Measuring the distance between quantum many-body wave functions

Xiao Chen\textsuperscript{1}, Tianci Zhou\textsuperscript{2} and Cenke Xu\textsuperscript{3}

\textsuperscript{1} Kavli Institute for Theoretical Physics, University of California at Santa Barbara, CA 93106, United States of America
\textsuperscript{2} Department of Physics and Institute for Condensed Matter Theory, University of Illinois at Urbana-Champaign, IL 61801-3080, United States of America
\textsuperscript{3} Department of Physics, University of California at Santa Barbara, CA 93106, United States of America
E-mail: xchen@kitp.ucsb.edu, tzhou13@illinois.edu and xucenke@physics.ucsb.edu

Received 10 April 2018
Accepted for publication 13 June 2018
Published 6 July 2018

Abstract. We study the distance of two wave functions under chaotic time evolution. The two initial states differ only by a local perturbation. To be entitled ‘chaos’ the distance should have a rapid growth afterwards. Instead of focusing on the entire wave function, we measure the distance $d^2(t)$ by investigating the difference of two reduced density matrices of the subsystem $A$ that is spatially separated from the local perturbation. This distance $d^2(t)$ grows with time and eventually saturates to a small constant. We interpret the distance growth in terms of operator scrambling picture, which relates $d^2(t)$ to the square of commutator $C(t)$ (out-of-time-order correlator) and shows that both these quantities measure the area of the operator wave front in subsystem $A$. Among various one-dimensional spin-$\frac{1}{2}$ models, we numerically show that the models with non-local power-law interaction can have an exponentially growing regime in $d^2(t)$ when the local perturbation and subsystem $A$ are well separated. This so-called Lyapunov regime is absent in the spin-$\frac{1}{2}$ chain with local interaction only. After sufficiently long time evolution, $d^2(t)$ relaxes to a small constant, which decays exponentially as we increase the system size and is consistent with eigenstate thermalization hypothesis. Based on these results, we demonstrate that $d^2(t)$ is a useful quantity to characterize both quantum chaos and quantum thermalization in many-body wave functions.
1. Introduction and motivation

Quantum chaos is a subject that has attracted a lot of attention and efforts from various subfields of physics. It is important in understanding the quantum nature of black hole dynamics and plays a key role in the process of thermalization in a fully isolated quantum many-body system. The traditional method of studying quantum chaos is to analyze the spectrum correlation of the Hamiltonian. According to Bohigas–Giannoni–Schmidt conjecture, the spectrum of the quantum chaotic model shows level repulsion statistics which is universally determined by the random matrix theory of the same symmetry class [1]. However, the level repulsion statistics is a static property of the Hamiltonian and does not directly help us understand the dynamics of quantum chaos.

In a classical chaotic system, the trajectories of two states with a small initial separation can diverge exponentially under time evolution. This phenomenon is called the butterfly effect and can be characterized by the Lyapunov exponent [2]. In quantum
mechanics, physical states are wave functions in a Hilbert space. Then a natural question to ask is how to define a ‘distance’ between the wave functions, and use the distance to characterize quantum chaotic systems. A sensible definition of distance should reproduce the diverging behaviors under a quantum chaotic evolution.

The goal of this work is to give a proper definition of the notion of distance between wave functions to characterize quantum chaos. More specifically, we prepare two identical wave functions and perturb one of them with a local operator $\hat{O}(x)$, then evolve them with the same chaotic Hamiltonian, then we expect a proper definition of ‘distance’ will grow with time. Of course, one simple definition of ‘distance’ is the overlap between the wave functions. However due to the unitarity of quantum evolution, the wave function overlap is time independent and hence can not display any growth under time evolution, including chaotic evolution. It is thus not a useful quantity to characterize quantum chaos.

On the other hand, the evolution of the reduced density matrix $\hat{\rho}_A$ is not unitary. We therefore measure the Hilbert–Schmidt distance $d^2(t)$ of the two reduced density matrices and use this to quantify the difference of two states. Our approach is based on the following observation: if we choose a subsystem $A$ that is spatially separated from the local operator $\hat{O}(x)$, then it will take some time for a local observer in $A$ to ‘feel’ the difference between the two initial states. $A$ as the starting point of our study. The same quantity was first proposed in [3], where the authors used $d^2(t)$ to study the chaos dynamics in some one dimensional spin-1/2 chain model with conserved energy. In this paper, we will continue to explore this quantity in more detail, discuss its physical interpretation in the language of operator spreading and also make connection with the square of commutator (out-of-time-order correlator) [4].

To demonstrate this idea, we compute the time dependence of $d^2(t)$ and observe the growth pattern and saturation value in spin-$\frac{1}{2}$ one-dimensional chains with a finite length. Not surprisingly, the scaling behavior of $d^2(t)$ is model dependent. In spin models with non-local interactions [5, 6], we observe a clear exponential growing regime after the perturbation has propagated to the subsystem. By contrast such exponential regime is absent in spin models with only local interactions. This is slightly different from classical chaos, where the separation of trajectories will always grow exponentially with time. In both cases, we find that the saturation values are the distance of two independent random pure states (Page states) [7], meaning that the initial similarities of the two states have been almost completely washed out after a long time chaotic evolution.

The growth behaviors of $d^2(t)$ can be understood in the operator spreading picture of the time evolved perturbation operator $\hat{O}(x,-t)$ developed in [8, 9]. The operator $\hat{O}(x,-t)$ comprises of operator basis that is spreading under the chaotic evolution. Using the tools in quantum information, we are able to relate the average distance $\bar{d}^2(t)$ over the initial Page states to the area of the wave front of the operator basis in $O(x,-t)$ in subsystem $A$. The interaction controls the shape of the wave front, and hence the growth pattern of $\bar{d}^2(t)$. 

https://doi.org/10.1088/1742-5468/aace1f
Based on this picture, we show that $d^2(t)$ is related to the square of commutator [4]

$$C(t) = -\text{Tr}\left\{\hat{O}(x,t)\hat{O}'(y)^2\right\}$$

(1.1)

which also measures the area of the wave front for one dimensional chaotic models. This quantity can be used to detect quantum chaos and is shown to grow exponentially with time in some large $N$ systems [10–16]. The growth rate is the quantum analog of the Lyapunov exponent and characterizes the quantum butterfly effect. Our argument relating $C(t)$ with $d^2(t)$ is supported by the numerical collapse of the two curves (after rescaling) in the models we study. Therefore $d^2(t)$ is a legitimate candidate to quantify quantum chaos.

Although the growth of $d^2(t)$ is model dependent, after sufficient long time evolution, once the wave front of $\hat{O}(x,t)$ fully spreads into region $A$, $d^2(t)$ saturates to a small constant, which decreases exponentially with the total system size. We further study $d^2(t)$ for two different initial states with the same energy in a model with static Hamiltonian. We find that $d^2(t)$ always relaxes to an exponentially small constant at late time. These results are consistent with eigenstate thermalization hypothesis [17, 18], which states that for a generic chaotic system, the reduced density matrix for a small subsystem can eventually approach a thermal form with the effective temperature set by the initial energy of the state.

We first briefly summarize the main results of this paper: (1) $d^2(t)$ measures the distance between the wave functions and has similar scaling behavior as the square of commutator. It can grow exponentially in time in spin-$1/2$ chain models with non-local power-law interactions. (2) The saturation value of $d^2(t \to \infty)$ is a small constant which characterizes the quantum thermalization in many-body systems.

