Atypicality of Most Few-Body Observables

Ryusuke Hamazaki
Department of Physics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Masahito Ueda
Department of Physics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan and
RIKEN Center for Emergent Matter Science (CEMS), Wako 351-0198, Japan

(Dated: February 27, 2018)

The eigenstate thermalization hypothesis (ETH), which dictates that all diagonal matrix elements within a small energy shell be almost equal, is a major candidate to explain thermalization in isolated quantum systems. According to the typicality argument, the maximum variations of such matrix elements should decrease exponentially with increasing the size of the system, which implies the ETH. We show, however, that the typicality argument does not apply to most few-body observables for few-body Hamiltonians when the width of the energy shell decreases at most polynomially with increasing the size of the system.

PACS numbers: 05.30.-d, 03.65.-w

Introduction. Thermalization in isolated quantum systems has long hovered over researchers [1–6] since von Neumann’s seminal work [7]. Recently, this problem has attracted growing interest [8–10] due to experimental advances in ultracold atoms [11–19], ions [20–23], and superconducting qubits [24]. These experiments have motivated theorists to identify the conditions under which thermalization occurs [25–37].

The long-time dynamics of isolated quantum systems can be analyzed through matrix elements of an observable in the energy eigenbasis. The eigenstate thermalization hypothesis (ETH) [4, 5] dictates that all diagonal matrix elements within a small energy shell be almost equal [38]. Then the expectation value of an observable in the steady state can be calculated from the microcanonical ensemble for all initial states with small energy fluctuations [39]. Meanwhile, off-diagonal matrix elements characterize autocorrelation functions and temporal fluctuations [10]. It is thus of fundamental importance to understand how such matrix elements generically behave in macroscopic systems.

To be specific, consider a set of eigenstates \{|E_\alpha\rangle\} of the Hamiltonian and introduce a projector \(\mathcal{P}_{sh} = \sum_{|E_\alpha - E| \leq 2\Delta E} |E_\alpha\rangle \langle E_\alpha|\) onto the Hilbert space \(\mathcal{H}_{sh}\) for an energy shell of median \(E\) and width \(2\Delta E\). Let the spectral decomposition of an observable \(\hat{O}\) projected onto \(\mathcal{H}_{sh}\) be \(\mathcal{P}_{sh} \hat{O} \mathcal{P}_{sh} = \sum_{i=1}^{d_{sh}} a_i |a_i\rangle \langle a_i|\), where \(d_{sh} = \text{dim}[\mathcal{H}_{sh}]\) is the dimension of the Hilbert space within the energy shell. Then the matrix elements of \(\hat{O}\) within the energy shell can be expressed as \(\mathcal{O}_{\alpha \beta} = \langle E_\alpha | \mathcal{P}_{sh} \hat{O} \mathcal{P}_{sh} | E_\beta\rangle = \sum_i a_i U_{\alpha i} U^*_\beta i\), where \(U_{\alpha i} := \langle E_\alpha | a_i\rangle\) constitutes the \(d_{sh} \times d_{sh}\) unitary matrix \(U\).

To investigate the ETH, let us consider the maximum variation of \(\mathcal{O}_{\alpha \alpha}\) within the energy shell. This quantity enables us to directly judge whether all diagonal matrix elements are almost equal, which is required for justifying thermalization from arbitrary initial states [40]. As shown in Ref. [31] (see Appendix I of the Supplemental Material [41]), the maximum deviation of \(\mathcal{O}_{\alpha \alpha}\) from its average value decreases exponentially with increasing the size of the system for almost all (typical) \(U\)'s over the unitary Haar measure. This mathematical property is referred to as the typicality with respect to the unitary Haar measure [42]. Based on the typicality, it is argued [31] that for actual \(\mathcal{H}\) and \(\hat{O}\) of our concern the variations of \(\mathcal{O}_{\alpha \alpha}\) are exponentially small. We refer to this conjecture as the typicality argument [31] to distinguish it from the above-mentioned (mathematically rigorous) typicality. Since exponentially small variations of \(\mathcal{O}_{\alpha \alpha}\) imply the ETH within the same energy shell, the typicality argument offers a possible scenario for the justification of the ETH [43]. Such an idea was originally put forth by von Neumann for microspaces [7, 44] and it has recently been generalized to arbitrary observables [31, 45]. Note that the spirit of the typicality argument is similar to that of applying random matrix theory (RMT) [31, 44, 46–48] to physics [49].

In this Letter, however, we show that the typicality argument cannot be applied to most few-body observables for lattice Hamiltonians. In fact, we show that diagonal matrix elements for most few-body observables do not behave typically even if the energy width decreases algebraically with increasing the size of the system. In other words, the maximum variation of \(\mathcal{O}_{\alpha \alpha}\) does not decrease exponentially. Our approach provides rigorous results without assuming the unitary Haar measure [31] nor the specific form of matrix elements proposed in Ref. [50].

Setup. We assume that the energy width \(\Delta E\) scales with the system size \(N\) as \(\Delta E \propto N^{-p}\) and that \(d_{sh}\) increases exponentially with \(N\), where \(-1 < p < 0\) (sub-extensive) for the energy width of the microcanonical ensemble and \(p = \frac{2}{3}\) for that of the diffusive energy (many-body Thouless energy) [10] with \(D\) being the spatial dimension.

\[\sum_{|E_\alpha - E| \leq 2\Delta E} |E_\alpha\rangle \langle E_\alpha|\]
FIG. 1. (Left) Space of all the unitary matrices $U$’s whose matrix elements are constituted from inner products between the eigenbases of the Hamiltonian of a system and those of an observable. Almost all $U$’s with respect to the unitary Haar measure belong to $U_{(a_i),\eta}$, a set of all unitary matrices for which the maximum variations of matrix elements $O_{\alpha\alpha}$ decrease exponentially with increasing the size of the system. (Right) In a system with few-body interactions, we consider observables which are mostly unitary matrices $B$ operators are defined as a linear combination of few-body operator bases \{\hat{A}_f\}. Then, for most realizations of $\hat{O}$, the corresponding $U$’s are atypical when the energy window under consideration decreases at most polynomially with increasing the size of the system.

In the following we consider a spin system on $N$ lattice sites. The entire Hilbert space can be written as $\mathcal{H} = \bigotimes_{x=1}^{N} \mathcal{H}_x$, where $\mathcal{H}_x$ is a local Hilbert space at site $x$. Let $\mathcal{L}(\mathcal{H})$ and $\mathcal{L}(\mathcal{H}_x)$ be operator spaces acting on $\mathcal{H}$ and $\mathcal{H}_x$, respectively. We take an orthonormal basis set for $\mathcal{L}(\mathcal{H}_x)$ as $\{ |\lambda_x^\mu = 1_x, \lambda_x^1, \ldots, \lambda_x^{S^2-1} \rangle \}$, where $S = \dim[\mathcal{H}_x]$ and $\lambda_x^\mu (0 \leq \mu \leq S^2 - 1)$ are $S \times S$ Hermitian matrices subject to $\text{Tr}_x [\lambda_x^\mu \lambda_x^{\mu'}] = S \delta_{\mu\mu'}$. Then, the basis set that spans $\mathcal{L}(\mathcal{H})$ is written as $\mathcal{B}_N = \{ \hat{A}_{\mu_1, \ldots, \mu_N} = \bigotimes_{x=1}^{N} \hat{\lambda}_x^{\mu_x} |0 \leq \mu_x \leq S^2 - 1 \},$ where $\text{Tr}[\hat{A}_{\mu_1, \ldots, \mu_N} \hat{A}_{\mu'_1, \ldots, \mu'_N}] = S^N \prod_{x=1}^{N} \delta_{\mu_x \mu'_x}.$

We next define $m$-body operators. For this purpose, we take a basis set $\mathcal{B}_m \subset \mathcal{B}_N$ whose elements act nontrivially on at most $m$ sites: $\mathcal{B}_m = \{ \bigotimes_{x=1}^{N} \hat{\lambda}_x^{\alpha_{x,q}} |1 \leq q \leq m, 1 \leq x_{\alpha_{x,q}} \leq N, 1 \leq \alpha_{x,q} \leq S^2 - 1 \}$ for $m \geq 1$ and $\mathcal{B}_0 = \{ \bigotimes_{x=1}^{N} \hat{0}_x \}$. Then, $m$-body operators are defined as a linear combination of elements in $\mathcal{B}_m$, but not in $\mathcal{B}_{m-1}$. If $m (m \ll N)$ does not depend on $N$, we call them few-body operators. We note that our few-body operators are defined in a much broader sense than usual.

To discuss characteristic behaviors of few- and many-body observables, we next consider observables which are randomly chosen from at most $m$-body operators.

**Definition** (Randomly chosen observables from $\mathcal{L}_m$). Let $\hat{A}_1, \ldots, \hat{A}_n$ be elements in $\mathcal{B}_m$, where $n = \sum_{q=0}^{m} \binom{N^2}{q} (S^2 - 1)^q$ is the number of the bases and $\text{Tr}[\hat{A}_f \hat{A}_g] = SN \delta_{fg}$. Let us consider a set $\mathcal{L}_m$ of at most $m$-body observables, which can be written as a linear combination of $\hat{A}_f$. Now, we take an observable $\hat{G} \in \mathcal{L}_m$ expressed as

$$\hat{G} = \sum_{f=1}^{n} G_f \hat{A}_f,$$

where real variables $\tilde{G} = (G_1, \ldots, G_f, \ldots, G_n)$ are randomly chosen according to an arbitrarily given probability distribution $P(\tilde{G})$. When $P(\tilde{G})$ is invariant under an arbitrary $n \times n$ orthogonal transformation, we call $\hat{G}$ an observable randomly chosen from $\mathcal{L}_m$ [51]. Note that we may arbitrarily choose $P(\tilde{G})$ to suit our purpose; in contrast, if we choose $U$ from a unitary Haar measure, it is unclear from what probability distribution an observable is chosen. In this sense, our scheme of sampling observables has a well-defined operational meaning.

