Hilbert space average of transition probabilities

Nico Hahn, Thomas Guhr, Daniel Waltner

Fakultät für Physik, Universität Duisburg-Essen, Lotharstraße 1, 47048 Duisburg, Germany.

February 24, 2020

Abstract

The typicality approach and the Hilbert space averaging method as its technical manifestation are important concepts of quantum statistical mechanics. Extensively used for expectation values we extend them in this paper to transition probabilities. In this context we also find that the transition probability of two random uniformly distributed states is connected to the spectral statistics of the considered operator. Furthermore, within our approach we are capable to consider distributions of matrix elements between states, that are not orthogonal. We will demonstrate our quite general result numerically for a kicked spin chain in the integrable resp. chaotic regime.

1. Introduction

Consider a set of pure states \( \{ |\phi\rangle \} \) initially featuring very similar expectation values concerning an observable \( M \). Then dynamical quantum typicality refers to the situation in which the majority of these states will also feature very similar expectation values to the same or another observable at any later time [1]. This statement holds true under quite general conditions as it is not even dependent on the details of the dynamics. A necessary requirement is that the dimension \( N \) of the finite dimensional Hilbert space \( \mathcal{H} \) is large. In [2] the construction of these sets was generalized and sufficient conditions were explored. It was found that besides a large Hilbert space dimension the preset expectation value \( \langle \phi | M | \phi \rangle \) as well as the spectrum of \( M \) play an important role.

The tool to develop these findings is the Hilbert space averaging method [3]. With its aid one can establish the average of a function depending on pure states, e.g. the expectation value of some arbitrary observable. The standard deviation as the square root of the Hilbert space variance describes the width of the distribution of the states around this value and is found to be small if certain conditions are met [2]. Using these methods it was possible to get more insight into the eigenstate thermalization hypothesis (ETH) [4–7], a premise that strives to explain the thermalization of isolated quantum systems. Furthermore the resulting identities have been utilized as a computational scheme, e.g. to calculate spin-correlations [8–11] as well as in the case of out-of-time-ordered correlations (OTOC’s) [12], which are able to diagnose many body localized phases [13]. The intention of the present work is to extend the Hilbert space averaging method and open it up to new possibilities by studying the transition probabilities \( |\langle \chi | A | \psi \rangle|^2 \) of an arbitrary operator \( A \).

On the technical level this averaging procedure is closely connected to the concept of random matrix theory [14,15]. This method substitutes the matrices, relevant for the description of the system (e.g. the Hamiltonian, scattering matrix etc.), by random matrices. It was applied quite successfully to systems sharing a sufficiently high amount of complexity. The only information about the system entering the model are its symmetries, revealing universal system features exclusively depending on the symmetry class. In previous works, where concepts of random matrix theory have been applied to study transition probabilities [16,17], the components of the states were drawn from a Gaussian distribution with zero mean, implying statistical orthogonality of the states in high dimensions. Thus, these states are able to model the eigenstates of an observable and the transition probability between those orthogonal eigenstates. A well known result of this approach is the Porter-Thomas distribution [20,21], which determines the statistics of the transition probabilities in the case of time reversal symmetry. However,
in this paper we follow a new route, allowing us to explicitly choose the overlap of the deployed states, putting statistical orthogonality aside. The situation of non-orthogonal states is relevant, when states, that are not eigenstates of the Hamiltonian or another observable, important for the system description, are prepared. For example within the doorway mechanism (see Refs \[22,23\] and Refs therein) a distinct state (the so-called doorway state), that is not an eigenstate of the Hamiltonian, but is distinguished by another feature (e.g. collective motion in many-body systems) is coupled to a background of states, that are not orthogonal to the doorway state. This mechanism has far reaching applications in nuclear physics (giant multipole resonance) \[24\], molecular physics \[25,26\] and mesoscopic physics \[27,28\].

The outline of the paper is as follows: in section 2 we will introduce the Hilbert space averaging method and give a short review on the concept of quantum typicality, focusing especially on the construction of typical states and on the aforementioned conditions of dynamical quantum typicality. In section 3 we will shortly introduce the model on which we present our findings of section 4, namely an expression for the Hilbert space average of transition probabilities. More involved calculations are relegated to the appendix.

2. Framework

We restrict ourselves to a finite dimensional Hilbert space $\mathcal{H} = \mathbb{C}^N$ with dimension $N$. In practice this will be due to the nature of the system or because of conserved quantities, e.g. the Hilbert space could be the energy shell of an isolated system. We now give an overview of the concept of quantum typicality.

We want to find the average value of a function $f : \mathcal{H} \to \mathbb{C}$. This is the Hilbert space average \[3\]

$$\text{HA} [f(\psi)] = \frac{\int d[\psi] f(\psi) \delta(\langle \psi|\psi \rangle - 1)}{\int d[\psi] \delta(\langle \psi|\psi \rangle - 1)}. \quad (2.1)$$

Thus, we choose normalized vectors as representative for the state, which allows us to restrict our integration to the unit sphere in $\mathcal{H} = \mathbb{C}^N$. Then we normalize this integral by the volume of the unit sphere.

We note that for a normalized vector there is still a relative phase $e^{i\lambda}$ left as a degree of freedom. However, the functions, treated on the next pages, fulfill $\forall \lambda \in [0, 2\pi) : f(\psi) = f(e^{i\lambda}\psi)$ (in other words they are $U(1)$ invariant) and our normalization makes up for that multiple occurrence of identical states in the integral over the sphere $S^{2N-1}$. The latter can be seen as a fibre bundle $S^{2N-1} \sim \mathbb{C}P^{N-1} \times U(1)$, whose base manifold is the complex projective space $\mathbb{C}P^{N-1}$, that is the space of all rays in $\mathbb{C}^N$ \[29\]. Likewise the Hilbert space variance can be defined as

$$\text{HV}[f(\psi)] = \text{HA} [f^2(\psi)] - \text{HA}^2 [f(\psi)]. \quad (2.2)$$

A first important application is the expectation value of an operator $M$. As shown in \[3\] this can be obtained without using the explicit form of (2.1). Because the integration regime is invariant under the action of the unitary group $U(N)$ (this is also known as Haar measure \[29\]) we can choose an arbitrary orthonormal basis $\{|i\rangle\}$ to represent our states and then use the linearity of the Hilbert space average

$$\text{HA} [\langle \psi|M|\psi \rangle] = \sum_{i,j} M_{ij} \text{HA} [\psi_i^* \psi_j] \quad (2.3)$$

with the coordinates $\psi_i = \langle i|\psi \rangle$. The unit sphere is invariant under $\psi_i \to -\psi_i$ for all $i$ and therefore the average will vanish for $i \neq j$. Now we can use the normalization of our states

$$\text{HA} \left[ \sum_i |\psi_i|^2 \right] = \sum_i \text{HA} [|\psi_i|^2] = 1 \quad (2.4)$$

to conclude

$$\text{HA} [|\psi_i|^2] = \frac{1}{N} \quad (2.5)$$

and

$$\text{HA} [\langle \psi|M|\psi \rangle] = \frac{\text{Tr} M}{N}. \quad (2.6)$$

2
The Hilbert space variance of the expectation value is

$$HV[(\psi|M|\psi)] = \frac{1}{N+1} \left( \frac{\text{Tr} M^2}{N} - \frac{\text{Tr}^2 M}{N^2} \right).$$

(2.7)

We observe that the variance is one order lower in $N$ than (2.6) implying sharp expectation values in the limit $N \to \infty$. For a specific $M$ one has to consider its $N$ dependence, but we can assume that its traces are $N$ independent with out loss of generality. The calculation of higher moments of $\langle \psi|M|\psi \rangle$ is more complicated. In appendix A.2 we present a formula, that achieves this for positive integer powers.