This paper is structured as follows. We first define the Hilbert–Schmidt distance $d^2(t)$ between two wave functions in section 2 and then numerically study $d^2(t)$ in two spin-$1/2$ chain models with local interactions in section 3. In section 4, we further study $d^2(t)$ in spin models with non-local power-law interaction. In section 5, we give a physical interpretation for $d^2(t)$ in terms of the operator scrambling picture and make the connection with the square of commutator. We summarize and conclude in section 6. The appendices are devoted to the details of the calculations and techniques used in this paper.

2. Definition of the Hilbert–Schmidt distance

As we explained in the introduction, we prepare two initial states $|\psi_1\rangle$ and $|\psi_2\rangle = \hat{O}(x)|\psi_1\rangle$, where $\hat{O}(x)$ is a local perturbation at position $x$. Our goal is to define a distance between these two states during the evolution under the same unitary operator $\hat{U}(t)$.

Intuitively, the initial distance should be small because these two states only have a local difference. Under unitary time evolution, we can write $|\psi_2(t)\rangle$ as $\hat{O}(x,-t)|\psi_1(t)\rangle$, where $\hat{O}(x,-t) = \hat{U}(t)\hat{O}(x)\hat{U}^\dagger(t)$ is the backward evolved Heisenberg operator. This operator becomes more and more non-local as evolution, suggesting a growing distance between the states. A properly defined distance should reflect these properties.
A naïve trial is to use the norm of the difference
\[ d^2 \sim \| |\psi_1(t)\rangle - |\psi_2(t)\rangle \|^2 \] (2.1)
where \(|\psi_i(t)\rangle = U(t)|\psi_i\rangle\) \((i = 1, 2)\) are time evolved states. The 2-norm is related to the overlap of \(\langle \psi_2(t)|\psi_1(t)\rangle\), from which we can see that the distance does not satisfy either of the requirements proposed above. First we notice that the initial states can be orthogonal even if \(\hat{O}(x)\) is only a 1-bit operation. This gives zero overlap and large initial separation. More importantly, the overlap does not change under the unitary time evolution and we end up with a constant distance.

These problems will not occur if we only pay attention to part of the wavefunction—the subsystem. A local perturbation outside the subsystem will give an identical initial reduced density matrix, thus giving a zero initial distance. Furthermore we note that the dynamics of the subsystem is not unitary: the time dependent reduced density matrices \(\hat{\rho}_1(t), \hat{\rho}_2(t)\) of \(|\psi_1(t)\rangle, |\psi_2(t)\rangle\) do not obey the Heisenberg equation. This motivates us to use Hilbert–Schmidt distance (square-norm distance) between these two reduced density matrices
\[ d^2(\hat{\rho}_1, \hat{\rho}_2) \equiv \text{Tr}(\hat{\rho}_1(t) - \hat{\rho}_2(t))^2 \] (2.2)
as a measure of separation of states. This quantity is semi-positive definite and satisfies the triangle inequality, thus it is a well-defined distance. We will use it to explore the quantum chaos dynamics.

We expect the following development of \(d^2(t)\) for the one dimensional system shown in figure 1. The local perturbation \(\hat{O}(x, t = 0)\) is outside of subsystem \(A\), so at \(t = 0\), the distance \(d^2(t = 0)\) is strictly equal to zero. As time evolves, the operator \(\hat{O}(x, -t)\) spreads out in the space and as it reaches the subsystem \(A\), \(d^2(t)\) becomes nonzero and increases with time. After sufficient long time, \(\hat{O}(x, -t)\) has spread over the entire space and \(d^2(t)\) will saturate to a constant.

3. The spin chain models with local interaction

3.1. Floquet spin-\(\frac{1}{2}\) chain

The dynamics of the Floquet (periodically driven) system is determined by the unitary time evolution operator over one period, namely the Floquet operator. Following [19], we consider the following Floquet operator:
\[ \hat{U}_F = \exp[-i\tau\hat{H}_z] \exp[-i\tau\hat{H}_x], \] (3.1)

where
\[ \hat{H}_x = \sum_{j=1}^{L} g\hat{\sigma}_j^x \]
\[ \hat{H}_z = \sum_{j=1}^{L-1} \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z + \sum_{j=1}^{L} h\hat{\sigma}_j^z. \] (3.2)
This model is a one dimensional periodically driven system with period \( T = 2\tau \). In the numerical calculation, we choose open boundary condition with number of sites \( L = 24 \). The system parameters are \((g, h) = (0.9045, 0.8090)\). The parameter \( \tau \) controls the period of the Floquet operator.

For both numerical and analytical convenience, we choose the initial state \( |\psi_1\rangle \) as the Page state, which is defined as,

\[
|\Psi(\{\alpha_i\})\rangle = \sum_i \alpha_i |C_i\rangle
\]

where the coefficients \( \alpha_i \) of the state in a fixed basis \( \{ |C_i\rangle \} \) are random complex numbers subject to the normalization constraint, with a probability distribution invariant under a unitary basis transformation. This state has (almost) maximal entanglement entropy and was first studied in detail by Page [7]. \( |\psi_2\rangle \) is obtained by acting a local Hermitian operator \( \hat{O}(x) = \hat{\sigma}_x(x) \) at location \( x \) on \( |\psi_1\rangle \). We study the Hilbert–Schmidt distance \( d_2(t) \) between \( |\psi_1(t)\rangle \) and \( |\psi_2(t)\rangle \) under the unitary time evolution, where the over-line denotes averaging over the ensemble of the initial Page states \( |\psi_1\rangle \).

Since we are considering a Floquet model, the evolution time \( t \) is an integer multiple of period, \( t = nT \) with \( n \in \mathbb{Z}_+ \). There are in total three parameters we can tune: the half period of Floquet system \( \tau \), the length \( L_A \) of the subsystem \( A \) and the location of the local operator \( x \). We first fix \( L_A = 6, x = 1 \) and change \( \tau \). As shown in figure 2(a), when \( t/T < 18 \), \( \hat{O}(x, -t) \) does not propagate to region \( A \) and we have \( \overline{d^2} = 0 \). When \( t \geq 18T \), \( \overline{d^2(t)} \) starts to increase and eventually saturates to a small constant independent of \( \tau \). How fast \( \overline{d^2(t)} \) relaxes to the constant is \( \tau \) dependent. When \( \tau = 0.8 \), \( \overline{d^2(t)} \) almost reaches the saturation value in one time step, suggesting rapid spreading of the local operator for this \( \tau \) value.

We present the dependence of \( \overline{d^2(t)} \) on \( x \) in figure 2(b). The time for \( \overline{d^2(t)} \) starting to increase is linearly proportional to the distance between \( \hat{\sigma}_x(x) \) and subsystem \( A \). \( \overline{d^2(t)} \) will eventually saturate to the same constant independent of \( x \). Here we provide an interpretation for this saturation value. After long time evolution, the information of \( \hat{O}(x, -t) \) is fully scrambled in the Hilbert space and we expect that \( |\psi_1(t)\rangle \) and \( |\psi_2(t)\rangle \) are close to two independent Page states. Since the reduced density matrix of Page state belongs to the fixed trace Wishart–Laguerre random matrix ensemble [20, 21], we can obtain an analytical result for \( \overline{d^2(\hat{\rho}_1, \hat{\rho}_2)} \) of two independent Page states

\[
\overline{d^2(\hat{\rho}_1, \hat{\rho}_2)} = \frac{2}{m}
\]
where the subsystem $A$ and $B$ have dimensions $n = 2^{L_A}$, $m = 2^{L_B}$ with ratio $\alpha = n/m$ (The detail of this calculation can be found appendix B). In figure 2(c), we present the saturation value of $\overline{d^2(t)}$ of the Floquet model and we find that they match well with $\overline{d^2(\hat{\rho}_1, \hat{\rho}_2)}$ of two independent Page states. In the thermodynamic limit $L \to \infty$ while keeping $L_A$ fixed, $\overline{d^2(\hat{\rho}_1, \hat{\rho}_2)}$ is exponentially small. We will explain why this constant is so small in the next section.