**Atypicality of most few-body observables.** We investigate the behavior of matrix elements of random observables defined above and compare it with what the Haar measure predicts. As defined above, the diagonal matrix elements $O_{\alpha\alpha}$ within the energy shell $| |E_\alpha - E| \leq \Delta E |$ are given by $O_{\alpha\alpha} = \sum_{i=1}^{d_{sh}} |a_i| |U_{\alpha i}|^2$. We define $U_{(a_i),\eta}$ as a set of all $U$’s that lead to the inequality $\max_{|E_\alpha - E| \leq \Delta E} | |O_{\alpha\alpha} - O_{\beta\beta}| | \leq ||\hat{O}||_\text{op} d_{sh}^{\eta}$ for given $\{a_i\}$, where $||\hat{O}||_\text{op}$ denotes the operator norm and $\eta > 0$. This inequality means that the maximum variation of $O_{\alpha\alpha}$ within the energy shell decreases exponentially as a function of $N$, which also implies the ETH of $O_{\alpha\alpha}$ within the shell. As illustrated in Fig. 1, almost all (typical) $U$’s with respect to the Haar measure belong to $U_{(a_i),\eta}$ in the thermodynamic limit for $0 < \eta < \frac{1}{2}$ (see Appendix I for the proof [41]).

We first consider a few-body Hamiltonian (i.e., Hamiltonian with few-body interactions) and few-body observables. We show that for most few-body observables, the corresponding $U$ is atypical in the sense that $U \notin U_{(a_i),\eta}$ (see Fig. 1). In fact, we can show the following theorem.

**Theorem** (Atypicality of most few-body observables). Let us consider a $k$-body Hamiltonian, and assume that $N$ is sufficiently large and that $m (k \leq m \ll N)$ is independent of $N$. Suppose that we randomly choose an observable $\hat{O} = \sum_f G_f \hat{A}_f$ from $\mathcal{L}_m$, from which we obtain the corresponding $\{a_i\}$ and $U$. Then,

$$\mathbb{P}_{\mathcal{L}_m} [U \in U_{(a_i),\eta}] \leq \frac{\sqrt{\pi n} ||\hat{H}||_\text{op} \Lambda}{2\Delta E} \frac{\Gamma \left( \frac{n}{2} \right)}{\Gamma \left( \frac{n-1}{2} \right)} d_{sh}^{-\eta},$$

for $\Delta E$ representing the energy width of the few-body observable.
FIG. 2. Schematic illustration of the idea behind the proof of Theorem. (Left) We first show that $|\vec{\Gamma}|$ does not decrease exponentially with increasing $N$ unless $|\vec{H}|$ is exponentially large. This can be seen from the fact that $|\vec{G}, \vec{H}|$ does not decrease exponentially as a function of $N$. (Right) Then, for $|\vec{G}, \vec{H}|$ to be exponentially small ($\leq ||\vec{O}||_{\text{op}} d_{\text{sh}}^{-\eta}$), $\vec{G}$ should be almost orthogonal to $\vec{\Gamma}$ (we assume that $|\vec{G}|$ is not exponentially large). The probability of such an event is exponentially small ($\propto d_{\text{sh}}^{-\eta}$) unless the dimension $n$ of the hypersphere is exponentially large. This is the case for few-body observables.

where $P_{\mathcal{L}_m}$ denotes the probability with respect to $P(\vec{G})$, and $\Lambda = \max_{f} ||\hat{A}_f||_{\text{op}} \leq S \frac{\pi}{2}$. When $||\vec{H}||_{\text{op}}$ does not grow exponentially with respect to $N$, the left-hand of the inequality (2) vanishes for large $N$. Note that the assumption of the scaling $\Delta E \propto N^{-p}$ is sufficient to bound the right-hand side.

The inequality (2) shows that, for physically relevant Hamiltonians and most few-body observables, $\max_{|E_{\alpha} - E_{\beta}|, |E_{\beta} - E_{\gamma}| \leq \Delta E} |\mathcal{O}_{\alpha \alpha} - \mathcal{O}_{\beta \beta}|$ does not decrease as a power of $d_{\text{sh}}$. This means that the corresponding $U$ is atypical. As long as $m$ satisfies $k \leq m$ and is independent of $N$ (i.e., few-body), atypicality holds true for every $m$.

Proof of Theorem (see Fig. 2). We first note that $\vec{H} \in \mathcal{L}_m$ satisfies the following condition for a $k$-body Hamiltonian:

$$\max_{|E_{\alpha} - E_{\beta}|, |E_{\beta} - E_{\gamma}| \leq \Delta E} |(\hat{H})_{\alpha \alpha} - (\hat{H})_{\beta \beta}| = \xi_d. \quad (3)$$

Here $\xi_d = 2\Delta E$ does not decrease faster than polynomial in $N$.

Let $\gamma$ and $\delta$ be labels of eigenstates that satisfy $(\hat{H})_{\gamma \gamma} - (\hat{H})_{\delta \delta} = \xi_d$. Define $\Gamma_f = (\hat{A}_f)_{\gamma \gamma} - (\hat{A}_f)_{\delta \delta}$. Then the expansion $\vec{H} = \sum_{f=1}^{n} H_f \hat{A}_f$ leads to $\vec{H} \cdot \vec{\Gamma} = \xi_d$, where $\vec{H} = (H_1, \cdots, H_n)$ and $\vec{\Gamma} = (\Gamma_1, \cdots, \Gamma_n)$. Since $|\vec{H}| = \sqrt{\frac{\text{tr}[\vec{H}^2]}{S^2}} \leq ||\vec{H}||_{\text{op}}$, we obtain

$$|\vec{\Gamma}| \geq \frac{\xi_d}{||\vec{H}||_{\text{op}}}. \quad (4)$$

Next, we evaluate the left-hand side of Eq. (2). Since $\max_{|E_{\alpha} - E_{\beta}|, |E_{\beta} - E_{\gamma}| \leq \Delta E} |\mathcal{O}_{\alpha \alpha} - \mathcal{O}_{\beta \beta}| \geq |\vec{G} \cdot \vec{\Gamma}|$, we obtain

$$P_{\mathcal{L}_m} \max_{|E_{\alpha} - E_{\beta}|, |E_{\beta} - E_{\gamma}| \leq \Delta E} |\mathcal{O}_{\alpha \alpha} - \mathcal{O}_{\beta \beta}| \leq ||\hat{\mathcal{O}}||_{\text{op}} \epsilon$$

$$
\leq P_{\mathcal{L}_m} \max_{|E_{\alpha} - E_{\beta}|, |E_{\beta} - E_{\gamma}| \leq \Delta E} |\vec{G} \cdot \vec{\Gamma}| \leq ||\hat{\mathcal{O}}||_{\text{op}} \epsilon
$$

(5)

for any $\epsilon > 0$, where we use the fact $P[a \leq c] \geq P[b \leq c]$ for $a \leq b$.

To evaluate Eq. (5), note that the probability $P(\vec{G} d \vec{G})$ can be written as $P((\vec{G}) d^{\vec{G}})$ because of the invariance under orthogonal transformations ($\Omega$ denotes the high-dimensional solid angle). Then, denoting the angle between $\vec{G}$ and $\vec{\Gamma}$ by $\theta$, we obtain

$$P_{\mathcal{L}_m} \left| \vec{G} \cdot \vec{\Gamma} \right| \leq ||\hat{\mathcal{O}}||_{\text{op}} \epsilon$$

$$\leq P_{\mathcal{L}_m} \left| \cos \theta \right| \leq \frac{\sqrt{n} ||\vec{H}||_{\text{op}} \Lambda}{\xi_d}$$

$$\leq \frac{\sqrt{n} \Lambda ||\vec{H}||_{\text{op}} \Lambda}{2\Delta E} \left( \frac{\pi}{2} \right) \left( \frac{\Lambda}{\Delta E} \right). \quad (6)$$

Here, in deriving the second line, we use $||\hat{\mathcal{O}}||_{\text{op}} \leq \Lambda ||\vec{G}|| \sqrt{n}$ that results from the property of an operator norm and the Cauchy-Schwartz inequality (see Eq. (37) in Supplementary Material [41]). With $\epsilon = d_{\text{sh}}^{-\eta}$, the left-hand side of (5) becomes $P_{\mathcal{L}_m}[U \in \mathcal{U}_{(a_\eta)}]$, which, together with (6), completes the proof of the theorem. □

We note that our Theorem holds true for an arbitrary $k$-body Hamiltonian. Thus, it is natural to take $\vec{H}$ as a Hamiltonian written as a sum of spatially local operators (such locality is expected to be necessary for standard statistical mechanics). In this case, $||\vec{H}||_{\text{op}} \propto N$ is expected. Thus, for the left-hand of (5) to vanish, we can take $\epsilon \propto N^{-p} \sim n^{-\frac{\eta}{2}} (z > 1 + p + m)$ if $\Delta E \propto N^{-p}$, since $\sqrt{n} \Lambda \left( \frac{\pi}{2} \right) \left( \frac{\Lambda}{\Delta E} \right) \rightarrow n \sim N^m$ for large $N$. This means that the maximum variation for most few-body observables decays slower than $n \sim N^{-p}$.

We emphasize that observables treated in the theorem can be either local or non-local observables as long as they are few-body. Thus, our theorem applies to momentum distributions [26, 45] and structure factors [52], which are non-local but nevertheless expected to obey standard statistical mechanics [10]. Our theorem implies that it is likely that $\vec{U}$ satisfies $\vec{U} \notin \mathcal{U}_{(a_\eta)}$ for local (or non-local but physical) observables, unless special reasons dictate otherwise. However, it is often the case that we are specifically interested in spatially local observables. An extension of our theorem to most local
observables for translation-invariant Hamiltonians with local interactions is given in Appendix IV [41].

A measure consistent with the typicality argument. As already noted, Theorem is meaningful only when $m$ does not grow in comparable proportion with $N$. Indeed, we can show that for most observables randomly chosen from $\mathcal{L}_{m=N}$, the corresponding $U$’s satisfy $U \in \mathcal{U}_{\{a_i\},\eta}$. This means that we can construct an operational measure consistent with the typicality argument if $N$-body observables are available, as stated in the following proposition.

