Supplemental Material for "Hilbert subspace imprint: a new mechanism for non-thermalization"

Hui Yu,[1] Jiangping Hu,[1,2,3,*] and Shi-Xin Zhang[1,†]

[1] *Beijing National Laboratory for Condensed Matter Physics and Institute of Physics, Chinese Academy of Sciences, Beijing 100910, China*
[2] *Kavli Institute of Theoretical Sciences, University of Chinese Academy of Sciences, Beijing 100190, China*
[3] *New Cornerstone Science Laboratory, Beijing 100190, China*

Contents

I. More numerical results for $U(1)$-breaking Hamiltonian with NNN terms

A. The time evolution of charge probability distribution for the antiferromagnetic state at $\gamma = 0.9$

Fig. S1 displays the time evolution of the charge probability distribution P_Q for the antiferromagnetic (AF) state at $\gamma = 0.9$. Despite the distribution being confined to only three sectors $Q = 0$ and $Q = \pm 4$, the AF state exhibits clear thermalization behavior, a consequence of the exponentially large Hilbert space dimensions associated with these sectors, consistent with our findings described in the main text.

B. The time evolution of charge probability distribution for ferromagnetic and antiferromagnetic states at $\gamma = 0.3$

Fig. S2(a) demonstrates that the charge probability P_Q spreads across all accessible sectors for the ferromagnetic (F) state. indicating the onset of thermalization as symmetry-breaking strength $1 - \gamma$ increases. Meanwhile, Fig. S2(b) reveals enhanced thermalization in the AF case, evidenced by a more pronounced shift of spectral weight to the $Q = \pm 4$ and ± 8 sectors compared to Fig. S1.

* jphu@iphy.ac.cn
† shixinzhang@iphy.ac.cn

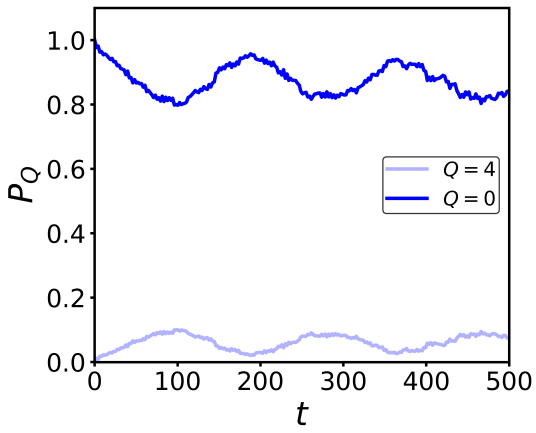

FIG. S1. The time evolution of charge distributions $P_Q(t)$ for AF initial state for $L = 12$. $Q = 12, 8, -8, -12$ sectors are omitted as they contain negligible population. Note that $P_{Q=-4}$ (not shown) equals $P_{Q=4}$ by symmetry.

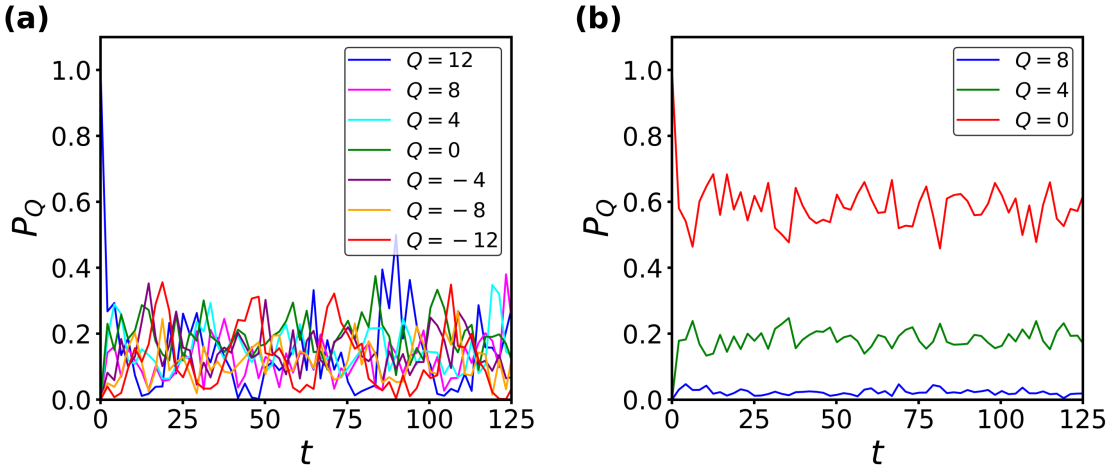

FIG. S2. The time evolution of charge distributions $P_Q(t)$ for (a) F and (b) AF initial state in a system of size $L = 12$ with symmetry-breaking strength $\gamma = 0.3$. For the AF case, the $Q = 12$ and $Q = -12$ sectors are omitted due to negligible population. By symmetry, the $P_{Q=-4}$ and $P_{Q=-8}$ (not shown) distributions are equal to $P_{Q=4}$ and $P_{Q=8}$ respectively.

