# Rigorous Bounds on Eigenstate Thermalization

Shoki Sugimoto,[1] Ryusuke Hamazaki,[2] and Masahito Ueda[1, 3]

[1]*Department of Physics, The University of Tokyo,*
*7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*[*]
[2]*Nonequilibrium Quantum Statistical Mechanics RIKEN Hakubi Research Team,*
*RIKEN Cluster for Pioneering Research (CPR), RIKEN iTHEMS, Wako, Saitama 351-0198, Japan*
[3]*RIKEN Center for Emergent Matter Science (CEMS), Wako 351-0198, Japan*

The eigenstate thermalization hypothesis (ETH), which asserts that every eigenstate of a many-body quantum system is indistinguishable from a thermal ensemble, plays a pivotal role in understanding thermalization of isolated quantum systems. Yet, no evidence has been obtained as to whether the ETH holds for *any* few-body operators in a chaotic system; such few-body operators include crucial quantities in statistical mechanics, e.g., the total magnetization, the momentum distribution, and their low-order thermal and quantum fluctuations. Here, we identify rigorous upper and lower bounds on $m_*$ such that *all* $m$-body operators with $m < m_*$ satisfy the ETH in fully chaotic systems. For arbitrary dimensional $N$-particle systems subject to the Haar measure, we prove that there exist $N$-independent positive constants $\alpha_*^{(L)}$ and $\alpha_*^{(U)}$ such that $\alpha_*^{(L)} \leq m_*/N \leq \alpha_*^{(U)}$ holds. The bounds $\alpha_*^{(L)}$ and $\alpha_*^{(U)}$ depend only on the spin quantum number for spin systems and the particle-number density for Bose and Fermi systems. Thermalization of *typical* systems for *any* few-body operators is thus rigorously proved.

Recent experiments in cold atoms and ions have demonstrated that quantum systems thermalize unitarily without heat reservoirs [1–6]. This finding brings up a striking possibility of incorporating statistical mechanics into a single framework of quantum mechanics – a scenario envisioned by von Neumann about a century ago [7]. It has been argued that a single pure quantum state becomes indistinguishable from a thermal ensemble as far as few-body observables are concerned [8–18] due to the interplay between quantum entanglement and physical constraints on observable quantities such as locality or few-bodiness [8, 9, 14–18]. Depending on the choice of operators used to distinguish between a quantum state and a thermal ensemble, varied notions of quantum-thermal equilibrium, such as microscopic thermal equilibrium (MITE) and macroscopic thermal equilibrium (MATE), have been introduced [8, 9, 13, 17, 19, 20].

The eigenstate thermalization hypothesis (ETH) [7, 21, 22] is considered to be the primary mechanism behind thermalization in isolated quantum systems. The ETH for an operator $\hat{A}$ means that (i) every energy eigenstate of a system is in thermal equilibrium regarding the expectation value of $\hat{A}$ and that (ii) off-diagonal elements of $\hat{A}$ in an energy eigenbasis is vanishingly small. The ETH ensures thermalization of $\hat{A}$ for any initial state with a macroscopically definite energy, given no massive degeneracy in the energy spectrum [14–18]. The ETH has been tested numerically for several local or few-body quantities [23–31]. However, whether the ETH holds for *all* few-body operators and whether it breaks down for many-body operators have yet to be fully addressed.

Von Neumann [7] and Reimann [32] proved the ETH for *almost all* Hermitian operators. However, their results do not imply the ETH for physically realistic operators because *almost all* operators considered in Refs. [7, 32] involve highly nonlocal correlations that are close to $N$-body [33] and therefore unphysical. Some works [17, 34] tested the ETH against several classes of few-body operators, such as local operators of a subsystem; however, their method cannot deal with generic few-body operators that act on an entire system., such as the total magnetization, the momentum distribution, and their thermal and quantum fluctuations.

In this Article, we derive upper and lower bounds for $m_*$ such that *all* $m$-body operators with $m < m_*$ satisfy the ETH. For fully chaotic $N$-particle systems in arbitrary spatial dimensions whose eigenstates are distributed according to the Haar measure, we prove that there exist $N$-independent constants $\alpha_*^{(L)}$ and $\alpha_*^{(U)}$ such that $\alpha_*^{(L)} \leq m_*/N \leq \alpha_*^{(U)}$ holds. Here, $\alpha_*^{(L)}$ and $\alpha_*^{(U)}$ depend only on the spin quantum number for spin systems or the particle-number density for Bose and Fermi systems (Figure 1). Our result is directly applicable to *any* few-body operators of interest in statistical mechanics and their thermal and quantum fluctuations without approximations, such as coarse-graining. In particular, our result implies that the ETH holds true even if we observe *any* low-order fluctuations of *any* few-body operators.

Our method also provides a quantitative criterion for deciding whether energy eigenstates of a system can be considered fully chaotic, which is applicable to systems with non-negligible finite-size effects, including ion and cold-atom systems, for which the thermodynamic limit cannot be taken [1–6, 35, 36].

*Unified measure for quantum-thermal equilibrium.—* While several measures and criteria of quantum-thermal equilibrium have been introduced in the literature [13, 17, 19, 20], they cannot be applied to generic few-body

(a) Spin
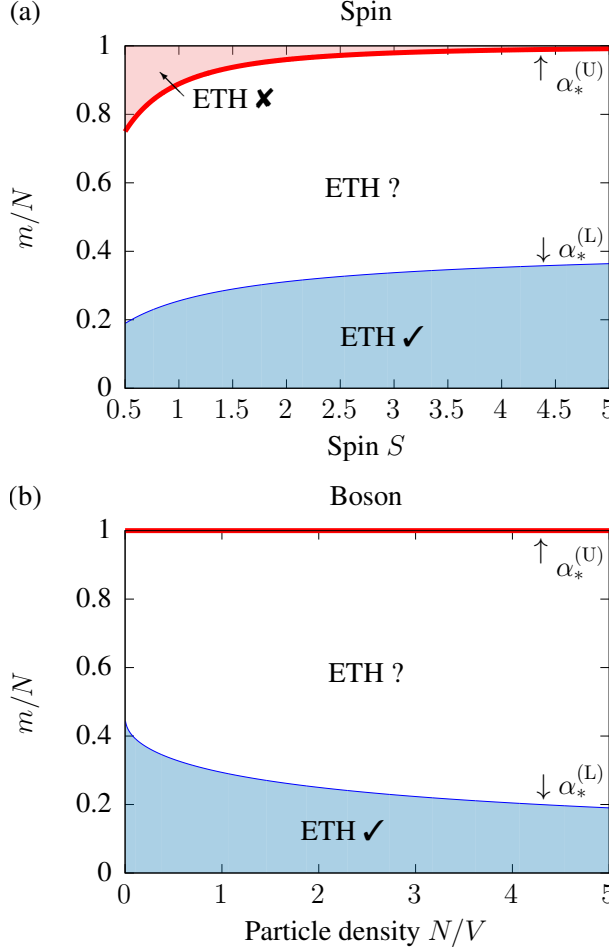


Table 1: **Definitions of the $m$-body operator space and its dimension.**

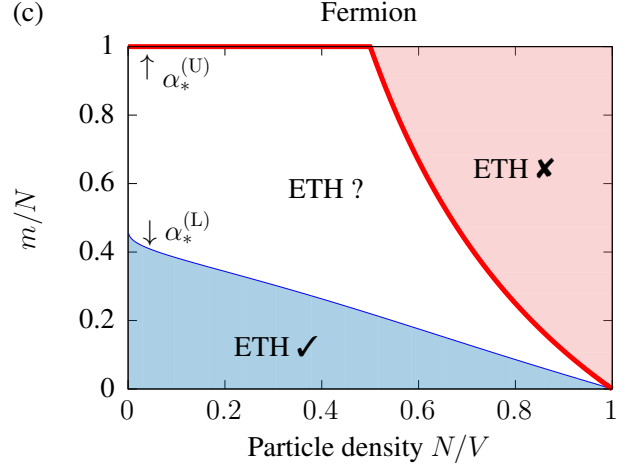| | |
|---|---|
| Spin | $M = \sum_{j=0}^{m} \binom{N}{j} [(2S+1)^2 - 1]^j, \quad D = (2S+1)^N$ |
| | $\mathcal{A}^{[0,m]} := \text{span}\left\{ \hat{\sigma}_{x_1}^{(p_1)} \cdots \hat{\sigma}_{x_j}^{(p_j)} \mid \begin{smallmatrix} j \leq m \\ 1 \leq x_1 < \cdots < x_m \leq N \\ 1 \leq p_j < (2S+1)^2 \end{smallmatrix} \right\}$ |
| Boson | $M = \binom{m+V-1}{m}^2, \quad D = \binom{N+V-1}{N}$ |
| | $\mathcal{A}^{[0,m]} := \text{span}\left\{ \hat{b}_{x_1}^{\dagger} \cdots \hat{b}_{x_m}^{\dagger} \hat{b}_{y_1} \cdots \hat{b}_{y_m} \mid 1 \leq x_j, y_j \leq V \right\}$ |
| Fermion | $M = \binom{V}{\min\{m, V-N\}}^2, \quad D = \binom{V}{N}$ |
| | $\mathcal{A}^{[0,m]} := \text{span}\left\{ \hat{f}_{x_1}^{\dagger} \cdots \hat{f}_{x_m}^{\dagger} \hat{f}_{y_1} \cdots \hat{f}_{y_m} \mid \begin{smallmatrix} 1 \leq x_1 < \cdots < x_m \leq V \\ 1 \leq y_1 < \cdots < y_m \leq V \end{smallmatrix} \right\}$ |

(b) Boson



(c) Fermion



FIG. 1. **Regions of $m/N$ where the ETH typically holds and breaks down for fully chaotic systems whose eigenstates are distributed according to the Haar measure.** The blue region (marked with "✓") shows where the ETH typically holds for *all* operators in the space $\mathcal{A}^{[0,m]}$ of $m$-body operators. This region is delimited by $\alpha_*^{(L)}$. The red region (marked with "✗") shows where the ETH typically breaks down for some $m$-body operators. This region is delimited by $\alpha_*^{(U)}$. There is an ETH-breaking operator within $\mathcal{O}(1/\sqrt{N})$ around the thick red curve for each $S$ and $N/V$. The integer $m_*$ such that the ETH holds for *all* $m$-body operators with $m < m_*$ lies somewhere in the white region (marked with "?"). The definitions of the $m$-body operator space and its dimension are shown in Table 1. For spin systems (a), we have $\alpha_*^{(U)} = 1 - (2S+1)^{-2}$, where $S$ is the spin quantum number. For Bose systems (b) and Fermi systems (c), $N$ and $V$ denote the numbers of particles and lattice sites, respectively. In1q (b), we have $\alpha_*^{(U)} = 1$, and the boundary of the blue region $\alpha_*^{(L)}$ is proportional to $\rho^{-1/2}$ for large $\rho$. In1q (c), we have $\alpha_*^{(U)} = (1-\rho)/\rho$, and the boundary of the blue region $\alpha_*^{(L)}$ is proportional to $(1-\rho)$ for $\rho \simeq 1$.

operators because they are defined for specific choices of operators, such as those acting only on a subsystem.

To quantitatively study how the ETH depends on physical constraints on observables, we introduce the following measure of the closeness to a thermal ensemble, which is applicable to an arbitrary set $\mathcal{A}$ of observable quantities.

**Definition 1** (Unified measure of quantum-thermal equilibrium). *To quantify the distance $\|\hat{\sigma} - \hat{\rho}_{\text{th}}\|$ between a quantum state $\hat{\sigma}$ and a thermal state $\hat{\rho}_{\text{th}}$, we introduce*

*the following (semi-)norm,*

$$\left\| \hat{X} \right\|_p^{(\mathcal{A})} := \sup_{\hat{A} \in \mathcal{A} + \mathbb{R}\hat{I}} \left| \text{tr}\left( \frac{\hat{A}}{\|\hat{A}\|_q} \hat{X} \right) \right|, \tag{1}$$

*where $p^{-1} + q^{-1} = 1$ with $p \geq 1$, and $\|\hat{A}\|_q$ is the Schatten $q$-norm of $\hat{A}$ [37, 38], i.e., the standard $q$-norm of the singular values of $\hat{A}$. Here, $\hat{X}$ is an arbitrary operator which is not necessarily Hermitian. Then, the (pseudo-)distance $\|\hat{\sigma} - \hat{\rho}_{\text{th}}\|_1^{(\mathcal{A})}$ with $\hat{\rho}_{\text{th}}$ being a thermal ensemble serves as a unified measure of quantum-thermal equilibrium.*

We say that a quantum state $\hat{\sigma}$ is in thermal equilib-

rium relative to $\mathcal{A}$ if $\|\hat{\sigma} - \hat{\rho}_{\mathrm{th}}\|_1^{(\mathcal{A})} < \epsilon$ holds for a sufficiently small $\epsilon\,(>0)$. This definition follows Refs. [19, 20]; however, we here concern only the expectation value of $\hat{A} \in \mathcal{A}$ and not the probability distribution over the spectrum of $\hat{A}$. Nonetheless, our framework can deal with the probability distribution by including sufficiently high powers of $\hat{A}$ in $\mathcal{A}$, offering finer control over the precision in observing the distribution. By appropriately choosing $\mathcal{A}$, our notion of quantum-thermal equilibrium unifies previously introduced ones, such as subsystem thermal equilibrium [8, 9, 34], microscopic thermal equilibrium (MITE), and macroscopic thermal equilibrium (MATE) [17, 19, 20] (see Supplemental Information I for details).

The normalization constant $\|\hat{A}\|_q$ in the inequality (1) has been chosen so that the (semi-)norm $\|\cdot\|_p^{(\mathcal{A})}$ serves as a generalization of the Schatten $p$-norm. For example, $\|\cdot\|_1^{(\mathcal{A})}$ and $\|\cdot\|_2^{(\mathcal{A})}$ generalize the trace norm and Hilbert-Schmidt norm, respectively. In general, the duality of the Schatten norm [37, 38] gives $\|\hat{X}\|_p = \|\hat{X}\|_p^{(\mathcal{L}(\mathcal{H}))}$ $(\geq \|\hat{X}\|_p^{(\mathcal{A})})$, where $\mathcal{L}(\mathcal{H})$ is the space of all Hermitian operators acting on a Hilbert space $\mathcal{H}$ of a system.

Among $\|\cdot\|_p^{(\mathcal{A})}$ with $p \geq 1$, only $\|\cdot\|_1^{(\mathcal{A})}$ can be used to define a measure of quantum-thermal equilibrium because it is (i) invariant under a linear transformation: $\hat{A} \mapsto a'\hat{A} + b'$, (ii) dimensionless, and (iii) thermodynamically intensive [30]. The others, $\|\cdot\|_p^{(\mathcal{A})}$ with $p > 1$, are not suitable because they are not thermodynamically intensive (see Supplemental Information II).

Many of the previous works that numerically tested the ETH with respect to several operators $\hat{A}_1, \cdots, \hat{A}_J$ [23, 25, 26, 29] essentially set $\mathcal{A} = \{\hat{A}_1, \cdots, \hat{A}_J\}$. However, if one really wants to distinguish an energy eigenstate from $\hat{\rho}_{\mathrm{th}}$ without any exception, one should use not only a single or a few quantities but $all$ the quantities compatible with physical constraints under consideration.