We further study the integrable system by tuning off the field in $z$ direction and compare it with the non-integrable system. The early time behaviors are rather similar for both $h = 0.809$ (non-integrable) and $h = 0$ ($z$ field off, integrable) as shown figure 3(a). However, in late times, $\overline{d^2(t)}$ saturates to constants in the non-integrable system while oscillates periodically in the integrable system. The oscillation behavior is caused by the ballistic quasi-particles created by the local operator $\hat{O}(x)$. Once the information has spread over the system, the quasi-particles will bounce back and forth periodically. As a result, the length of the system and the velocity of the quasi-particles determine the period $^{[22]}$.

We also consider $h = 0.05$ which weakly breaks the integrability and present the result in figure 3(b). Initially, the dynamics is similar to the case of $h = 0$. As time evolves, the amplitude of the oscillation becomes smaller and vanishes after sufficiently long time evolution. We consider a relatively smaller system size $L = 18$ since it takes very long time for $\overline{d^2(t)}$ to saturate. We find that the saturation value is again the same as that for two independent Page states. Therefore the integrability breaking term, albeit small, will eventually wash out the memory of the initial state and drive the system to the non-integrable behaviors.

**3.2. Quantum Ising model with a longitudinal field**

The Hamiltonian of the quantum Ising model with a longitudinal field is

$$\hat{H}_{Ising} = -\sum_i \hat{\sigma}_i^x \hat{\sigma}_{i+1}^x - h_x \sum_i \hat{\sigma}_i^x - h_z \sum_i \hat{\sigma}_i^z.$$  

We take $(h_x, h_z) = (1.05, 0.5)$ $^{[23]}$ and compute $\overline{d^2(t)}$ for Page state ensemble under unitary time evolution governed by the Ising Hamiltonian $\hat{H}_{Ising}$. The local perturbation is still $\hat{O}(x) = \hat{\sigma}_x(x)$.

We present the results in figure 4 for various $L_A$. As we move $\hat{O}(x)$ further away from subsystem $A$, it takes longer time, but eventually leads to a saturation of $\overline{d^2(t)}$ to a constant very close to values of two independent Page states. The growth of $\overline{d^2(t)}$ is always slower than an exponential function. This behavior is similar to that of the Floquet spin chain we studied in the previous section.

We also explore the initial state ensemble dependence of the saturation value. For the case of Ising model, we prepare another set of states that are random product states (RPS) with energy $E \in [-1, 1]$. A RPS is a tensor product of state on each site whose spin directions are uniformly distributed on Bloch sphere. It thus has zero initial entanglement entropy (EE). We apply a perturbation on the RPS state and then compute the time dependent distance $\overline{d^2(t)}$. The results are presented in figure 4(b). The
saturation values of $d^2(t)$ for this set of initial states is different from the ones of two random states, but it still decays exponentially as we increase the system size.

The small saturation value of $d^2(t \rightarrow \infty)$ can be understood in the following way.

For a generic system, the initial wavefunction will eventually thermalize under its own dynamics at long times \[17, 18\]. In particular, it has been proven that in $1+1d$ conformal field theory (satisfying generalized Gibbs ensemble \[24, 25\]), the reduced density matrix of a small subsystem will relax to the thermal density matrix up to exponentially small corrections with the temperature determined by the initial energy density \[26\]. Therefore, if the two initial states have the same energy density, after sufficiently long time evolution, the reduced density matrix of both states will thermalize at the same effective temperature with $d^2(t)$ approaching to a small constant. The

Figure 2. (a) $d^2(t)$ versus the time step $t/(2\tau)$ for various $\tau$ with fixed $L_A$ and $x$. (b) $d^2(t)$ versus $t/(2\tau)$ for various $x$ with fixed $L_A$ and $\tau$. (c) Comparison between the saturation value of $d^2(t)$ and $d^2(\hat{\rho}_1, \hat{\rho}_2)$ between two independent Page states in equation (3.4).

Figure 3. (a) $d^2(t)$ versus $t/(2\tau)$ for non-integrable system with $h = 0.809$ and integrable system with $h = 0$. (b) $d^2(t)$ versus $t/(2\tau)$ for a system which weakly breaks integrability with $h = 0.05$. The red dashed line on the top is $d^2$ between two independent Page states.

https://doi.org/10.1088/1742-5468/aaeclf
Ising model considered here satisfies ETH [23]. The local perturbation applied to $|\psi_2\rangle$ will not change energy density and gives rise to an exponentially small saturation value of $\bar{d}^2(t \to \infty)$.

Furthermore, we study the saturation value of $\bar{d}^2(t)$ when $|\psi_1\rangle$ and $|\psi_2\rangle$ are not connected by a local unitary transformation but have the same energy density. We choose $|\psi_1\rangle$ and $|\psi_2\rangle$ to be two different PRS states. At $t = 0$, $\bar{d}^2(t)$ for subsystem $A$ is a $O(1)$ constant. After sufficiently long time evolution, $\bar{d}^2(t)$ relaxes to a small constant, which decreases exponentially as $L$ increases and is also consistent with ETH.

4. The spin chain models with power-law interaction

4.1. Floquet spin-$\frac{1}{2}$ chain model

In this section, we consider Floquet model with power-law interaction

$$\hat{U}_F = \exp[-i\tau \hat{H}_z] \exp[-i\tau \hat{H}_x],$$

(4.1)

where

$$\hat{H}_x = \sum_{j=1}^{L} g \hat{\sigma}_j^x,$$

$$\hat{H}_z = \sum_{j>k} \hat{\sigma}_j^z \hat{\sigma}_k^z \alpha + \sum_{j=1}^{L} h \hat{\sigma}_j^z,$$

(4.2)

We fix parameter $(g, h, \tau) = (0.9045, 0.8090, 0.4)$ and compute $\bar{d}^2(t)$ averaged over the Page state ensemble. We take the power law exponent $\alpha = 2.0$.

As shown in figure 5(a), when the distance between region $A$ and $\hat{\sigma}_x(x)$ is large, we can observe a clear intermediate fast growing regime. It has positive curvature on the log-log scale in the inset, indicating that it is faster than any power law growth. The fact that it is a straight line on the semi-log plot confirms the existence of an exponential growth regime.

This is different from the spin chain models with local interaction, where an exponentially growing regime is invisible in the parameter range we choose. The physical interpretation for this regime will be given in section 5.

Notice that at early time, $\bar{d}^2(t)$ scales as $t^2$ (inset of figure 5(a)), which can be understood by taking short time expansion for $\hat{\sigma}_x(x, -t)$. Similar behavior is also found in [27] and we will not discuss it here.

For fixed $L_A$, as we decrease the distance between $\hat{\sigma}_x(x)$ and subsystem $A$, this intermediate exponentially growing regime becomes smaller and vanishes eventually. Nevertheless, $\bar{d}^2(t)$ always saturates to the same constant, which is also the same as that of two independent Page states. We also present $\bar{d}^2(t)$ for various $L_A$ with fixed $x = 1$ in figure 5(b). As we increase the subsystem length $L_A$, the distance between $x$ and the subsystem $A$ reduces, and therefore the exponentially growing regime becomes
smaller. We observe that the growth rate is not very sensitive to the distance between 
\[ \hat{\sigma}_x(x) \] 
and region \( A \). Notice that when \( L_A = 2 \), \( \overline{d^2(t)} \) is averaged of 200 Page states due to large fluctuation in the data. When the subsystem has \( L_A \geq 4 \), the fluctuation is much smaller and the ensemble average is unnecessary. In the numerical calculation, we take the average over 8 states.