Next we will review the typicality results of [2] and choose $M$ as an observable. Reimann constructs a statistical operator

$$\rho(m, M) = \frac{1}{N} \frac{1}{1 + y(m - M)}.$$  

(2.8)

where $m$ is the desired expectation value of $\rho(m, M)$ concerning $M$. The parameter $y \in \mathbb{R}$ is chosen so that the normalization $\text{Tr} \rho = 1$ and $\text{Tr} M \rho = m$ are fulfilled, and is determined by the roots of a rational function. The states

$$|\phi\rangle := \sqrt{N} \rho |\psi\rangle$$

(2.9)

exhibit the desired preset expectation value $m$ as their Hilbert space average

$$\text{HA}[\langle \phi|M|\phi \rangle] = N \text{HA}[\langle \sqrt{\rho} M \sqrt{\rho} |\psi\rangle] = \text{Tr} M \rho = m.$$  

(2.10)

They are not necessarily normalized, but in average they are

$$\text{HA}[\langle \phi|\phi \rangle] = N \text{HA}[\langle \sqrt{\rho} |\psi\rangle] = \text{Tr} \rho = 1.$$  

(2.11)

The Hilbert space variance indicates how much one can trust these results in the case of an individual $|\phi\rangle$. The purity

$$P = \text{Tr} \rho^2$$  

(2.12)

is an upper bound for the variances

$$HV[\langle \phi|\phi \rangle] = \frac{NP - 1}{N+1} \leq P \quad \text{and} \quad HV[\langle \phi|M|\phi \rangle] \leq \|M\|^2 P.$$  

(2.13)

where the operator norm $\|M\|$ is defined as the absolute value of the largest eigenvalue in modulus. Therefore one achieves typicality if the purity is low $P \ll 1$. While a large Hilbert space is a prerequisite for low purity, due to (2.8) $P$ is also strongly dependent on the observable $M$ and the expectation value $m$. As an example consider the magnetization in $z$-direction

$$M_z = \sum_{i=1}^{n} \sigma^z_i$$

(2.14)

of an $n$-particle spin-1/2 chain (c.f. section 3 and 4). In this case the eigenvalues of the local spin observables $\sigma^z_i$ are $\pm 1$. Therefore the eigenvalue density of $M_z$ is a binomial distribution, i.e. there are only few eigenstates near the maximal resp. minimal magnetization of $\pm n$, while the majority is located around a magnetization of zero in the center. In fact, there is only a single state that has the maximal resp. minimal magnetization as its expectation value, making it highly untypical. This leads to a high purity of statistical operators with expectation values near those values. It is exactly the other way around for statistical operators with Tr $\rho M_z \approx 0$.

Finally we note that the above equations remain valid if $M$ is replaced by any other observable $B(t)$, that does not fulfill (2.10). Thus dynamical quantum typicality is retrieved, namely that states $\{|\phi\rangle\}$ initially featuring $\langle \phi|M|\phi \rangle \approx m$ with small variance will also center around a common expectation value for another observable $B$ with small variance. As the deployed observables enter only in terms of their operator norm, the upper bound of the variance (2.13) is time independent for unitary time evolution.

We want to give a geometrical interpretation of what is done here, illustrated in Fig. 1 for $N = 3$. We uniformly sampled 100 vectors (in $\mathbb{R}^3$) on the unit sphere and applied $\sqrt{N} \rho$ with random positive eigenvalues of $\rho$, so that Tr $\rho = 1$. Because $\sqrt{N} \rho$ is positive and Hermitian one can comprehend it as the
deformation of the former integration regime (the unit sphere) into an ellipsoid. The new states \{|\phi\rangle\} are still uniformly distributed, but now over the ellipsoid. This ellipsoid is then described by the quadric

$$\langle \psi | (N \rho)^{-1} | \psi \rangle - 1 = 0.$$  

A subsequent normalization of each state will generate a nonuniform probability distribution on the unit sphere, which is given by the norm of the states on the surface of the ellipsoid

$$p(\psi) := \|\sqrt{N \rho} | \psi \rangle\|.$$  

In practice we interpolated our data \cite{30} to generate the density plot shown in the figure.

We want to make a short remark about (2.9). Inserting the fully mixed state \(\rho_m = 1/N\) (the one with the lowest purity) yields the uniform distribution from which we drew the states \(|\psi\rangle\).

3. Kicked Ising Chain

Now we introduce the model, on which we want to present the findings of the next section. The kicked Ising chain (KIC) is a many-body quantum system that entails rich dynamics, reaching from integrable to fully chaotic. It consists of a closed ring of \(n\) spins (here spin-1/2 thus the Hilbert space dimension is \(N = 2^n\)) with nearest neighbour interaction and an on site magnetic field. Additionally the system is periodically kicked by a transverse magnetic field leading to a discretized time evolution. The Hamiltonian is thus a sum of two parts

$$H = H_I + H_K \sum_{\tau=-\infty}^{\infty} \delta(t - \tau).$$  

(3.1)

The interaction part is

$$H_I = \sum_{i=1}^{n} (J \sigma_i^z \sigma_{i+1}^z + h \sigma_i^x),$$  

(3.2)

where \(J\) is the coupling strength and \(h\) the on site magnetic field. The kick part is

$$H_K = \sum_{i=1}^{n} b \sigma_i^x$$  

(3.3)

with the transverse magnetic field \(b\). In the present case of a spin-1/2 chain, \(\sigma_i^\alpha\) with \(\alpha \in \{x, y, z\}\) are the Pauli matrices acting at site \(i\). Due to the boundary condition we have \(\sigma_i^\alpha = \sigma_{i+n}^\alpha\). The time evolution operator for one period (Floquet operator) also splits in two parts and is given by

$$U = T \exp \left[ -i \int_0^1 H(t) dt \right] = U_I U_K$$  

(3.4)

with

$$U_I = \exp[-i H_I], \quad U_K = \exp[-i H_K]$$  

(3.5)

and the time ordering operator \(T\). Here we use units in which \(\hbar = 1\). The system becomes integrable for \(b = 0\) as the Hamiltonian will be diagonal in the basis of \(\sigma^z\) and also if \(h = 0\). In the latter case it can be mapped onto a chain of non-interacting spinless fermions via a Jordan-Wigner transformation \cite{31,32}.
In quantum chaos the spectral statistics of the operators governing the dynamics, i.e. the Hamiltonian and the time evolution operator, is used to identify regular and chaotic behaviour. In this context the spectral form factor
\[ K(T) = \frac{\left| \text{Tr} \ U^T \right|^2}{N} \]
(3.6)
is a popular quantity. In recent years the KIC form factor was well examined. We will reconsider it in the next section, where we find a connection between spectral statistics and the Hilbert space average.

4. Hilbert space average of transition probabilities

We now turn to transition probabilities describing the probability to find \( |\psi\rangle \) in \( |\chi\rangle \) after applying \( A \). Assuming again that \( \{|\psi\rangle\} \) is random uniformly distributed the Hilbert space average evaluates to
\[ \text{HA}_{\psi} \left[ |\langle \chi | A | \psi \rangle|^2 \right] = \text{HA}_{\chi} \left[ \frac{\langle \chi | A A^\dagger | \chi \rangle}{N} \right] = \frac{\langle \chi | A A^\dagger | \chi \rangle}{N}. \]
(4.1)

Here we used (2.6) and henceforth we indicate which state will be averaged by a subscript if ambiguities cannot be excluded. The order of \( A \) and its adjoint operator on the right hand side changes if \( |\psi\rangle \) and \( |\chi\rangle \) on the left hand side are interchanged. This can be disregarded if \( A \) commutes with its adjoint which is the case for Hermitian or unitary operators.

To understand this result, it helps to consider, for the time being, a unitary operator \( A \), which does not change the relative orientation of the states and thus has no effect on their distribution. Then the average probability to find \( |\psi\rangle \) parallel to one of the \( N \) axes of the vector space is
\[ \text{HA}_{\psi} \left[ |\langle \chi | A | \psi \rangle|^2 \right] = \frac{1}{N}, \]
(4.2)
leading to their statistical orthogonality in the case of large \( N \). Returning to a general operator \( A \) we can average the remaining state \( |\chi\rangle \) and arrive at
\[ \text{HA}_{\psi,\chi} \left[ |\langle \chi | A | \psi \rangle|^2 \right] = \text{HA}_{\chi} \left[ \frac{\langle \chi | A A^\dagger | \chi \rangle}{N} \right] = \frac{\text{Tr} A A^\dagger}{N^2}. \]
(4.3)

Endowing \( |\psi\rangle \) as well as \( |\chi\rangle \) with arbitrary statistical operators like in (2.9) we find
\[ \text{HA}_{\psi,\chi} \left[ |\langle \chi | \sqrt{N\rho} A \sqrt{N\rho} | \psi \rangle|^2 \right] = \text{Tr} \rho' A\rho A^\dagger \]
(4.4)
and especially for \( \rho' = \rho \) and \( A = I \)
\[ \text{HA}_{\psi,\chi} \left[ |\langle \chi | N\rho | \psi \rangle|^2 \right] = P. \]
(4.5)

Comparing (4.2) and (4.5) this motivates the effective dimension
\[ d_{\text{eff}} := \frac{1}{P} \]
(4.6)
used by various authors, which describes the variety of different states appearing in the mixture \( \rho \). Applying (2.7) for the corresponding Hilbert space variances yields
\[ \text{HV}_{\psi} \left[ |\langle \chi | A | \psi \rangle|^2 \right] = \frac{N - 1}{N^2(N+1)} \langle \chi | A A^\dagger | \chi \rangle^2 \]
(4.7)
\[ \text{HV}_{\psi,\chi} \left[ |\langle \chi | A | \psi \rangle|^2 \right] = \frac{N - 1}{N^3(N+1)^2} \left( \text{Tr} (A A^\dagger)^2 + \text{Tr}^2 A A^\dagger \right), \]
which rapidly decrease in \( N \).
4.1. Transition probability of uniformly distributed states