For general $\mathcal{A}$, it is difficult or even impossible to calculate the maximum in $\|\cdot\|_1^{(\mathcal{A})}$ with respect to $\hat{A} \in \mathcal{A}$, both numerically and analytically. To overcome this difficulty, we place upper and lower bounds on $\|\cdot\|_1^{(\mathcal{A})}$ in terms of the quantity $\|\cdot\|_2^{(\mathcal{A})}$, which is computable given an orthonormal basis of $\mathcal{A} + \mathbb{R}\hat{I}$.

*Measure of the ETH.—* The measure of the ETH, which requires (i) $all$ energy eigenstates to be in thermal equilibrium and (ii) $all$ off-diagonal elements of an observable in an energy eigenbasis is vanishingly small, is obtained as

$$\Lambda_1^{(\hat{H},\mathcal{A})}(E) := \max_{|E_\alpha\rangle, |E_\beta\rangle \in \mathcal{H}_{E,\Delta E}} \left\| \hat{\rho}_{\alpha\beta} - \hat{\rho}_{\delta E}^{(\mathrm{mc})}(E_\alpha)\delta_{\alpha\beta} \right\|_1^{(\mathcal{A})},$$

where $\hat{\rho}_{\alpha\beta} := |E_\alpha\rangle\langle E_\beta|$ with $|E_\alpha\rangle$ being an energy eigenstate belonging to the eigenenergy $E_\alpha$, and $\hat{\rho}_{\delta E}^{(\mathrm{mc})}(E_\alpha)$ is the microcanonical ensemble within the energy shell $\mathcal{H}_{E_\alpha,\delta E}$. The ETH holds for $all$ operators in $\mathcal{A}$ if and

only if $\lim_{N\to\infty} \Lambda_1^{(\hat{H},\mathcal{A})} = 0$. Indeed, if $\Lambda_1^{(\hat{H},\mathcal{A})}$ is sufficiently small, $all$ the operators in $\mathcal{A}$ give almost the same expectation values for $\hat{\rho}_{\alpha\alpha}$ and $\hat{\rho}_{\delta E}^{(\mathrm{mc})}(E_\alpha)$, and their off-diagonal elements within the energy shell $\mathcal{H}_{E_\alpha,\delta E}$ becomes sufficiently small. In that sense, $\Lambda_1^{(\hat{H},\mathcal{A})}$ is the most sensitive ETH measure ever. If $\Lambda_1^{(\hat{H},\mathcal{A})}$ remains finite in the limit $N \to \infty$, there exists an operator $\hat{A} \in \mathcal{A}$ such that the expectation values of $\hat{A}$ for $|E_\alpha\rangle$ and $\hat{\rho}_{\delta E}^{(\mathrm{mc})}$ are different, or some off-diagonal elements of $\hat{A}$ within $\mathcal{H}_{E_\alpha,\delta E}$ remain nonnegligible. Therefore, the ETH with respect to $\mathcal{A}$ breaks down in this case.

*Distribution of $\Lambda_1^{(\hat{H},\mathcal{A})}(E)$ for fully chaotic systems.—* Having introduced the measure of the ETH and the bounds for $\|\cdot\|_1^{(\mathcal{A})}$ in terms of the computable quantity $\|\cdot\|_2^{(\mathcal{A})}$, we are in a position to discuss the validity of the ETH relative to a set of observables $\mathcal{A}$.

In fully chaotic systems, whose eigenvectors distribute according to the unitary Haar measure, we can derive the following theorem by using the concentration inequality for the Haar measure on $\mathbb{SU}(D)$ [39–42], which is a stronger result than the commonly used Levy's lemma (see Methods for the outline of the proof).

**Theorem 1.** *Let $\mathcal{G}_{\mathrm{inv}}$ be an invariant random matrix ensemble, the eigenvectors of whose matrices are distributed according to the unitary Haar measure. We set $D := \dim \mathcal{H}$ and $M := \dim \mathcal{A}$. Then, for any $\epsilon > 0$, we have*

$$\frac{M}{4D^2} \leq (\Lambda_1^{(\hat{H},\mathcal{A})})^2 + \mathcal{O}\left(\frac{D^\epsilon}{\sqrt{D}}\right) \leq \frac{M}{D} \qquad (2)$$

*for almost all $\hat{H} \in \mathcal{G}_{\mathrm{inv}}$. Here, "for almost all $\hat{H} \in \mathcal{G}_{\mathrm{inv}}$" means that the ratio of exceptional Hamiltonians $\hat{H}$ in $\mathcal{G}_{\mathrm{inv}}$ for which the inequality (2) does not hold is bounded from above by $\exp\left(-\mathcal{O}(D^{2\epsilon})\right)$.*

With Theorem 1, we can test whether or not the ETH with respect to $\mathcal{A}$ typically holds in $\mathcal{G}_{\mathrm{inv}}$ by counting the dimension of the operator space $\mathcal{A}$. Theorem 1 also provides a quantitative criterion, applicable to finite-size systems, for deciding whether a realistic Hamiltonian $\hat{H}$ can be considered to have fully chaotic energy eigenstates. For example, the existence of an operator $\hat{A}$ that violates the upper bound of the inequality (2) implies that the eigenstate of $\hat{H}$ has a "structure" that can be detected by $\hat{A}$, suggesting that $\hat{H}$ is not fully chaotic. Furthermore, by comparing the $N$-dependence of $\Lambda_1^{(\hat{H},\mathcal{A})}$ for a realistic Hamiltonian $\hat{H}$ with that of the upper and lower bounds in the inequality (2), we can infer the $N$-dependence of $\Lambda_1^{(\hat{H},\mathcal{A})}$ for large $N$ that is difficult to access.

*Rigorous upper and lower bounds for the ETH.—* We apply Theorem 1 to test the ETH for few- and many-body observables. For $N$-site spin-$S$ systems, we define the $m$-body operator space $\mathcal{A}_N^{[0,m]}$ as the space

of operators that can be expressed as a linear combination of operators acting nontrivially on at most $m$ spins. The $m$-body operator space $\mathcal{A}_N^{[0,m]}$ includes the set $\mathcal{S}_{\text{few}}^{(m)}$ of "few-body" operators [17], which act nontrivially on at most $m$ spins. For example, we have $\hat{M}_z := \sum_{j=1}^{N} \hat{\sigma}_j^{(z)}, (\hat{M}_z)^2, \cdots, (\hat{M}_z)^m \in \mathcal{A}_N^{[0,m]}$, but none of these operators are included in $\mathcal{S}_{\text{few}}^{(m)}$ unless $m = N$. For Bose and Fermi systems, we define $\mathcal{A}_N^{[0,m]}$ to be the space of operators that can be written as a linear combination of products of $m$ annihilation operators $\hat{b}$ and $m$ creation operators $\hat{b}^\dagger$. We also define $\mathcal{A}_N^{[m_-,m_+]}$ to be the orthogonal complement of $\mathcal{A}_N^{[0,m_--1]}$ with respect to $\mathcal{A}_N^{[0,m_+]}$.

By counting the dimension of $\mathcal{A}_N^{[m_-,m_+]}$ and that of the total Hilbert space, we can apply Theorem 1 to obtain the following theorem (see Methods for the proof outline).

**Theorem 2** (Upper and lower bounds for the ETH). *Let $m_*$ be the largest number such that the ETH with respect to $\mathcal{A}^{[0,m]}$ typically holds in $\mathcal{G}_{\text{inv}}$ for all $m < m_*$. Then, there exist $N$-independent constants $\alpha_*^{(\text{L})} \in (0, 1/2]$ and $\alpha_*^{(\text{U})} > 0$ that satisfy $\alpha_*^{(\text{L})} \leq m_*/N \leq \alpha_*^{(\text{U})}$.*

*More specifically,*

1. *for $\mathcal{A} = \mathcal{A}_N^{[0,m]}$ with $m/N < \alpha_*^{(\text{L})}$, we have*

$$\left(\Lambda_1^{(\hat{H},\mathcal{A})}\right)^2 \leq e^{-\mathcal{O}(N)}, \tag{3}$$

*for almost all $\hat{H} \in \mathcal{G}_{\text{inv}}$, and therefore the ETH with respect to $\mathcal{A}_N^{[0,m]}$ typically holds in $\mathcal{G}_{\text{inv}}$;*

2. *for $\mathcal{A} = \mathcal{A}_N^{[m_-,m_+]}$ with $m_\pm = \alpha_*^{(\text{U})} N \pm c_\pm \sqrt{N}$, where $c_\pm$ are arbitrary positive constants, there exists a constant $C > 0$ such that*

$$\lim_{N\to\infty} \left(\Lambda_1^{(\hat{H},\mathcal{A})}\right)^2 \geq C \tag{4}$$

*holds for almost all $\hat{H} \in \mathcal{G}_{\text{inv}}$. Therefore, the ETH with respect to $\mathcal{A}_N^{[m_-,m_+]}$ typically breaks down in $\mathcal{G}_{\text{inv}}$.*

*Here, "for almost all $\hat{H} \in \mathcal{G}_{\text{inv}}$" means that the fraction of exceptional Hamiltonians $\hat{H}$ in $\mathcal{G}_{\text{inv}}$ for which the inequality (3) or (4) does not hold is double-exponentially small with respect to $N$.*

*For spin systems, we have $\alpha_*^{(\text{L})}(S) \geq 0.1892\cdots$ and $\alpha_*^{(\text{U})}(S) = 1 - (2S+1)^{-2}$. For Bose and Fermi systems with particle density given by $\rho$, we have $\alpha_*^{(\text{U})}(\rho) = 1$ and $\alpha_*^{(\text{U})}(\rho) = (1-\rho)/\rho$, respectively (see Figure 1).*

The first part (3) of Theorem 2 suggests that for fully chaotic quantum many-body systems, whose eigenstates can be considered sufficiently random with respect to the Haar measure, the ETH can hold even if we observe $\mathcal{O}(N)$-body operators. Because our result is not restricted to the operators acting only on subsystems, we conclude that the decomposition of the total system into a subsystem and the rest is not essential for the ETH to hold. For the same reason, our result directly applies to any operators acting on the whole system, which cannot exactly be dealt with in previous works [17, 34]. These operators include extensive sums of local operators (e.g., total magnetization), few-body operators (e.g., momentum distributions), and low-order powers of these quantities. Since the central moments of a few-body operator $\hat{A}$ are polynomials of $\hat{A}$, and their values scale polynomially in $N$, the exponential decay of the upper bound (3) for the ETH measure implies that *the ETH typically holds in $\mathcal{G}_{\text{inv}}$ including any low-order fluctuations of any few-body operators.*

Given a probable argument based on the comparison of the reduced density operators of energy eigenstates and a thermal ensemble on a subsystem [34], we believe that the ETH typically breaks down for $\mathcal{A}^{[0,N/2]}$, i.e., we expect $m_*/N \leq 1/2$. For high-density Fermi systems with $\rho > 2/3$, our upper bound gives a better upper bound $m_*/N \leq \alpha_*^{(\text{U})} = (1-\rho)/\rho < 1/2$. However, for spin and Bose systems, our method does not provide $\alpha_*^{(\text{U})} < 0.5$. Nonetheless, the second part (4) of Theorem 2 is somewhat stronger than the expectation $m_*/N \leq 1/2$ in that it identifies a smaller region of $m$ than $[0, N/2]$ where ETH-breaking operators exist, namely, $[m_-, m_+]$ with $m_\pm = \alpha_*^{(\text{U})} N \pm \mathcal{O}(\sqrt{N})$.

Apart from quantum thermalization, the (semi-)norm $\|\cdot\|_1^{(\mathcal{A})}$ introduced in the inequality (1) serves as the measure of the closeness between two quantum states $\hat{\sigma}$ and $\hat{\rho}$ relative to $\mathcal{A}$. Thus, it can be used in various situations other than the ETH, such as a comparison between the state during time evolution and the steady state. In particular, it can be used to construct the space of macroscopic states from that of quantum states by identifying quantum states that are very close to each other in terms of $\|\cdot\|_1^{(\mathcal{A})}$. Rigorously formulating the correspondence between microscopic and macroscopic states in this direction and deriving the macroscopic dynamics from the microscopic one are important future problems.

---

[1] S. Trotzky, Y.-A. A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch, Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional Bose gas, Nature Physics **8**, 325 (2012).

[2] T. Langen, R. Geiger, M. Kuhnert, B. Rauer, and J. Schmiedmayer, Local emergence of thermal correlations in an isolated quantum many-body system, Nature Physics **9**, 640 (2013).

[3] G. Clos, D. Porras, U. Warring, and T. Schaetz, Time-Resolved Observation of Thermalization in an Isolated Quantum System, Physical Review Letters **117**, 170401 (2016).

[4] A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, and M. Greiner, Quantum thermalization through entanglement in an isolated many-body system, Science **353**, 794 (2016).

[5] C. Neill, P. Roushan, M. Fang, Y. Chen, M. Kolodrubetz, Z. Chen, A. Megrant, R. Barends, B. Campbell, B. Chiaro, A. Dunsworth, E. Jeffrey, J. Kelly, J. Mutus, P. J. J. O'Malley, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. Polkovnikov, and J. M. Martinis, Ergodic dynamics and thermalization in an isolated quantum system, Nature Physics **12**, 1037 (2016).

[6] Y. Tang, W. Kao, K.-Y. Li, S. Seo, K. Mallayya, M. Rigol, S. Gopalakrishnan, and B. L. Lev, Thermalization near Integrability in a Dipolar Quantum Newton's Cradle, Physical Review X **8**, 021030 (2018).

[7] J. von Neumann, Proof of the ergodic theorem and the H-theorem in quantum mechanics, The European Physical Journal H **35**, 201 (2010).

[8] S. Popescu, A. J. Short, and A. Winter, Entanglement and the foundations of statistical mechanics, Nature Physics **2**, 754 (2006).

[9] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, Canonical Typicality, Physical Review Letters **96**, 050403 (2006).

[10] P. Reimann, Typicality for Generalized Microcanonical Ensembles, Physical Review Letters **99**, 160404 (2007).

[11] S. Sugiura and A. Shimizu, Thermal Pure Quantum States at Finite Temperature, Physical Review Letters **108**, 240401 (2012).

[12] S. Sugiura and A. Shimizu, Canonical Thermal Pure Quantum State, Physical Review Letters **111**, 010401 (2013).

[13] H. Tasaki, Typicality of Thermal Equilibrium and Thermalization in Isolated Macroscopic Quantum Systems, Journal of Statistical Physics **163**, 937 (2016).

[14] A. Polkovnikov, K. Sengupta, A. Silva, and M. Vengalattore, Colloquium: Nonequilibrium dynamics of closed interacting quantum systems, Reviews of Modern Physics **83**, 863 (2011).

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

[16] C. Gogolin and J. Eisert, Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems, Reports on Progress in Physics **79**, 056001 (2016).