C. Log-linear scaling plot for ferromagnetic and antiferromagnetic states

Here, we replot the data from Fig. 2(a) and (b) on a log-linear scale. For the ferromagnetic initial state (Fig. S3(a)), the data at $\gamma = 0.3$ exhibit a linear trend, whereas the data at $\gamma = 0.9$ show a clear deviation from linearity. In contrast, the antiferromagnetic state (Fig. S3(b)) exhibits a consistent linear relationship for both values of γ, indicating exponential scaling with system size.

D. Entanglement entropy scaling in singly spin-flipped ferromagnetic states

Fig. S4 shows the system-size scaling of late-time entanglement entropy for a F state with a single spin flip at the $L/2$-th site under different γ. At $\gamma = 0.9$ (weakly symmetry breaking), the EE remains approximately constant with increasing system size L, demonstrating the characteristic area-law scaling of non-thermal states. In contrast, for stronger symmetry breaking $\gamma = 0.3$, the EE exhibits volume law linear scaling with system size, a clear signature of thermalization.

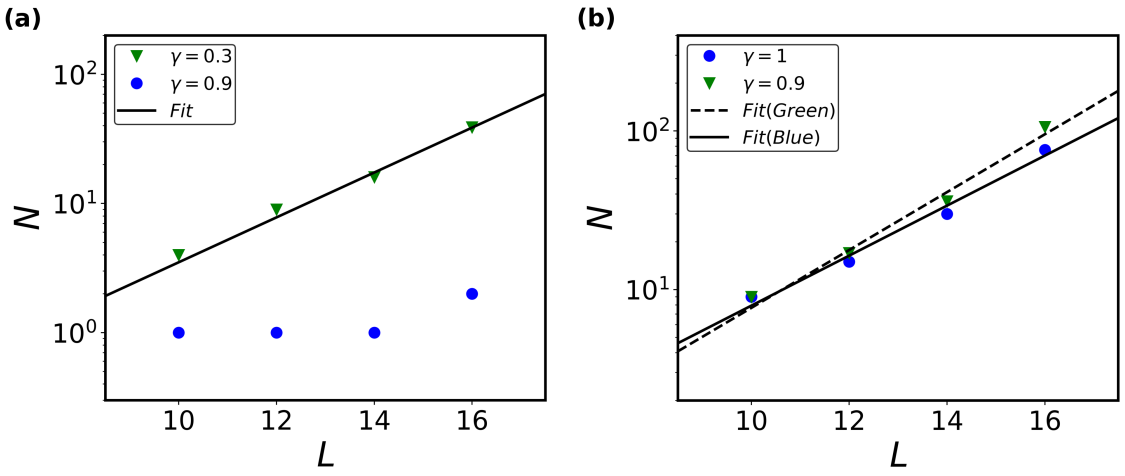

FIG. S3. Log-linear plots of N versus system size for (a) F and (b) AF initial states at different symmetry-breaking strengths γ of H_{U1}. The black solid and dashed lines are the corresponding linear fits.

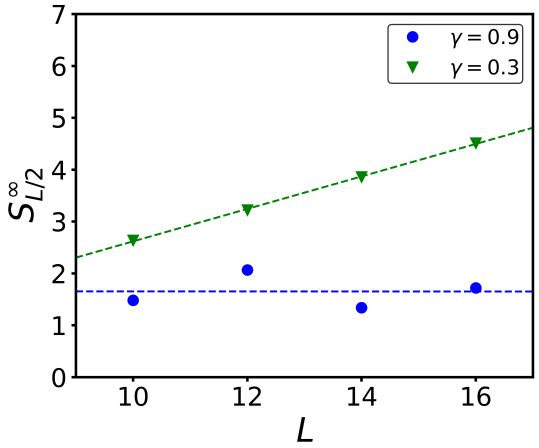

FIG. S4. The steady-state half-chain entanglement entropy $S^{\infty}_{L/2}$ is plotted as a function of system size L for symmetry breaking strengths $\gamma = 0.9$ and 0.3. Numerical simulation results are shown as data points, while dashed lines indicate the corresponding linear fits to the data.