However, now we want to impose a further condition onto \(|\psi\rangle\) and \(|\chi\rangle\) by fixing their overlap \(z = \langle\chi|\psi\rangle\). A more suitable way to describe their relation would be the angle \(\theta = \arccos|\langle\chi|\psi\rangle|\) as it would live up to the spherical nature of their habitat, but the first option appears computationally to be much easier to handle. But later on we will use the angle to parametrize our numerics. The detailed derivation of the following results is found in appendix A.1. We find

\[
\text{HA}_\psi \left[|\langle\chi|A|\psi\rangle|^2 \delta(\langle\chi|\psi\rangle - z)\right] = \frac{1 - |z|^2}{N - 1} \langle\chi|AA^\dagger|\chi\rangle + \frac{N|z|^2 - 1}{N - 1} |\langle\chi|A|\chi\rangle|^2
\]  

(4.8)

with the variance

\[
\text{HV}_\psi \left[|\langle\chi|A|\psi\rangle|^2 \delta(\langle\chi|\psi\rangle - z)\right] = \text{HA}_\psi \left[|\langle\chi|A|\psi\rangle|^4 \delta(\langle\chi|\psi\rangle - z)\right] - \text{HA}_\psi \left[|\langle\chi|A|\psi\rangle|^2 \delta(\langle\chi|\psi\rangle - z)\right]^2
\]

\[
= \lambda_1 (N,|z|) \langle\chi|AA^\dagger|\chi\rangle^2 + \lambda_2 (N,|z|) \langle\chi|AA^\dagger|\chi\rangle \langle\chi|A^\dagger A|\chi\rangle^2
\]

\[
+ \lambda_3 (N,|z|) |\langle\chi|A|\chi\rangle|^4.
\]

(4.9)

Here the \(\delta\)-function of a complex argument has to be understood in terms of a product of two \(\delta\)-functions concerning the real and imaginary parts. The prefactors

\[
\lambda_1 (N,|z|) = \frac{2 \left(1 - |z|^2\right)^2}{N(N - 1)} - \frac{(1 - |z|^2)^2}{(N - 1)^2}
\]

(4.10)

\[
\lambda_2 (N,|z|) = \frac{2 \left(1 - |z|^2\right)^2}{(N - 1)^2} + \frac{2 \left(1 - |z|^2\right) |z|^2}{N - 1} - \frac{4 \left(1 - |z|^2\right)^2}{N(N - 1)}
\]

\[
\lambda_3 (N,|z|) = \frac{2 \left(1 - |z|^2\right)^2}{N(N - 1)} - \frac{(1 - |z|^2)^2}{(N - 1)^2} - \frac{2 \left(1 - |z|^2\right) |z|^2}{N - 1}
\]

are rational functions in \(N\) and \(|z|\) that go as \(\lambda_1 \sim \mathcal{O}(N^{-2})\) resp. \(\lambda_2, \lambda_3 \sim \mathcal{O}(N^{-1})\). This is important because we want the variance to decrease for increasing Hilbert space dimensions for later typicality observations. We note that (4.8) and (4.9) only depend on \(|z|^2\), i.e. the phase relation of \(|\psi\rangle\) and \(|\chi\rangle\) does not matter as it is expected for the transition probability.

While the second term in (4.9) is the square of (4.8), the first term is the second moment of \(|\langle\chi|A|\psi\rangle|^2\) given by

\[
\text{HA}_\psi \left[|\langle\chi|A|\psi\rangle|^4 \delta(\langle\chi|\psi\rangle - z)\right] = \frac{2 \left(1 - |z|^2\right)^2}{N(N - 1)} \langle\chi|AA^\dagger|\chi\rangle^2
\]

\[
+ \left[ \frac{2 \left(1 - |z|^2\right)^2}{N(N - 1)} - \frac{4 \left(1 - |z|^2\right) |z|^2}{N - 1} + |z|^4 \right] \langle\chi|A|\chi\rangle^4
\]

\[
- \left[ \frac{4 \left(1 - |z|^2\right)^2}{N(N - 1)} - \frac{4 \left(1 - |z|^2\right) |z|^2}{N - 1} \right] \langle\chi|AA^\dagger|\chi\rangle \langle\chi|A^\dagger A|\chi\rangle^2.
\]

(4.11)

It is easy to check that the special cases give the correct results. Choosing \(A = \mathbb{I}\) yields

\[
\text{HA}_\psi \left[|\langle\chi|\psi\rangle|^2 \delta(\langle\chi|\psi\rangle - z)\right] = |z|^2
\]

(4.12)

\[
\text{HV}_\psi \left[|\langle\chi|\psi\rangle|^2 \delta(\langle\chi|\psi\rangle - z)\right] = 0.
\]

Thus, the \(\delta\)-function acts as expected and the average yields the desired overlap of \(|\chi\rangle\) with \(|\psi\rangle\), whereas the variance vanishes. This also yields the relation

\[
\lambda_1 (N,|z|) + \lambda_2 (N,|z|) + \lambda_3 (N,|z|) = 0.
\]

(4.13)
Furthermore we find
\[
\text{HA}_{\psi} \left[ \langle |A|\psi \rangle^2 \delta(\langle \psi \rangle - 1) \right] = \langle |A|\psi \rangle^2 
\]
because by setting \( z = 1 \) (or any other complex number with unit magnitude) we fix \( |\psi\rangle \) along \( |\chi\rangle \). The set of states parallel to \( |\chi\rangle \) contains only one element (that being \( |\chi\rangle \) itself) which explains the variance of zero.

Averaging over the remaining state yields
\[
\text{HA}_{\psi,\chi} \left[ \langle |A|\psi \rangle^2 \delta(\langle \psi \rangle - z) \right] = \frac{N}{N-1} \text{Tr} AA^\dagger + \frac{N|z|^2 - 1}{N^2 - N} |\text{Tr} A|^2. \tag{4.15}
\]
The appropriate variance is
\[
\text{HV}_{\psi,\chi} \left[ \langle |A|\psi \rangle^2 \delta(\langle \psi \rangle - z) \right] = \text{HA}_{\psi,\chi} \left[ \langle |A|\psi \rangle^4 \delta(\langle \psi \rangle - z) \right] - \text{HA}_{\psi,\chi} \left[ \langle |A|\psi \rangle^2 \delta(\langle \psi \rangle - z) \right]. \tag{4.16}
\]
The first term is again the second moment of \( \langle |A|\psi \rangle^2 \) as in (4.11). In a moment we will provide an explicit expression for unitary \( A \), which is the case we are considering now. We want to test these results by using the Floquet operator of the KIC, \( A = U \). Using its unitarity, equations (4.8) and (4.9), which are valid in the case of a fixed state \( |\chi\rangle \), yield
\[
\text{HA}_{\psi} \left[ \langle |U|\psi \rangle^2 \delta(\langle \psi \rangle - z) \right] = \frac{1 - |z|^2}{N-1} + \frac{N|z|^2 - 1}{N-1} \langle |U|\chi \rangle^2 \tag{4.17}
\]
and
\[
\text{HV}_{\psi} \left[ \langle |U|\psi \rangle^2 \delta(\langle \psi \rangle - z) \right] = \lambda_1 (N, |z|) + \lambda_2 (N, |z|) \langle |U|\chi \rangle^2 + \lambda_3 (N, |z|) \langle |U|\chi \rangle^4. \tag{4.18}
\]

In Fig. 2 we illustrate (4.17) and (4.18). We chose \( |\chi\rangle \) as a random sample from the uniform distribution. In the following we present a short protocol on how we find the states \( |\psi\rangle \), that fulfill \( z = \langle \chi | \psi \rangle \), but else are random. Let \( |\xi\rangle \) be uniformly distributed, then
\[
|\chi\rangle_{\perp} = \frac{|\xi\rangle - \langle \chi | \xi \rangle |\chi\rangle}{\sqrt{1 - \langle \chi |\xi\rangle^2}}. \tag{4.19}
\]
is orthogonal to $|\chi\rangle$ and normalized. Thus $|\psi\rangle$ is in dependence of $\theta = \arccos |z|$ resp. $|z|$ 

$$|\psi\rangle = \cos\theta |\chi\rangle + \sin\theta |\chi\rangle_\perp \quad (4.20)$$

$$= |z| |\chi\rangle + \sqrt{1 - |z|^2} |\chi\rangle_\perp.$$

It should be clarified, that unlike $|\chi\rangle$, $|\chi\rangle_\perp$ is not fixed as it is dependent on the fully random state $|\xi\rangle$ from above. Furthermore we used the KIC Floquet operator in the chaotic regime ($J = b = \pi/4$, $h = \pi/5$).