[17] T. Mori, T. N. Ikeda, E. Kaminishi, and M. Ueda, Thermalization and prethermalization in isolated quantum systems: a theoretical overview, Journal of Physics B: Atomic, Molecular and Optical Physics **51**, 112001 (2018).

[18] J. M. Deutsch, Eigenstate thermalization hypothesis, Reports on Progress in Physics **81**, 082001 (2018).

[19] S. Goldstein, D. A. Huse, J. L. Lebowitz, and R. Tumulka, Thermal Equilibrium of a Macroscopic Quantum System in a Pure State, Physical Review Letters **115**, 100402 (2015).

[20] S. Goldstein, D. A. Huse, J. L. Lebowitz, and R. Tumulka, Macroscopic and microscopic thermal equilib-rium, Annalen der Physik **529**, 1600301 (2017).

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

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

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

[24] M. Rigol, Breakdown of Thermalization in Finite One-Dimensional Systems, Physical Review Letters **103**, 100403 (2009).

[25] M. Rigol, Quantum quenches and thermalization in one-dimensional fermionic systems, Physical Review A **80**, 053607 (2009).

[26] G. Biroli, C. Kollath, and A. M. Läuchli, Effect of Rare Fluctuations on the Thermalization of Isolated Quantum Systems, Physical Review Letters **105**, 250401 (2010).

[27] L. F. Santos and M. Rigol, Localization and the effects of symmetries in the thermalization properties of one-dimensional quantum systems, Physical Review E **82**, 031130 (2010).

[28] R. Steinigeweg, J. Herbrych, and P. Prelovšek, Eigenstate thermalization within isolated spin-chain systems, Physical Review E **87**, 012118 (2013).

[29] W. Beugeling, R. Moessner, and M. Haque, Finite-size scaling of eigenstate thermalization, Physical Review E **89**, 042112 (2014).

[30] S. Sugimoto, R. Hamazaki, and M. Ueda, Test of the Eigenstate Thermalization Hypothesis Based on Local Random Matrix Theory, Physical Review Letters **126**, 120602 (2021).

[31] S. Sugimoto, R. Hamazaki, and M. Ueda, Eigenstate Thermalization in Long-Range Interacting Systems, Physical Review Letters **129**, 030602 (2022).

[32] P. Reimann, Generalization of von Neumann's Approach to Thermalization, Physical Review Letters **115**, 010403 (2015).

[33] R. Hamazaki and M. Ueda, Atypicality of Most Few-Body Observables, Physical Review Letters **120**, 080603 (2018).

[34] J. R. Garrison and T. Grover, Does a Single Eigenstate Encode the Full Hamiltonian?, Physical Review X **8**, 021026 (2018).

[35] M. Gring, M. Kuhnert, T. Langen, T. Kitagawa, B. Rauer, M. Schreitl, I. Mazets, D. Adu Smith, E. Demler, J. Schmiedmayer, D. A. Smith, E. Demler, and J. Schmiedmayer, Relaxation and prethermalization in an isolated quantum system, Science **337**, 1318 (2012).

[36] B. Neyenhuis, J. Zhang, P. W. Hess, J. Smith, A. C. Lee, P. Richerme, Z.-X. Gong, A. V. Gorshkov, and C. Monroe, Observation of prethermalization in long-range interacting spin chains, Science Advances **3**, e1700672 (2017).

[37] R. Bhatia, *Matrix Analysis*, Graduate Texts in Mathematics, Vol. 169 (Springer New York, New York, NY, 1997).

[38] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 3rd ed. (Cambridge University Press, 2012).

[39] M. Ledoux, *The concentration of measure phenomenon*, 89 (American Mathematical Society, 2001).

[40] V. D. Milman and G. Schechtman, *Asymptotic Theory of Finite Dimensional Normed Spaces: Isoperimetric Inequalities in Riemannian Manifolds*, Vol. 1200 (Springer, 2001).

[41] M. Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Modern Birkhäuser Classics (Springer Science & Business Media, Boston, MA, 2007).

[42] E. S. Meckes, *The Random Matrix Theory of the Classical Compact Groups* (Cambridge University Press, 2019)

## METHODS

*Proof outline of Theorem 1.* Theorem 1 follows from (i) the inequality $\|\cdot\|_2^{(\mathcal{A})} \leq \|\cdot\|_1^{(\mathcal{A})} \leq \sqrt{D}\|\cdot\|_2^{(\mathcal{A})}$ resulting from the inequality $D^{-1/2}\|\hat{A}\|_2 \leq \|\hat{A}\|_\infty \leq \|\hat{A}\|_2$ [37, 38] and (ii) the concentration inequality for the Haar measure on $\mathbb{SU}(D)$ [39–42]; for any $\delta > 0$ and any Lipschitz continuous function $f$ on $\mathbb{SU}(D)$ with Lipschitz constant $\eta_f$,

$$\mathbb{P}\Big[\big|f(\hat{U}) - \mathbb{E}_{\hat{U}}[f]\big| \geq \delta\Big] \leq 2\exp\left[-\frac{\delta^2 D}{4\eta_f^2}\right], \quad (5)$$

where $\mathbb{E}_{\hat{U}}$ denotes the average over the Haar measure for $\hat{U}$, and $a, b$ are positive constants.

We can show that both $(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2$, where $\delta\hat{\rho}_{\alpha\beta} := \hat{\rho}_{\alpha\beta} - \hat{\rho}_{\delta E}^{(mc)}(E_\alpha)\delta_{\alpha\beta}$, and $(\Lambda_1^{(\hat{H},\mathcal{A})})^2$ are Lipschitz continuous concerning $\hat{U}$ which diagonalizes the Hamiltonian $\hat{H}$. Their Lipschitz constants are bounded from above by an $N$-independent constant $\eta$ (see Supplemental Information III.C and III.D). Therefore, we can apply the concentration inequality (5) to $(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2$. In addition, the expectation value of $(\|\delta\hat{\rho}_{\alpha\beta}\|_2^{(\mathcal{A})})^2$ with respect to the Haar measure is calculated to be $\mathbb{E}[(\|\delta\hat{\rho}_{\alpha\beta}\|_2^{(\mathcal{A})})^2] \simeq M/D^2$ (see Supplemental Information III.A) which leads to $M/D^2 \lesssim \mathbb{E}[(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2] \lesssim M/D$.

By applying the inequality (5) to $f(\hat{U}) = (\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2$, we obtain

$$\mathbb{P}\Big[(\Lambda_1^{(\hat{H},\mathcal{A})})^2 - \mathbb{E}_{\hat{U}}\Big[(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2\Big] \geq \delta\Big]$$
$$\leq d_{E,\Delta E}^2 \mathbb{P}\Big[(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2 - \mathbb{E}_{\hat{U}}\Big[(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2\Big] \geq \delta\Big]$$
$$\leq 2d_{E,\Delta E}^2 \exp\left[-\frac{\delta^2 D}{4\eta^2}\right], \quad (6)$$

where the first inequality follows from the definition of $(\Lambda_1^{(\hat{H},\mathcal{A})})^2 = \max_{|E_\alpha\rangle,|E_\beta\rangle \in \mathcal{H}_{E,\Delta E}}(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2$. From the inequality (6), it follows that

$$\mathbb{E}_{\hat{U}}[(\Lambda_1^{(\hat{H},\mathcal{A})})^2] \leq \mathbb{E}_{\hat{U}}\Big[(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2\Big] + \mathcal{O}\left(\sqrt{\frac{\log d_{E,\Delta E}}{D}}\right). \quad (7)$$

(See Supplemental Information III.B.)

Then, we apply the inequality (5) to $f(\hat{U}) = (\Lambda_1^{(\hat{H},\mathcal{A})})^2$ and set $\delta = D^{-1/2+\epsilon}$ for an arbitrary $\epsilon > 0$ to conclude

$$(\Lambda_1^{(\hat{H},\mathcal{A})})^2 = \mathbb{E}_{\hat{U}}[(\Lambda_1^{(\hat{H},\mathcal{A})})^2] + \mathcal{O}\Big(D^{-1/2+\epsilon}\Big) \quad (8)$$

for almost all $\hat{U}$, where the fraction of exceptional $\hat{U}$ for which Eq. (8) does not hold is no larger than $\exp\big(-\mathcal{O}(D^{2\epsilon})\big)$.

Finally, Eqs. (7) and (8) combined with a trivial bound $\mathbb{E}_{\hat{U}}[(\Lambda_1^{(\hat{H},\mathcal{A})})^2] \geq \mathbb{E}_{\hat{U}}[(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2]$ give

$$\Big|(\Lambda_1^{(\hat{H},\mathcal{A})})^2 - \mathbb{E}_{\hat{U}}\Big[(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2\Big]\Big| \leq \mathcal{O}\Big(D^{-1/2+\epsilon}\Big) \quad (9)$$

for almost all $\hat{U}$. This inequality proves Theorem 1.

*Proof outline of Theorem 2.* For spin-$S$ systems, it is straightforward to show that

$$\dim \mathcal{A}_N^{[m_-,m_+]} = D^2 \sum_{j=m_-}^{m_+} P_j, \quad (10)$$

where $D := \dim \mathcal{H}$, and $P_j := \binom{N}{j}\frac{(d_{loc}^2-1)^j}{d_{loc}^{2N}}$ with $d_{loc} := 2S + 1$ is the probability mass for the binomial distribution $\mathcal{B}(N,p)$ with $p = 1 - (d_{loc})^{-2}$. As a property of the binomial distribution, we have $P_{m-1} < P_m$ for $m < (N+1)p$. Therefore, we have $\dim \mathcal{A}_N^{[0,m]} \leq (m+1)P_m$ for $m/N < p$.

To obtain the lower bound (3), we employ Stirling's formula, obtaining

$$\frac{\dim \mathcal{A}_N^{[0,m]}}{D} \leq \exp\Big[NG\Big(\frac{m}{N}\Big) + \mathcal{O}(\log N)\Big] \quad (11)$$

for $m/N < p$, where $G(x) := H(x) + x\log(d_{loc}^2 - 1) - \log d_{loc}$, and $H(x) = -x\log x - (1-x)\log(1-x)$. The root $\alpha_*^{(L)}$ of $G(x)$ lies in $(0, 1/2)$, and $G(x)$ satisfies $G(x) < 0$ for $x < \alpha_*^{(L)}$. Therefore, we obtain the inequality (3) from the upper bound of the inequality (2) in Theorem 1.

The upper bound (4) follows immediately from the fact that $\mathcal{B}(N,p)$ converges to a Gaussian distribution $\mathcal{N}(Np, Np(1-p))$ for large $N$ in distribution.

For $N$-particle Bose and Fermi systems on a $V$-site lattice, by setting $\alpha := m/N$ and $\rho := N/V$, we have

$$\dim \mathcal{A}_N^{[0,m]} = \binom{V+m-1}{m}^2$$
$$= \exp\left[2V(1+\alpha\rho)H\Big(\frac{1}{1+\alpha\rho}\Big) + \mathcal{O}(\log V)\right], \quad (12)$$

for Bose systems and

$$\dim \mathcal{A}_N^{[0,m]} = \binom{V}{\min\{m, V-N\}}^2$$
$$= \begin{cases} \exp[2VH(\alpha\rho) + \mathcal{O}(\log V)] & (m < V-N); \\ D^2 & (m \geq V-N), \end{cases} \quad (13)$$

for Fermi systems. The dimension of $\mathcal{A}_N^{[m_-,m_+]}$ is given by $\dim \mathcal{A}_N^{[0,m_+]} - \dim \mathcal{A}_N^{[0,m_--1]}$ (see Supplemental Information IV for details). It follows from $D^2 = \dim \mathcal{A}_N^{[0,N]}$ and some straightforward calculations that Theorem 2 is proved.

## AUTHOR CONTRIBUTIONS

All authors contributed equally to this work.

# Supplemental Information:
# Rigorous Bounds on Eigenstate Thermalization

Shoki Sugimoto,[1] Ryusuke Hamazaki,[2] and Masahito Ueda[1,3]

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

[2]*Nonequilibrium Quantum Statistical Mechanics RIKEN Hakubi Research Team,*
*RIKEN Cluster for Pioneering Research (CPR),*
*RIKEN iTHEMS, Wako, Saitama 351-0198, Japan*

[3]*RIKEN Center for Emergent Matter Science (CEMS), Wako 351-0198, Japan*

(Dated: March 20, 2023)

## I. $\|\cdot\|_1^{(\mathcal{A})}$ AS A UNIFIED MEASURE FOR QUANTUM-THERMAL EQUILIBRIUM

As mentioned in the main text, the seminorm $\|\cdot\|_1^{(\mathcal{A})}$ with an apt choice of $\mathcal{A}$ serves as a unified measure for various notions of quantum-thermal equilibrium. In this section, we provide some examples.

### A. Subsystem thermal equilibrium

For any subsystem $\mathcal{S}$ and any $\hat{A}_{\mathcal{S}} \otimes \mathrm{id}_{\mathcal{S}^c} \in \mathcal{L}(\mathcal{H}_{\mathcal{S}}) \otimes \mathrm{id}_{\mathcal{S}^c}$, we have $\|\hat{A}_{\mathcal{S}} \otimes \mathrm{id}_{\mathcal{S}^c}\|_\infty = \|\hat{A}_{\mathcal{S}}\|_\infty$. By setting $\mathcal{A} = \mathcal{L}(\mathcal{H}_{\mathcal{S}}) \otimes \mathrm{id}_{\mathcal{S}^c}$, we have

$$
\begin{aligned}
\left\|\hat{X}\right\|_1^{(\mathcal{A})} &= \max_{\hat{A} \in \mathcal{L}(\mathcal{H}_{\mathcal{S}}) \otimes \mathrm{id}_{\mathcal{S}^c}} \left| \mathrm{tr}\left( \frac{\hat{A}}{\|\hat{A}\|_\infty} \hat{X} \right) \right| \\
&= \max_{\hat{A}_{\mathcal{S}} \in \mathcal{L}(\mathcal{H}_{\mathcal{S}})} \left| \mathrm{tr}\left( \frac{\hat{A}_{\mathcal{S}}}{\|\hat{A}_{\mathcal{S}}\|_\infty} \mathrm{tr}_{\mathcal{S}^c}(\hat{X}) \right) \right| \\
&= \left\| \mathrm{tr}_{\mathcal{S}^c}(\hat{X}) \right\|_1.
\end{aligned}
\tag{S1}
$$

Thus, the seminorm $\|\cdot\|_1^{(\mathcal{A})}$ reduces to the trace norm on a subsystem $\mathcal{S}$ for $\mathcal{A} = \mathcal{L}(\mathcal{H}_{\mathcal{S}}) \otimes \mathrm{id}_{\mathcal{S}^c}$, and the smallness of $\|\hat{\sigma} - \hat{\rho}_{\mathrm{th}}\|_1^{(\mathcal{A})}$ for a thermal ensemble $\hat{\rho}_{\mathrm{th}}$ defines those quantum states $\hat{\sigma}$ that are in thermal equilibrium in a subsystem $\mathcal{S}$.

## B.  Microscopic thermal equilibrium (MITE) [1–3]

The notion of microscopic thermal equilibrium (MITE) is introduced in Refs. [1, 2] as follows.