**Proposition.** Consider randomly choosing an observable $\hat{O} = \sum_f G_f \hat{A}_f$ from $\mathcal{L}_N$, from which we obtain the corresponding $\{a_i\}$ and $U$. Then, we can show that (see Appendix II [41])

$$
\mathbb{P}_{\mathcal{L}_N}[U \notin \mathcal{U}_{\{a_i\},\eta}] \leq 2d \exp \left( -\frac{d d_{sh}^{-2\eta}}{72\pi^3} \right),
$$

where $d = \dim[\mathcal{H}] = S^N$. The right-hand side vanishes for sufficiently large $N$ when $\eta < \frac{1}{2}$.

This proposition suggests that most random observables chosen from $\mathcal{L}_N$ satisfy the ETH within the energy shell under the normalization $|\hat{O}|_{\text{op}} = 1$. In Appendix III [41], we show numerical results indicating that many-body correlations can satisfy the ETH (a similar numerical result was presented in Ref. [53]).

We note that we can make similar analyses for off-diagonal matrix elements $\hat{O}_{\alpha\beta} (E_\alpha \neq E_\beta)$, which are related to the Fourier transform of the autocorrelation function [54–56]. In Appendix V [41], we show that the magnitudes of off-diagonal matrix elements for most few-body observables fluctuate within the energy shell more than what the uniform Haar measure predicts. This can be proven from the fact that an operator written in the form of $i[\hat{A},\hat{H}]$ has atypical off-diagonal matrix elements. On the other hand, we can construct an operational measure consistent with the typicality argument if many-body observables are available.

We briefly comment on previous investigations on whether the typicality (or RMT) argument applies to a realistic setup. They mainly investigate the behavior of $U$ in light of the complexity of the Hamiltonian, motivated by analyses in semiclassical systems [5, 57]. When the Hamiltonian commutes with many local conserved quantities due to, e.g., many-body localization [58, 59], matrix elements of few-body observables are not typical (i.e., atypical) [55, 56, 60–66]. In contrast, the typicality argument has been conjectured to be applicable to a generic nonintegrable Hamiltonian [67] and few-body observables, for which the matrix elements are expected to be calculated by RMT within a sufficiently small energy shell [10]. For some statistics (e.g., variances of matrix elements), this conjecture has been numerically tested in Refs. [54, 55, 62, 63, 68]. On the other hand, our results show that the Haar measure cannot predict the maximum variation of $O_{\alpha\alpha}$ for most few-body or local observables in macroscopic systems (the latter requires a translation-invariant $\hat{H}$), when the width of the energy shell decreases at most polynomially with increasing the size of the system.

Conclusions and discussions. We have reexamined the typicality argument that relies on the unitary Haar measure by focusing on few-body observables. By considering an arbitrary few-body Hamiltonian (which can be local) and random few-body observables, we have shown that matrix elements do not behave typically for most few-body observables even if the energy width decays algebraically with increasing the size of the system (Theorem). We have also constructed an operational measure consistent with the typicality argument on diagonal matrix elements (Proposition). This is possible if many-body observables are available.

Our approach provides rigorous results without assuming any specific form of matrix elements. In fact, if we assume that all $O_{\alpha\alpha}$’s are written as $\mathcal{A}(E_\alpha)$ with a smooth function $\mathcal{A}$ of energy in the thermodynamic limit [50] and that $\frac{d\mathcal{A}(E)}{dE}$ is not exponentially small, the atypicality of diagonal matrix elements is expected. Namely, under such assumptions the maximum deviation will be $\sim \frac{d\mathcal{A}(E)}{dE} \times 2\Delta E$ in the thermodynamic limit, which is not exponentially small if $\Delta E \propto N^{-\eta}$. However, our proof of Theorem does not rely on these assumptions. Moreover, Theorem and Proposition show that the few-body property of observables is crucial in considering statistics of matrix elements, which was not addressed in previous literature [69]. Our results suggest that the above assumption for $\frac{d\mathcal{A}(E)}{dE}$ often seems to hold in numerics [60, 70] because few-body observables are mainly concerned [71].

Our results indicate that the typicality argument based on the Haar measure does not apply to realistic Hamiltonians and most few-body observables, if the width of the energy shell decreases at most polynomially. For the diagonal matrix elements, the typicality argument (see Appendix I for detail [41]) cannot be used to justify the ETH. However, we have not excluded the possibility that the maximum variation of diagonal matrix elements decreases algebraically with increasing the size of the system. If this is the case, the ETH still holds true and thermalization occurs in the thermodynamic limit. We also note that we cannot judge the validity of von Neumann’s original argument [7] on the basis of the present study. In fact, he took a coarse-grained procedure of the original macroscopic observables, which adds subextensive corrections to these observables. Such corrections are negligible for discussing thermalization in macroscopic systems, but make the inequality (2) inapplicable.

While we have mainly considered the maximum variation of diagonal matrix elements to investigate whether
the typicality argument can explain the ETH, our technique of random few-body observables may be applied to investigate other properties of thermalization. In particular, it is worthwhile to investigate whether the equilibration timescales of generic few-body observables are explicitly shown to differ from what the Haar measure predicts [72].

We are grateful to Zongping Gong and Tomohiro Shitara for reading the manuscript with valuable comments. We also thank Marcos Rigol and Takahiro Sagawa for fruitful discussions and comments. This work was supported by KAKENHI Grant No. JP26287088 from the Japan Society for the Promotion of Science, a Grant-in-Aid for Scientific Research on Innovative Areas Topological Materials Science (KAKENHI Grant No. JP15H05855), and the Photon Frontier Network Program from MEXT of Japan, and the Mitsubishi Foundation. R. H. was supported by the Japan Society for the Promotion of Science through Program for Leading Graduate Schools (ALPS) and JSPS fellowship (JSPS KAKENHI Grant No. JP17J03189).

[1] D. T. Haar, Rev. Mod. Phys. 27, 289 (1955).
[2] A. Peres, Phys. Rev. A 30, 504 (1984).
[3] R. V. Jensen and R. Shankar, Phys. Rev. Lett. 54, 1879 (1985).
[4] J. M. Deutsch, Phys. Rev. A 43, 2046 (1991).
[5] N. Linden, S. Popescu, A. J. Short, and A. Winter, Phys. Rev. E 79, 061103 (2009).
[6] G. Clos, D. Porras, U. Warring, and T. Schaetz, Phys. Rev. Lett. 117, 170401 (2016).
[7] C. Smith, A. B. Lee, P. Richerme, B. Neyenhuis, P. W. Hess, P. Hauke, M. Heyl, D. A. Huse, and C. Monroe, Nature Physics 12, 907 (2016).
[8] P. Reimann, M. Rigol, V. Dunjko, and M. Olshanii, Nature Physics 4, 854 (2008).
[9] L. F. Santos and M. Rigol, Phys. Rev. E 81, 036206 (2010).
[10] N. Linden, S. Popescu, A. J. Short, and A. Winter, Phys. Rev. E 79, 061103 (2009).
[11] C. Gogolin, M. P. Müller, and J. Eisert, Phys. Rev. Lett. 106, 040401 (2011).
[12] L. P. R. Gomes, M. Rigol, and J. Eisert, Phys. Rev. Lett. 110, 080402 (2012).
[13] L. P. R. Gomes, M. Rigol, and J. Eisert, Phys. Rev. Lett. 110, 080402 (2012).
[14] L. P. R. Gomes, M. Rigol, and J. Eisert, Phys. Rev. Lett. 110, 080402 (2012).
[15] L. P. R. Gomes, M. Rigol, and J. Eisert, Phys. Rev. Lett. 110, 080402 (2012).
[16] L. P. R. Gomes, M. Rigol, and J. Eisert, Phys. Rev. Lett. 110, 080402 (2012).
[17] L. P. R. Gomes, M. Rigol, and J. Eisert, Phys. Rev. Lett. 110, 080402 (2012).
[18] L. P. R. Gomes, M. Rigol, and J. Eisert, Phys. Rev. Lett. 110, 080402 (2012).
to satisfy the weak ETH. It is known that the weak ETH does not imply thermalization [73].

[41] See Supplemental Material for detailed derivations, numerical calculations, and generalization of the results, which includes Ref. [74].

[42] This definition is different from that of the typicality of pure states [75–78].

[43] Note that their argument essentially uses the same measure for matrix elements as ours to quantify the ETH (see also Appendix I).

[44] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, The European Physical Journal H 35, 173 (2010).

[45] M. Rigol and M. Srednicki, Phys. Rev. Lett. 108, 110601 (2012).

[46] E. P. Wigner, in Mathematical Proceedings of the Cambridge Philosophical Society, Vol. 47 (Cambridge Univ Press, 1951) pp. 790–798.

[47] T. A. Brody, J. Flores, J. B. French, P. A. Mello, and S. S. M. Wong, Rev. Mod. Phys. 53, 385 (1981).

[48] O. Bohigas, M. J. Giannoni, and C. Schmit, Phys. Rev. Lett. 52, 1 (1984).

[49] Some of the spectral statistics in semiclassically chaotic or disordered systems are proven to have the universality predicted by RMT using the nonlinear sigma model or the periodic-orbit theory [79–81].

[50] M. Srednicki, Journal of Physics A: Mathematical and General 32, 1163 (1999).

[51] This definition depends on the choice of \( \{ \hat{A}_f \} \) and \( P(\hat{G}) \). Our discussion holds true for any choice as long as the orthonormality, Hermiticity of \( \{ \hat{A}_f \} \) and the invariance of \( P(\hat{G}) \) under orthogonal transformations hold true.

[52] E. J. Torres-Herrera and L. F. Santos, Phys. Rev. E 89, 062110 (2014).

[53] P. Hosur and X.-L. Qi, Phys. Rev. E 93, 042138 (2016).

[54] E. Khatami, G. Pupillo, M. Srednicki, and M. Rigol, Phys. Rev. Lett. 111, 050403 (2013).

[55] W. Beugeling, R. Moessner, and M. Haque, Phys. Rev. E 91, 012144 (2015).

