eigenvalues of ρ are contaminated by machine precision, we define the width as twice the distance from the median to the 25th percentile of the entanglement spectrum (i.e., median to 75th percentile of the spectrum of the RDM). This matches the interquartile range when both can be reliably computed.
 [31] V. Oganesyan and D. A. Huse, *Phys. Rev. B* **75**, 155111 (2007).
 [32] T. Guhr, A. MüllerGroeling, and H. A. Weidenmüller, *Physics Reports* **299**, 189 (1998).
 [33] X. Chen and A. W. W. Ludwig, *Phys. Rev. B* **98**, 064309 (2018).
 [34] Y. Y. Atas, E. Bogomolny, O. Giraud, and G. Roux, *Phys. Rev. Lett.* **110**, 084101 (2013).
 [35] F. Pollmann, private communication.
 [36] W. W. Ho and D. A. Abanin, *Phys. Rev. B* **95**, 094302 (2017).
 [37] D. A. Rowlands and A. Lamacraft, arXiv preprint arXiv:1806.01723 (2018).
 [38] M. Knap, arXiv preprint arXiv:1806.04686 (2018).
 [39] S. Xu and B. Swingle, arXiv preprint arXiv:1802.00801 (2018).
 [40] T. Rakovszky, S. Gopalakrishnan, S. A. Parameswaran, and F. Pollmann, Unpublished.
 [41] R. Islam, R. Ma, P. M. Preiss, M. E. Tai, A. Lukin, M. Rispoli, and M. Greiner, *Nature* **528**, 77 (2015).