The most notable thing one can infer from Fig. 2 is the increase in variance with growing angle $\theta$. The case where both states are identical and which we already discussed in (4.14) corresponds to $\theta = 0$ resp. $|z| = 1$. Here the variance vanishes. The maximum variance is reached at $\theta = \pi/2$, where the states are orthogonal. At this value the measure of accessible states $|\psi\rangle$, that fulfill $|\langle\chi|\psi\rangle| = |z|$, is also maximized. Moreover we detect some slightly untypical states, that do not lie in the standard deviation. From the Hilbert space dimensions $2^8$ (left) and $2^{10}$ (right) we can see, when noting the scales in Fig. 2, a decrease in variance with larger dimension.

Continuing with an average over both states we arrive at (4.15) for the Floquet operator 

$$\mathcal{H}_A^{\psi,\chi} \left[ \langle \langle \chi|U|\psi\rangle \rangle^2 \delta(\langle \chi|\psi\rangle - z) \right] = \frac{N - |z|^2}{N^2 - 1} + \frac{N |z|^2 - 1}{N^3 - N} |\text{Tr} U|^2, \quad (4.21)$$

where the spectral form factor $K(1) = |\text{Tr} U|^2 / N$ emerges. An extension to arbitrary $T$ is straightforward. The Hilbert space variance (4.16) for the Floquet operator is quite lengthy. Exploiting the unitarity of $U$ we obtain for the second moment (4.11) 

$$\mathcal{H}_A^{\psi,\chi} \left[ \langle \langle \chi|U|\psi\rangle \rangle^4 \delta(\langle \chi|\psi\rangle - z) \right] = \frac{2 \left(1 - |z|^2\right)^2}{N(N - 1)} \quad (4.22)$$

$$+ \left[ \frac{2 \left(1 - |z|^2\right)^2}{N(N - 1)} - \frac{4 \left(1 - |z|^2\right) |z|^2}{N - 1} + |z|^4 \right] \mathcal{H}_A \left[ \langle \langle \chi|U|\chi\rangle \rangle^4 \right]$$

$$+ \left[ \frac{4 \left(1 - |z|^2\right) |z|^2}{N - 1} - \frac{4 \left(1 - |z|^2\right)^2}{N(N - 1)} \right] \mathcal{H}_A \left[ \langle \langle \chi|U|\chi\rangle \rangle^2 \right].$$

Figure 3: Blue line: Analytical result for the Hilbert space average (4.21). Red dots: Transition probabilities for 100 individual realizations. Green diamonds: Their arithmetic mean and the analytical result for the standard deviation according to (4.16) plotted as a tube around the average.
Ne converges to an exponential distribution

We present a detailed derivation of this result is found in appendix A.3. After rescaling which replaces the Porter-Thomas distribution [20, 21] in the case of broken time reversal invariance.

We observe the interesting feature that the Hilbert space average of the transition probabilities are fully determined by the spectral statistics of the deployed operator. However, this is not possible for the variance, because of the term \( \text{Re} \left( \text{Tr} U^2 \text{Tr} U^\dagger \right) \) occurring in \( \text{HA}_{\psi,\chi} \left[ |\langle \chi | U | \psi \rangle |^2 \delta (|\langle \chi | \psi \rangle | - z) \right] \).

The second term in the Hilbert space variance (4.16) is given by the square of (4.21)

\[
\text{HA}_{\psi,\chi} \left[ |\langle \chi | U | \psi \rangle |^2 \delta (|\langle \chi | \psi \rangle | - z) \right]^2 = \left( \frac{N - |z|^2}{N^2 - 1} \right)^2 \left( \frac{N |z|^2 - 1}{N^3 - N} \right)^2 |\text{Tr} U|^4 \quad (4.24)
\]

\[
+ 2 \left( \frac{N - |z|^2}{N^3 - N} \right) \left( \frac{N |z|^2 - 1}{N^3 - N} \right) |\text{Tr} U|^2.
\]

We observe the interesting feature that the Hilbert space average of the transition probabilities are fully determined by the spectral statistics of the deployed operator. However, this is not possible for the variance, because of the term \( \text{Re} \left( \text{Tr} U^2 \text{Tr} U^\dagger \right) \) occurring in \( \text{HA}_{\psi,\chi} \left[ |\langle \chi | U | \psi \rangle |^4 \right] \).

In Fig. 3 the Hilbert space average is shown for a completely chaotic system \((J = b = \pi/4, h = \pi/5)\) on the left and for an integrable system \((J = h = 0, b = \pi/4)\) on the right. As opposed to the case, in which we kept one state fixed, we do not observe a large dynamics of the transition probability in dependence of \( \theta \). The derivative of (4.21)

\[
\partial_{|z|} \text{HA}_{\psi,\chi} \left[ |\langle \chi | U | \psi \rangle |^2 \delta (|\langle \chi | \psi \rangle | - z) \right] = \frac{2 |z|}{N^2 - 1} (NK(1) - 1) \quad (4.25)
\]

determines this behaviour. We can infer, that for \( K(1) \gg 1/N \) the probability should grow with \( |z| \) and decrease for \( K(1) \ll 1/N \). However, the prefactor \( 1/(N^2 - 1) \) ensures that the derivative is practically zero for large Hilbert space dimensions as can be seen for the chaotic case, shown in the left panel of Fig. 3.

In the integrable case we chose the Floquet operator as \( U = \bigotimes_{j=1}^N \exp \left[ -i \frac{\theta}{N} \pi \sigma^z \right] \), i.e. we rotate every spin by \( \pi/2 \) around the \( x \)-axis, while there is no interaction between the spins. Here a large value for the spectral form factor, \( K(1) = 1 \), is found for all times by straightforward, exact calculation. And indeed we find a positive slope in \( |z| \), as the expression on the right hand side of (4.25) is one order larger in \( N \) than for \( K(1) \ll 1/N \).

Furthermore we study the distribution of transition probabilities \( |\langle \chi | U | \psi \rangle|^2 \) under the secondary condition \( |\langle \chi | \psi \rangle| = z \). Without this condition one finds a distribution of the Kumaraswamy type

\[
p(s) = (N - 1)(1 - s)^{N-2}, \quad (4.26)
\]

which replaces the Porter-Thomas distribution [20, 21] in the case of broken time reversal invariance. We present a detailed derivation of this result is found in appendix A.3. After rescaling \( s \to s/N \) it converges to an exponential distribution \( Ne^{-Ns} \) for large \( N \). Hence, the latter distribution and (4.26) share the same properties of equal mean and standard deviation

\[
\mu = \sqrt{\mu_2} = \frac{1}{N}, \quad (4.27)
\]
and parameter independent skewness and kurtosis
\[ \mu_3 = 2 \quad \mu_4 = 9 \] (4.28)
in this limit. In Fig. 4 we show histograms of the transition probability \( |\langle \chi|U|\psi\rangle|^2 \) for 10000 individual realizations of \( |\chi\rangle \) and \( |\psi\rangle \) and under the condition \( |\langle \chi|\psi\rangle| = 0 \), i.e. \( \theta = \pi/2 \). Once more the left side is obtained from a chaotic Floquet operator \( J = b = \pi/4, h = \pi/5 \) and the right side from the one, that shows no interaction and only rotates every spin \( J = h = 0, b = \pi/4 \). Our numerically obtained distribution are convincingly close to (4.26). This is because the assumption that the states are orthogonal is well fulfilled, because of the high dimensional space. For different values of \( z \) we observe also an exponential distribution, but with deviating parameters \( \mu \), \( \mu_2 \), \( \mu_3 \) and \( \mu_4 \). Thus, the information about the initial \( z \) is erased by the time evolution. However choosing the special case \( U = I \) would yield the distribution \( \delta \left(s - |z|^2\right) \).

4.2. Transition probability of nonuniformly distributed states

Now we want to extend the typicality considerations of [1, 2] introduced in section 2 to the transition probabilities. To be more precise we want to change the distribution of the states according to (2.9).

Again we start with a fixed \( |\chi\rangle \) and only take the average with respect to \( |\psi\rangle \). Henceforth we use the abbreviation \( \Lambda = \sqrt{N\rho} \) and set
\[ |\psi\rangle \to \Lambda |\psi\rangle. \] (4.29)
If the purity of \( \rho \) is low, the states typically stay close to a unit norm. We want to normalize the argument of the delta function \( \delta (\langle \chi|\Lambda|\psi\rangle - z) \). Because \( \Lambda \) is Hermitian, we can normalize it by \( \|\Lambda|\chi\| = \sqrt{|\Lambda|^2|\chi|} \).