[56] D. J. Luitz and Y. Bar Lev, Phys. Rev. Lett. 117, 170404 (2016).

[57] M. V. Berry, Journal of Physics A: Mathematical and General 10, 2083 (1977).

[58] D. Basko, I. Aleiner, and B. Altshuler, Annals of physics 321, 1126 (2006).

[59] A. Pal and D. A. Huse, Phys. Rev. B 82, 174411 (2010).

[60] M. Rigol, Phys. Rev. Lett. 103, 100403 (2009).

[61] T. N. Ikeda, Y. Watanabe, and M. Ueda, Phys. Rev. E 87, 012125 (2013).

[62] R. Steinigeweg, J. Herbrych, and P. Prelovšek, Phys. Rev. E 87, 012118 (2013).

[63] W. Beugeling, R. Moessner, and M. Haque, Phys. Rev. E 89, 042112 (2014).

[64] V. Alba, Phys. Rev. B 91, 155123 (2015).

[65] R. Hamazaki, T. N. Ikeda, and M. Ueda, Phys. Rev. E 93, 032116 (2016).

[66] Z. Lan and S. Powell, arXiv preprint arXiv:1706.02601 (2017).

[67] As explicit counterexamples, some nonintegrable systems have been proposed for which the ETH does not hold even without localization [65, 82, 83].

[68] R. Mondaini and M. Rigol, Phys. Rev. E 96, 012157 (2017).

[69] In other words, our work rigorously proves and refines the observation in Ref. [10] that the prediction of the RMT seems inapplicable for a non-small energy shell.

[70] H. Kim, T. N. Ikeda, and D. A. Huse, Phys. Rev. E 90, 052105 (2014).

[71] We note that in Floquet nonintegrable systems, where the energy is not conserved and our theorem does not apply, \( \frac{dA(E)}{dE} \) (\( E \) represents the quasienergy) is found to be small in general even for few-body observables [84, 85].

[72] P. Reimann, Nature communications 7 (2016).

[73] G. Birolì, C. Kollath, and A. M. Läuchli, Phys. Rev. Lett. 105, 250401 (2010).

[74] S. Goldstein, J. L. Lebowitz, C. Mastrodonato, R. Tumulka, and N. Zanghì, Phys. Rev. E 81, 011109 (2010).

[75] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, Phys. Rev. Lett. 96, 050403 (2006).

[76] S. Popescu, A. J. Short, and A. Winter, Nature Physics 2, 754 (2006).

[77] P. Reimann, Phys. Rev. Lett. 99, 160404 (2007).

[78] H. Tasaki, Journal of Statistical Physics 163, 937 (2016).

[79] A. D. Mirlin, Physics Reports 326, 259 (2000).

[80] S. Müller, S. Heusler, P. Braun, F. Haake, and A. Altland, Phys. Rev. Lett. 93, 014103 (2004).

[81] F. Haake, Quantum signatures of chaos, Vol. 54 (Springer Science & Business Media, 2010).

[82] N. Shiraishi and T. Mori, Phys. Rev. Lett. 119, 030601 (2017).

[83] T. Mori and N. Shiraishi, arXiv preprint arXiv:1707.05921 (2017).

[84] L. DAlessio and M. Rigol, Physical Review X 4, 041048 (2014).

[85] A. Lazarides, A. Das, and R. Moessner, Physical Review E 90, 012110 (2014).
Supplementary Material for “Atypicality of Most Few-Body Observables”

Ryusuke Hamazaki

Department of Physics, University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Masahito Ueda

Department of Physics, University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan and
RIKEN Center for Emergent Matter Science (CEMS), Wako 351-0198, Japan

(Dated: February 27, 2018)

PACS numbers: 05.30.-d, 03.65.-w
I. TYPICALITY OF DIAGONAL MATRIX ELEMENTS

We begin by recalling that diagonal matrix elements $O_{\alpha \alpha}$ within the energy shell ($|E_{\alpha} - E| \leq \Delta E$) are determined by a unitary matrix $U$ through $O_{\alpha \alpha} = \sum_{i=1}^{d_{sh}} a_i |U_{i\alpha}|^2$; if eigenvalues $\{a_i\}$ are fixed. We define $U_{(a_{i}, \eta)}$ as a set of all $U$’s that lead to the inequality \[ \max_{|E_{\alpha} - E|, |E_{\beta} - E| \leq \Delta E} |O_{\alpha \alpha} - O_{\beta \beta}| \leq \|\hat{O}\|_{op} d_{sh}^{-\eta} \] for given $\{a_i\}$. Here $\|\hat{O}\|_{op}$ denotes an operator norm and $\eta > 0$. This inequality means that the maximum difference among diagonal elements within the energy shell is exponentially small in $N$, which also implies the ETH of $\hat{O} / \|\hat{O}\|_{op}$.

There have been a number of attempts [1–4] to justify the ETH by proving $U \in U_{(a_{i})}$ for almost all $U$’s with respect to the Haar measure. We present here a slightly modified version of Reimann’s result [4] to compare it with our results in the main text:

\[ \mathbb{P}_U[U \notin U_{(a_{i}), \eta}] \leq 2d_{sh} \exp \left[ -\frac{d_{sh}^{1-2\eta}}{72\pi^3} \right], \quad (1) \]

where $\mathbb{P}_U$ denotes a probability distribution with respect to the unitary Haar measure.

**Proof.** In Ref. [4], Reimann has shown that (see Eq. (40) in the supplementary material of the reference)

\[ \mathbb{P}_U \left[ \max_{|E_{\alpha} - E|, |E_{\beta} - E| \leq \Delta E} |O_{\alpha \alpha} - O_{\beta \beta}| > \epsilon \right] \leq 2d_{sh} \exp \left[ -\frac{2\epsilon^2 d_{sh}}{9\pi^3 \Delta_{O}^2} \right], \quad (2) \]

where $\Delta_{O} = \max_i a_i - \min_i a_i$ and $O_{sh} = \frac{1}{d_{sh}} \sum_{|E_{\alpha} - E| \leq \Delta E} O_{\alpha \alpha}$. Since

\[ \frac{1}{2} \max_{|E_{\alpha} - E|, |E_{\beta} - E| \leq \Delta E} |O_{\alpha \alpha} - O_{\beta \beta}| \leq \max_{|E_{\alpha} - E| \leq \Delta E} |O_{\alpha \alpha} - O_{\beta \beta}|, \quad (3) \]

we obtain

\[ \mathbb{P}_U \left[ \max_{|E_{\alpha} - E|, |E_{\beta} - E| \leq \Delta E} |O_{\alpha \alpha} - O_{\beta \beta}| > 2\epsilon \right] \leq \mathbb{P}_U \left[ \max_{|E_{\alpha} - E| \leq \Delta E} |O_{\alpha \alpha} - O_{\beta \beta}| > \epsilon \right], \quad (4) \]

where we use the fact that $\mathbb{P}[a > c] \leq \mathbb{P}[b > c]$ for $a \leq b$. Substituting $2\epsilon = \|\hat{O}\|_{op} d_{sh}^{-\eta}$ into (4), we obtain

\[ \mathbb{P}_U \left[ \max_{|E_{\alpha} - E|, |E_{\beta} - E| \leq \Delta E} |O_{\alpha \alpha} - O_{\beta \beta}| > \|\hat{O}\|_{op} d_{sh}^{-\eta} \right] \leq \mathbb{P}_U \left[ \max_{|E_{\alpha} - E| \leq \Delta E} |O_{\alpha \alpha} - O_{\beta \beta}| > \frac{1}{2} \|\hat{O}\|_{op} d_{sh}^{-\eta} \right] \leq 2d_{sh} \exp \left[ -\frac{\|\hat{O}\|_{op}^2 d_{sh}^{1-2\eta}}{18\pi^3 \Delta_{O}^2} \right]. \quad (5) \]

Finally, using $\Delta_{O} \leq 2\|\hat{P}_{sh} \hat{O} \hat{P}_{sh}\|_{op} \leq 2\|\hat{O}\|_{op}$, we obtain

\[ \mathbb{P}_U \left[ \max_{|E_{\alpha} - E|, |E_{\beta} - E| \leq \Delta E} |O_{\alpha \alpha} - O_{\beta \beta}| > \|\hat{O}\|_{op} d_{sh}^{-\eta} \right] \leq 2d_{sh} \exp \left[ -\frac{d_{sh}^{1-2\eta}}{72\pi^3} \right], \quad (6) \]
II. PROOF OF THE PROPOSITION IN THE MAIN TEXT

We first show that a randomly chosen observable \( \hat{G} = \sum_f G_f \hat{A}_f \) from \( \mathcal{L}_N \) has eigenstates that are uniformly distributed with respect to the Haar measure. For an arbitrary \( S^N \times S^N \) unitary transformation \( \hat{R} \), we obtain

\[
\hat{R} \hat{G} \hat{R}^\dagger = \sum_{f=1}^{S^{2N}} G_f \hat{R} \hat{A}_f \hat{R}^\dagger = \sum_{f=1}^{S^{2N}} \sum_{g=1}^{S^{2N}} \mathcal{R}_{fg} \hat{A}_g
\]

where \( \mathcal{R}_{fg} \) is defined through the operator expansion \( \hat{R} \hat{A}_f \hat{R}^\dagger = \sum_{g=1}^{S^{2N}} \mathcal{R}_{fg} \hat{A}_g \). We define \( \hat{G} = \sum_{g=1}^{S^{2N}} G_g \mathcal{R}_{gf} \). From the normalization condition of \( \hat{A}_f \), namely \( \text{Tr}[\hat{A}_f \hat{A}_g] = S^N \delta_{fg} \), we obtain \( \sum_h \mathcal{R}_{fh} \mathcal{R}_{gh} = \delta_{fg} \) by considering the operator expansion of \( \hat{R} \hat{A}_f \hat{R}^\dagger = \hat{A}_g \hat{R}^\dagger \). Moreover, from the Hermiticity \( \hat{A}_f^\dagger = \hat{A}_f \), we obtain \( \mathcal{R}_{fg} = \mathcal{R}_{fg}^* \) by considering the operator expansion of \( \hat{R} \hat{A}_f \hat{R}^\dagger \). Thus, we can show that \( \mathcal{R} \) is an \( S^{2N} \times S^{2N} \) orthogonal matrix. Then, if we pick up observables randomly from \( \mathcal{L}_N \), the probabilities of choosing \( \hat{G} \) and \( \hat{R} \hat{G} \hat{R}^\dagger \) are equal due to the invariance assumption \( P(\hat{G}) = P(\mathcal{R} \hat{G}) \). Consequently, if we diagonalize a randomly chosen \( \hat{G} \), its eigenstates are uniformly distributed with respect to the unitary Haar measure.