---

**Definition I.1: Microscopic thermal equilibrium (MITE) [2]**

A state $\hat{\sigma}$ is said to be in microscopic thermal equilibrium (MITE) on a length scale $l_0$ if it satisfies

$$\|\mathrm{tr}_{\mathcal{S}^c}(\hat{\sigma}) - \mathrm{tr}_{\mathcal{S}^c}(\hat{\rho}_{\mathrm{th}})\|_1 < \epsilon \tag{S2}$$

for every subsystem $\mathcal{S}$ with $\mathrm{Diam}(\mathcal{S}) \leq l_0$, where $\epsilon \ll 1$ and the diameter of $\mathcal{S}$ is defined by $\mathrm{Diam}(\mathcal{S}) := \sup_{x,y \in \mathcal{S}} d(x,y)$ for some distance $d$.

---

As mentioned in Ref. [2], MITE can be regarded as the thermal equilibrium relative to

$$\mathcal{A}_{\mathrm{MITE}} = \bigcup_{\mathcal{S}:\ \mathrm{Diam}(\mathcal{S}) \leq l_0} \mathcal{L}(\mathcal{H}_{\mathcal{S}}) \otimes \mathrm{id}_{\mathcal{S}^c}. \tag{S3}$$

Then, we have the following proposition.

---

**Proposition I.1**

Let $\hat{\sigma}$ be an arbitrary quantum state and $\hat{\rho}_{\mathrm{th}}$ be a thermal ensemble. Then,

$$\hat{\sigma} \text{ is in MITE} \iff \|\hat{\sigma} - \hat{\rho}_{\mathrm{th}}\|_1^{(\mathcal{A}_{\mathrm{MITE}})} \leq \epsilon, \tag{S4}$$

where $\epsilon \ll 1$.

---

*Proof.* The proof follows immediately from the following equation:

$$
\begin{aligned}
\|\hat{\sigma} - \hat{\rho}_{\mathrm{th}}\|_1^{(\mathcal{A}_{\mathrm{MITE}})} &= \max_{\hat{A} \in \mathcal{A}_{\mathrm{MITE}}} \left| \mathrm{tr}\left( \frac{\hat{A}}{\|\hat{A}\|_\infty}(\hat{\sigma} - \hat{\rho}_{\mathrm{th}}) \right) \right| \\
&= \max_{\mathcal{S}:\ \mathrm{Diam}(\mathcal{S}) \leq l_0} \max_{\hat{A}_S \in \mathcal{L}(\mathcal{H}_S)} \left| \mathrm{tr}_S\left( \frac{\hat{A}_S}{\|\hat{A}_S\|_\infty}\left( \mathrm{tr}_{S^c}(\hat{\sigma}) - \mathrm{tr}_{S^c}(\hat{\rho}_{\mathrm{th}}) \right) \right) \right| \\
&= \max_{\mathcal{S}:\ \mathrm{Diam}(\mathcal{S}) \leq l_0} \|\mathrm{tr}_{S^c}(\hat{\sigma}) - \mathrm{tr}_{S^c}(\hat{\rho}_{\mathrm{th}})\|_1. 
\end{aligned} \tag{S5}
$$

∎

Mori et al. [3] extended the notion of MITE by lifting the spatial constraints $\mathrm{Diam}(\mathcal{S}) \leq l_0$ to

a "few-body" constraint as $|\mathcal{S}| = k$ with an integer $k$ of $\mathcal{O}(1)$, i.e., they consider

$$\mathcal{A}_{\mathrm{MITE}}^{(\mathrm{few})} = \bigcup_{\mathcal{S}:\,|\mathcal{S}|=k} \mathcal{L}(\mathcal{H}_{\mathcal{S}}) \otimes \mathrm{id}_{\mathcal{S}^c} \tag{S6}$$

in addition to $\mathcal{A}_{\mathrm{MITE}}$.

The same proof for Proposition I.1 applies to the MITE with respect to $\mathcal{A}_{\mathrm{MITE}}^{(\mathrm{few})}$, and we have the following proposition;

---

**Proposition I.2**

Let $\hat{\sigma}$ be an arbitrary quantum state and $\hat{\rho}_{\mathrm{th}}$ be a thermal ensemble. Then,

$$\hat{\sigma} \text{ is in MITE with respect to } \mathcal{A}_{\mathrm{MITE}}^{(\mathrm{few})} \iff \|\hat{\sigma} - \hat{\rho}_{\mathrm{th}}\|_1^{(\mathcal{A}_{\mathrm{MITE}}^{(\mathrm{few})})} \leq \epsilon, \tag{S7}$$

where $\epsilon \ll 1$.

---

### C. Macroscopic thermal equilibrium (MATE) [1–3]

The notion of macroscopic thermal equilibrium (MATE) is introduced in Refs. [1, 2] as follows.

---

**Definition I.2: Macroscopic thermal equilibrium (MATE) [2]**

Consider a collection of macro observables $\hat{M}_1, \cdots, \hat{M}_K$. By suitably coarse graining the operators $\hat{M}_1, \cdots, \hat{M}_K$, it is expected that we obtain a set of mutually commuting operators $\tilde{M}_1, \cdots, \tilde{M}_K$ with $\tilde{M}_j \approx \hat{M}_j$ for all $j = 1, \cdots, K$. We take $\tilde{M}_1$ as the coarse-grained Hamiltonian, whose eigenspaces are energy shells $\mathcal{H}_{E,\Delta E}$.

Since $\tilde{M}_1, \cdots, \tilde{M}_K$ commute with each other, we can simultaneously diagonalize them, and the energy shell $\mathcal{H}_{E,\Delta E}$ can be decomposed accordingly as $\mathcal{H}_{E,\Delta E} = \bigoplus_\nu \mathcal{H}_\nu$. Here, $\mathcal{H}_\nu$ is called macro-spaces, and we denote the projector onto $\mathcal{H}_\nu$ by $\hat{P}_\nu$.

In each $\mathcal{H}_{E,\Delta E}$, it is expected that one macro-space called thermal equilibrium macro-space $\mathcal{H}_{\mathrm{eq}}$ covers the most of the dimensions of $\mathcal{H}_{E,\Delta E}$, i.e.,

$$\frac{\dim \mathcal{H}_{\mathrm{eq}}}{\dim \mathcal{H}_{E,\Delta E}} = 1 - \tilde{\epsilon} \tag{S8}$$

with $\tilde{\epsilon} \ll 1$. Without loss of generality, we set $\mathcal{H}_{\mathrm{eq}} = \mathcal{H}_{\nu=1}$.

Under these setups, a state $\hat{\sigma} \in \dim \mathcal{H}_{E,\delta E}$ is said to be in macroscopic thermal equilib-

rium (MATE) if and only if

$$\mathrm{tr}(\hat{\sigma}\hat{P}_{\mathrm{eq}}) \geq 1 - \delta \tag{S9}$$

for a suitably small tolerance $\delta > 0$.

As mentioned in Ref. [2], MATE can be regarded as the thermal equilibrium relative to the (coarse-grained) macroscopic observables $\tilde{M}_1, \cdots, \tilde{M}_K$. Because we focus on the joint distribution of the observed values of $\tilde{M}_1, \cdots, \tilde{M}_K$ in MATE, $\mathcal{A}$ for MATE is given by

$$\mathcal{A}_{\mathrm{MATE}} = \{\hat{P}_{\mathrm{eq}}\}. \tag{S10}$$

Then, we have the following proposition.

**Proposition I.3**

Let $\hat{\sigma}$ be an arbitrary quantum state. Then,

$$\hat{\sigma} \text{ is in MATE with tolerance } \delta \, (\geq 2\tilde{\epsilon}) \iff \left\| \hat{\sigma} - \hat{\rho}_{\delta E}^{(\mathrm{mc})} \right\|_1^{(\mathcal{A}_{\mathrm{MATE}})} \leq \delta - \tilde{\epsilon}. \tag{S11}$$

*Proof.* For $\mathcal{A}_{\mathrm{MATE}} = \{\hat{P}_{\mathrm{eq}}\}$, we have

$$\left\| \hat{\sigma} - \hat{\rho}_{\delta E}^{(\mathrm{mc})} \right\|_1^{(\mathcal{A}_{\mathrm{MATE}})} = \left| \mathrm{tr}(\hat{\sigma}\hat{P}_{\mathrm{eq}}) - \frac{\dim \mathcal{H}_{\mathrm{eq}}}{\dim \mathcal{H}_{E,\Delta E}} \right|. \tag{S12}$$

Therefore, $\left\| \hat{\sigma} - \hat{\rho}_{\delta E}^{(\mathrm{mc})} \right\|_1^{(\mathcal{A}_{\mathrm{MATE}})} \leq \epsilon$ is equivalent to

$$1 - (\tilde{\epsilon} + \epsilon) \leq \mathrm{tr}(\hat{\sigma}\hat{P}_{\mathrm{eq}}) \leq 1 + (\epsilon - \tilde{\epsilon}). \tag{S13}$$

By setting $\epsilon := \delta - \tilde{\epsilon} \, (\geq \tilde{\epsilon})$, we obtain

$$\left\| \hat{\sigma} - \hat{\rho}_{\delta E}^{(\mathrm{mc})} \right\|_1^{(\mathcal{A}_{\mathrm{MATE}})} \leq \delta - \tilde{\epsilon} \iff 1 - \delta \leq \mathrm{tr}(\hat{\sigma}\hat{P}_{\mathrm{eq}}), \tag{S14}$$

which is the desired result. ∎

## II.   SCALING BEHAVIOR OF THE NORMALIZATION CONSTANT $\|\hat{A}\|_q$

Here, we demonstrate that the quantity $\hat{A}/\|\hat{A}\|_q$ in the definition of $\|\cdot\|_1^{(\mathcal{A})}$ in Eq. (1) in the main text is themodynamically intensive only for $q = \infty$. For that purpose, we derive the $N$-dependence of the normalization constant $\|\hat{A}\|_q$ for an extensive operator $\hat{M}_z := \sum_{j=1}^{N} \hat{\sigma}_j^{(z)}$, where $\hat{\sigma}^{(z)}$ is the Pauli $z$-operator. The eigenstates of $\hat{M}_z$ are given by tensor products of those of $\hat{\sigma}^{(z)}$, and the eigenvalues are given by $-N + 2j$ $(j = 0, \cdots, N)$, where $j$ is the number of spins where the local state is the eigenstate of $\hat{\sigma}^{(z)}$ with eigenvalue $+1$. Therefore, we have

$$
\|\hat{M}_z\|_q^q = \sum_{j=0}^{N} \binom{N}{j} |-N + 2j|^q
$$

$$
= N^q \sum_{j=0}^{N} \left| -1 + 2\frac{j}{N} \right|^q \exp\left[ NH\left(\frac{j}{N}\right) + \frac{1}{2} \log \frac{N}{2\pi j(N-j)} + \mathcal{O}(N^{-1}) \right], \tag{S15}
$$

where $H(x) := -x \log x - (1-x) \log(1-x)$ is the binary entropy.

Since we are interested in the $N$-dependence of $\|\hat{M}_z\|_q$ for large $N$, we approximate the sum with the integral, obtaining

$$
\|\hat{M}_z\|_q^q \simeq N^{q+1} \int_0^1 dx \, |-1 + 2x|^q \exp\left[ NH(x) - \frac{1}{2} \log\left(x(1-x)\right) - \frac{1}{2} \log(2\pi N) + \mathcal{O}(N^{-1}) \right]
$$

$$
\simeq \frac{N^{q+\frac{1}{2}}}{\sqrt{2\pi}} \int_0^1 dx \, |-1 + 2x|^q \exp\left[ NH(x) - \frac{1}{2} \log\left(x(1-x)\right) + \mathcal{O}(N^{-1}) \right]
$$

$$
\simeq \frac{N^{q+\frac{1}{2}}}{\sqrt{2\pi}} \int_{-\frac{1}{2}}^{\frac{1}{2}} dx \, |2x|^q \exp\left[ NH\left(\frac{1}{2} + x\right) - \frac{1}{2} \log\left(\frac{1}{4} - x^2\right) + \mathcal{O}(N^{-1}) \right]. \tag{S16}
$$

We then employ the saddle point method, which leads to

$$
\|\hat{M}_z\|_q^q \simeq \frac{N^{q+\frac{1}{2}}}{\sqrt{2\pi}} \int_{-\frac{1}{2}}^{\frac{1}{2}} dx \, |2x|^q \exp\left[ N \log 2 - \frac{4Nx^2}{2} + \log 2 + \frac{x^2}{2} + \mathcal{O}(x^4, N^{-1}) \right]
$$

$$
= 2^N \frac{N^{\frac{q}{2}}}{\sqrt{2\pi}} \int_{-\sqrt{N}}^{\sqrt{N}} dx \, |x|^q \exp\left[ -\frac{x^2}{2} + \mathcal{O}(N^{-1}) \right]
$$

$$
\simeq 2^N \frac{N^{\frac{q}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dx \, |x|^q \exp\left[ -\frac{x^2}{2} \right]
$$

$$
= 2^N N^{\frac{q}{2}} \frac{2^{\frac{q}{2}}}{\sqrt{\pi}} \Gamma\left(\frac{q+1}{2}\right). \tag{S17}
$$

Therefore, for any fixed $q$ and sufficiently large $N$, the normalization constant $\|\hat{M}_z\|_q$ scales as $\sim \sqrt{N} \, D^{\frac{1}{q}}$, where $D = 2^N$ is the dimension of the total Hilbert space.

## III.   PROOF OF THEOREM 1 IN THE MAIN TEXT

Theorem 1 in the main text is the consequence of the concentration inequality for the Haar measure on $\mathbb{SU}(D)$ [4–7] applied to the quantities $(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2$, where $\delta\hat{\rho}_{\alpha\beta} := \hat{\rho}_{\alpha\beta} - \hat{\rho}_{\delta E}^{(\mathrm{mc})}(E_\alpha)\delta_{\alpha\beta}$, and $(\Lambda_1^{(\hat{H},\mathcal{A})})^2$ as functions of the unitary matrix $\hat{U}$ that diagonalizes the Hamiltonian, i.e., $\hat{U}|E_\alpha^{(0)}\rangle = |E_\alpha\rangle$ for an arbitrarily fixed orthonormal basis $\{|E_\alpha^{(0)}\rangle\}_{\alpha=1}^D$ of $\mathcal{H}$. The concentration inequality for the Haar measure on $\mathbb{SU}(D)$ [4–7] implies that, for any $\delta > 0$ and an arbitrary Lipschitz function $f(\hat{U})$ of $\hat{U} \in \mathbb{SU}(D)$ with Lipschitz constant $\eta_f$,

$$\mathbb{P}\Big[f(\hat{U}) - \mathbb{E}[f] \geq \delta\Big] \leq \exp\left(-\frac{\delta^2 D}{4\eta_f^2}\right) \tag{S18}$$

$$\text{and} \quad \mathbb{P}\Big[f(\hat{U}) - \mathbb{E}[f] \leq -\delta\Big] \leq \exp\left(-\frac{\delta^2 D}{4\eta_f^2}\right), \tag{S19}$$

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

To apply the inequalities (S18) and (S19), we need (i) to evaluate the averages of $(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2$ and $(\Lambda_1^{(\hat{H},\mathcal{A})})^2$ and (ii) to show the Lipschitz continuity of $(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2$ and $(\Lambda_1^{(\hat{H},\mathcal{A})})^2$.