**Figure 4.** (a) \( \overline{d^2(t)} \) versus time \( t \) for Ising model defined in equation (3.5) on the semi-log scale, where the average is performed over 200 Page states. The dashed line is \( \overline{d^2} \) between two independent Page states. (b) \( \overline{d^2(t)} \) of the same model versus time \( t \) for various \( L \) on the semi-log scale, where the average is performed over 200 random product states. (c) \( \overline{d^2(t)} \) versus time \( t \) of two RPS states with energy close to zero under the same unitary time evolution.

**Figure 5.** (a) \( \overline{d^2(t)} \) versus \( t/(2\tau) \) for various \( x \) with fixed \( L_A \) on the semi-log scale for Floquet operator in equation (4.1). The average is over 200 Page states. (b) \( \overline{d^2(t)} \) versus \( t/(2\tau) \) for various \( L_A \) with fixed \( x \) on the semi-log scale. We perform 200 Pages states average for \( L_A = 2 \) and 8 Page states average for \( L_A \geq 4 \).
4.2. Ising model with power-law interaction

We now consider a static Hamiltonian with a non-local power-law interaction,

\[ \hat{H} = -\sum_{j>k} \frac{\hat{\sigma}^z_j \hat{\sigma}^z_k}{(j-k)^\alpha} - h_x \sum_i \hat{\sigma}^x_i - h_z \sum_i \hat{\sigma}^z_i \]  

(4.3)

where we fix the parameter \((h_x, h_z) = (1.05, 0.5)\) to be the same as in equation (3.5). We take the power law exponent \(\alpha = 2.1\), where we find the longest exponentially growing regime for \(d^2(t)\).

We still take \(|\psi_1\rangle\) as a Page state and \(|\psi_2\rangle = \hat{\sigma}^x(x)|\psi_1\rangle\). As we can observe in figure 6, the non-local power-law interaction leads to an exponentially growing regime in \(d^2(t)\), which becomes longer as we increase the distance between \(\hat{\sigma}^x(x)\) and region \(A\). The saturation value of \(d^2(t)\) is still the same as that for two independent Page states.

5. Connection between \(d^2(t)\) and the square of commutator

In this section, we use the operator spreading picture to interpret the behaviors of \(d^2(t)\) and relate it to the square of commutator. We will show that they are essentially the same quantity after a proper rescaling in the models we consider.

5.1. \(d^2(t)\) in operator spreading picture

For simplicity, we will restrict our discussion to the spin-\(\frac{1}{2}\) chain that is numerically studied in this paper, but a generalization to models with \(q > 2\) dimensional on-site Hilbert space is straightforward.

On each site of the spin-\(\frac{1}{2}\) chain, we can define a local operator basis which consists of the identity operator \(\hat{I}_2\) and the three Pauli matrices \(\hat{\sigma}^x, \hat{\sigma}^y, \hat{\sigma}^z\). Using the tensor product of the local basis, we can build a complete operator basis \(\{\hat{H}_j\}\) for the entire Hilbert space, in which each operator \(\hat{H}_j\) is a string of \(\hat{\sigma}^{x,y,z}\) and \(\hat{I}_2\). This basis has a natural inner product \(\text{Tr}(\hat{H}_i \hat{H}_j) = \delta_{ij}|H|\), where \(|H|\) is the size of the Hilbert space. We can choose \(4^L\) operators \(\{B_j\}\) from \(\{H_j\}\) which consists of only identity operators in region \(A\). They form an operator basis completely confined in region \(B\). The average over these operators can be used to take a partial trace on any operator \(\hat{O}\) in the full Hilbert space [28],

\[ \frac{1}{|B|} \sum_i \hat{B}_i \hat{O} \hat{B}_i = \text{Tr}_B(\hat{O} \otimes \hat{I}_B) \]  

(5.1)

where \(|B|\) is the Hilbert space dimension of region \(B\). A corollary of this identity is that for any operator \(\hat{O}\)

\[ \text{Tr}_A(\text{Tr}_B \hat{O} \text{Tr}_B \hat{O}) = \text{Tr}_A \text{Tr}_B((\text{Tr}_B \hat{O} \otimes \hat{I}_B)\hat{O}) = \frac{1}{|B|} \sum_i \text{Tr}(B_i \hat{O} B_i \hat{O}). \]  

(5.2)
This equation has been used to show the connection between the OTO correlator and operator entanglement entropy [28–30].

Using the above identity, we are going to rewrite $d^2(t)$ in terms of spreading of local operator in real space. For the pure state under unitary time evolution, the reduced density matrix for region $A$ (after tracing out region $B$) takes the form,

$$\hat{\rho}(t) = \text{Tr}_B \left[ \hat{U}(t) \hat{\psi} \langle \hat{\psi} | \hat{U}^\dagger(t) \right].$$

The difference of the time dependent density matrices in region $A$ is $\hat{\rho}_1(t) - \hat{\rho}_2(t) = \hat{U}(t) \hat{V} \hat{U}^\dagger(t)$, where $\hat{V} = |\psi_1\rangle \langle \psi_1| - |\psi_2\rangle \langle \psi_2|$. Therefore the time dependent distance is

$$d^2(t) = \text{Tr}_A \left[ \text{Tr}_B \hat{U}(t) \hat{V} \hat{U}^\dagger(t) \right]^2 = \frac{1}{|B|} \sum_i \text{Tr} \left[ \hat{B}_i(t) \hat{V} \hat{B}_i(t) \hat{V} \right]$$

(5.4)

where $\hat{B}_i(t) = \hat{U}^\dagger(t) \hat{B}_i \hat{U}$. The above equation is obtained by using the operator identity in equation (5.2) and is also diagrammatically interpreted in figure 7.

By plugging the explicit form of $\hat{V}$ into equation (5.4), we can expand $d^2(t)$ as

$$d^2(t) = \frac{1}{|B|} \sum_i \left[ \langle \psi_1 | \hat{B}_i(t) | \psi_1 \rangle^2 + \langle \psi_2 | \hat{B}_i(t) | \psi_2 \rangle^2 - 2 \langle \psi_1 | \hat{B}_i(t) | \psi_2 \rangle \langle \psi_2 | \hat{B}_i(t) | \psi_1 \rangle \right].$$

(5.5)

Taking $|\psi_1\rangle$ as a Page state, we can use its properties (equations (A.16) and (A.17) in appendix A) and obtain the following expression,
Measuring the distance between quantum many-body wave functions

\[ d^2(t) = \frac{2}{|B|} \sum_i \text{Tr}^2(\hat{B}_i) - \text{Tr}^2(\hat{B}_i \hat{O}(x, -t)) \]

\[ = \frac{2}{|B|} \frac{|H|}{|H| + 1} \left\{ 1 - \frac{1}{|H|^2} \sum_i \text{Tr}^2(\hat{B}_i \hat{O}(x, -t)) \right\}. \tag{5.6} \]

This is one of the most important results in this paper. Notice that this averaged result is state independent and depends only on the spreading of \( \hat{O}(x, -t) \) in the Hilbert space.

Specifically, \( \hat{O}(x, -t) \) can always be expanded in the operator basis as

\[ \hat{O}(x, -t) = \sum_j \alpha_j(t) \hat{H}_j. \tag{5.7} \]

For Hermitian operators on Hermitian basis \( \{ \hat{H}_j \} \), the components \( \alpha_j(t) \) are real numbers and they satisfy the normalization constraint \( \sum_j \alpha_j(t)^2 = 1 \). The trace with the basis \( \{ \hat{B}_i \} \) evaluates the component of \( \hat{O}(x, -t) \) that is completely confined in region \( B \).