Similarly to Eq. (1), we can show the bound on the right-hand side in Eq. (7) in the main text. By considering the unitary Haar measure for the entire Hilbert space, we obtain

\[
\mathbb{P}_{\mathcal{L}_N} \left[ \max_{\alpha, \beta} |O_{\alpha \alpha} - O_{\beta \beta}| > ||\hat{O}||_{op} d_{sh}^{-\eta} \right] \leq 2d \exp \left[ -\frac{d d_{sh}^{-2\eta}}{72 \pi^3} \right].
\]

which completes the proof of the inequality in (1).
Since \( \max_{\alpha,\beta} |O_{\alpha\alpha} - O_{\beta\beta}| \geq \max_{|E_{\alpha} - E|, |E_{\beta} - E| \leq \Delta E} |O_{\alpha\alpha} - O_{\beta\beta}| \), we obtain

\[
P_L N \left[ \max_{|E_{\alpha} - E|, |E_{\beta} - E| \leq \Delta E} |O_{\alpha\alpha} - O_{\beta\beta}| > ||\hat{O}||_{op \eta} \right] \leq \mathbb{P}_L N \left[ \max_{\alpha,\beta} |O_{\alpha\alpha} - O_{\beta\beta}| > ||\hat{O}||_{op \eta} \right].
\] (9)

This completes the proof of Eq. (7) in the main text.

III. EIGENSTATE THERMALIZATION HYPOTHESIS FOR MANY-BODY CORRELATIONS

In this Appendix, we numerically show that the eigenstate thermalization hypothesis (ETH) is expected to hold true even for many-body correlations in a one-dimensional non-integrable spin-1/2 system.

In Fig. 1, we show the eigenstate expectation values (EEVs) \( \langle E_{\alpha}|\hat{O}_N|E_{\alpha}\rangle \) for integrable and nonintegrable systems. Here, many-body correlations \( \hat{O}_N \) are defined as

\[
\hat{O}_N = \prod_{l=1}^{N} \hat{\sigma}_l^z,
\] (10)

where \( \hat{\sigma}_l^z \) is the \( z \) component of the Pauli operator at site \( l \).

For an integrable system, we take a transverse-field Ising model with the open boundary condition whose Hamiltonian can be written as

\[
\hat{H} = -\sum_{l=1}^{N-1} J \hat{\sigma}_l^z \hat{\sigma}_{l+1}^z - \sum_{l=1}^{N} h' \hat{\sigma}_l^x,
\] (11)

where we take \( J = 1 \) and \( h' = -1.05 \). For a nonintegrable system, we take a Hamiltonian

\[
\hat{H} = -\sum_{l=1}^{N-1} J (1 + \epsilon_l) \hat{\sigma}_l^z \hat{\sigma}_{l+1}^z - \sum_{l=1}^{N} h' \hat{\sigma}_l^z - \sum_{l=1}^{N} h \hat{\sigma}_l^z,
\] (12)

where \( h = 0.5 \) and a random variable \( \epsilon_l \) is uniformly chosen from \([-0.1, 0.1]\) at each bond.

Figure 1 shows that the fluctuations of the EEVs rapidly decrease with increasing \( N \) for nonintegrable systems, whereas they remain large for integrable systems. This result implies that the ETH does (does not) hold true for nonintegrable (integrable) systems, even for many-body correlations given by Eq. (10). We note, however, that Proposition in the main text holds true regardless of whether the system is integrable or not. From this proposition, we expect that the ETH holds true for more complex many-body observables even for integrable systems.
FIG. 1. $N$-dependences of eigenstate expectation values (EEVs) of $\hat{O}_N$ for integrable (upper row) and nonintegrable (lower row) systems with $N = 8$ (left), $N = 10$ (middle), and $N = 12$ (right). The fluctuations of the EEVs decrease with increasing $N$ for nonintegrable systems, whereas they remain large for integrable systems.

IV. ATYPICALITY OF MOST LOCAL OBSERVABLES

We are often interested in the property of spatially local observables, rather than general few-body observables. We can actually show atypicality of $O_{\alpha\alpha}$ for most of such local observables in a manner similar to the case with Theorem in the main text, if the Hamiltonian is local and translationally invariant. In the following, we consider spins on one-dimensional $N$ lattice sites ($x = 1, \ldots, N$) for simplifying the notation (generalization to higher dimensions is straightforward).

We take a subsystem $S_l$ composed of $l$ neighboring lattices, which we call $x = 1, 2, \ldots, l$ (note that we can start from $x = 1$ because we will consider a translation-invariant sys-
tem below). We assume that $l \ll N$ does not depend on $N$, which makes $S_l$ a local subsystem. We define $\mathcal{H}_{S_l}$ and $\mathcal{L}^{\text{loc}}_l \equiv \mathcal{L}(\mathcal{H}_{S_l})$ as the Hilbert space of $S_l$ and the operator space acting on $S_l$, respectively. An orthonormal basis set for $\mathcal{L}^{\text{loc}}_l$ can be written as $B^{\text{loc}}_l = \{ \hat{\lambda}^{x}_{f} | 0 \leq \mu_{x} \leq S^2 - 1 \} \equiv \{ \hat{\lambda}^{x}_{f} \}_{f=1}^{r}$, where $r = S^{2l}$. We choose the orthonormality condition as $\text{Tr}_{S_l}[\hat{\lambda}^{x}_{f} \hat{\lambda}^{y}_{g}] = S^{l} \delta_{fg}$.

Next we define random local observables from $\mathcal{L}^{\text{loc}}_l$ as follows:

**Definition 1** (Randomly chosen local observables from $\mathcal{L}^{\text{loc}}_l$). We take an observable $\hat{G} \in \mathcal{L}^{\text{loc}}_l$ expressed as

$$
\hat{G} = \sum_{f=1}^{r} g_{f} \hat{\lambda}^{x}_{f},
$$

where real variables $\vec{g} = (g_{1}, \cdots, g_{r})$ are randomly chosen according to an arbitrarily specified probability distribution $P^{\text{loc}}(\vec{g})$. When $P^{\text{loc}}(\vec{g})$ is invariant under an arbitrary $r \times r$ orthogonal transformation, we call $\hat{G}$ an observable randomly chosen from $\mathcal{L}^{\text{loc}}_l$.

We consider an $l'$-local, translation-invariant Hamiltonian $\hat{H}$ in the following sense. It is written as $\hat{H} = \sum_{i=1}^{N} \hat{h}_{i,l'}$, where $\hat{h}_{i,l'}$ only acts on $x = 1, 2, \cdots, l'$ and $T[\hat{h}_{i,l'}] = \hat{h}_{i+1,l'}$ ($T$ is the translation of one lattice).

Under this setup, we obtain the following theorem:

**Theorem 1** (Atypicality of most local observables). Let us consider an $l'$-local translation-invariant Hamiltonian, and assume that $N$ is sufficiently large and that $l$ ($l' \leq l \ll N$) is independent of $N$. Suppose that we randomly choose an observable $\hat{O} = \sum_{f} g_{f} \hat{\lambda}^{x}_{f}$ from $\mathcal{L}^{\text{loc}}_l$, from which we obtain the corresponding $\{a_{i}\}$ and $U$. Then,

$$
\mathbb{P}_{\mathcal{L}^{\text{loc}}_l} [ U \in \mathcal{U}_{\{a_{i}\},\eta} ] \leq \frac{\sqrt{\pi r N ||\hat{h}||_{\text{op}} \Lambda}}{2\Delta E} \frac{\Gamma \left( \frac{5}{2} \right)}{\Gamma \left( \frac{r+1}{2} \right)} d_{\text{sh}}^{-\eta},
$$

where $\mathbb{P}_{\mathcal{L}^{\text{loc}}_l}$ denotes a probability with respect to $P^{\text{loc}}(\vec{g})$, and $\Lambda = \max_{f} ||\hat{\lambda}_{f}||_{\text{op}} \leq S^{\frac{1}{2}} = r^{1/4}$. When $||\hat{h}||_{\text{op}}$ does not grow exponentially in $N$, the left-hand side vanishes for large $N$. Note that the assumption of the scaling $\Delta E \propto N^{-\nu}$ is sufficient to bound the right-hand side.