E. Log-linear scaling plot for singly spin-flipped ferromagnetic state

The data from Fig. 3(a) are replotted on a log-linear scale. As shown in Fig. S5, the single spin-flipped ferromagnetic state exhibits a linear trend only at $\gamma = 0.3$, while the data for $\gamma = 0.9$ shows that the trend is much slower than an exponential scaling.

F. The time evolution of charge probability distribution for singly spin-flipped ferromagnetic state at $\gamma = 0.9$

Fig. S6 displays the time evolution of the charge probability distribution P_Q for a single spin-flipped F state in an $L = 16$ chain at $\gamma = 0.9$, with the flipped spin located at 8th site. We observe that the charge distribution $P_Q(t)$ in the $Q = 14$ also exhibits strong localization, closely resembling the behavior previously seen in the $Q = 10$ sector for the $L = 12$ from the main text. Specifically, the charge probability distribution of $Q = 14$ in $L = 16$ fluctuates around 0.7, indicating that the system remains predominantly confined within this charge sector. This behavior persists for extended durations up to $t > 60000$, confirming the HSI phenomenon is consistent across the system sizes we have studied.

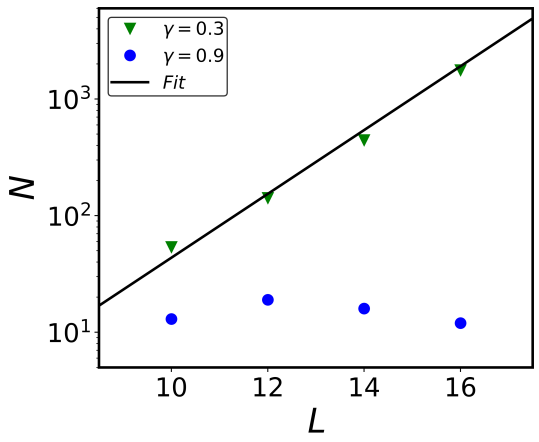

FIG. S5. Log-linear plots of N versus system size for ferromagnetic state with a single spin flip at different symmetry-breaking strengths γ of H_{U1}. The black solid is the linear fit.

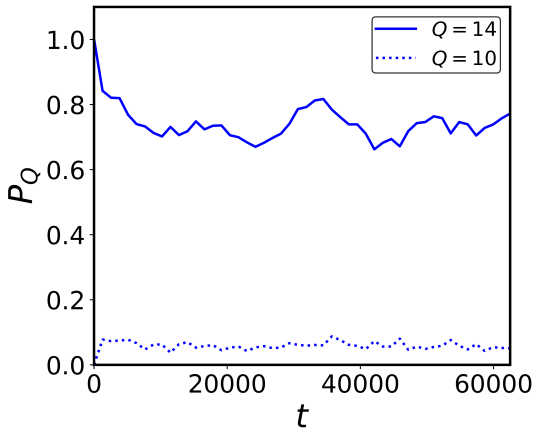

FIG. S6. Time evolution of the charge distribution $P_Q(t)$ for a single spin-flipped F state under the Hamiltonian H_{U1} with $L = 16$ at $\gamma = 0.9$. Sectors with negligible population ($Q = 6, 2, -2, -6, -10, -14$) have been omitted for clarity.

II. More numerical results for disordered $U(1)$-breaking Hamiltonian

We now analyze the thermalization properties of F and AF states under a disordered Hamiltonian with weakly broken $U(1)$-symmetry and periodic boundary conditions:

$$H = -\sum_{j=1}^{L}\left[\sigma_j^x\sigma_{j+1}^x + \gamma\sigma_j^y\sigma_{j+1}^y + \mu\sigma_j^z\sigma_{j+1}^z\right] - \sum_{j=1}^{L} h_j\sigma_j^z. \tag{S1}$$

where σ_j^α ($\alpha = x, y, z$) are Pauli matrices acting on site j, with $\mu = 0.8$ sets the nearest-neighbor coupling strengths. Disorder is introduced via random fields h_j along the z-axis, sampled uniformly from $[-W, W]$ with $W = 3.2$. The parameter γ controls the $U(1)$ symmetry-breaking strength. Notably, for $\gamma = 1, \mu = 0$, this Hamiltonian reduces to the one-dimensional Anderson model [1], which exhibits a well-known localization transition. To ensure that the system resides in the quantum chaotic regime, we compute the level spacing ratio r [2] within the half-filling sector, which requires the presence of the $U(1)$ symmetry. Therefore, this calculation is performed at $\gamma = 1$, not at 0.9 where $U(1)$ symmetry is explicitly broken. Also, symmetry breaking is expected to drive the system toward more thermalized behavior. The n-th level spacing ratio is defined as:

$$r_n = \frac{\min(\Delta_n, \Delta_{n+1})}{\max(\Delta_n, \Delta_{n+1})} \tag{S2}$$

where $\Delta_n = E_{n+1} - E_n$ and E_n is the n-th eigenenergy in ascending order. The level spacing ratio r is obtained by averaging r_n over different energy levels n and across an ensemble of disorder configurations. Our calculation shows $r \approx 0.51$ at $\mu = 0.8$ and $\gamma = 1$, which aligns with the theoretical value for the quantum chaotic regime, thereby confirming that the model is in the thermal phase for the studied parameters. All physical quantities are obtained by averaging over multiple disorder realizations.

A. Thermalization properties for the ferromagnetic state at different γ

Fig. S7(a) demonstrates that for $\gamma = 0.9$, the F state exhibits clear non-thermal behavior, as characterized by the nearly system-size-independent of N. The time evolution of the charge probability distribution in Fig. S7(b) further confirms the non-thermal nature, showing approximately 60% of the population remains confined to the Q=14 sector of dimension 1, with an additional 20% in the Q=10 sector. Upon increasing the symmetry-breaking strength $1-\gamma$ to 0.7, the F state begins to thermalize, as supported by the exponential growth of N with L and a significantly broader charge distribution shown in Fig. S7(c).

B. Thermalization properties for the antiferromagnetic state at different γ

In contrast to the F state, AF initial state shows thermalization under weak symmetry breaking, as confirmed by the exponential growth of N with L in Fig. S8(a). This growth becomes more pronounced with stronger symmetry breaking. Despite thermalization, Fig. S8(b) confirms the charge distribution remains dominated by the $Q = 0$ and $Q = \pm 4$ sectors, a consequence of their large Hilbert space dimensions. As the symmetry breaking becomes more pronounced, Fig. S8(c) displays a systematic charge redistribution of spectral weight toward the $Q = \pm 4$ and $Q = \pm 8$ sectors.

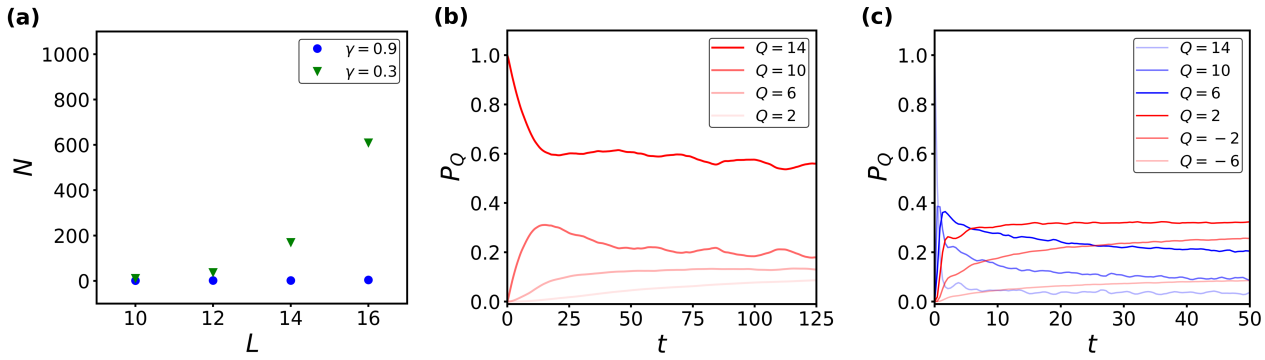

FIG. S7. Panel (a) displays the scaling of N with system size for F initial state at varying γ. Panels (b) and (c) show the time evolution of charge distributions $P_Q(t)$ initialized in F states for $L = 14$ at $\gamma = 0.9$ and $\gamma = 0.3$. Sectors $Q = 12, 8, -8, -12$ and $Q = 12, -12$ are omitted in panels (b) and (c) due to negligible weight. By symmetry, the $P_Q = -4$ and $P_Q = -8$ (not shown) distribution are equal to $P_Q = 4$ and $P_Q = 8$ respectively. All results are averaged over 400 independent realizations.