Therefore \( d^2(t) \) is proportional to the weights of the operator basis that is not entirely in region \( B \).

The spreading of operator \( \hat{O}(x, -t) \) depends on the behavior of the coefficients \( \alpha_j(t) \) and gives rise to an effective hydrodynamical description in one dimensional chaotic systems \([8, 9, 31, 32]\). As time evolves, the string operators dominating the sum in equation (5.7) grow spatially and we can define the end point distribution to measure the spreading of \( \hat{O}(x, -t) \),

\[ f_R(s, t) \equiv \sum_j \alpha_j(t)^2 \delta(\text{right most site}(\hat{H}_j) = s) \tag{5.8} \]

which counts all the components of the basis that ends at site \( s \) at time \( t \). It is subjected to the normalization

\[ \int_0^L f_R(s, t) ds = 1 \tag{5.9} \]

and can be interpreted as the (coarse grained) probability distribution of operators in \( \hat{O}(-x, t) \) that ends at location \( s \) at time \( t \). One can similarly define the left end distribution \( f_L(s, t) \).
Initially, the local operator $\hat{O}(x)$ only occupies one operator basis on one site. In early time, almost all operators contained in $\hat{O}(x, -t)$ are confined in region $B$, hence $d^2(t)$ is small. Under the chaotic evolution, the end points distribution should move outwards on the left and right ends. Figure 8 is the schematic of the shape of the distribution at late times. If the right end points of most operators are moving to the right, the distribution $f_R(s, t)$ should have a sharp wave front. Once the wave front crosses the cut at $y$, there will be an appreciable fraction of operators in $\hat{O}(x, -t)$ that has spread into region $A$, which contributes to the rapid growth of $d^2(t)$. In fact, the area of the wave front cut by the boundary of $A$ and $B$ is proportional to $d^2(t)$,

$$d^2(t) = \frac{2}{|B|} \int_y^L f_R(s, t) ds \sim \frac{2}{|B|} \int_y^L f_R(s, t) ds. \quad (5.10)$$

After sufficiently long time evolution, we expect that all the string operators in $\hat{O}(x, -t)$ have spanned the entire system. This means that $f_R(s, t)$ is localized at $s = L$ and $d^2(t)$ saturates to

$$d^2(t) \sim \frac{2}{|B|}. \quad (5.11)$$

This is consistent with analytical result of the distance of two independent Page states (see appendix B).

### 5.2. $d^2(t)$ and square of commutator

We would like to compare $d^2(t)$ with another interesting quantity, the square of commutator

$$C(t) = -\frac{1}{|H|} \text{Tr}[\hat{O}(x, -t), \hat{O}'(y)]^2 \quad (5.12)$$

where we take $\hat{O}'(y)$ as a local Pauli operator sitting at position $y$ that does not evolve. We will just take it to be $\hat{\sigma}^x$. In the operator basis we choose, only those do not commute with $\hat{O}'(y)$ will contribute to $C(t)$, i.e.

$$C(t) = -\sum_j \alpha_j^2(t) \frac{1}{|H|} \text{Tr}[\hat{H}_j, \hat{\sigma}^x(y)]^2 = 4 \sum_j \alpha_j^2(t) \delta(\hat{H}_j|_y = \hat{\sigma}^y) \quad (5.13)$$

Under the chaotic evolution, it is reasonable to assume that the probability of having $\hat{\sigma}^x, y, z$ and $\hat{I}$ are essentially the same$^4$. As a result, roughly half (in total weight) of the operators going across the position $y$ will contribute and hence

$$C(t) = \frac{1}{2} 4 \sum_j \alpha_j^2(t) \delta(\text{right most site} \geq y) = 2 \int_y^L ds f_R(s, t). \quad (5.14)$$

---

$^4$ This assumption may not be valid in the presence of conservation law (see figure 9(c)). Moreover, it may also break down in the chaotic system which weakly breaks integrability. All these problems deserve to be further explored in the future.

https://doi.org/10.1088/1742-5468/aace1f
We therefore show that $d^2(t)$ is essentially the rescaled version of $C(t)$ with the same distance $|y - x|$ in these chaotic systems

$$d^2(t) = \frac{1}{|B|} C(t).$$  \hfill (5.15)

The numerical verification of the collapse is shown in figures 9 and 10(a). Notice that $C_s(t)$ we compute is averaged over Page ensemble,

$$C_s(t) = -\langle \psi | [\hat{O}(x, -t), \hat{O}'(y)]^2 | \psi \rangle$$ \hfill (5.16)

which is equivalent to $C(t)$ defined in equation (5.12) (using the identity equation (A.16)). The collapse is an evidence that both $d^2(t)$ and $C(t)$ measure the area of the wave front in the chaotic many-body systems.

One interesting question is: how does $d^2(t)$ or $C(t)$ grow with the time? Based on the arguments above, we can see that it depends on the shape of the operator wave front and the way it propagates, which is ultimately determined by the form of the interaction. For spin-$\frac{1}{2}$ chain with local interaction, we do not find an obvious exponentially growing regime in $d^2(t)$ and $C(t)$. The absence of Lyapunov regime in spin-$1/2$ chain model is also discussed in [33] and is different from the classical chaotic spin chain model where we effectively have $S \to \infty$ [34]. This is consistent with recent study in the random tensor network with local unitary time evolution, where $f_R(s,t)$ in the coarse grained picture is the probability distribution generated by a biased random walk [8, 9]. The distribution $f_R(s,t)$ is a moving Gaussian packet with the width of front broadened diffusively and there is no clear exponentially growing regime in $C(t)$.

However, non-local power-law interaction changes the story. As shown in figures 9(b) and (d), when the distance between $x$ and $y$ is large, we find a clear exponentially growing regime in both $d^2(t)$ and $C(t)$. This is because the power-law interaction generates a wider wave packet in $f_R(s,t)$. We expect that when $|y - x|$ is large, the front has a near-exponential shape. To verify this point, we investigate $d^2(t)$ and $C(t)$ for different distances $|y - x|$ at the same time $t$. As shown in figure 10(b), the area of the wave front in regime $A$ decreases exponentially as we increase the distance $|y - x|$. The slope seems to be invariant as we increase time $t$.

The near-exponential shape of the wave front is responsible for the exponentially growing regime in $d^2(t)$ or $C(t)$, with the growth rate as the so-called quantum Lyapunov exponent, which is also found in some quantum many-body models in the large $N$ limit.
Figure 9. (a) \( \hat{d}^2(t) \) and \( \mathcal{C}_s(t) \) versus \( t/(2\tau) \) on the semi-log scale for Floquet operator in equation (3.1) with local interaction \( (L_A = 2) \). (b) \( \hat{d}^2(t) \) and \( \mathcal{C}_s(t) \) versus \( t/(2\tau) \) on the semi-log scale for Floquet operator in equation (4.1) with power-local interaction \( (L_A = 2) \). The slope of the curve (for the exponentially growing regime) is 0.244. (c) \( \hat{d}^2(t) \) and \( \mathcal{C}_s(t) \) versus \( t \) on the semi-log scale for the static Hamiltonian in equation (3.5) with local interaction. (d) \( \hat{d}^2(t) \) and \( \mathcal{C}_s(t) \) versus \( t \) on the semi-log scale for the static Hamiltonian in equation (4.3). The slope of the curve (for the exponentially growing regime) is 0.902. In (a),(b) and (d), we take \( \hat{O}(x, t = 0) = \hat{\sigma}^z \) in \( \hat{d}^2(t) \) and \( \hat{O}(x, t = 0) = \hat{O}'(y) = \hat{\sigma}^z \) in \( \mathcal{C}_s(t) \). In (c), for \( \mathcal{C}_s(t) \), we fix \( \hat{O}(x, t = 0) = \hat{\sigma}^z \) and consider \( \hat{O}'(y) = \hat{\sigma}^z \) (red curve) and \( \hat{O}'(y) = \hat{\sigma}^z \) (yellow curve). The difference between these two curves at early time is because \( \sigma^z \) has a larger overlap with energy density operator. The \( \hat{d}^2(t) \) of \( \hat{O}(x, t = 0) = \hat{\sigma}^z \) (dashed black curve) lies in the middle between the red and blue curves.
Measuring the distance between quantum many-body wave functions