*Proof of Theorem 1.* The proof is similar to that of Theorem in the main text. We first note that $\hat{h}_{i,l'} \in \mathcal{L}^{\text{loc}}_l$ and that $\hat{h}_{i,l'}$ satisfies the following condition:

$$
(\hat{h}_{i,l'})_{\gamma\gamma} - (\hat{h}_{i,l'})_{\delta\delta} = \frac{1}{N} \sum_{i} [(\hat{h}_{i,l'})_{\gamma\gamma} - (\hat{h}_{i,l'})_{\delta\delta}]
$$

$$
= \frac{1}{N} [(\hat{H})_{\gamma\gamma} - (\hat{H})_{\delta\delta}] = \frac{2\Delta E}{N},
$$

6
where $\gamma$ and $\delta$ are labels of maximum/minimum eigenvalues within the energy shell. Define
\[
\gamma_f = (\hat{\Lambda}^\text{loc}_f)_{\gamma\gamma} - (\hat{\Lambda}^\text{loc}_f)_{\delta\delta}.
\]
Then the expansion $\hat{h}_{1,f} = \sum_{f=1}^r h_f \hat{\Lambda}^\text{loc}_f$ leads to
\[
\tilde{h} \cdot \tilde{\gamma} = \frac{2\Delta E}{N},
\]
where $\tilde{h} = (h_1, \cdots, h_r)$ and $\tilde{\gamma} = (\gamma_1, \cdots, \gamma_r)$. Since $|\tilde{h}| = \sqrt{\sum_{i} [h_i]^2} \leq ||\hat{h}_{1,f}||_{\text{op}} \equiv ||\hat{h}||_{\text{op}}$ (note that $||\hat{h}_{i,f}||_{\text{op}} = ||\hat{h}||_{\text{op}}$ for all $i$), we obtain
\[
|\tilde{\gamma}| \geq \frac{2\Delta E}{N||\hat{h}||_{\text{op}}}. \tag{16}
\]

Next, we evaluate the left-hand side of Eq. (14). Since $\max_{|E_{\alpha} - E|, |E_{\beta} - E| \leq \Delta E} |\mathcal{O}_{\alpha\alpha} - \mathcal{O}_{\beta\beta}| \geq |\tilde{g} \cdot \tilde{\gamma}|$, we obtain
\[
\mathbb{P}_{L_i}^{\text{loc}} \left[ \max_{|E_{\alpha} - E|, |E_{\beta} - E| \leq \Delta E} |\mathcal{O}_{\alpha\alpha} - \mathcal{O}_{\beta\beta}| \leq ||\hat{O}||_{\text{op}} \epsilon \right] \leq \mathbb{P}_{L_i}^{\text{loc}} \left[ |\tilde{g} \cdot \tilde{\gamma}| \leq ||\hat{O}||_{\text{op}} \epsilon \right]. \tag{17}
\]

Next, denoting the angle between $\tilde{g}$ and $\tilde{\gamma}$ by $\phi$, we obtain
\[
\mathbb{P}_{L_i}^{\text{loc}} \left[ |\tilde{g} \cdot \tilde{\gamma}| \leq ||\hat{O}||_{\text{op}} \epsilon \right] \leq \mathbb{P}_{L_i}^{\text{loc}} \left[ |\cos \phi| \leq \frac{\sqrt{\pi r} N ||\hat{h}||_{\text{op}} \Lambda \epsilon}{2 \Delta E} \right]
\]
\[
\leq \frac{\sqrt{\pi r} N ||\hat{h}||_{\text{op}} \Lambda \epsilon}{2 \Delta E} \frac{\Gamma \left( \frac{r}{2} \right)}{\Gamma \left( \frac{r+1}{2} \right)}. \tag{18}
\]

Here, in deriving the second line, we have used $||\hat{O}||_{\text{op}} \leq \Lambda ||\tilde{g}|| \sqrt{r}$ and the invariance condition of $P^{\text{loc}}(\tilde{g})$ (see also the main text). Combining (17) and (18), and taking $\epsilon = d_{sh}^{-n}$, the left-hand side of (17) becomes $\mathbb{P}_{L_i}^{\text{loc}}[U \in U_{(a_1),n}]$, which completes the proof of the theorem. \qed

From (17) and (18), we can further estimate how slowly the maximum variation of diagonal matrix elements decays. To do this, we take $\Delta E = ||\hat{h}||_{\text{op}} N^{-p}$ ($-1 < p$). Since $\frac{\sqrt{\pi r} \Lambda}{\Gamma \left( \frac{r}{2} \right)}$ does not depend on $N$, we can take $\epsilon = N^{-z}$ ($z > 1 + p$) for the left-hand side of (17) to vanish in the thermodynamic limit. We note that this does not contradict the result of the weak ETH [5]. The weak ETH states that the variance of $\mathcal{O}_{\alpha\alpha}$ within the energy shell decreases faster than $N^{-1}$ if the equilibrium correlation function of $\hat{O}$ rapidly decreases with the distance. Since the maximum variation is always larger than the variance, our result and the weak ETH can be compatible (indeed, the weak ETH is known to hold for integrable systems, for which the strong ETH breaks down).

\section{Off-Diagonal Matrix Elements}

We here analyze off-diagonal matrix elements. For energy eigenstates $|E_{\alpha}\rangle, |E_{\beta}\rangle \in \mathcal{H}_{\text{sh}}$, we obtain $E - \Delta E \leq \frac{E_{\alpha} + E_{\beta}}{2} \leq E + \Delta E$ and $-2\Delta E \leq E_{\beta} - E_{\alpha} = \omega_{\alpha\beta} \leq 2\Delta E$, where we set
FIG. 2. Region I and II for off-diagonal matrix elements. Region I satisfies $E - 2\delta E < E_\alpha < E$ and $E \leq E_\beta \leq E + 2\delta E$, where $\delta E (\ll \Delta E)$ is a small constant. Region II satisfies $E - \Delta E < E_\alpha < E - \Delta E + 2\delta E$ and $E + \Delta E - 2\delta E \leq E_\beta \leq E + \Delta E$.

$h = 1$. If the off-diagonal matrix elements have almost the same order of magnitude over the energy shell $H_{sh}$, no characteristic timescale of thermalization for $\hat{O}$ should appear after $\sim \frac{1}{2\Delta E}$.

We define the following regions I and II for matrix elements (see Fig.2). Region I satisfies $E - 2\delta E < E_\alpha < E$, $E \leq E_\beta \leq E + 2\delta E$, where $\delta E (\ll \Delta E)$ is a small constant. We denote the numbers of energy eigenstates $|E_\alpha\rangle$ and $|E_\beta\rangle$ that satisfy this condition by $p_I$ and $q_I$, respectively. Then the total number of relevant matrix elements is $p_I q_I$. Note that $0 < \omega_{\alpha\beta} < 4\delta E$ is satisfied. Similarly, region II satisfies $E - \Delta E < E_\alpha < E - \Delta E + 2\delta E$ and $E + \Delta E - 2\delta E \leq E_\beta \leq E + \Delta E$. We denote the numbers of energy eigenstates $|E_\alpha\rangle$ and $|E_\beta\rangle$ that satisfy this condition by $p_{II}$ and $q_{II}$, respectively. Then the total number of relevant matrix elements is $p_{II} q_{II}$. Note that the condition $2\Delta E - 4\delta E < \omega_{\alpha\beta} < 2\Delta E$ is satisfied.
As a quantity of our interest, we define the spectral average of off-diagonal matrix elements for each region as follows:

\[ \langle O_{\text{off}}^2 \rangle_I = \frac{1}{p_I q_I} \sum_{\alpha} |O_{\alpha\beta}|^2, \]  
(19)

\[ \langle O_{\text{off}}^2 \rangle_{II} = \frac{1}{p_{II} q_{II}} \sum_{\alpha} |O_{\alpha\beta}|^2. \]  
(20)

Here each sum is taken over all matrix elements in I or II. In the following discussions, we consider \( d_{\text{sh}} \langle O_{\text{off}}^2 \rangle_{I/II} \), since \( \langle O_{\text{off}}^2 \rangle_{I/II} \) is expected to decrease as \( \sim d_{\text{sh}}^{-1} \) [6] (see Ref. [7] for an exception).

First, we can prove a statement similar to Eq. (1) for the off-diagonal spectral average. For this purpose, we recall that matrix elements \( O_{\alpha\beta} \) within the energy shell are determined by a unitary matrix \( U \) through \( O_{\alpha\beta} = \sum_{i=1}^{d_{\text{sh}}} a_i U_{\alpha i} U_{\beta i}^* \) for fixed eigenvalues \( \{a_i\} \). Similarly, \( \langle O_{\text{off}}^2 \rangle_I \) and \( \langle O_{\text{off}}^2 \rangle_{II} \) are also calculated from \( U \). Then we have the following proposition.

**Proposition 1** (Typicality with respect to the Haar measure). We define \( U'_{\{a_i\}, \eta} \) as a set of all \( U \)'s that lead to the inequality \( |d_{\text{sh}} \langle O_{\text{off}}^2 \rangle_I - d_{\text{sh}} \langle O_{\text{off}}^2 \rangle_{II}| \leq ||\hat{O}||_{\text{op}} d_{\text{sh}}^{-\eta} \) for given \( \{a_i\} \), where \( \eta > 0 \). We then have

\[ P_U [U \notin U'_{\{a_i\}, \eta}] \leq \frac{4B_I}{d_{\text{sh}}^{2\eta} d_I} + \frac{4B_{II}}{d_{\text{sh}}^{2\eta} d_{II}}, \]  
(21)

where \( B_I \) and \( B_{II} \) are some constants, \( d_I = \min\{p_I, q_I\} \), and \( d_{II} = \min\{p_{II}, q_{II}\} \). If \( d_I, d_{II} \gg d_{\text{sh}}^{2\eta} \) in the thermodynamic limit, the right-hand side vanishes in this limit. This means that the variations of typical magnitudes of off-diagonal matrix elements decrease exponentially within an energy shell.

**Proof.** First we show

\[ P_U \left[ |d_{\text{sh}} \langle O_{\text{off}}^2 \rangle_I - d_{\text{sh}} \langle O_{\alpha\beta} \rangle^2| > ||\hat{O}||_{\text{op}} d_{\text{sh}}^{-\eta} \right] \leq \frac{B_I}{d_{\text{sh}}^{2\eta} d_I} \]  
(22)

and a similar inequality for II, where the overline denotes the average with respect to the Haar measure. We begin by calculating the ensemble average of the spectral average over the Haar measure:

\[ d_{\text{sh}} \langle O_{\text{off}}^2 \rangle_I = \frac{d_{\text{sh}}}{p_I q_I} \sum_{\alpha} |O_{\alpha\beta}|^2 \]

\[ = d_{\text{sh}} |O_{\alpha\beta}|^2. \]  
(23)
Here we have used the fact that $|\mathcal{O}_{\alpha\beta}|^2$ is independent of $\alpha$ and $\beta$ in the energy shell. Next we consider the ensemble variance of the spectral average. When $d_1(<d_{sh})$ is large enough, we obtain
\begin{equation}
\left( d_{sh} \langle \mathcal{O}_{\text{off}}^2 \rangle_I - d_{sh} \langle \mathcal{O}_{\text{off}}^2 \rangle_{II} \right)^2 \leq B_1 \| \hat{O} \|_{\text{op}}^4 d_1^{-1},
\end{equation}
where $B_1$ is a constant. The proof is given in a similar manner as in Ref. [7]. By Chebyshev’s inequality, we obtain Eq. (22). A similar proof can be made for the region II.