### A.   Estimation of $\mathbb{E}\Big[(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2\Big]$

As mentioned in the main text, it is in general difficult to calculate the maximum with respect to $\hat{A} \in \mathcal{A} + \mathbb{R}\hat{I}$ in $\|\cdot\|_1^{(\mathcal{A})}$. Therefore, we employ the inequality $\|\cdot\|_2^{(\mathcal{A})} \leq \|\cdot\|_1^{(\mathcal{A})} \leq \sqrt{D}\|\cdot\|_2^{(\mathcal{A})}$. Here, we can calculate $\|\cdot\|_2^{(\mathcal{A})}$ given an orthonormal basis of $\mathcal{A} + \mathbb{R}\hat{I}$ as in the following proposition.

---

**Proposition III.1**

Let $\mathcal{A}$ be a space of Hermitian operators and $\{\hat{\Lambda}^{(s)}\}_{s=1}^M$ be an orthonormal basis of $\mathcal{A} + \mathbb{R}\hat{I}$, where $M := \dim(\mathcal{A} + \mathbb{R}\hat{I})$. For an arbitrary linear operator $\hat{X}$ (not necessarily Hermitian), we introduce a column vector $\vec{X}$ with elements $X_s := \mathrm{tr}(\hat{\Lambda}^{(s)}\hat{X})$ $(s = 1, \cdots, M)$ and denote its Euclidean norm by $\|\vec{X}\|_2$.

Then, we have

$$\left\|\hat{X}\right\|_2^{(\mathcal{A})} := \max_{\hat{A}\in\mathcal{A}+\mathbb{R}\hat{I}}\left|\mathrm{tr}\left(\frac{\hat{A}}{\|\hat{A}\|_2}\hat{X}\right)\right| = \sqrt{\frac{\|\vec{X}\|_2^2 + |\vec{X}^T \cdot \vec{X}|}{2}}. \tag{S20}$$

In particular, we have

$$\frac{1}{\sqrt{2}}\|\vec{X}\|_2 \leq \left\|\hat{X}\right\|_2^{(\mathcal{A})} \leq \|\vec{X}\|_2. \tag{S21}$$

*Proof.* For an arbitrary operator $\hat{A} \in \mathcal{A} + \mathbb{R}\hat{I}$, we expand it as

$$\hat{A} = \sum_{s=1}^{M} c_s \hat{\Lambda}^{(s)}, \quad c_s := \mathrm{tr}\left(\hat{\Lambda}^{(s)}\hat{A}\right) \in \mathbb{R}. \tag{S22}$$

Then, we have

$$\begin{aligned}
\left\|\hat{X}\right\|_2^{(\mathcal{A})} &= \max_{\hat{A} \in \mathcal{A} + \mathbb{R}\hat{I}} \left| \mathrm{tr}\left(\frac{\hat{A}}{\|\hat{A}\|_2}\hat{X}\right)\right| \\
&= \max_{\hat{A} \in \mathcal{A} + \mathbb{R}\hat{I}} \frac{\left|\sum_{s=1}^{M} c_s \mathrm{tr}\left(\hat{\Lambda}^{(s)}\hat{X}\right)\right|}{\|\vec{c}\|_2} \\
&= \max_{\vec{c}:\ \|\vec{c}\|_2=1} \left|\vec{c}\cdot\left(\mathrm{Re}\,\vec{X} + i\,\mathrm{Im}\,\vec{X}\right)\right| \\
&= \max_{\vec{c}:\ \|\vec{c}\|_2=1} \sqrt{\vec{c}^T \cdot \left[(\mathrm{Re}\,\vec{X})(\mathrm{Re}\,\vec{X})^T + (\mathrm{Im}\,\vec{X})(\mathrm{Im}\,\vec{X})^T\right] \cdot \vec{c}}, \tag{S23}
\end{aligned}$$

where we introduced the column vectors $\vec{c} := (c_1, c_2, \cdots)^T$ and $\vec{X} := (\mathrm{tr}(\hat{\Lambda}^{(1)}\hat{X}), \mathrm{tr}(\hat{\Lambda}^{(2)}\hat{X}), \cdots)^T$.

If $\mathrm{Im}\,\vec{X} \propto \mathrm{Re}\,\vec{X}$, the maximum in Eq. (S23) is attained when $\vec{c} \propto \mathrm{Re}\,\vec{X}$, and we obtain

$$\left\|\hat{X}\right\|_2^{(\mathcal{A})} = \sqrt{\|\mathrm{Re}\,\vec{X}\|_2^2 + \|\mathrm{Im}\,\vec{X}\|_2^2} = \sqrt{\sum_{s=1}^{M}\left|\mathrm{tr}\left(\hat{\Lambda}^{(s)}\hat{X}\right)\right|^2}. \tag{S24}$$

When $\mathrm{Im}\,\vec{X} \not\propto \mathrm{Re}\,\vec{X}$, we define $\cos\theta := (\mathrm{Re}\,\vec{X}^T \cdot \mathrm{Im}\,\vec{X})/\|\mathrm{Re}\,\vec{X}\|_2\|\mathrm{Im}\,\vec{X}\|_2$, and introduce the following orthonormal vectors:

$$\vec{e}_1 := \frac{\mathrm{Re}\,\vec{X}}{\|\mathrm{Re}\,\vec{X}\|_2}, \quad \vec{e}_2 := \frac{1}{\sin\theta}\left(\frac{\mathrm{Im}\,\vec{X}}{\|\mathrm{Im}\,\vec{X}\|_2} - \vec{e}_1\cos\theta\right). \tag{S25}$$

We then obtain an orthonormal basis of $\mathbb{R}^M$ by extending the orthonormal set $\{\vec{e}_1, \vec{e}_2\}$. The matrix $[(\mathrm{Re}\,\vec{X})(\mathrm{Re}\,\vec{X})^T + (\mathrm{Im}\,\vec{X})(\mathrm{Im}\,\vec{X})^T]$ in this basis is given by

$$\begin{aligned}
&(\mathrm{Re}\,\vec{X})(\mathrm{Re}\,\vec{X})^T + (\mathrm{Im}\,\vec{X})(\mathrm{Im}\,\vec{X})^T \\
&= \left(\begin{array}{cc|c}
\|\mathrm{Re}\,\vec{X}\|_2^2 + \|\mathrm{Im}\,\vec{X}\|_2^2\cos^2\theta & \|\mathrm{Im}\,\vec{X}\|_2^2\cos\theta\sin\theta & \\
\|\mathrm{Im}\,\vec{X}\|_2^2\cos\theta\sin\theta & \|\mathrm{Im}\,\vec{X}\|_2^2\sin^2\theta & \mathbf{0}_{2,M-2} \\
\hline
\multicolumn{2}{c|}{\mathbf{0}_{M-2,2}} & \mathbf{0}_{M-2,M-2}
\end{array}\right). \tag{S26}
\end{aligned}$$

where $\mathbf{0}_{m,n}$ denotes the $m \times n$ zero matrix. The nontrivial eigenvalues of the matrix (S26) are

$$\lambda_{\pm} = \frac{\|\vec{X}\|_2^2 \pm |\vec{X}^T \cdot \vec{X}|}{2}, \tag{S27}$$

where

$$\vec{X}^T \cdot \vec{X} = \sum_{s=1}^{M} \left[ \mathrm{tr} \left( \hat{\Lambda}^{(s)} \hat{X} \right) \right]^2 = \|\mathrm{Re}\,\vec{X}\|_2^2 - \|\mathrm{Im}\,\vec{X}\|_2^2 + 2i\|\mathrm{Re}\,\vec{X}\|_2\|\mathrm{Im}\,\vec{X}\|_2 \cos\theta. \tag{S28}$$

The maximum in Eq. (S23) is attained when $\vec{c}$ is an eigenvector belonging to the eigenvalue $\lambda_+$. Therefore, we obtain

$$\left\| \hat{X} \right\|_2^{(\mathcal{A})} = \sqrt{\frac{\|\vec{X}\|_2^2 + |\vec{X}^T \cdot \vec{X}|}{2}}. \tag{S29}$$

The equation (S29) reduces to Eq. (S24) when $\mathrm{Im}\,\vec{X} \propto \mathrm{Re}\,\vec{X}$. ∎

In our application, we set

$$\hat{X} = \delta\hat{\rho}_{\alpha\beta} \quad \left( := \hat{\rho}_{\alpha\beta} - \hat{\rho}_{\delta E}^{(\mathrm{mc})}(E_\alpha)\delta_{\alpha\beta} \right). \tag{S30}$$

Recall that $\hat{U}$ is the unitary operator that diagonalizes the Hamiltonian, i.e., $\hat{U} |E_\gamma^{(0)}\rangle = |E_\gamma\rangle$. Therefore, $\hat{X}_0 := \hat{U}^\dagger \hat{X} \hat{U}$ is independent of $\hat{U}$. Then, our task is to estimate the Haar average of the quantity $\|\vec{X}\|_2^2 = \sum_{s=1}^{M} \left| \mathrm{tr} \left( \hat{\Lambda}^{(s)} \hat{U} \hat{X}_0 \hat{U}^\dagger \right) \right|^2$.

By explicitly calculating the fourth moments of the unitary Haar measure, we obtain

$$\mathbb{E}\left[ \|\vec{X}\|_2^2 \right] = \frac{M-1}{D^2-1} \|\hat{X}_0\|_2^2. \tag{S31}$$

It is also straightforward to show

$$\|\hat{X}_0\|_2^2 = 1 - \frac{1}{\dim \mathcal{H}_{E_\alpha, \delta E}} \delta_{\alpha\beta}, \tag{S32}$$

which implies $\frac{1}{2} \leq \|\hat{X}_0\|_2^2 \leq 1$ except for the trivial case of $\dim \mathcal{H}_{E_\alpha, \delta E} = 1$. We do not consider this trivial case. Then, Proposition III.1 gives

$$\frac{1}{4}\frac{M}{D^2} \leq \mathbb{E}\left[ (\|\delta\hat{\rho}_{\alpha\beta}\|_2^{(\mathcal{A})})^2 \right] + \mathcal{O}\left( \frac{1}{D^2} \right) \leq \frac{M}{D^2}, \tag{S33}$$

as claimed in the Methods section of the main text. Combining Proposition III.1 and the bound $\|\cdot\|_2^{(\mathcal{A})} \leq \|\cdot\|_1^{(\mathcal{A})} \leq \sqrt{D}\|\cdot\|_2^{(\mathcal{A})}$, we obtain

$$\frac{1}{4}\frac{M}{D^2} \leq \mathbb{E}\left[ (\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2 \right] + \mathcal{O}\left( \frac{1}{D^2} \right) \leq \frac{M}{D}. \tag{S34}$$

## B.   Estimation of $\mathbb{E}[(\Lambda_1^{(\hat{H},\mathcal{A})})^2]$

The remaining part of the proof of Theorem 1 in the main text applies to $(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})})^n$ for arbitrary $p > 0$ and $n > 0$ without any change, so we consider the general case $(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})})^n$, including $(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^2$. Accordingly, we introduce $\Lambda_p^{(\hat{H},\mathcal{A})} := \max_{|E_\alpha\rangle,|E_\beta\rangle \in \mathcal{H}_{E,\Delta E}} \|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})}$.

To estimate $\mathbb{E}[(\Lambda_1^{(\hat{H},\mathcal{A})})^2]$, we need the Lipchitz continuity of $(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})})^n$, which we prove in the next subsection. In this subsection, we assume that $(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})})^n$ as a function of $\hat{U}$ is Lipschitz continuous and that its Lipschitz constant is bounded from above by a constant $\eta_n$ independent of $D$. Then, the concentration inequality (S18) for $(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})})^n$ gives

$$\mathbb{P}\left[(\Lambda_p^{(\hat{H},\mathcal{A})})^n - \mathbb{E}\left[(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})})^n\right] \geq \delta\right] \leq d_{E,\delta E}^2 \mathbb{P}\left[(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})})^n - \mathbb{E}\left[(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})})^n\right] \geq \delta\right]$$
$$\leq \exp\left(-\frac{\delta^2 D}{4\eta_n^2} + 2\log d_{E,\Delta E}\right), \tag{S35}$$

where $d_{E,\Delta E} := \dim \mathcal{H}_{E,\Delta E}$. We introduce

$$\delta_0 := \sqrt{8\eta_n^2 \frac{\log d_{E,\Delta E}}{D}}, \quad \left(\implies -\frac{\delta_0^2 D}{4\eta_n^2} + 2\log d_{E,\Delta E} = 0\right), \tag{S36}$$

and obtain

$$\begin{aligned}
\mathbb{E}[(\Lambda_p^{(\hat{H},\mathcal{A})})^n] &= \int_0^\infty x \frac{d}{dx}\left(-\mathbb{P}[(\Lambda_p^{(\hat{H},\mathcal{A})})^n \geq x]\right) dx \\
&= \int_0^\infty \mathbb{P}[(\Lambda_p^{(\hat{H},\mathcal{A})})^n \geq x] dx \\
&\leq \int_0^{\mathbb{E}[(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})})^n]+\delta_0} dx + \int_{\delta_0}^\infty \exp\left(-\frac{x^2 D}{4\eta_n^2} + 2\log d_{E,\Delta E}\right) dx \\
&= \mathbb{E}\left[\left(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})}\right)^n\right] + \delta_0 + \int_{\delta_0}^\infty \exp\left(-\frac{(x^2 - \delta_0^2)D}{4\eta_n^2}\right) dx \\
&\leq \mathbb{E}\left[\left(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})}\right)^n\right] + \delta_0 + \int_0^\infty \exp\left(-\frac{x^2 D}{4\eta_n^2}\right) dx \\
&\leq \mathbb{E}\left[\left(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})}\right)^n\right] + \delta_0 + \sqrt{\frac{\pi\eta_n^2}{D}} \\
&= \mathbb{E}\left[\left(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})}\right)^n\right] + \mathcal{O}\left(\sqrt{\frac{\log d_{E,\Delta E}}{D}}\right).
\end{aligned} \tag{S37}$$

This result together with a trivial inequality $\mathbb{E}[(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})})^n] \leq \mathbb{E}[(\Lambda_1^{(\hat{H},\mathcal{A})})^n]$ leads to

$$\mathbb{E}[(\Lambda_p^{(\hat{H},\mathcal{A})})^n] = \mathbb{E}\left[\left(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})}\right)^n\right] + \mathcal{O}\left(\sqrt{\frac{\log d_{E,\Delta E}}{D}}\right). \tag{S38}$$

## C.   Lipschitz continuity of $\left(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})}\right)^n$

To apply the concentration inequality to $(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})})^n$, we need to show its Lipschitz continuity as a function of a unitary operator $\hat{U} \in \mathbb{SU}(D)$. We also derive a $D$-independent upper bound for its Lipschitz constant $\eta_n$.