17https://doi.org/10.1088/1742-5468/aace1f

J. Stat. Mech.
(2018) 073101

but is absent in the spin-\(\frac{1}{2}\) chain with local interaction where \(N = 2\). It seems that the long-range power-law interaction effectively increases \(N\) and gives rise to this Lyapunov regime. In the future, it would be interesting to have a better understanding on this Lyapunov regime and explore the dependence of Lyapunov exponent on the distance \(|y - x|\) in some random non-local unitary tensor networks.

In most of the results, we use the Page states to compute \(d^2(t)\) for both simple numerical implementation and analytical control. They are however difficult to be generated in experiments. To resolve this issue, we also consider \(d^2(t)\) averaged over thermal states, which in comparison are more physical and much easier to be realized in the experiments. We prepare the thermal states by evolving the RPS states under unitary time evolution. After sufficiently long time evolution, these states will thermalize and can be used as the initial state for \(d^2(t)\). These thermal states share similar properties with the Page states but are subjected to some conservation laws determined by the Hamiltonian. In figure 9, we see that the thermal state average matches well with \(d^2(t)\) of the Page states for a large regime of time with both the Floquet evolution and time independent Hamiltonian. Notably, in spin model with non-local interactions, the thermal state average \(d^2(t)\) also grows exponentially with time, showing the same behavior as Page ensemble.

6. Conclusion

In conclusion, we study Hilbert–Schmidt distance \(d^2(t)\) between reduced density matrices of two many-body wave functions initially differed by a local perturbation. Under unitary time evolution governed by a chaotic Hamiltonian, \(d^2(t)\) grows with time and
eventually saturates to a small constant. In contrast to the exponential divergence of nearby trajectories in classical chaos, the growth of $d^2(t)$ depends on the form of the interactions in the non-integrable models. We find that for spin-1/2 chain model with non-local power-law interaction, $d^2(t)$ has a clear exponential growing regime if the local perturbation is far away from the subsystem. While for spin-1/2 chain with local interaction, $d^2(t)$ always grows slower than the exponential function.

Moreover, we show that after sufficiently long time evolution, $d^2(t)$ will approach a small constant, indicating the thermalization of a small subsystem in a many-body quantum state.

We further use operator scrambling picture to show that $d^2(t)$ is the rescaled area of the wave front of the evolved local perturbation operator. The same interpretation also applies to the square of the commutator $C(t)$ under chaotic evolution. Hence the growth scaling of $d^2(t)$ is determined by the shape of the wave front, which is ultimately determined by the form of interactions. From the collapse of the numerical data of $C(t)$ and $d^2(t)$, we conclude that the exponent in the exponential growing region of $d^2(t)$ is the quantum Lyapunov exponent which was previously observed in some large $N$ models.

Acknowledgments

We acknowledge useful discussion with Thomas Faulkner, David Huse, Yevgeny Bar Lev, Andreas W W Ludwig, Shinsei Ryu and Xueda Wen. XC was supported by a post-doctoral fellowship from the Gordon and Betty Moore Foundation, under the EPiQS initiative, Grant GBMF4304, at the Kavli Institute for Theoretical Physics. TZ was supported by the National Science Foundation under grant numbers DMR-1306011. Cenke Xu is supported by the David and Lucile Packard Foundation, and NSF Grant No. DMR-1151208. We acknowledge support from the Center for Scientific Computing from the CNSI, MRL: an NSF MRSEC (DMR-1121053).

Appendix A. Average over Page ensemble

In this section, we compute the expectation values $\langle \psi | \hat{O} | \psi \rangle$ and $\langle \psi | \hat{O}_1 | \psi \rangle \langle \psi | \hat{O}_2 | \psi \rangle$ over the Page ensemble. For the ease of notation, we will reserve $n$ to denote the complex dimension of the Hilbert space. When comparing the result with main text, $n$ should be identified with $|H|$. 

A.1. Probability measure

The random pure state (Page state) has a probability measure that is uniform in the Hilbert space. Such a uniform measure has $U(n)$ invariance, meaning that the average is invariant under the $U(n)$ action, e.g.

$$\langle \psi | \hat{O} | \psi \rangle = \langle \psi | U^\dagger \hat{O} U | \psi \rangle, \quad \forall U \in U(n).$$ (A.1)
Such invariance can be constructed by choosing complex Gaussian random variables as the coefficients in an orthonormal basis. For vector \( c \in \mathbb{C}^n \), whose component \( c_i \) are sampled from standard complex normal distribution, we have the required invariance,

\[
E(Uc) = Uc \quad \text{and} \quad E(U(c - \mu)(c - \mu)^\dagger U^\dagger) = Uc(c)U^\dagger = cc^\dagger.
\] (A.2)

This is how we generate the Page state in the numerical calculation.

Writing \( c_j = x_{2j-1} + ix_{2j} \), the \( N = 2n \)-dimensional real vector \( x \) has independent real Gaussian random variable as its component. With the normalization constraint, the probability measure is

\[
d\mu \propto \exp\left(-\frac{1}{2} |x|^2 \right) \delta(1 - |x|^2) \prod_i dx_i \propto \delta(1 - |x|^2) \prod_i dx_i
\] (A.3)

which is the uniform measure on the \( 2n - 1 \) dimensional sphere. The expectation value of the state \( \langle \psi | \hat{O} | \psi \rangle \) is a function (polynomial) \( f(x) \) of real and imaginary parts of the elements. Averaging over the random states is given by the integral

\[
\langle \psi | \hat{O} | \psi \rangle = \int_{S^{2n-1}} d\mu f(x).
\] (A.4)

A.2. Moments of the probability measure

The \( \langle \psi | \hat{O} | \psi \rangle \) and \( \langle \psi | \hat{O}_1 | \psi \rangle \langle \psi | \hat{O}_2 | \psi \rangle \) can be converted to the average over the coefficients of the Page state, e.g.

\[
\langle \psi | \hat{O} | \psi \rangle = O_{ij}c_i^\dagger c_j = O_{ij}(x_{2i-1} - ix_{2i})(x_{2j-1} + ix_{2j})
\]

\[
\langle \psi | \hat{O}_1 | \psi \rangle \langle \psi | \hat{O}_2 | \psi \rangle = O_{ij}^1 O_{kl}^2 (x_{2i-1} - ix_{2i})(x_{2j-1} + ix_{2j})(x_{2k-1} - ix_{2k})(x_{2l-1} + ix_{2l})
\] (A.5)

hence we need up to 4th order moments of \( x \).

By the reflection symmetry, there are only 3 types of moments that need to be computed,

\[
A_1 = \int d\mu x_i^2, \quad \int d\mu x_i^2 x_j^2 = \begin{cases} A_2 & i = j \\ B_2 & i \neq j \end{cases}
\] (A.6)

The 2nd moment

\[
A_1 = \int d\mu x_i^2 = \frac{1}{N} \int d\mu |x|^2 = \frac{1}{N}
\] (A.7)

is be obtained by \( O(N) \) symmetry, which also relates \( A_2 \) and \( B_2 \),

\[
1 = \int d\mu |x|^2 \times |x|^2 = NA_2 + N(N - 1)B_2.
\] (A.8)

So we only need \( A_2 \) or \( B_2 \).