Now we show Proposition 1. Since
\begin{equation}
|d_{sh} \langle \mathcal{O}_{\text{off}}^2 \rangle_I - d_{sh} \langle \mathcal{O}_{\text{off}}^2 \rangle_{II}| \leq 2 \max \left\{ |d_{sh} \langle \mathcal{O}_{\text{off}}^2 \rangle_I - d_{sh} \langle \mathcal{O}_{\alpha\beta} \rangle^2 |, |d_{sh} \langle \mathcal{O}_{\text{off}}^2 \rangle_{II} - d_{sh} \langle \mathcal{O}_{\alpha\beta} \rangle^2 | \right\},
\end{equation}
we obtain
\begin{align}
\mathbb{P}_U \left[ |d_{sh} \langle \mathcal{O}_{\text{off}}^2 \rangle_I - d_{sh} \langle \mathcal{O}_{\text{off}}^2 \rangle_{II}| > \| \hat{O} \|_{\text{op}}^2 d_{sh}^{-n} \right] & \leq \mathbb{P}_U \left[ 2 \max \left\{ |d_{sh} \langle \mathcal{O}_{\text{off}}^2 \rangle_I - d_{sh} \langle \mathcal{O}_{\alpha\beta} \rangle^2 |, |d_{sh} \langle \mathcal{O}_{\text{off}}^2 \rangle_{II} - d_{sh} \langle \mathcal{O}_{\alpha\beta} \rangle^2 | \right\} > \| \hat{O} \|_{\text{op}}^2 d_{sh}^{-n}/2 \right] \\
& \leq \mathbb{P}_U \left[ |d_{sh} \langle \mathcal{O}_{\text{off}}^2 \rangle_I - d_{sh} \langle \mathcal{O}_{\alpha\beta} \rangle^2 | > \| \hat{O} \|_{\text{op}}^2 d_{sh}^{-n}/2 \right] + \mathbb{P}_U \left[ |d_{sh} \langle \mathcal{O}_{\text{off}}^2 \rangle_{II} - d_{sh} \langle \mathcal{O}_{\alpha\beta} \rangle^2 | > \| \hat{O} \|_{\text{op}}^2 d_{sh}^{-n}/2 \right] \\
& \leq \frac{4B_1}{d_{sh}^{-2n}d_1} + \frac{4B_{II}}{d_{sh}^{-2n}d_{II}},
\end{align}
where we have used $\mathbb{P}[\max\{a,b\} > \epsilon] \leq \mathbb{P}[a > \epsilon] + \mathbb{P}[b > \epsilon]$.

Just as we have done for diagonal matrix elements, we can show that most few-body observables are atypical in the sense that the corresponding $U$ does not belong to $\mathcal{U}_n$. This can be stated in the form of the following theorem.

**Theorem 2.** Suppose that there exists an observable $\hat{A} \in \mathcal{L}_m$ such that $d_{sh} \langle A_{\text{off}}^2 \rangle_{II}$ and $\| \hat{A} \|_{\text{op}}^{-1}$ decrease no faster than polynomials in $N$. We also assume that the Hamiltonian is a $k$-body operator ($k \leq m$) and $\| \hat{H} \|_{\text{op}}$ does not increase exponentially in $N$.

For sufficiently small $\delta E$, we can show that
\begin{equation}
\mathbb{P}_{\mathcal{L}_m} \left[ U \in \mathcal{U}_{(a),n} \right] \leq \sqrt{\frac{2\pi n^{\frac{3}{2}} \| \hat{B} \|_{\text{op}} \Lambda \Gamma \left( \frac{n}{2} \right)}{\zeta_{\text{od}}}} d_{sh}^{-\frac{n}{2}}.
\end{equation}
Here $\hat{B}$ is either $\hat{A}$ or $i[\hat{H}, \hat{A}]$, and $\zeta_{\text{od}}$ is some constant that decreases no faster than polynomials in $N$. The right-hand side vanishes in the thermodynamic limit when $m$ does not depend on $N$. This means the following: the typical homogeneous variance inside the
energy shell, which is predicted by the unitary Haar measure, cannot be observed for the off-diagonal matrix elements of most few-body observables.

Proof. We first seek for an observable \( \hat{B} \in \mathcal{L}_m \) that satisfies the following condition:

\[
|d_{sh} \langle B_{\text{off}}^2 \rangle_I - d_{sh} \langle B_{\text{off}}^2 \rangle_{II}| \geq \xi_{\text{od}},
\]

where \( \xi_{\text{od}} \) does not decrease faster than polynomials in \( N \). We show that either \( \hat{A} \) or \( i[\hat{H}, \hat{A}] \) satisfy this condition.

Let us take a positive constant \( c \). If \( |d_{sh} \langle A_{\text{off}}^2 \rangle_{II} - d_{sh} \langle A_{\text{off}}^2 \rangle_I| \geq c \), we can take \( \hat{B} = \hat{A} \) and \( \xi_{\text{od}} = c \); otherwise, we take \( \hat{B} = i[\hat{H}, \hat{A}] \). Then

\[
|d_{sh} \langle B_{\text{off}}^2 \rangle_I - d_{sh} \langle B_{\text{off}}^2 \rangle_{II}| = \frac{d_{sh}}{p_{II}q_{II}} \sum_i (E_\alpha - E_\beta)^2 |A_{\alpha\beta}|^2 - \frac{d_{sh}}{p_{II}q_{II}} \sum_{II} (E_\alpha - E_\beta)^2 |A_{\alpha\beta}|^2.
\]

We note that

\[
\frac{d_{sh}}{p_{II}q_{II}} \sum_{II} (E_\alpha - E_\beta)^2 |A_{\alpha\beta}|^2 \geq 4(\Delta E - 2\delta E)^2 d_{sh} \langle A_{\text{off}}^2 \rangle_{II},
\]

\[
\frac{d}{p_{II}q_{II}} \sum_i (E_\alpha - E_\beta)^2 |A_{\alpha\beta}|^2 \leq 16\delta E^2 d_{sh} \langle A_{\text{off}}^2 \rangle_I \leq 16\delta E^2(c + d_{sh} \langle A_{\text{off}}^2 \rangle_{II}).
\]

Thus, if we take \( \delta E \) such that

\[
\delta E^2 \leq \frac{\Delta E^2 d_{sh} \langle A_{\text{off}}^2 \rangle_{II}}{64(c + d_{sh} \langle A_{\text{off}}^2 \rangle_{II})} \left( \leq \frac{\Delta E^2}{64} \right)
\]

for a fixed \( \Delta E \), we obtain

\[
|d_{sh} \langle B_{\text{off}}^2 \rangle_I - d \langle B_{\text{off}}^2 \rangle_{II}| = \frac{d_{sh}}{p_{II}q_{II}} \sum_i (E_\alpha - E_\beta)^2 |A_{\alpha\beta}|^2 - \frac{d_{sh}}{p_{II}q_{II}} \sum_{II} (E_\alpha - E_\beta)^2 |A_{\alpha\beta}|^2
\]

\[\geq 4(\Delta E - 2\delta E)^2 d_{sh} \langle A_{\text{off}}^2 \rangle_{II} - 16\delta E^2(c + d_{sh} \langle A_{\text{off}}^2 \rangle_{II})\]

\[\geq \frac{9}{4}\Delta E^2 d_{sh} \langle A_{\text{off}}^2 \rangle_{II} - \frac{1}{4}\Delta E^2 d_{sh} \langle A_{\text{off}}^2 \rangle_{II}\]

\[= 2\Delta E^2 d_{sh} \langle A_{\text{off}}^2 \rangle_{II}.\]

Since the last term decreases no faster than polynomials in \( N \), we can take \( \xi_{\text{od}} = 2\Delta E^2 d_{sh} \langle A_{\text{off}}^2 \rangle_{II} \).

We note that for \( \hat{B} = i[\hat{H}, \hat{A}] \), \( \sqrt{d_{sh} \langle B_{\text{off}}^2 \rangle_{II}} - \sqrt{d_{sh} \langle B_{\text{off}}^2 \rangle_I} \propto \Delta E \) approximately holds for sufficiently small \( \delta E \). Thus, for such a few-body operator, we do not see the plateau-like structure of off-diagonal matrix elements suggested in nonintegrable systems [8] even for small \( \Delta E \).
Now we give the proof of the atypicality of most few-body observables. Let us define

\[ 2Z_{fg} = d_{sh} \langle \hat{\Lambda}_f : \hat{\Lambda}_g \rangle_{II} - d_{sh} \langle \hat{\Lambda}_f : \hat{\Lambda}_g \rangle_I + \text{c.c.}, \quad (33) \]

\[ \langle \hat{\Lambda}_f : \hat{\Lambda}_g \rangle_{II} = \frac{1}{p_{II} q_{II}} \sum_{II} (\hat{\Lambda}_f)_{\alpha\beta} (\hat{\Lambda}_g)_{\beta\alpha}, \quad (34) \]

\[ \langle \hat{\Lambda}_f : \hat{\Lambda}_g \rangle_I = \frac{1}{p_{I} q_{I}} \sum_{I} (\hat{\Lambda}_f)_{\alpha\beta} (\hat{\Lambda}_g)_{\beta\alpha}. \quad (35) \]

Then \( \hat{B} = \sum_f B_f \hat{\Lambda}_f \) leads to

\[ \left| \sum_{fg} B_f B_g Z_{fg} \right| \geq \xi_{od}. \quad (36) \]

Here, \( Z_{fg} \) is real and symmetric, and can be diagonalized by an orthogonal transformation \( R \). By writing \( R Z R^T = \text{diag}(D_1, \ldots, D_f, \ldots, D_n) \) with \( D_1 \leq \cdots \leq D_n \) and \( \tilde{B} = R \tilde{B} \), we have

\[ \left| \sum_f B_f^2 D_f \right| \geq \xi_{od}. \quad (37) \]