We introduce the function

$$f_p^{(A)}(\hat{U}) := \mathrm{tr}\left(\frac{\hat{A}}{\|\hat{A}\|_q}\hat{U}\hat{X}_0\hat{U}^\dagger\right), \tag{S39}$$

where $\hat{X}_0 := \hat{U}^\dagger\delta\hat{\rho}_{\alpha\beta}\hat{U}$ does not depend on $\hat{U}$. We omit the indices $\alpha$ and $\beta$ in the definition of $\hat{X}_0$ because they play no role in the following argument. The Hölder inequality and the well-known inequality $\|\hat{A}\|_\infty \le \|\hat{A}\|_q$ for any $q \ge 1$ give

$$|f_p^{(A)}(\hat{U})| \le \frac{\|\hat{A}\|_\infty}{\|\hat{A}\|_q}\|\hat{U}\hat{X}_0\hat{U}^\dagger\|_1 \le \|\hat{X}_0\|_1. \tag{S40}$$

The quantity $\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})}$ can be rewritten in terms of $f_p^{(A)}(\hat{U})$ as

$$\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})} = \max_{\hat{A}\in\mathcal{A}+\mathbb{R}\hat{I}}\max_{\sigma=\pm}(-1)^\sigma f_p^{(A)}(\hat{U}) =: F_p^{(\mathcal{A})}(\hat{U}). \tag{S41}$$

The Lipschitz continuity of $(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})})^n$ will be established if we can show that the Lipschitz constant of $[f_p^{(A)}(\hat{U})]^n$ are bounded from above by a constant independent of $\hat{A}$. To see this, we consider the following inequality:

$$\left|[F_p^{(\mathcal{A})}(\hat{U}_1)]^n - [F_p^{(\mathcal{A})}(\hat{U}_2)]^n\right| = \left|\max_{\hat{A}\in\mathcal{A}+\mathbb{R}\hat{I}}\max_{\sigma=\pm}(-1)^\sigma[f_p^{(A)}(\hat{U}_1)]^n - \max_{\hat{A}\in\mathcal{A}+\mathbb{R}\hat{I}}\max_{\sigma=\pm}(-1)^\sigma[f_p^{(A)}(\hat{U}_2)]^n\right|$$

$$\le \max_{\hat{A}\in\mathcal{A}+\mathbb{R}\hat{I}}\max_{\sigma=\pm}\left|[f_p^{(A)}(\hat{U}_1)]^n - [f_p^{(A)}(\hat{U}_2)]^n\right|$$

$$= \max_{\hat{A}\in\mathcal{A}+\mathbb{R}\hat{I}}\left|\int_{U_{1\to2}} n[f_p^{(A)}(\hat{U})]^{n-1}\left(\nabla_U f_p^{(A)}(\hat{U})\cdot\mathrm{d}\hat{U}\right)\right|, \tag{S42}$$

where $U_{1\to2}$ is a straight path connecting $\hat{U}_1$ and $\hat{U}_2$ defined by

$$\hat{U}_{1\to2}(t) := t\hat{U}_1 + (1-t)\hat{U}_2, \quad t \in [0,1], \tag{S43}$$

and

$$\nabla_U f_p^{(A)}(\hat{U})\cdot\mathrm{d}\hat{U} = \sum_{ij}\left(\frac{\partial f_p^{(A)}}{\partial U_{ij}}\mathrm{d}U_{ij} + \frac{\partial f_p^{(A)}}{\partial U_{ij}^*}\mathrm{d}U_{ij}^*\right) \tag{S44}$$

with $U_{ij}$ being a matrix element of $\hat{U}$ with respect to an orthonormal basis $\{|j\rangle\}_{j=1}^{D}$. Here, we have $\|\hat{U}_{1\to 2}(t)\|_{\infty} \leq t\|\hat{U}_1\|_{\infty} + (1-t)\|\hat{U}_2\|_{\infty} = 1$. We then obtain

$$\left|[F_p^{(\mathcal{A})}(\hat{U}_1)]^n - [F_p^{(\mathcal{A})}(\hat{U}_2)]^n\right|$$

$$\leq \max_{\hat{A}\in\mathcal{A}+\mathbb{R}\hat{I}} \max_{t\in[0,1]} \left(n\left|f_p^{(A)}(\hat{U}_{1\to 2}(t))\right|^{n-1} \left\|\nabla_U f_p^{(A)}(\hat{U}_{1\to 2}(t))\right\|_2\right) \int_{U_{1\to 2}} \|\mathrm{d}\hat{U}\|_2$$

$$= \max_{\hat{A}\in\mathcal{A}+\mathbb{R}\hat{I}} \max_{t\in[0,1]} \left(n\left|f_p^{(A)}(\hat{U}_{1\to 2}(t))\right|^{n-1} \left\|\nabla_U f_p^{(A)}(\hat{U}_{1\to 2}(t))\right\|_2\right) \|\hat{U}_1 - \hat{U}_2\|_2. \tag{S45}$$

The derivative of $f_p^{(A)}(\hat{U})$ with respect to the matrix elements of $\hat{U}$ is given by

$$\frac{\partial f_p^{(A)}}{\partial U_{ij}} = \frac{1}{\|\hat{A}\|_q} \langle j|\hat{X}_0\hat{U}^\dagger\hat{A}|i\rangle, \quad \frac{\partial f_p^{(A)}}{\partial U_{ij}^*} = \frac{1}{\|\hat{A}\|_q} \langle i|\hat{A}\hat{U}\hat{X}_0|j\rangle, \tag{S46}$$

and therefore

$$\sup_{\hat{U}:\,\|\hat{U}\|_{\infty}\leq 1} \left\|\nabla f_p^{(A)}(\hat{U})\right\|_2 = \sup_{\hat{U}:\,\|\hat{U}\|_{\infty}\leq 1} \sqrt{\sum_{ij}\left(\left|\frac{\partial f_p^{(A)}}{\partial U_{ij}}\right|^2 + \left|\frac{\partial f_p^{(A)}}{\partial U_{ij}^*}\right|^2\right)}$$

$$= \sup_{\hat{U}:\,\|\hat{U}\|_{\infty}\leq 1} \sqrt{\frac{1}{\|\hat{A}\|_q^2}\sum_{ij}\left(\left|\langle j|\hat{X}_0\hat{U}^\dagger\hat{A}|i\rangle\right|^2 + \left|\langle i|\hat{A}\hat{U}\hat{X}_0|j\rangle\right|^2\right)}$$

$$= \sup_{\hat{U}:\,\|\hat{U}\|_{\infty}\leq 1} \sqrt{\frac{1}{\|\hat{A}\|_q^2}\left(\|\hat{X}_0\hat{U}^\dagger\hat{A}\|_2^2 + \|\hat{A}\hat{U}\hat{X}_0\|_2^2\right)}. \tag{S47}$$

By applying the Hölder inequality $\|\hat{X}_0\hat{U}^\dagger\hat{A}\|_2 \leq \|\hat{X}_0\|_2\|\hat{U}^\dagger\hat{A}\|_{\infty}(=\|\hat{X}_0\|_2\|\hat{A}\|_{\infty})$ and the well-known inequality $\|\hat{A}\|_{\infty} \leq \|\hat{A}\|_q$ for any $q \geq 1$, we obtain

$$\sup_{\hat{U}:\,\|\hat{U}\|_{\infty}\leq 1} \left\|\nabla f_p^{(A)}(\hat{U})\right\|_2 = \sqrt{2}\|\hat{X}_0\|_2. \tag{S48}$$

Finally, by combining Eqs. (S45), (S48) and $\|\hat{X}_0\|_2 \leq \|\hat{X}_0\|_1 \leq 2$, we obtain

$$\left|[F_p^{(\mathcal{A})}(\hat{U}_1)]^n - [F_p^{(\mathcal{A})}(\hat{U}_2)]^n\right| \leq n2^{n+\frac{1}{2}}\|\hat{U}_1 - \hat{U}_2\|_2, \tag{S49}$$

which gives a $D$-independent upper bound $n2^{n+\frac{1}{2}}$ for the Lipschitz constant $\eta_n$ of $(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})})^n$.

## D. Lipschitz continuity of $\left(\Lambda_1^{(\hat{H},\mathcal{A})}\right)^n$

The Lipschitz continuity of $\left(\Lambda_1^{(\hat{H},\mathcal{A})}\right)^n$ follows from that of $(\|\delta\hat{\rho}_{\alpha\beta}\|_p^{(\mathcal{A})})^n$ in the same way as in the inequality in Eq. (S42):

$$\left|\left(\Lambda_1^{(\hat{H}_1,\mathcal{A})}\right)^n - \left(\Lambda_1^{(\hat{H}_2,\mathcal{A})}\right)^n\right| = \left|\max_{|E_\alpha\rangle,|E_\beta\rangle\in\mathcal{H}_{E,\Delta E}} [F_p^{(\mathcal{A})}(\hat{U}_1)]^n - \max_{|E_\alpha\rangle,|E_\beta\rangle\in\mathcal{H}_{E,\Delta E}} [F_p^{(\mathcal{A})}(\hat{U}_2)]^n\right|$$

$$\leq \max_{|E_\alpha\rangle,|E_\beta\rangle\in\mathcal{H}_{E,\Delta E}} \left|[F_p^{(\mathcal{A})}(\hat{U}_1)]^n - [F_p^{(\mathcal{A})}(\hat{U}_2)]^n\right|$$

$$\leq \eta_n\|\hat{U}_1 - \hat{U}_2\|_2, \tag{S50}$$

where in the last inequality, we used the Lipschitz continuity of $[F_p^{(\mathcal{A})}(\hat{U}_1)]^n$ which is proved in the previous subsection. Therefore, $\left(\Lambda_1^{(\hat{H}_1,\mathcal{A})}\right)^n$ is Lipschitz continuous with Lipschitz constant no larger than $\eta_n$.

## E. Derivation of the bounds for $(\Lambda_p^{(\hat{H},\mathcal{A})})^n$ in terms of $\mathbb{E}\left[(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})})^n\right]$

By applying the concentration inequalities (S18) and (S19) to $(\Lambda_p^{(\hat{H},\mathcal{A})})^n$, we obtain

$$\mathbb{P}\left[\left|(\Lambda_p^{(\hat{H},\mathcal{A})})^n - \mathbb{E}[(\Lambda_p^{(\hat{H},\mathcal{A})})^n]\right| \geq \delta\right] \leq \exp\left(-\mathcal{O}(\delta^2 D)\right). \tag{S51}$$

By substituting Eq. (S38) into Eq. (S51) and suitably setting $\delta = \frac{D^\epsilon}{\sqrt{D}} + \mathcal{O}\left(\sqrt{\frac{\log d_{E,\Delta E}}{D}}\right)$, we obtain

$$\mathbb{P}\left[\left|(\Lambda_p^{(\hat{H},\mathcal{A})})^n - \mathbb{E}\left[\left(\|\delta\hat{\rho}_{\alpha\beta}\|_1^{(\mathcal{A})}\right)^n\right]\right| \geq \frac{D^\epsilon}{\sqrt{D}}\right] \leq \exp\left(-\mathcal{O}(D^{2\epsilon})\right). \tag{S52}$$

Finally, Eq. (S52) for $n = 2$ and $p = 1$ combined with Eq. (S34) yields

$$\mathbb{P}\left[(\Lambda_p^{(\hat{H},\mathcal{A})})^n - \frac{M}{D} \geq \frac{D^\epsilon}{\sqrt{D}}\right] \leq \exp\left(-\mathcal{O}(D^{2\epsilon})\right)$$

$$\text{and} \quad \mathbb{P}\left[(\Lambda_p^{(\hat{H},\mathcal{A})})^n - \frac{M}{4D^2} \leq \frac{D^\epsilon}{\sqrt{D}}\right] \leq \exp\left(-\mathcal{O}(D^{2\epsilon})\right), \tag{S53}$$

which is the precise meaning of the statement "$\frac{M}{4D^2} \leq (\Lambda_p^{(\hat{H},\mathcal{A})})^n + \mathcal{O}\left(\frac{D^\epsilon}{\sqrt{D}}\right) \leq \frac{M}{D}$ for almost all $\hat{H} \in \mathcal{G}_{\text{inv}}$" of Theorem 1 in the main text, where we introduced $\mathcal{G}_{\text{inv}}$ as an invariant random matrix ensemble [1]. However, our proof shows that Theorem 1 applies to any random matrix ensemble whose probability measure is invariant under any unitary transformation $\hat{H} \mapsto \hat{V}\hat{H}\hat{V}^\dagger$ with $\hat{V}$ being a unitary operator.

---

[1] The random matrix ensemble $\mathcal{G}_{\text{inv}}$ is called an invariant random matrix ensemble if the probability measure of $\mathcal{G}_{\text{inv}}$ is proportional to $\exp\left(-\operatorname{tr}V(\hat{H})\right)$ ($\hat{H} \in \mathcal{G}_{\text{inv}}$) for a real function $V$.

## IV. PROOF OF THEOREM 2 IN THE MAIN TEXT

### A. Definition of the $m$-body operator space
### and proof of Theorem 2 in the main text for spin systems

#### 1. Definition of the m-body operator space

For $N$-site spin-$S$ systems, the total Hilbert space is given by $\mathcal{H}_N = (\mathcal{H}_{\mathrm{loc}})^{\otimes N}$ with $\mathcal{H}_{\mathrm{loc}}$ being the local Hilbert space of each spin. We denote $d_{\mathrm{loc}} := \dim \mathcal{H}_{\mathrm{loc}} = 2S - 1$ and $D_N := \dim \mathcal{H}_N = (d_{\mathrm{loc}})^N$.

We define the *exactly $m$-body* operator space $\mathcal{A}_N^{(m)}$ to be the space of operators that can be expressed as a linear combination of operators acting nontrivially on *exactly $m$* spins, i.e,

$$\mathcal{A}_N^{(m)} := \mathrm{span}_{\mathbb{R}} \left\{ \hat{\sigma}_{x_1}^{(p_1)} \cdots \hat{\sigma}_{x_m}^{(p_m)} \mid {}^{1 \leq x_1 < \cdots < x_m \leq N}_{1 \leq p_j < d_{\mathrm{loc}}} \right\}, \tag{S54}$$

where $\{\hat{\sigma}^{(p)}\}_{p=0}^{d_{\mathrm{loc}}-1}$ with $d_{\mathrm{loc}} := 2S + 1$ is an orthonormal basis of $\mathcal{L}(\mathcal{H}_N)$ with $\hat{\sigma}^{(0)} \propto \hat{I}$, and $\hat{\sigma}_x^{(p)}$ is the operator $\hat{\sigma}^{(p)}$ acting on the site $x$.