We first compute the normalization constant, which is proportional to the area of sphere

\[
https://doi.org/10.1088/1742-5468/aace1f
\]
\[
\int_{\mathbb{R}^N} \delta(1 - |\mathbf{x}|^2) \prod_i dx_i = \int_0^\infty r^{N-1} dr \delta(1 - r^2) S_{N-1} = \frac{1}{2} S_{N-1} \int_0^\infty \delta(1 - r^2) r^{N-2} dr = \frac{1}{2} S_{N-1}. \tag{A.9}
\]

With this we can compute \(A_2\) with the remaining \(O(N-1)\) symmetry,
\[
A_2 = \frac{2}{S_{N-1}} \int x_i^2 \delta(1 - |\mathbf{x}|^2) \prod_i dx_i = \frac{S_{N-2}}{S_{N-1}} \int_0^1 dx x^\frac{2}{N-1} = \frac{S_{N-2} \sqrt{\pi}}{S_{N-1}} \frac{\Gamma(\frac{N-1}{2})}{\Gamma(\frac{N}{2} + 1)}
\]
\[
= \frac{3}{4} \frac{\Gamma(\frac{N}{2})}{\Gamma(\frac{N}{2} + 2)} = \frac{1}{N(N+2)} \quad \Rightarrow \quad B_2 = \frac{1}{N(N+2)} \tag{A.10}.
\]

In fact, the general pure \(2k\) moments can be computed similarly
\[
A_{2k} = \frac{2}{S_{N-1}} \int x_i^{2k} \delta(1 - |\mathbf{x}|^2) \prod_i dx_i = \frac{(2k-1)!!}{(N+2k-2)(N+2k-4)\cdots N}. \tag{A.11}
\]

Now we apply these results to the quantity we need to compute in equation (A.5), where \(N = 2n\). We have the 2nd moments
\[
\overline{c_i^* c_j} = (x_{2i-1} - ix_{2i})(x_{2j-1} + ix_{2j}) = \delta_{ij} \frac{2}{N} = \frac{1}{n} \delta_{ij}. \tag{A.12}
\]

The general 4th moments reduces to three types of contractions
\[
\overline{c_i^* c_j c_k^* c_l^*} = (\delta_{ij} \delta_{kl} + \delta_{il} \delta_{jk}) \overline{c_i^* c_k^* c_l^*} (1 - \delta_{ik}) + \delta_{ik} \delta_{jl} \overline{c_i^* c_j c_l^*} (1 - \delta_{ij}) + \delta_{ij} \delta_{jk} \delta_{kl} \overline{c_i^* c_j^* c_k^*} \tag{A.13}
\]
where each of them can be represented by \(A_2\) and \(B_2\)
\[
\overline{c_i^* c_k^* c_k c_i^*} (1 - \delta_{ik}) = (1 - \delta_{ik})(x_{2i-1}^2 + x_{2i}^2)(x_{2k-1}^2 + x_{2k}^2)
\]
\[
= (1 - \delta_{ik}) 4 B_2 = (1 - \delta_{ik}) \frac{1}{n(n+1)}
\]
\[
\overline{c_i^* c_j^* c_j c_i^*} (1 - \delta_{ij}) = (1 - \delta_{ij})(x_{2i-1}^2 - x_{2i}^2 - 2ix_{2i-1}x_{2i})(x_{2j-1}^2 - x_{2j}^2 + 2ix_{2j-1}x_{2j})
\]
\[
= (1 - \delta_{ij})(2B_2 - 2B_2) = 0
\]
\[
\overline{c_i^* c_j^* c_i^* c_l^*} = (x_{2i-1}^2 + x_{2i}^2)^2 = \frac{2}{n(n+1)}. \tag{A.14}
\]

Therefore we have ‘Wick theorem’ for the moments of \(c_i\)
\[
\overline{c_i^* c_j c_k c_l} = (\delta_{ij} \delta_{kl} + \delta_{il} \delta_{jk}) \frac{1}{n(n+1)} (1 - \delta_{ik}) + \frac{2}{n(n+1)} \delta_{ik} \delta_{jl} \delta_{ij}
\]
\[
= (\delta_{ij} \delta_{kl} + \delta_{il} \delta_{jk}) \frac{1}{n(n+1)}. \tag{A.15}
\]
A.3. Average of expectation values

We apply the results of the previous sections to the average of expectation values. For operator $\hat{O}$, we have

$$\langle \psi | \hat{O} | \psi \rangle = O_{ij} c_i^* c_j = \frac{1}{n} \delta_{ij} O^{ij} = \frac{1}{n} \text{Tr}(\hat{O}).$$

(A.16)

For two operators $\hat{O}_1$ and $\hat{O}_2$, we have

$$\langle \psi | \hat{O}_1 | \psi \rangle \langle \psi | \hat{O}_2 | \psi \rangle = O_{ij}^{1} O_{kl}^{2} c_i^* c_j c_k^* c_l = \frac{1}{n(n+1)} O_{ij}^{1} O_{kl}^{2} (\delta_{ij} \delta_{kl} + \delta_{il} \delta_{jk})$$

$$= \frac{1}{n(n+1)} [\text{Tr}(\hat{O}_1 \hat{O}_2) + \text{Tr}(\hat{O}_1) \text{Tr}(\hat{O}_2)]$$

(A.17)

where $n$ is the Hilbert space dimension $|H|$ in the main text.

Appendix B. Averaged distance of two independent Page states

As discussed in appendix A, the random pure state (Page state) can be constructed by assigning complex Gaussian random variables as the coefficients in any orthonormal basis.

We divide the Hilbert space into subsystem $A$ with dimension $n$ and $B$ with dimension $m \geq n$. In the same orthonormal basis, we can obtain the decomposition coefficients $\psi_{ij}$ of the wavefunction on the subsystem basis. Call this $n \times m$ matrix $Y$, the reduced density matrix is $\hat{\rho}_A = YY^\dagger$. The reduced density matrix consists of products of Gaussian random variables, and thus is a Wishart matrix with unit trace constraint.

The element of $Y$ matrix satisfies a Gaussian probability distribution $[35]$

$$P(\{Y_{ij}\}) \sim \exp \left[ \frac{\beta}{2} \text{Tr} YY^\dagger \right]$$

(B.1)

where the Dyson index $\beta = 2$ in this case. Therefore the eigenvalue of $YY^\dagger$ scales as $n$, and the (rescaled) spectral density is defined as

$$\varrho(x, \rho) = \frac{1}{n} \sum_i \delta(x - \frac{\lambda_i}{n}).$$

(B.2)

In the large $n$ limit (keeping the ratio $\alpha = \frac{n}{m}$ fixed), the spectral density reduces to the Marchenko–Pastur distribution with parameters

$$\varrho_w(x) = \frac{1}{2\pi x} \sqrt{(\alpha_+ - x)(x - \alpha_-)}$$

(B.3)

where $\alpha_\pm = (1 \pm \sqrt{\alpha})^2$.

Wishart matrix with the trace constraint is called fixed trace Wishart–Laguerre ensemble. Due to the constraint, the eigenvalues scales as $\frac{1}{n}$, so we shall instead use $[21, 35]$
\[ \rho(x, \rho) = \frac{1}{n} \sum_i \delta(x - n\lambda_i). \]  

The corresponding asymptotic distribution is

\[ \varrho_{w, u}(x) = \frac{1}{2\pi \alpha x} \sqrt{(\alpha_+ - x)(x - \alpha_-)} \]  

it is different from the one without the trace constraint.