We note that

\[ \sum_f B_f^2 \sqrt{\sum_f D_f^2} \geq \sqrt{\sum_f B_f^4 \sum_f D_f^2} \geq \left| \sum_f B_f^2 D_f \right| \geq \xi_{od}, \quad (38) \]

which is obtained by using \( \left( \sum_f B_f^2 \right)^2 \geq \sum_f B_f^4 \) and the Cauchy-Schwarz inequality. We also note that

\[ \sum_f B_f^2 = \sum_f B_f^2 = \frac{\text{Tr}[\tilde{B}^2]}{S^N} \leq ||\tilde{B}||_{op}^2. \quad (39) \]

Then

\[ |\tilde{D}| = \sqrt{\sum_f D_f^2} \geq \frac{\xi_{od}}{||\tilde{B}||_{op}^2}, \quad (40) \]

where \( \tilde{D} = (D_1, \ldots, D_n) \). We also define \( D_M := \max_f |D_f| \geq \frac{\xi_{od}}{\sqrt{n}||\tilde{B}||_{op}^2}. \)

We now show that

\[ \mathbb{P}_{L_m} \left[ d \langle \mathcal{O}^2_{\text{off}} \rangle_I - d \langle \mathcal{O}^2_{\text{off}} \rangle_{II} \right] \leq ||\hat{\mathcal{O}}||_{op}^2 d_{sh}^{-\eta} \]

\[ \mathbb{P}_{L_m} \left[ \sum_{fg} G_f G_g Z_{fg} \right] \leq ||\hat{\mathcal{O}}||_{op}^2 d_{sh}^{-\eta} \quad (41) \]
is small (note that \( \hat{O} = \sum f G_f \hat{\Lambda}_f \)). Since the probability is invariant under the orthogonal transformation \( \mathcal{R} \), we have

\[
P_{\mathcal{L}_m} \left[ \sum_{fg} G_f G_g Z_{fg} \right] \leq \| \hat{O} \|^2_{op} d_{sh}^{-n} = P_{\mathcal{L}_m} \left[ \sum_f G_f^2 D_f \right] \leq \| \hat{O} \|^2_{op} d_{sh}^{-n}. \tag{42}
\]

To evaluate the right-hand side, we first fix \( |\vec{G}| \) and consider the probability on the hypersphere \( \sum f G_f^2 = |\vec{G}|^2 \). We note that

\[
\| \hat{O} \|^2_{op} \leq \left( \sum_{f=1}^n |G_f| \cdot \| \hat{\Lambda}_f \|_{op} \right)^2 \leq \Lambda^2 \left( \sum_{f=1}^n |\vec{G}_f| \right)^2 \leq \Lambda^2 |\vec{G}|^2 n, \tag{43}
\]

where the last inequality comes from the Cauchy-Schwarz inequality. Thus, the right-hand side in Eq. (42) is further bounded from above by

\[
P_{\mathcal{L}_m} \left[ |\vec{G}|^2 \delta \leq \sum_f G_f^2 D_f \leq |\vec{G}|^2 \delta \right], \tag{44}
\]

where \( \delta = \Lambda^2 n d_{sh}^{-\eta} \) decreases exponentially as a function of \( N \).

Without loss of generality, we can assume \( D_1 = \min_f D_f = -D_M \) (the following discussion holds true for the case with \( D_n = \max_f D_f = D_M \)). Then, Eq. (44) is equivalent to

\[
P_{\mathcal{L}_m} \left[ |\vec{G}|^2 (-\delta + D_M) \leq \sum_{f=2}^n G_f^2 E_f \leq |\vec{G}|^2 (\delta + D_M) \right] \tag{45}
\]

under the constraint

\[
\sum_{f=1}^n G_f^2 = |\vec{G}|^2, \tag{46}
\]

where \( E_f = D_f + D_M > 0 \) for \( f \geq 2 \).

Equations (45) and (46) allow a geometrical interpretation that we should evaluate an overlap of the \((n-1)\)-dimensional hypersphere and the \(n\)-dimensional thin elliptic hypercylinder shell. For fixed \( |\vec{G}| \), the volume of the overlap can be evaluated by integrating out \( G_1 \):

\[
\mathcal{N} = \int_{|\vec{G}|^2} B^\mathcal{E} \frac{dG_2 \cdots dG_n}{\sqrt{|\vec{G}|^2 - G_2^2 - \cdots - G_n^2}}, \tag{47}
\]
FIG. 3. (a) Two possible configurations of overlaps of an \((n-1)\)-dimensional ball \(B\) and an \((n-1)\)-dimensional thin elliptic hypercylinder shell \(E\) for \(n = 3\). (b) Evaluation of the volume of the overlap \(\mathcal{N}'\) via \(W\) (Eq. (49)). (c) Circular configuration for which both \(W\) and \(\mathcal{N}'\) are maximized. In this case, \(\mathcal{N}\) is also maximized.

where \(B\) denotes an \((n-1)\)-dimensional ball with the radius \(|\vec{G}|\) and \(E\) denotes an \((n-1)\)-dimensional thin elliptic shell (see Figure 3).

We consider those configurations of \(E\) which maximize the overlap volume \(\mathcal{N}\). We first note that the quantity

\[
\mathcal{N}' = \int_{B \cap E} dG_2 \cdots dG_n,
\]

satisfies

\[
\mathcal{N}' \leq \left[1 - \left(\frac{D_M - \delta}{D_M + \delta}\right)^{\frac{n-1}{2}}\right] W,
\]

where \(W\) is the overlap volume between \(B\) and the region \(\left\{ \sum_{f=2}^{n} G_f^2 E_f \leq |\vec{G}|^2 (\delta + D_M) \right\}\) (see Fig. 3(b)). We note that \(W\) is maximized when \(B \subset \left\{ \sum_{f=2}^{n} G_f^2 E_f \leq |\vec{G}|^2 (\delta + D_M) \right\}\).
Furthermore, the equality of Eq. (49) is attained when \( \sum_{f=2}^{n} G_f^2 E_f \leq |\vec{G}|^2 (\delta + D_M) \) \( \subset \mathcal{B} \). Thus, \( \mathcal{N}' \) is maximized when \( \mathcal{B} = \left\{ \sum_{f=2}^{n} G_f^2 E_f \leq |\vec{G}|^2 (\delta + D_M) \right\} \) (see Fig. 3(c)). In this case, \( \mathcal{N} \) is also maximized, since the factor \( \sqrt{|\vec{G}|^2 - G_2^2 - \cdots - G_n^2} \) becomes smaller as \( (G_2, \cdots, G_n) \) approaches the edge of \( \mathcal{B} \). Thus, Eq. (45) is maximized when \( \vec{E} = (D_M + \delta, \cdots, D_M + \delta) \).

Going back to the original problem, we obtain

\[
P_{\mathcal{L}_m} \left[ -|\vec{G}|^2 \delta \leq \sum_{f=1}^{n} G_f^2 D_f \leq |\vec{G}|^2 \delta \right] = \frac{4B_II d^2_{sh} + 4B_II d^2_{sh} - 2}{d^2 d_{sh} - 2} \leq \frac{2\pi}{\xi_{od}} ||\hat{B}||_{\text{op}} \Lambda d^{-\frac{2}{3}} \cdot \Gamma \left( \frac{n}{2} \right) \cdot \Gamma \left( \frac{n-1}{2} \right),
\]

which completes the proof of Theorem 2.

Finally, we consider observables randomly chosen from \( \mathcal{L}_{m=N} \), which include many-body observables. Then, most of them satisfy \( U \in \mathcal{U}_{(a_i),\eta} \) as stated in the following proposition.

**Proposition 2.** Let \( P_{\mathcal{L}_m} \) be a probability with respect to \( P(\vec{G}) \). If we take \( m = N \), the following inequality holds.

\[
P_{\mathcal{L}_N} \left[ U \notin \mathcal{U}_{(a_i),\eta} \right] \leq \frac{4B_II d^2_{sh} + 4B_II d^2_{sh} - 2}{d^2 d_{sh} - 2} \cdot \frac{d_{sh}^{2+2\eta}}{d^2 d_{sh}^{2+2\eta}}.
\]

The right-hand side vanishes when \( d_{sh}^{2+2\eta} \cdot d_{sh}^{2+2\eta} \ll 1 \) for sufficiently large \( N \). The proof goes similarly to that of Proposition in the main text. We can show that \( U \) distributes uniformly over the unitary Haar measure over the entire Hilbert space, and the same method for proving Proposition 1 can be used. That is, \( d_{sh} \) in the left-hand side of Eq. (24) can be replaced by \( d \). Consequently, we can use the Chebyshev’s inequality to
$\mathbb{P}_{\mathcal{L}_N} \left[ |d \langle O^2 \rangle_I - d \langle O^2 \rangle_{II}| > \|\hat{O}\|_\text{op}^2 d_{\text{sh}}^{-1-\eta} \right]$, which leads to Eq. (54).

[1] J. v. Neumann, Zeitschrift für Physik 57, 30 (1929), English translation (by R. Tumulka), The European Physical Journal H 35, 201 (2010).
[2] S. Goldstein, J. L. Lebowitz, C. Mastrodonato, R. Tumulka, and N. Zanghi, Phys. Rev. E 81, 011109 (2010).
[3] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghi, The European Physical Journal H 35, 173 (2010).
[4] P. Reimann, Phys. Rev. Lett. 115, 010403 (2015).
[5] G. Biroli, C. Kollath, and A. M. Läuchli, Phys. Rev. Lett. 105, 250401 (2010).
[6] W. Beugeling, R. Moessner, and M. Haque, Phys. Rev. E 91, 012144 (2015).
[7] T. A. Brody, J. Flores, J. B. French, P. A. Mello, A. Pandey, and S. S. M. Wong, Rev. Mod. Phys. 53, 385 (1981).
[8] L. D’Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, Advances in Physics 65, 239 (2016).