III. More numerical results for disordered Z_2-breaking Hamiltonian

A. Log-linear scaling plot for ferromagnetic and antiferromagnetic states

Here, we replot the data from Fig. 4(a) and (b) on a log-linear scale. For both the ferromagnetic and antiferromagnetic initial states, Fig. S9(a) and (b) exhibit a linear trend for each value of γ, indicating exponential scaling with system size.

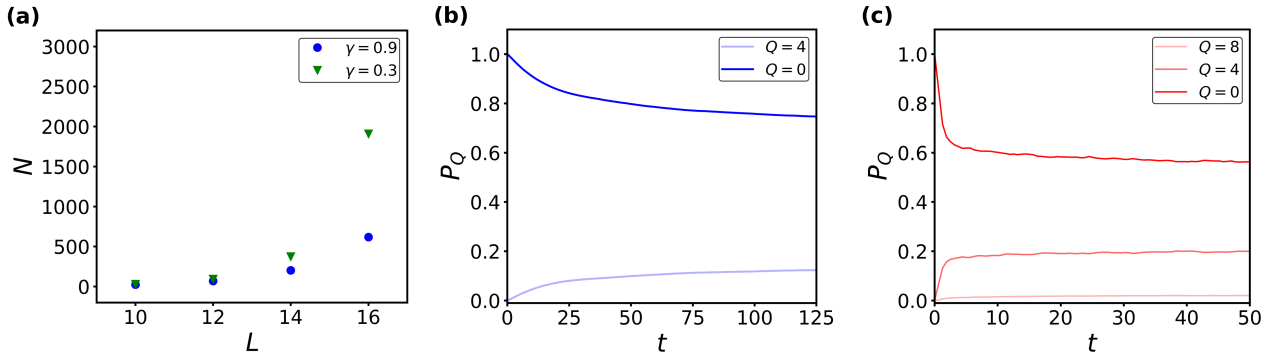

FIG. S8. Panel (a) displays the scaling of N with system size for AF initial state at varying γ. Panels (b) and (c) show the time evolution of charge distributions $P_Q(t)$ initialized in AF states for $L = 14$ at $\gamma = 0.9$ and $\gamma = 0.3$, respectively. Sectors $Q = -2, -6, -10, -14$ and $Q = -10, -14$ are omitted in panels (b) and (c) due to negligible weight. By symmetry, the $P_{Q=-4}$ and $P_{Q=-8}$ (not shown) distributions are equal to $P_{Q=4}$ and $P_{Q=8}$ respectively. All results are averaged over 400 independent realizations.

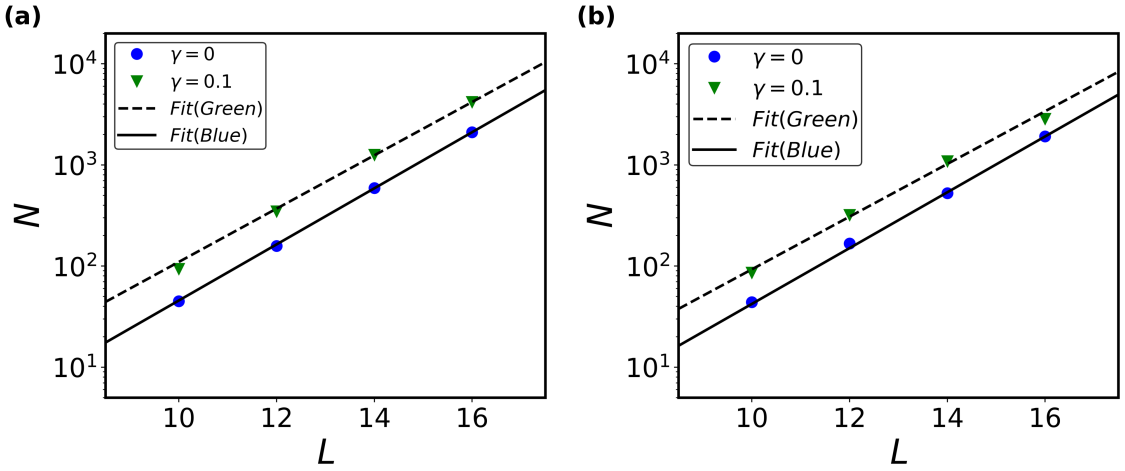

FIG. S9. Log-linear plots of N versus system size for (a) F and (b) AF initial states at different symmetry-breaking strengths γ of H_{Z2}. The black solid and dashed lines are the corresponding linear fits.