The distance we want to compute is the second moments of the eigenvalues of \( \hat{\rho}_1 - \hat{\rho}_2 \)

\[ d^2(t) = \text{Tr}(\hat{\rho}_1 - \hat{\rho}_2)^2 = \sum_i \lambda_i^2 \]  

which is the distribution of the difference of two independent variables sampled from the Marchenko–Pastur distribution. Because of the constraint \( \sum_i \lambda_i = 0 \), the eigenvalues also scale as \( \frac{1}{n} \),

\[ \rho(x, \hat{\rho}_1 - \hat{\rho}_2) = \frac{1}{n} \sum_i \delta(x - n\lambda_i). \]  

Taking a function \( f(x) \), the average of the sum \( \sum_i f(\lambda_i) \) can be carried over the spectral density function

\[ \frac{1}{n} \sum_{i=1}^{n} f(\lambda_i) = \int f(\lambda) \frac{1}{n} \sum_{i=1}^{n} \delta(\lambda - \lambda_i) P(\{\lambda_i\}) \, d\lambda \]
\[ = \int f(\lambda) \frac{1}{n} \sum_{i=1}^{n} \delta(n\lambda - n\lambda_i) P(\{\lambda_i\}) \, dn\lambda \]
\[ = \int f(\frac{x}{n}) \varrho(x, \hat{\rho}_2 - \hat{\rho}_1) \, dx. \]  

Mejía et al [21] computes the asymptotic eigenvalue distribution and its absolute moments

\[ m_z = \int |x|^z \varrho(x) \, dx = \frac{\Gamma(z + 1)(2\alpha)^{\frac{z}{2}}}{\Gamma(\frac{z}{2} + 1)\Gamma(\frac{z}{2} + 2)} \, _2F_1(1 - \frac{z}{2}, -\frac{z}{2}; \frac{3z}{2} + 2; \frac{\alpha}{2}) \]  

where \( _2F_1 \) is the hypergeometric function. Hence

\[ d^2 = n \int \frac{1}{n^2} x^2 \varrho(x) \, dx = \frac{1}{n} m_2 = \frac{1}{n} (2\alpha)_2F_1(0, -1; 3; \frac{\alpha}{2}) \]
\[ = \frac{2\alpha}{n} = \frac{2}{m}. \]  

References

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

[2] Gutzwiller M C 2013 Chaos in Classical and Quantum Mechanics vol 1 (New York: Springer)
[3] Leviatan E, Pollmann F, Bardarson J H, Huse D A and Altman E 2017 Quantum thermalization dynamics with matrix-product states (arXiv:1702.08994)

[4] Larkin A I and Ovchinnikov Yu N 1969 Quasiclassical method in the theory of superconductivity Sov. J. Exp. Theor. Phys. 28 1200

[5] Hastings M B and Koma T 2017 Out-of-time-ordered density correlators in luttinger liquids J. High Energy Phys. 01083

[6] Gong Z-X, Foss-Feig M, Michalakis S and Gorshkov A V 2014 Persistence of locality in systems with power-law interactions Phys. Rev. Lett. 113 030602

[7] Page D N 1993 Average entropy of a subsystem Phys. Rev. Lett. 71 1291–4

[8] von Keyserlingk C W, Pollmann F and Sondhi S L 2018 Operator hydrodynamics, OTOCs, and entanglement growth in systems without conservation laws Phys. Rev. X 8 021013

[9] Nalam A, Vijay S and Haah J 2018 Operator spreading in random unitary circuits Phys. Rev. X 8 021014

[10] Shenker S H and Stanford D 2014 Black holes and the butterfly effect J. High Energy Phys. JHEP03(2014)067

[11] Shenker S H and Stanford D 2014 Multiple shocks J. High Energy Phys. JHEP12(2014)046

[12] Shenker S H and Stanford D 2015 Stringy effects in scrambling J. High Energy Phys. JHEP05(2015)132

[13] Maldacena J and Stanford D 2016 Remarks on the Sachdev–Ye–Kitaev model Phys. Rev. D 94 106002

[14] Kitaev A 2014 Talks Given at the Fundamental Physics Prize Symp. (10 November 2014–12 February 2015) (KITP)

[15] Kitaev A 2015 Talks at KITP (7 April 2015–27 May 2015)

[16] Sachdev S and Ye J 1993 Gapless spin-fluid ground state in a random quantum Heisenberg magnet Phys. Rev. Lett. 70 3339–42

[17] Srednicki M 1999 Chaos and quantum thermalization Phys. Rev. E 50 888–901

[18] Deutsch J M 1991 Quantum statistical mechanics in a closed system Phys. Rev. A 43 2046–9

[19] Kim H, Ikeda T N and Huse D A 2014 Testing whether all eigenstates obey the eigenstate thermalization hypothesis Phys. Rev. E 90 052105

[20] Forrester P J 2010 Log-Gases and Random Matrices (LMS-34) (Princeton, NJ: Princeton University Press)

[21] Mejia J, Zapata C and Botero A 2017 The difference between two random mixed quantum states: exact and asymptotic spectral analysis J. Phys. A: Math. Theor. 50 025301

[22] Calabrese P and Cardy J 2007 Quantum quenches in extended systems J. Stat. Mech. P06008

[23] Bañuls M C, Cirac J I and Hastings M B 2011 Strong and weak thermalization of infinite nonintegrable quantum systems Phys. Rev. Lett. 106 050405

[24] Rigol M, Muramatsu A and Oshamii M 2006 Hard-core bosons on optical superlattices: dynamics and relaxation in the superfluid and insulating regimes Phys. Rev. A 74 053616

[25] Rigol M, Dunjko V, Yurovsky V and Oshamii M 2007 Relaxation in a completely integrable many-body quantum system: an ab initio study of the dynamics of the highly excited states of 1d lattice hard-core bosons Phys. Rev. Lett. 98 050405

[26] Cardy J 2016 Quantum quenches to a critical point in one dimension: some further results J. Stat. Mech. 023103

[27] Dóra B and Moessner R 2017 Out-of-time-ordered density correlators in luttinger liquids Phys. Rev. Lett. 119 026802

[28] Hosur P, Qi X-L, Roberts D A and Yoshida B 2016 Chaos in quantum channels J. High Energy Phys. JHEP2(2016)004

[29] Fan R, Zhang P, Shen H and Zhai H 2017 Out-of-time-order correlation for many-body localization Sci. Bull. 62 707–11

[30] Harrow A W and Low R A 2009 Random quantum circuits are approximate 2-designs Commun. Math. Phys. 291 257

[31] Khemani V, Vishwanath A and Huse D A 2017 Operator spreading and the emergence of dissipation in unitary dynamics with conservation laws (arXiv:1710.09835)

[32] Rakovszky T, Pollmann F and von Keyserlingk C 2017 Diffusive hydrodynamics of out-of-time-ordered correlators with charge conservation (arXiv:1710.09827)

[33] Luitz D J and Lev Y B 2017 Information propagation in isolated quantum systems Phys. Rev. B 96 020406

[34] de Wijn A S, Hess B and Fine B V 2013 Lyapunov instabilities in lattices of interacting classical spins at infinite temperature J. Phys. A: Math. Theor. 46 254012

[35] Nadal C, Majumdar S N and Vergassola M 2011 Statistical distribution of quantum entanglement for a random bipartite state J. Stat. Phys. 142 403–38

https://doi.org/10.1088/1742-5468/aace1f