Furthermore as the normalization is not dependent on \( |\psi\rangle \), we use the scaling property of the delta function
\[ \delta \left( \frac{\langle \chi|\Lambda|\psi\rangle}{\sqrt{|\Lambda|^2|\chi|}} - z \right) = \sqrt{|\Lambda|^2|\chi|} \delta \left( \langle \chi|\Lambda|\psi\rangle - z\sqrt{|\Lambda|^2|\chi|} \right). \] (4.30)

In this way we extracted the \( \Lambda \)-dependence of the absolute value of \( \langle \chi|\Lambda|\psi\rangle \) and are thus able to retain our choice \( z \in [0,1] \) by rescaling. The square root prefactor will cancel out, because the same is done in the denominator. Therefore we arrive at an integral which is equivalent to (4.8) for nonuniformly distributed \( |\psi\rangle \). The Hilbert space average will be
\[ \text{HA}_{\psi} \left[ |\langle \chi|\Lambda|\psi\rangle|^2 \delta \left( \frac{\langle \chi|\Lambda|\psi\rangle}{\sqrt{|\Lambda|^2|\chi|}} - z \right) \right] = \frac{1 - |z|^2}{N - 1} \langle \chi|\Lambda^2 A^\dagger A |\chi\rangle + \frac{N |z|^2 - 1}{N - 1} \frac{\langle \chi|\Lambda^2 A^\dagger A |\chi\rangle^2}{\langle \chi|\Lambda^2 |\chi\rangle}. \] (4.31)

Once again the reader may consider appendix A.1 for a detailed derivation. The variance is
\[ \text{HV}_{\psi} \left[ \langle \chi|\Lambda A|\psi\rangle^2 \delta \left( \frac{\langle \chi|\Lambda A|\psi\rangle}{\sqrt{|\Lambda|^2|\chi|}} - z \right) \right] = \lambda_1 (N,|z|) \langle \chi|\Lambda A^2 A^\dagger |\chi\rangle^2 \] (4.32)
\[ \quad + \lambda_2 (N,|z|) \langle \chi|\Lambda^2 A^\dagger A |\chi\rangle \frac{\langle \chi|\Lambda^2 |\chi\rangle^2}{\langle \chi|\Lambda^2 |\chi\rangle}, \]
\[ \quad + \lambda_3 (N,|z|) \frac{|\langle \chi|\Lambda A^2 |\chi\rangle|^4}{\langle \chi|\Lambda^2 |\chi\rangle^2}, \]
where the prefactors are the same as in (4.10). It is easy to check that (4.31) and (4.32) are equal to (4.8) and (4.9) for \( \Lambda = I \).

Now we want to average over the remaining state in (4.31). But before we set
\[ |\chi\rangle \to \Lambda'|\chi\rangle \] (4.33)
with \( \Lambda' = \sqrt{N\rho'} \). The average
\[ \text{HA}_{\chi} \left[ \langle \chi|\Lambda' A^\dagger A^\dagger |\chi\rangle \right] = \frac{\text{Tr} \Lambda' A^\dagger A^\dagger \Lambda'}{N} \] (4.34)
The assumption as for paper. Therefore we are using the geometric series 1/H^A shows the purities of the deployed statistical operators. Result for the standard deviation according to (4.32) plotted as a tube around the average. The inset shows the purities of the deployed statistical operators.

is known. However, an exact analytical calculation of Hilbert space averages of the kind HA_x \( \left[ \langle \chi |o|\chi \rangle^2 / \langle \chi |\beta|\chi \rangle \right] \), which is needed for the second term in (4.31), is beyond the scope of this paper. Therefore we are using the geometric series \( 1/(1-x) = 1 + x + x^2 + O(x^3) \) as an approximation for \( |x| \ll 1 \). In our case it is

\[
\frac{1}{\langle \chi |\Lambda^A |\Lambda^A |\chi \rangle} = \frac{1}{1 - (1 - \langle \chi |\Lambda^A |\Lambda^A |\chi \rangle)} = 1 + (1 - \langle \chi |\Lambda^A |\Lambda^A |\chi \rangle) \]

\[= 3 - 3 \langle \chi |\Lambda^A |\Lambda^A |\chi \rangle + \langle \chi |\Lambda^A |\Lambda^A |\chi \rangle^2 + O \left( (1 - \langle \chi |\Lambda^A |\Lambda^A |\chi \rangle)^3 \right).\]

The assumption \( \langle \chi |\Lambda^A |\Lambda^A |\chi \rangle \approx 1 \) is justified in the case of low purities of \( \rho \) and \( \rho' \). It is

\[
HA_x \left[ \langle \chi |\Lambda^A |\Lambda^A |\chi \rangle \right] = \frac{\text{Tr} \Lambda^A |\Lambda^A |}{N} = N \text{ Tr} \rho \rho'.
\]

Therefore in the case of \( \rho = \rho' \) we have \( HA_x \left[ \langle \chi |\Lambda^A |\Lambda^A |\chi \rangle \right] = NP \) and thus \( P \) needs to be close to its lower bound, which is \( 1/N \). Generally we have

\[
HA_x \left[ \langle \chi |\Lambda^A |\Lambda^A |\chi \rangle \right] = N \sum_{i,j=1}^N p_i p_j' |\langle i|j \rangle|^2,
\]

where \( p_i \) and \( p_j' \) are the eigenvalues of \( \rho \) and \( \rho' \) in their respective eigenbasis. Assuming that \( \rho \) and \( \rho' \) are chosen as fully random, the expected overlap of their eigenvectors is \( |\langle i|j \rangle|^2 = 1/N \) (cf. (4.2)) and \( HA_x \left[ \langle \chi |\sqrt{N\rho} |\sqrt{N\rho'} |\chi \rangle \right] \approx 1 \) will follow. Thus we can write the approximated Hilbert space average as

\[
HA_x \left[ \langle \chi |\Lambda^A |\Lambda^A |\chi \rangle \right] \approx 3 HA_x \left[ \langle \chi |\Lambda^A |\Lambda^A |\chi \rangle \right]^2 - 3 HA_x \left[ \langle \chi |\Lambda^A |\Lambda^A |\chi \rangle \right] \langle \chi |\Lambda^A |\Lambda^A |\chi \rangle^2 + HA_x \left[ \langle \chi |\Lambda^A |\Lambda^A |\chi \rangle \right] \langle \chi |\Lambda^A |\Lambda^A |\chi \rangle^2,
\]

where we deploy formula (A.18) of the appendix to calculate the occurring Hilbert space averages. The
appropriate variance is

\[
\text{HV}_{\psi,\chi} \left[ |\langle \chi | A' A | \psi \rangle|^2 \delta \left( \frac{\langle \chi | A' A | \psi \rangle}{\sqrt{\langle \chi | A' A | \chi \rangle}} - z \right) \right] = \text{HA}_{\psi,\chi} \left[ |\langle \chi | A' A | \psi \rangle|^2 \delta \left( \frac{\langle \chi | A' A | \psi \rangle}{\sqrt{\langle \chi | A' A | \chi \rangle}} - z \right) \right]
\]

Here again we find \( \text{HA}_{\chi,\chi} \left[ \left| \langle \chi | A' A | \chi \rangle / \langle \chi | \beta \chi \rangle \right|^2 \right] \) and also \( \text{HA}_{\chi,\chi} \left[ \left| \langle \chi | A' A | \chi \rangle / \langle \chi | \beta \chi \rangle \right|^2 \right] \), which is why we confine ourselves to a numerical calculation of the Hilbert space variance, when testing our approximation.

For the purpose of a numerical illustration we start with (4.31) and (4.32), where one of the states is still fixed. As operator \( A \) we use again the KIC Floquet operator \( U \). Furthermore we will use \( \Lambda \) to center the states \( |\psi\rangle \) around a preset expectation value \( m_z \) of the magnetization in \( z \)-direction

\[
M_z = \sum_{i=1}^{n} \sigma_i^z.
\]

The expectation value of the uniform ensemble, ergo the fully mixed state \( \rho_m \), is \( \text{Tr} M_z \rho_m = 0 \). We choose now instead \( \rho_{m_z} \), defined by \( \text{Tr} M_z \rho_{m_z} = m_z \) with the operator (4.38).

We find the states \( |\psi\rangle \), that fulfill \( \langle \chi | A | \psi \rangle = z \sqrt{\langle \chi | A^2 | \chi \rangle} \), but else are uniformly distributed, by using \( |\chi\rangle \) from (4.19)

\[
|\psi\rangle = \cos \theta \frac{\Lambda |\chi\rangle}{\| \Lambda |\chi\rangle \|} + \sin \theta \frac{\Lambda^{-1} |\chi\rangle}{\| \Lambda^{-1} |\chi\rangle \|} = |z| \frac{\Lambda |\chi\rangle}{\| \Lambda |\chi\rangle \|} + \sqrt{1 - |z|^2} \frac{\Lambda^{-1} |\chi\rangle}{\| \Lambda^{-1} |\chi\rangle \|}.
\]

This procedure is completely analogue to that in the last section. We want to mention again that while the vector \( \Lambda |\chi\rangle / \| \Lambda |\chi\rangle \| \) is fixed, \( \Lambda^{-1} |\chi\rangle / \| \Lambda^{-1} |\chi\rangle \| \) is not as we draw \( |\chi\rangle \) uniformly from the set of all states, that are perpendicular to \( |\chi\rangle \).