Then, we define the $m$-body operator space $\mathcal{A}^{[0,m]}$ and $\mathcal{A}_N^{[m_-,m_+]}$ by

$$\mathcal{A}_N^{[0,m]} := \bigoplus_{\tilde{m}=0}^m \mathcal{A}^{(\tilde{m})}, \quad \mathcal{A}_N^{[m_-,m_+]} := \bigoplus_{\tilde{m}=m_-}^{m_+} \mathcal{A}^{(\tilde{m})}. \tag{S55}$$

With these definitions, it is clear that $\mathcal{A}_N^{[m_-,m_+]}$ is the orthogonal complement of $\mathcal{A}_N^{[0,m_--1]}$ with respect to $\mathcal{A}_N^{[0,m_+]}$. The dimension of $\mathcal{A}_N^{(m)}$ is given by

$$\dim \mathcal{A}_N^{(m)} := \binom{N}{m} (d_{\mathrm{loc}}^2 - 1)^m = D_N^2 P_m, \tag{S56}$$

where

$$P_m := \binom{N}{m} \left( 1 - \frac{1}{d_{\mathrm{loc}}^2} \right)^m \left( \frac{1}{d_{\mathrm{loc}}^2} \right)^{N-m} \tag{S57}$$

is the probability mass function for the binomial distribution $\mathcal{B}(N, p)$ with $p := 1 - (d_{\mathrm{loc}})^{-2}$.

### 2. Proof of the first part (3) of Theorem 2 in the main text for spin systems

As a property of the binomial distribution, we have $P_{m-1} < P_m$ for $m \leq (N+1)p$. Therefore, for $0 < \alpha \leq p$, we have

$$
\begin{aligned}
\frac{\dim \mathcal{A}_N^{[0,\alpha N]}}{D_N} &= \sum_{m=0}^{\alpha N} \frac{\dim \mathcal{A}_N^{(m)}}{D_N} \\
&\leq (\alpha N + 1) D_N P_{\alpha N} \\
&= \alpha N D_N \exp\left[ N\left( H(\alpha) + \alpha \log(d_{\mathrm{loc}}^2 - 1) - 2\log d_{\mathrm{loc}} \right) + \mathcal{O}(\log N) \right] \\
&= \exp\left[ N\left( H(\alpha) + \alpha \log(d_{\mathrm{loc}}^2 - 1) - \log d_{\mathrm{loc}} \right) + \mathcal{O}(\log N) \right] \\
&= \exp\left[ N G_{d_{\mathrm{loc}}}^{(\mathrm{L})}(\alpha) + \mathcal{O}(\log N) \right],
\end{aligned}
\tag{S58}
$$

where we employed Stirling's formula in deriving the first equality, $G_{d_{\mathrm{loc}}}^{(\mathrm{L})}(\alpha) := H(\alpha) + \alpha \log(d_{\mathrm{loc}}^2 - 1) - \log d_{\mathrm{loc}}$, and $H(\alpha) := -\alpha \log \alpha - (1-\alpha)\log(1-\alpha)$ is the binary entropy. Here, we have

$$
G_{d_{\mathrm{loc}}}^{(\mathrm{L})}\left(\frac{1}{2}\right) = \log 2 + \frac{1}{2}\log(d_{\mathrm{loc}}^2 - 1) - \log d_{\mathrm{loc}} = \frac{1}{2}\log 4\left(1 - \frac{1}{d_{\mathrm{loc}}^2}\right),
\tag{S59}
$$

which is non-negative for $d_{\mathrm{loc}} \geq 2$. We also have $G_{d_{\mathrm{loc}}}^{(\mathrm{L})}(0) = -\log d_{\mathrm{loc}} < 0$. Therefore, the root $\alpha_*^{(\mathrm{L})}$ of $G_{d_{\mathrm{loc}}}^{(\mathrm{L})}$ lies in the range $(0, 1/2)$. Moreover, we have $G_{d_{\mathrm{loc}}}^{(\mathrm{L})}(\alpha) < 0$ for $\alpha < \alpha_*^{(\mathrm{L})}$. Thus, the upper bound in Eq. (S58) vanishes in the limit $N \to \infty$ when $\alpha < \alpha_*^{(\mathrm{L})}$. Then the upper bound of Thereom 1 in the main text implies that the ETH typically holds for all operators in $\mathcal{A}_N^{[0,\alpha N]}$ when $\alpha < \alpha_*^{(\mathrm{L})}$. Therefore, we obtain the lower bound $\alpha_*^{(\mathrm{L})} \leq m_*/N$ for $m_*$, and the first part (3) of Theorem 2 in the main text for spin systems is thus proved. Here, $\alpha_*^{(\mathrm{L})}$ is a monotonically increasing function of $d_{\mathrm{loc}}$, and we have $\alpha_*^{(\mathrm{L})} = 0.1892\cdots$ for $d_{\mathrm{loc}}$ and $\alpha_*^{(\mathrm{L})} \to 1/2$ as $d_{\mathrm{loc}} \to \infty$.

### 3. Proof of the second part (4) of Theorem 2 in the main text for spin systems

The binomial distribution $\mathcal{B}(N, p)$ with $p = 1 - (d_{\mathrm{loc}})^{-2}$ converges to the Gaussian distribution $\mathcal{N}(Np, Np(1-p))$ for a sufficiently large $N$. Therefore, if we set $m_\pm = Np \pm c_\pm \sqrt{N}$ with positive constants $c_\pm$, we have

$$
\frac{\dim \mathcal{A}_N^{[m_-, m_+]}}{D_N^2} \simeq \frac{1}{\sqrt{2\pi}} \int_{-c_-/\sqrt{p(1-p)}}^{c_+/\sqrt{p(1-p)}} e^{-\frac{x^2}{2}} \, \mathrm{d}x \geq \frac{1}{\sqrt{2\pi}} \int_{-c_-}^{c_+} e^{-\frac{x^2}{2}} \, \mathrm{d}x,
\tag{S60}
$$

for sufficiently large $N$. Here, we have used $p(1-p) < 1$. Then, the lower bound of Theorem 1 in the main text gives the upper bound $m_*/N \leq \alpha_*^{(\mathrm{U})}$ for $m_*$ and proves the second part (4) of Theorem 2 in the main text with $\alpha_*^{(\mathrm{U})} = p = 1 - (d_{\mathrm{loc}})^{-2}$.

## B. Definition of the $m$-body operator space
## and proof of Theorem 2 in the main text for Bose systems

For $N$-particle Bose systems on a lattice with $V$ sites, the total Hilbert space is given by

$$\mathcal{H}_{N,V} := \mathrm{span}_{\mathbb{C}}\left\{\hat{b}_{x_1}^{\dagger}\cdots\hat{b}_{x_N}^{\dagger}|0\rangle \mid 1 \le x_1 \le \cdots \le x_N \le V\right\}. \tag{S61}$$

Its dimension is given by

$$D_{N,V} := \dim\mathcal{H}_{N,V} = \binom{N+V-1}{N}. \tag{S62}$$

We define $\mathcal{A}_{N,V}^{[0,m]}$ to be the space of operators that can be expressed as a linear combination of products of $m$ annihilation operators $\hat{b}$ and $m$ creation operators $\hat{b}^{\dagger}$, i.e.,

$$\mathcal{A}_{N,V}^{[0,m]} := \left\{\hat{A} + \hat{A}^{\dagger},\ i(\hat{A} - \hat{A}^{\dagger}) \mid \hat{A} \in \tilde{\mathcal{A}}_{N,V}^{[0,m]}\right\}, \tag{S63}$$

$$\text{where}\quad \tilde{\mathcal{A}}_{N,V}^{[0,m]} := \mathrm{span}_{\mathbb{C}}\left\{\hat{P}_N(\hat{b}_{x_1}^{\dagger}\cdots\hat{b}_{x_m}^{\dagger}\hat{b}_{y_1}\cdots\hat{b}_{y_m})\hat{P}_N \mid 1 \le x_j \le V,\ 1 \le y_j \le V\right\}. \tag{S64}$$

Here, $\hat{P}_N$ is the projection operator onto $\mathcal{H}_{N,V}$, which is introduced to explicitly indicate that we are working within the sector $\mathcal{H}_{N,V}$ with a definite particle number. The space $\mathcal{A}_{N,V}^{[0,m]}$ contains the space $\mathcal{A}_{N,V}^{[0,m-1]}$, i.e., $\mathcal{A}_{N,V}^{[0,m-1]} \subset \mathcal{A}_{N,V}^{[0,m]}$. This is because the particle number operator $\hat{N} := \sum_{x=1}^{V}\hat{b}_x^{\dagger}\hat{b}_x$ is essentially equal to the identity operator due to the particle-number conservation. Indeed, for an arbitrary basis operator $\hat{P}_N(\hat{b}_{x_1}^{\dagger}\cdots\hat{b}_{x_{m-1}}^{\dagger}\hat{b}_{y_1}\cdots\hat{b}_{y_{m-1}})\hat{P}_N$ of $\tilde{\mathcal{A}}_{N,V}^{[0,m-1]}$, we have

$$\hat{P}_N(\hat{b}_{x_1}^{\dagger}\cdots\hat{b}_{x_{m-1}}^{\dagger}\hat{b}_{y_1}\cdots\hat{b}_{y_{m-1}})\hat{P}_N = \frac{1}{N-m+1}\hat{P}_N(\hat{b}_{x_1}^{\dagger}\cdots\hat{b}_{x_{m-1}}^{\dagger}\hat{N}\hat{b}_{y_1}\cdots\hat{b}_{y_{m-1}})\hat{P}_N$$

$$= \frac{1}{N-m+1}\sum_{x_m=1}^{V}\hat{P}_N(\hat{b}_{x_1}^{\dagger}\cdots\hat{b}_{x_{m-1}}^{\dagger}\hat{b}_{x_m}^{\dagger}\hat{b}_{x_m}\hat{b}_{y_1}\cdots\hat{b}_{y_{m-1}})\hat{P}_N, \tag{S65}$$

where we used $m \le N$ in deriving the first equality. The last equation in Eq. (S65) shows that $\hat{P}_N(\hat{b}_{x_1}^{\dagger}\cdots\hat{b}_{x_{m-1}}^{\dagger}\hat{b}_{y_1}\cdots\hat{b}_{y_{m-1}})\hat{P}_N \in \tilde{\mathcal{A}}_{N,V}^{[0,m]}$, which proves $\mathcal{A}_{N,V}^{[0,m-1]} \subset \mathcal{A}_{N,V}^{[0,m]}$. This fact also justifies the definition in Eq. (S64) as the $m$-body operator space $\mathcal{A}_{N,V}^{[0,m]}$ rather than the *exactly* $m$-body operator space $\mathcal{A}_{N,V}^{(m)}$.

Since $[\hat{b}_x, \hat{b}_y] = 0$, we can assume $x_1 \le x_2 \le \cdots \le x_m$ and $y_1 \le y_2 \le \cdots \le y_m$ in Eq. (S64). Therefore, the dimension of the $m$-body operator space $\mathcal{A}_{N,V}^{[0,m]}$ is given by

$$\dim\mathcal{A}_{N,V}^{[0,m]} = \binom{m+V-1}{m}^2. \tag{S66}$$

We introduce the particle density $\rho := N/V$, and $\alpha := m/N$. Then, Stirling's formula gives

$$\dim\mathcal{A}_{N,V}^{[0,m]} = \exp\left[2V(1+\alpha\rho)H\left(\frac{1}{1+\alpha\rho}\right) - \frac{1}{2}\log V + \mathcal{O}(1)\right]. \tag{S67}$$

Here, the function $xH(1/x)$ is a monotonically increasing function of $x$.

### 1. Proof of the first part (3) of Theorem 2 in the main text for Bose systems

From Eq. (S67), the upper bound of Theorem 1 in the main text is calculated to be

$$\frac{\dim \mathcal{A}_{N,V}^{[0,m]}}{D_{N,V}} = \exp\left[V G_\rho^{(\mathrm{L})}(\alpha) + \mathcal{O}(\log V)\right], \tag{S68}$$

where

$$G_\rho^{(\mathrm{L})}(\alpha) := 2(1+\alpha\rho)H\left(\frac{1}{1+\alpha\rho}\right) - (1+\rho)H\left(\frac{1}{1+\rho}\right). \tag{S69}$$

Here, $G_\rho^{(\mathrm{L})}$ is a monotonically increasing function of $\alpha$ and $G_\rho^{(\mathrm{L})}(0) < 0 < G_\rho^{(\mathrm{L})}(1/2)$ for $\rho > 0$. Indeed, we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\left[xH\left(\frac{1}{x}\right)\right] &= \frac{\mathrm{d}}{\mathrm{d}x}\left[\log x - (x-1)\log\left(1-\frac{1}{x}\right)\right] \\
&= \frac{1}{x} - \log\left(1-\frac{1}{x}\right) - \frac{1}{x} \\
&= -\log\left(1-\frac{1}{x}\right) \\
&> 0,
\end{aligned} \tag{S70}$$

and

$$\begin{aligned}
G_\rho^{(\mathrm{L})}\left(\frac{1}{2}\right) &= 2\log\left(1+\frac{\rho}{2}\right) - \rho\log\left(\frac{\rho}{2+\rho}\right) - \log(1+\rho) + \rho\log\frac{\rho}{1+\rho} \\
&= \log\frac{\left(1+\frac{\rho}{2}\right)^2}{1+\rho} + \rho\log\frac{2+\rho}{1+\rho} \\
&> 0.
\end{aligned} \tag{S71}$$

Hence, there is a root $\alpha_*^{(\mathrm{L})}$ of $G_\rho^{(\mathrm{L})}$ in the range $(0, 1/2)$, and we have $G_\rho^{(\mathrm{L})}(\alpha) < 0$ for $\alpha < \alpha_*^{(\mathrm{L})}$. Therefore, we obtain the lower bound $\alpha_*^{(\mathrm{L})} \leq m_*/N$ for $m_*$ and conclude the first part (3) of Theorem 2 in the main text for Bose systems.

### 2. Proof of the second part (4) of Theorem 2 in the main text for Bose systems

As mentioned in the main text, we define $\mathcal{A}_N^{[m_-,m_+]}$ to be the orthogonal complement of $\mathcal{A}_N^{[0,m_--1]}$ with respect to $\mathcal{A}_N^{[0,m_+]}$ so that we have $\mathcal{A}_N^{[0,m_+]} = \mathcal{A}_N^{[m_-,m_+]} \oplus \mathcal{A}_N^{[0,m_--1]}$ [2]. The dimension of $\mathcal{A}_N^{[m_-,m_+]}$ is given by

$$\dim \mathcal{A}_{N,V}^{[m_-,m_+]} = \dim \mathcal{A}_{N,V}^{[0,m_+]} - \dim \mathcal{A}_{N,V}^{[0,m_--1]}, \tag{S72}$$

---

[2] This definition of $\mathcal{A}_N^{[m_-,m_+]}$ depends on the choice of the inner product. However, $\dim \mathcal{A}_N^{[m_-,m_+]}$ is independent of its choice. Therefore, for our purpose, the existence of the space $\mathcal{A}_N^{[m_-,m_+]}$ is sufficient, and we do not need to worry about it.

where $\alpha_+ := m_+/N$.