In Fig. 5 the transition probability is shown for a low purity \( (m_z = 0.5) \) on the left side and for a high purity \( (m_z = 7) \) on the right. In both cases we work within the chaotic regime \( (J = b = \pi/4, \ldots) \).
$h = 2\pi/5$) of an $n = 8$-particle KIC. Again we find a good agreement between our analytical results and the numerical data. We observe that for a low purity there is not much difference compared to the results in Fig. 5. In the case of high purity we notice two things. First, there is a drastic change of the transition probability in dependence of $\theta$ and second, the variance does not follow the same scheme as before. Namely, it does not grow larger with increasing $\theta$, but stays small for all angles. This is because $\Lambda |\psi\rangle$ basically consists of eigenvectors, for which the corresponding eigenvalues are centered around the expectation value of the deployed statistical operator. In the present case this is $m_z = 7$, where the eigenvalue density of our specific observable is quite sparse (c.f. end of section 2). In other words the amount of states that fulfill $\langle \psi|\Lambda M_z \Lambda|\psi\rangle \approx 7$ and $\langle \chi|\Lambda|\psi\rangle/\sqrt{\langle \chi|\Lambda^2|\chi\rangle} = z$ is small compared to those that solely fulfill the latter condition, which leads to the overall small variance.

Next we consider the average with respect to both states and test our approximation (4.38). Here we use $\Lambda'$, just like $\Lambda$, to center the states $|\chi\rangle$ around a certain expectation value of $M_z$. In Fig. 6 we present again a realization with low purities ($m_z = 0.5, m'_z = -0.3$) and one with higher purities ($m_z = -m'_z = 3$), but still within the range of validity of our approximation. As expected the low purity case does not differ noticeable from its uniform counterpart in Fig. 3. In the high purity case we observe a similar effect as in Fig. 3 where we left one state fixed. The variance at $\theta = 0$ is maximized as this condition forces the states to be composed of similar eigenvectors. This happens at the cost of their expectation values concerning $M_z$, which are still centered around the preset values, but feature also a large variance. As the eigenvalue density becomes more dense in the center of the spectrum, there is a large amount of different options to fulfill $\theta = 0$, leading to the large variance of the transition probability. This is exactly the other way around for $\theta = \pi/2$, where the states can be arranged in such way, that both conditions are fulfilled.

5. Conclusion and Outlook

Typicality is a statistical pattern, that occurs with immense probability under the requirement of a large Hilbert space in quantum mechanics. Thus, it belongs to the high dimensionality phenomena, that makes it ideally suited for many-body systems, in which the Hilbert space dimension usually grows exponentially in the particle number.

We adopted the methods of quantum typicality [1–3], namely the Hilbert space average, and applied them to transition probabilities, particularly those where the initial overlap between the deployed states is fixed. By doing this we put the statistical orthogonality, commonly used in random matrix theory, aside. We explored the limits of the Hilbert space averaging methods analytically by averaging first over one state and subsequently over the remaining one. While doing so we used different distributions of states, uniform distributions as well as nonuniform ones, where the states initially feature an expectation value of a given observable close to a desired preset value. We compared the resulting equations for the average and the variance with numerical calculations, using the unitary Floquet operator of a kicked spin chain. Furthermore we could connect the transition probability for uniformly distributed states to the spectral statistics of the deployed operator.

We studied also the statistics of the transition probabilities under the additional condition $|\langle \chi|\psi\rangle| = z$. For $z = 0$ we confirmed the Kumaraswamy distribution from random matrix theory in the case of no constraint on $|\langle \chi|\psi\rangle|$

More generally our results apply to transition probabilities involving an arbitrary operator $A$. One interesting future application is the doorway mechanism [22,23]. It describes the coupling of the doorway state, that is not an eigenstate of the Hamiltonian, but is distinguished by another distinct feature (e.g. collective motion in many-body systems) to a background of states. Another application appears in describing decoherence features in open systems or in the case of eigenstate thermalization, in which the off-diagonal elements of an observable in the energy eigenbasis are expected to be centered around zero with small variance, while the diagonal elements are centered around a common non-vanishing value. Finally we would also consider it interesting to apply the methods developed here to different quantities. One popular example are reduced statistical operators, that emerge after tracing out pure typical states, that allow to compute for example the entanglement entropy between subsystems.
6. Acknowledgement

We thank Petr Braun mainly for fruitful discussions and a careful reading of the manuscript.
A. Appendix

A.1. Hilbert space average of transition probabilities

In this appendix we will derive (4.31) and (4.32) of the main text. Equation (4.8) and (4.9) follow by setting Λ = 1. So we want to calculate

$$\text{HA}_\psi \left[ |\langle \chi | A \Lambda | \psi \rangle |^2 \delta \left( \frac{|\langle \chi | A \Lambda | \psi \rangle|^2 - z}{\sqrt{|\Lambda|^2}} \right) \right] = \frac{\int d[\psi] |\langle \chi | A \Lambda | \psi \rangle|^2 \delta \left( \frac{|\langle \chi | A \Lambda | \psi \rangle - z \sqrt{|\Lambda|^2}}{2i} \right) \delta(|\psi| - 1)}{\int d[\psi] \delta \left( |\langle \chi | A \Lambda | \psi \rangle - z \sqrt{|\Lambda|^2}| \right) \delta(|\psi| - 1)}. \quad (A.1)$$

Our strategy is that we convert the integrals into a multidimensional Gaussian integrals. To this end we write the delta functions as Fourier integrals

$$\delta(|\psi| - 1) = \frac{1}{2\pi} \int d t_1 \exp \{ i t_1 \langle \psi | \psi \rangle \} \exp \{- i t_1 \} \quad (A.2)$$

$$\delta(|\langle \chi | A \Lambda | \psi \rangle - z') = \delta \left( \frac{|\langle \chi | A \Lambda | \psi \rangle + |\psi| A |\chi \rangle}{2} \right) \delta \left( \frac{|\langle \chi | A \Lambda | \psi \rangle - |\psi| A |\chi \rangle}{2i} \right) \quad (A.3)$$

with $z' = z \sqrt{|\Lambda|^2}$. The second delta function comes with a complex valued argument and thus has to be understood as product of two delta functions concerning the real resp. the imaginary part. We also make use here of $\delta(ax) = \frac{1}{|a|} \delta(x)$, where the prefactor 4 stems from. Furthermore the transition probability in the numerator of (A.1) is written in terms of a generating function

$$\left| \langle \chi | A \Lambda | \psi \rangle \right|^{2n} = \frac{1}{i^n} \partial_{J^n} \exp \{ i J \langle \chi | A \Lambda | \psi \rangle \langle \psi | A \Lambda^\dagger | \chi \rangle \} \bigg|_{J=0} \quad (A.4)$$

The entire numerator of (A.1) is then

$$\frac{4}{(2\pi)^{3/2}} \frac{\partial}{\partial J} \int d[\psi] [d[t] \exp \{ i \langle \psi | (J A A^\dagger | \chi \rangle A A + t_1 \mathbb{1}) | \psi \rangle + i (t_2 - i t_3) \langle \chi | A | \psi \rangle + i (t_2 + i t_3) \langle \psi | A | \chi \rangle \} ] \exp \{- 2i(t_2 \text{Re } z' + t_3 \text{Im } z') \} \bigg|_{J=0} \quad (A.5)$$

with $d[t] = dt_1 dt_2 dt_3$ and the boundaries of integration like in (A.2). Now the Gaussian integration can be performed. We use the well known formula

$$\int_{\mathbb{C}^N} d^N x \exp \{ i x D x + i S x + i x S \} = \frac{(i \pi)^N}{\det D} \exp \{- i S D^{-1} S \} \quad (A.6)$$

for the Hermitian matrix $D$ and $S \in \mathbb{C}^N$. In our case $D$ corresponds to $J A A^\dagger | \chi \rangle \langle \chi | A A + t_1 \mathbb{1}$, which is Hermitian as it consists only of a projector and an identity, and $S$ corresponds to $(t_2 + i t_3) | \chi \rangle$. This yields

$$\frac{4(i \pi)^N}{(2\pi)^{3/2}} \frac{\partial}{\partial J} \int d[t] \exp \left[ - i \left( \frac{t_2^2}{2} + \frac{t_3^2}{2} \right) \langle \chi | A (J A A^\dagger | \chi \rangle \langle \chi | A A + t_1 \rangle^{-1} A | \chi \rangle \right] \frac{\det [J A A^\dagger | \chi \rangle \langle \chi | A A + t_1 \rangle]}{\det [J A A^\dagger | \chi \rangle \langle \chi | A A + t_1 \rangle]} \left. \exp \{- 2i(t_2 \text{Re } z + t_3 \text{Im } z) - i t_1 \} \right|_{J=0} \quad (A.7)$$

The next step is taking the derivative. The derivative of a determinant dependent on a real parameter is

$$\left. \frac{\partial}{\partial J} \det D_J \right|_{J=0} = \det D_J \text{ Tr} \left( D_J^{-1} \partial_J D_J \right) \quad (A.8)$$

and therefore

$$\left. \frac{\partial}{\partial J} \left[ \exp \left[ - i \left( \frac{t_2^2}{2} + \frac{t_3^2}{2} \right) \langle \chi | A (J A A^\dagger | \chi \rangle \langle \chi | A A + t_1 \rangle^{-1} A | \chi \rangle \right] \frac{\det [J A A^\dagger | \chi \rangle \langle \chi | A A + t_1 \rangle]}{\det [J A A^\dagger | \chi \rangle \langle \chi | A A + t_1 \rangle]} \right] \right|_{J=0} = \quad (A.9)$$

$$\exp \left[ - i \frac{t_2^2}{2} + \frac{t_3^2}{2} \langle \chi | A^2 | \chi \rangle \right] \left[ - \frac{\langle \chi | A A^2 | \chi \rangle}{t_1^{N+1}} + i \frac{t_2^2 + t_3^2}{t_1^{N+2}} \left| \langle \chi | A A^2 | \chi \rangle \right|^2 \right]. \quad (A.10)$$
In the upcoming integrals we use the substitution $t_{2,3} = \sqrt{\langle \chi | \Lambda^2 | \chi \rangle}$ and $z' = z \sqrt{\langle \chi | \Lambda^2 | \chi \rangle}$ like in the main text
\[
\int dt |t|^{N+1} \exp \left[ -i t \frac{t_2^2 + t_3^2}{t_1^2} \langle \chi | \Lambda^2 | \chi \rangle - 2i \left( t_2 \text{Re } z' + t_3 \text{Im } z' \right) - it_1 \right] = i \frac{N!}{(N-1)!} \frac{(z^2 - 1)^{N-1}}{\langle \chi | \Lambda^2 | \chi \rangle} \quad (A.9)
\]
\[
\int dt |t|^N \exp \left[ -i t \frac{t_2^2 + t_3^2}{t_1^2 + z'^2} \langle \chi | \Lambda^2 | \chi \rangle - 2i \left( t_2 \text{Re } z' + t_3 \text{Im } z' \right) - it_1 \right] = i \frac{N!}{(N-1)!} \frac{(z^2 - 1)^{N-1}}{\langle \chi | \Lambda^2 | \chi \rangle} - i \frac{N!}{(N-2)!} \frac{(z^2 - 1)^{N-2}}{\langle \chi | \Lambda^2 | \chi \rangle} |z|^2.
\]

The integral in $t_2$ and $t_3$ is Gaussian again, whereas the $t_1$ integral can be reduced to
\[
\int dt \frac{e^{it\alpha}}{t} = i \pi \frac{\alpha}{(N-1)!} |\alpha|^{N-1} \text{sgn } \alpha.
\]

We get to this result by writing $\frac{1}{t^2}$ as a derivative and using the differentiation rule of the Fourier transform
\[
\mathcal{F} \left( \frac{1}{t^N} \right) = \frac{(-1)^{N-1}}{(N-1)!} \mathcal{F} \left( \frac{\partial^{N-1}}{\partial \alpha^{N-1}} \right) = \frac{(i\alpha)^{N-1}}{(N-1)!} \mathcal{F} \left( \frac{1}{\alpha} \right).
\]

Now we only have to set in
\[
\mathcal{F} \left( \frac{1}{t} \right) = i \pi \text{sgn } \alpha,
\]
which exists in the sense of its principal value [39]. All in all the numerator $\langle A \rangle$ evaluates to
\[
\int d[\psi] |\langle \chi | AA | \psi \rangle|^2 \delta \left( \langle \chi | AA | \psi \rangle - z' \right) \delta(\langle \psi | \psi \rangle - 1) = \frac{\pi^N}{2} \left[ \frac{1 - |z|^2}{(N-1)!} \left( \frac{\langle \chi | AA^t | \chi \rangle}{\langle \chi | AA | \chi \rangle} - \frac{|\langle \chi | AA | \chi \rangle|^2}{\langle \chi | AA^2 | \chi \rangle} \right) + \frac{1 - |z|^2}{(N-2)!} \frac{|z|^2}{\langle \chi | AA^2 | \chi \rangle} \right].
\]

The denominator is calculated in the same manner. But as there is no generating function for the transition probability a derivative is not necessary
\[
\int d[\psi] \delta \left( \langle \psi | \psi \rangle - 1 \right) \delta \left( \langle \chi | \psi \rangle - z' \right) = \frac{\pi^N}{2 \pi (\langle \chi | \Lambda^2 | \chi \rangle)} \frac{1 - |z|^2}{(N-2)!} (\langle \chi | AA^2 | \chi \rangle) \quad (A.14)
\]

In fact one retrieves this result already by setting $\Lambda = 1$ and dividing by the preset overlap $|\langle \chi | \psi \rangle|^2 = |z|^2$ in $A.13$.

We consider $\Lambda = 1$ for now in order to make a geometrical interpretation. In this case it corresponds to the volume of the intersection of the unit sphere, described by $\langle \psi | \psi \rangle = 1$, and a cone with its symmetry axis along $|\chi\rangle$ and an aperture of $2\theta = 2 \arccos |z|$, described by $\langle \chi | \psi \rangle = z$. It vanishes for $|z| = 1$ as the intersection will be merely a point. For $|z| = 0$ we obtain the volume of the $2N - 3$-dimensional unit sphere up to a prefactor of $1/4$
\[
\frac{\pi^{N-1}}{2} \frac{1}{(N-2)!} = \frac{1}{4} \text{Vol } S^{2N-3}.
\]

The origin of this prefactor lies in our choice for the arguments of the delta function, i.e. we chose $\delta(\langle \psi | \psi \rangle - 1)$ instead of $\delta(\| \psi \| - 1)$. But this is taken care of in the normalization $\langle A \rangle$. Returning to the general case and combining (A.13) and (A.14) we finally arrive at
\[
HA_\psi \left[ |\langle \chi | AA | \psi \rangle|^2 \delta \left( \frac{\langle \chi | AA^t | \chi \rangle}{\sqrt{\langle \chi | AA | \chi \rangle}} - z \right) \right] = \frac{1 - |z|^2}{N-1} \langle \chi | AA^2 A^t | \chi \rangle + \frac{N |z|^2 - 1}{N-1} \frac{|\langle \chi | AA^2 | \chi \rangle|^2}{\langle \chi | AA^2 | \chi \rangle}.
\]
Likewise we can calculate the second moment of $|\langle \psi | \Lambda | \psi \rangle|^2$, which is necessary for the variance, by taking the second derivative in (A.3). This yields

$$
\text{HA} \left[ |\langle \psi | \Lambda | \psi \rangle|^2 \right] = \frac{2 (1 - |z|^2)^2}{N(N - 1)} |\langle \psi | \Lambda | \psi \rangle|^2 
+ \left[ \frac{2 (1 - |z|^2)^2}{N(N - 1)} - \frac{4 (1 - |z|^2) |z|^2}{N - 1} + |z|^4 \right] |\langle \psi | \Lambda | \psi \rangle|^2 
- \left[ \frac{4 (1 - |z|^2)^2}{N(N - 1)} - \frac{4 (1 - |z|^2) |z|^2}{N - 1} \right] |\langle \psi | \Lambda | \psi \rangle|^2 
\frac{|\langle \psi | \Lambda^2 | \psi \rangle|^2}{|\langle \psi | \Lambda | \psi \rangle|^2}.
$$

(A.17)

A.2. Hilbert space average of positive integer powers of matrix elements

In this appendix we will present a formula for the higher moments of matrix elements using the methods of (3). For $M$ (not necessarily) different matrices $B_m$ and uniformly distributed $|\psi\rangle$ we found

$$
\text{HA} \left[ \langle \psi | B_1 | \psi \rangle \cdots \langle \psi | B_M | \psi \rangle \right] = \sum_{m_1 \ldots m_M=1}^{N} B_{i_1 j_1} \cdots B_{i_M j_M}^{M} \text{ HA} \left[ \prod_{m=1}^{M} \chi_{m}^* \chi_{m} \right] \text{ Per} \left( \begin{array}{cc}
\delta_{i_1 j_1} & \cdots \delta_{i_M j_M} \\
\vdots & \vdots \\
\delta_{i_M j_M} & \cdots \delta_{i_M j_M}
\end{array} \right),
$$

(A.18)

where the superscripts of the matrices denote their elements and Per the permanent. In the first step we merely used the definitions of the Hilbert space average, making it possible to express $|\chi\rangle$ in an arbitrary basis with the coordinates $\chi_i = \langle i | \chi \rangle$. For convenience of the reader we give the explicit expressions for $M = 2$ and $M = 3$

$$
\text{HA} \left[ \langle \psi | B_1 | \psi \rangle \langle \psi | B_2 | \psi \rangle \right] = \frac{1}{N(N + 1)} (\text{Tr} B_1 B_2 + \text{Tr} B_1 \text{Tr} B_2)
$$

(A.19)

$$
\text{HA} \left[ \langle \psi | B_1 | \psi \rangle \langle \psi | B_2 | \psi \rangle \langle \psi | B_3 | \psi \rangle \right] = \frac{(N - 1)!}{(N + 2)!} (\text{Tr} B_1 \text{ Tr} B_2 \text{ Tr} B_3 + \text{Tr} B_1 B_2 B_3 + \text{Tr} B_1 B_2 \text{ Tr} B_3 
+ \text{Tr} B_1 \text{ Tr} B_2 B_3 + \text{Tr} B_1 B_3 \text{ Tr} B_2 + \text{Tr} B_1 B_2 B_3 + \text{Tr} B_1 B_3 B_2).
$$

So (A.18) contains the sum of all possible combinations of traces of the matrices $B_1, \cdots, B_M$. Thus the average of the product of $M$ different matrix elements consists of $M!$ terms. We carefully verified this formula up to $M = 4$ by explicit calculation.