The lower bound of Theorem 1 in the main text applied to $\mathcal{A}_{N,V}^{[m_-,m_+]}$ is calculated to be

$$\frac{\dim \mathcal{A}_{N,V}^{[m_-,m_+]}}{(D_{N,V})^2} = \frac{\dim \mathcal{A}_{N,V}^{[0,m_+]}}{\dim \mathcal{A}_{N,V}^{[0,N]}} \left( 1 - \frac{\dim \mathcal{A}_{N,V}^{[0,m_--1]}}{\dim \mathcal{A}_{N,V}^{[0,m_+]}} \right), \tag{S73}$$

where we used $\dim \mathcal{A}_{N,V}^{[0,N]} = (D_{N,V})^2$. Since the function $xH(1/x)$ is a monotinically increasing function of $x$, the first factor $(\dim \mathcal{A}_{N,V}^{[0,m_+]} / \dim \mathcal{A}_{N,V}^{[0,N]})$ in Eq. (S73) vanishes in the limit $N \to \infty$ unless $\alpha_+ = 1 - o(1/V)$. The second factor $1 - \dim \mathcal{A}_{N,V}^{[0,m_--1]} / \dim \mathcal{A}_{N,V}^{[0,m_+]}$ remains finite when $m_+ - m_- \geq c$ for a suitably chosen constant $c$. In particular, choosing $m_+ = N$ and $m_- = m_+ - c_-\sqrt{N}$ for an positive constant $c_-$ is sufficient to ensure that the left-hand side of Eq. (S73) remains finite. This fact proves the second part (4) of Theorem 2 in the main text with $\alpha_*^{(\mathrm{U})} = 1$ for Bose systems.

## C. Definition of the $m$-body operator space
## and proof of Theorem 2 in the main text for Fermi systems

For $N$-particle Fermi systems on a lattice with $V$ sites, the total Hilbert space is given by

$$\mathcal{H}_{N,V} := \mathrm{span}_\mathbb{C}\left\{ \hat{f}_{x_1}^\dagger \cdots \hat{f}_{x_N}^\dagger \ket{0} \mid 1 \leq x_1 < \cdots < x_N \leq V \right\}, \tag{S74}$$

where $\hat{f}$ is the annihilation operator. The dimension of $\mathcal{H}_{N,V}$ is given by

$$D_{N,V} := \dim \mathcal{H}_{N,V} = \binom{V}{N}. \tag{S75}$$

We define $\mathcal{A}_{N,V}^{[0,m]}$ to be the space of operators that can be expressed as a linear combination of products of $m$ annihilation operators $\hat{f}$ and $m$ creation operators $\hat{f}^\dagger$, i.e.,

$$\mathcal{A}_{N,V}^{[0,m]} := \left\{ \hat{A} + \hat{A}^\dagger, \ i(\hat{A} - \hat{A}^\dagger) \mid \hat{A} \in \tilde{\mathcal{A}}_{N,V}^{[0,m]} \right\} \tag{S76}$$

$$\text{where} \quad \tilde{\mathcal{A}}_{N,V}^{[0,m]} := \mathrm{span}_\mathbb{C}\left\{ \hat{P}_N (\hat{f}_{x_1}^\dagger \cdots \hat{f}_{x_m}^\dagger \hat{f}_{y_1} \cdots \hat{f}_{y_m}) \hat{P}_N \mid {}^{1 \leq x_1 < \cdots < x_m \leq V,}_{1 \leq y_1 < \cdots < y_m \leq V} \right\}. \tag{S77}$$

Here, $\hat{P}_N$ is the projection operator onto $\mathcal{H}_{N,V}$. For the same reason as for Bose systems, the space $\mathcal{A}_{N,V}^{[0,m]}$ includes the space $\mathcal{A}_{N,V}^{[0,m-1]}$, i.e., $\mathcal{A}_{N,V}^{[0,m-1]} \subset \mathcal{A}_{N,V}^{[0,m]}$. In addition, we have $\mathcal{A}_{N,V}^{[0,m]} \subseteq \mathcal{L}(\mathcal{H}_{N,V})$, and the dimension of $\mathcal{L}(\mathcal{H}_{N,V})$ is $D_{N,V}^2$.

By considering the particle-hole transformation $\hat{C}$ defined by $\hat{C}\hat{f}\hat{C}^\dagger = \hat{f}^\dagger$ (equivalently $\hat{C}\hat{f}^\dagger\hat{C}^\dagger = \hat{f}$) and $\hat{C}\ket{0} = \hat{f}_1^\dagger \cdots \hat{f}_V^\dagger \ket{0}$, we have

$$\hat{C}\tilde{\mathcal{A}}_{N,V}^{[0,m]}\hat{C}^\dagger = \mathrm{span}_\mathbb{C}\left\{ \hat{C}\hat{P}_N\hat{C}^\dagger \hat{C}(\hat{f}_{x_1}^\dagger \cdots \hat{f}_{x_m}^\dagger \hat{f}_{y_1} \cdots \hat{f}_{y_m})\hat{C}^\dagger \hat{C}\hat{P}_N\hat{C}^\dagger \mid {}^{1 \leq x_1 < \cdots < x_m \leq V,}_{1 \leq y_1 < \cdots < y_m \leq V} \right\}$$

$$= \mathrm{span}_\mathbb{C}\left\{ \hat{P}_{V-N}(\hat{f}_{x_1} \cdots \hat{f}_{x_m}\hat{f}_{y_1}^\dagger \cdots \hat{f}_{y_m}^\dagger)\hat{P}_{V-N} \mid {}^{1 \leq x_1 < \cdots < x_m \leq V,}_{1 \leq y_1 < \cdots < y_m \leq V} \right\}. \tag{S78}$$

Here, for any $x_1, \cdots, x_m$ and $y_1, \cdots, y_m$, we have

$$\hat{f}_{x_1} \cdots \hat{f}_{x_m} \hat{f}_{y_1}^\dagger \cdots \hat{f}_{y_m}^\dagger = (-1)^{m^2} \hat{f}_{y_1}^\dagger \cdots \hat{f}_{y_m}^\dagger \hat{f}_{x_1} \cdots \hat{f}_{x_m} + \Big( (m-1)\text{-body terms} \Big). \tag{S79}$$

Therefore, Eq. (S65) implies $\hat{C} \mathcal{A}_{N,V}^{[0,m]} \hat{C}^\dagger \subset \mathcal{A}_{V-N,V}^{[0,m]}$ for $m \leq V - N$, and we obtain $\dim \mathcal{A}_{N,V}^{[0,m]} \leq \dim \mathcal{A}_{V-N,V}^{[0,m]}$ ($m \leq V - N$). By exchanging the roles of $\mathcal{A}_{N,V}^{[0,m]}$ and $\mathcal{A}_{V-N,V}^{[0,m]}$ in the above discussion, we obtain another inequality $\dim \mathcal{A}_{V-N,V}^{[0,m]} \leq \dim \mathcal{A}_{N,V}^{[0,m]}$ ($m \leq N$). Combining these results, we obtain

$$\dim \mathcal{A}_{N,V}^{[0,m]} = \dim \mathcal{A}_{V-N,V}^{[0,m]}, \quad (m \leq \min\{N, V-N\}). \tag{S80}$$

The equation (S80) together with the inclusion relation $\mathcal{A}_{N,V}^{[0,m-1]} \subset \mathcal{A}_{N,V}^{[0,m]}$ implies

$$\dim \mathcal{A}_{N,V}^{[0,m]} = D_{N,V}^2, \quad (m \geq \min\{N, V-N\}), \tag{S81}$$

where we used $\dim \mathcal{A}_{N,V}^{[0,N]} = D_{N,V}^2$ and $\dim \mathcal{A}_{V-N,V}^{[0,V-N]} = D_{V-N,V}^2 = D_{N,V}^2$.

For $m < \min\{N, V-N\}$, it suffices to compute $\dim \mathcal{A}_{N,V}^{[0,m]}$ for $N \leq V/2$ because of Eq. (S80). There are $\binom{V}{m}$ choices for both $\{x_1, \cdots, x_m\}$ and $\{y_1, \cdots, y_m\}$, and different choices give independent basis operators $\hat{P}_N(\hat{f}_{x_1}^\dagger \cdots \hat{f}_{x_m}^\dagger \hat{f}_{y_1} \cdots \hat{f}_{y_m})\hat{P}_N$ when $N \leq V/2$ [3]. Therefore, we obtain $\dim \mathcal{A}_{N,V}^{[0,m]} = \binom{V}{m}^2$ for $m \leq \min\{N, V-N\}$. Combining this result with Eq. (S81) and $m \leq N$, we finally obtain

$$\dim \mathcal{A}_{N,V}^{[0,m]} = \binom{V}{\min\{m, V-N\}}^2. \tag{S82}$$

We introduce the particle density $\rho := N/V$, and $\alpha := m/N$ and obtain

$$\dim \mathcal{A}_{N,V}^{[0,m]} = \begin{cases} \exp\left[2VH(\alpha\rho) - \frac{1}{2}\log V + \mathcal{O}(1)\right] & \left(\alpha < \min\left\{1, \frac{1-\rho}{\rho}\right\}\right); \\ D_{N,V}^2 & \left(\alpha \geq \min\left\{1, \frac{1-\rho}{\rho}\right\}\right). \end{cases} \tag{S83}$$

### 1. Proof of the first part (3) of Theorem 2 in the main text for Fermi systems

From Eq. (S83), the upper bound of Theorem 1 in the main text is calculated to be

$$\frac{\dim \mathcal{A}_{N,V}^{[0,m]}}{D_{N,V}} = \exp\left[V G_\rho^{(\mathrm{L})}(\alpha) + \mathcal{O}(\log V)\right], \tag{S84}$$

---

[3] This can be confirmed by considering the kernel and image of $\hat{P}_N(\hat{f}_{x_1}^\dagger \cdots \hat{f}_{x_m}^\dagger \hat{f}_{y_1} \cdots \hat{f}_{y_m})\hat{P}_N$.

where

$$G_\rho^{(\mathrm{L})}(\alpha) := \begin{cases} 2H(\alpha\rho) - H(\rho) & \left(\alpha < \min\left\{1, \frac{1-\rho}{\rho}\right\}\right); \\[2mm] H(\rho) & \left(\alpha \geq \min\left\{1, \frac{1-\rho}{\rho}\right\}\right). \end{cases} \tag{S85}$$

It is straightforward to confirm that $G_\rho^{(\mathrm{L})}(0) < 0 < G_\rho^{(\mathrm{L})}(1/2)$ for $\rho > 0$. Indeed, for $\rho \leq 1/2$, we have

$$\begin{aligned} G_\rho^{(\mathrm{L})}\left(\frac{1}{2}\right) &= -\rho \log \frac{\rho}{2} - (2-\rho)\log\left(1 - \frac{\rho}{2}\right) + \rho \log \rho + (1-\rho)\log(1-\rho) \\ &= \rho \log 2 - \log\left(1 - \frac{\rho}{2}\right) + (1-\rho)\log\frac{1-\rho}{1-\frac{\rho}{2}} \\ &> 0. \end{aligned} \tag{S86}$$

For $\rho > 1/2$, we have $G_\rho^{(\mathrm{L})}(1/2) = H(\rho) > 0$. Because $G_\rho^{(\mathrm{L})}$ is a monotonically increasing function of $\alpha$, there is a root $\alpha_*^{(\mathrm{L})}$ of $G_\rho^{(\mathrm{L})}$ in the range $(0, 1/2)$, and we have $G_\rho^{(\mathrm{L})}(\alpha) < 0$ for $\alpha < \alpha_*^{(\mathrm{L})}$. Therefore, we obtain the lower bound $\alpha_*^{(\mathrm{L})} \leq m_*/N$ for $m_*$, and the first part (3) of Theorem 2 in the main text for Fermi systems is thus proved.

## 2. Proof of the second part (4) of Theorem 2 in the main text for Fermi systems

As in the case of Bose systems discussed in Sect. IV B 2, we define $\mathcal{A}_N^{[m_-, m_+]}$ to be the orthogonal complement of $\mathcal{A}_N^{[0, m_- - 1]}$ with respect to $\mathcal{A}_N^{[0, m_+]}$ so that we have $\mathcal{A}_N^{[0, m_+]} = \mathcal{A}_N^{[m_-, m_+]} \oplus \mathcal{A}_N^{[0, m_- - 1]}$. The dimension of $\mathcal{A}_{N,V}^{[m_-, m_+]}$ $(m_- \leq m_+)$ is given by

$$\dim \mathcal{A}_{N,V}^{[m_-, m_+]} = \dim \mathcal{A}_{N,V}^{[0, m_+]} - \dim \mathcal{A}_{N,V}^{[0, m_- - 1]}, \tag{S87}$$

where $\alpha_+ := m_+/N$.

The lower bound of Theorem 1 in the main text applied to $\mathcal{A}_{N,V}^{[m_-, m_+]}$ is calculated to be

$$\frac{\dim \mathcal{A}_{N,V}^{[m_-, m_+]}}{(D_{N,V})^2} = \frac{\dim \mathcal{A}_{N,V}^{[0, m_+]}}{\dim \mathcal{A}_{N,V}^{[0, N]}}\left(1 - \frac{\dim \mathcal{A}_{N,V}^{[0, m_- - 1]}}{\dim \mathcal{A}_{N,V}^{[0, m_+]}}\right), \tag{S88}$$

where we used $\dim \mathcal{A}_{N,V}^{[0, N]} = (D_{N,V})^2$. Since the function $H(x)$ is a monotinically increasing function of $x$, the first factor $(\dim \mathcal{A}_{N,V}^{[0, m_+]}/\dim \mathcal{A}_{N,V}^{[0, N]})$ in Eq. (S88) vanishes in the limit $N \to \infty$ unless $\alpha_+ = \min\left\{1, \frac{1-\rho}{\rho}\right\} - o(1/V)$. The second factor $(1 - \dim \mathcal{A}_{N,V}^{[0, m_- - 1]}/\dim \mathcal{A}_{N,V}^{[0, m_+]})$ remains finite when $m_+ - m_- \geq c$ for a suitably chosen constant $c$. In particular, choosing $m_+ = \alpha_*^{(\mathrm{U})} N$ with $\alpha_*^{(\mathrm{U})} := \min\left\{1, \frac{1-\rho}{\rho}\right\}$ and $m_- = m_+ - c_-\sqrt{N}$ for a positive constant $c_-$ is sufficient to ensure that

the left-hand side of Eq. (S88) remains finite. This fact proves the second part (4) of Theorem 2 in the main text with $\alpha_*^{(\mathrm{U})} := \min\left\{1, \frac{1-\rho}{\rho}\right\}$ for Fermi systems.

[1] S. Goldstein, D. A. Huse, J. L. Lebowitz, and R. Tumulka, Physical Review Letters **115**, 100402 (2015).

[2] S. Goldstein, D. A. Huse, J. L. Lebowitz, and R. Tumulka, Annalen der Physik **529**, 1600301 (2017).

[3] T. Mori, T. N. Ikeda, E. Kaminishi, and M. Ueda, Journal of Physics B: Atomic, Molecular and Optical Physics **51**, 112001 (2018).

[4] M. Ledoux, *The concentration of measure phenomenon*, 89 (American Mathematical Society, 2001).

[5] V. D. Milman and G. Schechtman, *Asymptotic Theory of Finite Dimensional Normed Spaces: Isoperimetric Inequalities in Riemannian Manifolds*, Vol. 1200 (Springer, 2001).

[6] M. Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Modern Birkhäuser Classics (Springer Science & Business Media, Boston, MA, 2007).

[7] E. S. Meckes, *The Random Matrix Theory of the Classical Compact Groups* (Cambridge University Press, 2019)