A.3. Distribution of transition probabilities

We are interested in the distribution $p(s)$ of transition probabilities $|\langle \psi | U | \psi \rangle|^2$ with uniformly distributed states and unitary $U$. We follow the same strategy as before, i.e. converting the upcoming integral into a Gaussian integral with subsequent Fourier transformation

$$
p(s) = \int_{CN} d[\chi] d[\psi] \delta \left( s - |\langle \chi | U | \psi \rangle|^2 \right) \delta \left( \langle \psi | \chi \rangle - 1 \right) \delta \left( \langle \psi | \psi \rangle - 1 \right),
$$

(A.20)

$$
= \frac{1}{(2\pi)^{2N}} \int_{CN} d[\chi] d[\psi] d[t] \text{ Exp}[-it_1 s - it_2] \text{ Exp} \left[ - \langle \psi | \left( t_1 U | \psi \rangle \langle \psi | U^\dag + t_2 \right) | \chi \rangle \right] \delta \left( \langle \psi | \psi \rangle - 1 \right).
$$

The determinant yields

$$
\text{Det} \left[ t_1 U | \psi \rangle \langle \psi | U^\dag + t_2 \right] = t_2^{N-1} (t_2 + t_1 \langle \psi | \psi \rangle).
$$

(A.21)
Thus we can apply the remaining δ-function and arrive at the following Fourier integral
\[ \frac{(i\pi)^N}{(2\pi)^N} \int dt \frac{\text{Exp} \left[ -it_1 s - it_2 \right]}{t_2^{N-1}(t_2 + t_1)} = \frac{(i\pi)^N}{(2\pi)^N} \int dt \frac{\text{Exp} \left[ -it_1 s + i(s - 1)t_2 \right]}{t_2^{N-1}t_1}, \] (A.22)
which can, after the substitution in the second step, be evaluated in terms of (A.10). Thus, we find
\[ \pi^N \frac{4}{(N - 2)!} (1 - s)^{N-2}, \] (A.23)
and are able to extract the s-dependence. We denote \( \langle f(s) \rangle = \int_0^1 ds f(s)p(s) \). The normalization condition \( \langle 1 \rangle = 1 \) yields
\[ p(s) = (N - 1)(1 - s)^{N-2}, \] (A.24)
which is a probability distribution on \([0, 1]\) of the Kumaraswamy type. The mean and the variance are
\[ \langle s \rangle = \frac{1}{N} \quad \text{and} \quad \sigma^2 = \langle s^2 \rangle - \langle s \rangle^2 = \frac{N - 1}{N^2(N + 1)} \sim O(N^{-2}). \] (A.25)
Therefore it is \( \langle s \rangle \sim \sigma \). Furthermore for the skewness and the kurtosis we find
\[ \left\langle \left( \frac{s - \langle s \rangle}{\sigma} \right)^3 \right\rangle \xrightarrow{N \to \infty} 2 \quad \text{and} \quad \left\langle \left( \frac{s - \langle s \rangle}{\sigma} \right)^4 \right\rangle \xrightarrow{N \to \infty} 9. \] (A.26)
The probability density (A.24) shares these properties with the exponential distribution \( Ne^{-Ns} \). In fact it converges to the exponential distribution after rescaling \( s \to s/N \) for large \( N \). Furthermore replacing the integration regime in (A.20) by \( \mathbb{R}^N \) would yield the well known Porter-Thomas distribution.

References
[1] C. Bartsch, J. Gemmer. Phys. Rev. Lett. 102 110403 (2009).
[2] P. Reimann. Phys. Rev. E 97 062129 (2018).
[3] J. Gemmer, M. Michel, G. Mahler. Quantum Thermodynamics. Springer-Verlag Berlin Heidelberg (2009).
[4] P. Reimann. Phys. Rev. Lett. 120 230601 (2018).
[5] R. Steinigeweg, A. Khodja, H. Niemeyer, C. Gogolin, J. Gemmer. Phys. Rev. Lett. 112 130403 (2014).
[6] A. Khodja, R. Steinigeweg, J. Gemmer. Phys. Rev. E 91 012120 (2015).
[7] R. Steinigeweg, J. Herbrych, P. Prelovšek. Phys. Rev. E 87 012118 (2013).
[8] R. Steinigeweg, J. Gemmer, W. Brenig. Phys. Rev. Lett. 112 120601 (2014).
[9] R. Steinigeweg, J. Herbrych, F. Pollmann, W. Brenig. Phys. Rev. B 94 180401 (2016).
[10] T. A. Elsayed, B. V. Fine. Phys. Rev. Lett. 110 070404 (2013).
[11] J. Richter, R. Steinigeweg. Phys. Rev. B 99 094419 (2019).
[12] I. Kukuljan, S. Grozdanov, T. Prosen. Phys. Rev. B 96 060301 (2017).
[13] X. Chen, T. Zhou, D. A. Huse, E. Fradkin. Annalen der Physik 529 7 1600332 (2017).
[14] T. Guhr, A. Müller-Groeling, H. A. Weidenmüller. Phys. Rep. 299 4 189–425 (1998).
[15] M. L. Mehta. Random Matrices. Academic Press (2004).
[16] C. I. Barbosa, T. Guhr, H. L. Harney. Phys. Rev. E 62 1936–1949 (2000).
[17] A. Hamoudi, R. G. Nazmitdinov, E. Shahaliev, Y. Alhassid. Phys. Rev. C 65 064311 (2002).
[18] N. Ullah. Journal of Mathematical Physics 4 10 1279–1282 (1963).
[19] N. Ullah. Journal of Mathematical Physics 8 5 1095–1096 (1967).
[20] C. E. Porter, R. G. Thomas. Phys. Rev. 104 483–491 (1956).
[21] Y. Alhassid, R. D. Levine. Phys. Rev. Lett. 57 2879–2882 (1986).
[22] H. Kohler, T. Guhr, S. Åberg. New Journal of Physics 12 7 073026 (2010).
[23] H. Kohler, H.-J. Sommers, S. Åberg, T. Guhr. Phys. Rev. E 81 050103 (2010).
[24] V. Zelevinsky. Annual Review of Nuclear and Particle Science 46 1 237–279 (1996).
[25] I. Kawata, H. Kono, Y. Fujimura, A. D. Bandrauk. Phys. Rev. A 62 031401 (2000).
[26] E. Peyroula, R. Jost, M. Lombardi, P. Dupré. Chemical Physics 102 3 417–435 (1986).
[27] S. Åberg, T. Guhr, M. Miski-Oglu, A. Richter. Phys. Rev. Lett. 100 204101 (2008).
[28] M. Hussein, V. Kharchenko, L. Canto, R. Donangelo. Annals of Physics 284 1 178 – 194 (2000).
[29] I. Bengtsson, K. Zyczkowski. Geometry of Quantum States: An Introduction to Quantum Entanglement. Cambridge University Press (2006).
[30] P. J. Diggle, N. I. Fisher. Computers & Geosciences 11 6 725–766 (1985).
[31] E. Lieb, T. Schultz, D. Mattis. Ann. Phys. 16 3 407466 (1961).
[32] M. Akila, D. Waltner, B. Gutkin, T. Guhr. J. Phys. A 49 37 375101 (2016).
[33] H.-J. Stöckmann. Quantum Chaos - an introduction. Cambridge University Press (1999).
[34] F. Haake. Quantum Signatures of Chaos. Springer-Verlag Berlin Heidelberg (2010).
[35] B. Bertini, P. Kos, T. Prosen. Phys. Rev. Lett. 121 264101 (2018).
[36] P. Braun, D. Waltner, M. Akila, B. Gutkin, T. Guhr. arXiv:1902.06265 (2019).
[37] N. Linden, S. Popescu, A. J. Short, A. Winter. Phys. Rev. E 79 061103 (2009).
[38] K. Scharnhorst. Acta Applicandae Mathematicae 69 1 95–103 (2001).
[39] I. N. Sneddon. The use of integral transforms. McGraw-Hill Book Company (1972).