IV. HSI beyond the weak symmetry-breaking limit

Having demonstrated non-ergodic dynamics for a ferromagnetic state with one flipped spin in a near-symmetric case ($\gamma = 0.9$), we now show that a small symmetry-breaking parameter in the Hamiltonian is not a prerequisite for HSI. We employ the Hamiltonian with $\gamma = 0.6$ in this section and demonstrate that HSI can still emerge with initial state engineering. We initialize the circuit in a ferromagnetic state with a single spin flipped and define the target eigenstate subspace as the 15 eigenstates with the highest overlap to this initial state. As shown in Fig. S10, even for this moderate symmetry-breaking strength, the variational circuit successfully localizes the trained state within this target subspace, achieving an overlap around 0.8. This confirms that HSI can be generally realized with an initial state prepared by the shallow variational circuit, without requiring a small parameter or perturbation for the Hamiltonian.

V. Analytical derivation of the bounds on the expectation value $\langle \hat{O} \rangle$

Suppose an initial wavefunction $|\phi_0\rangle$ with significant overlap with an eigenstate in a small non-thermal Hilbert subspace, along with a small leakage to state outside this subspace that can thermalize. We aim to determine the

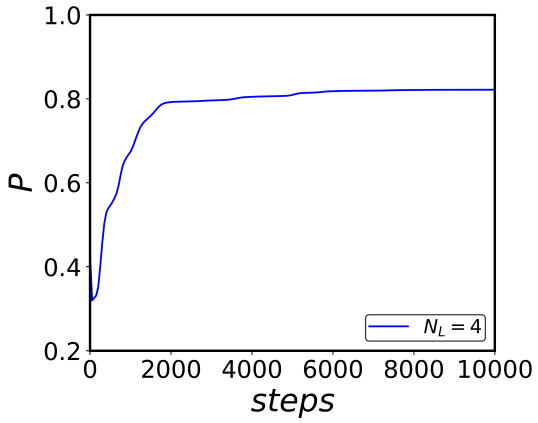

FIG. S10. Overlap probability P as a function of training steps for a variational quantum circuit. The circuit comprises 4-composite layers, with each composite layer structured identically to that depicted in Fig. **??**. This 12-qubit circuit operates on a ferromagnetic initial state containing a single spin flip.

bounds for the expectation value of an observable $\hat{O}$. The initial state can be expressed as:

$$|\psi_0\rangle = \sqrt{1-\Delta^2}|\phi\rangle + \Delta|\phi_\perp\rangle \tag{S3}$$

where $|\phi\rangle$ is the normalized state in the non-thermal subspace, $|\phi_\perp\rangle$ is the normalized state outside that space, which ultimately thermalize. Δ^2 quantifies the leakage probability. And we omit the irrelevant relative phase. The time evolution of $|\psi_0\rangle$ is given by:

$$|\psi(t)\rangle = \sqrt{1-\Delta^2}|\phi(t)\rangle + \Delta|\phi_\perp(t)\rangle \tag{S4}$$

with $|\phi(t)\rangle$ and $|\phi_\perp(t)\rangle$ still inside and outside the Hilbert subspace, respectively.

Next, we calculate the expectation value of an observable $\hat{O}$ with respect to the state $|\psi(t)\rangle$ as:

$$\langle\hat{O}\rangle = (1-\Delta^2)O_{non-th}(t) + \Delta^2 O_{th} + \Delta\sqrt{1-\Delta^2}\left[\langle\phi(t)|\hat{O}|\phi_\perp(t)\rangle + \langle\phi_\perp(t)|\hat{O}|\phi(t)\rangle\right] \tag{S5}$$

where $O_{non-th}(t) = \langle\phi(t)|\hat{O}|\phi(t)\rangle$ is the expectation value of $\hat{O}$ in the non-thermal subspace and $O_{th} = \langle\phi_\perp(t)|\hat{O}|\phi_\perp(t)\rangle$ is the thermal expectation value at late times according to ETH. To bound the cross terms, we apply the Cauchy-Schwarz inequality, which states that for any state $|\phi(t)\rangle$, $|\phi_\perp(t)\rangle$, and an operator $\hat{O}$, we have:

$$|\langle\phi(t)|\hat{O}|\phi_\perp(t)\rangle| \leq ||\hat{O}|| \cdot ||\phi(t)|| \cdot ||\phi_\perp(t)|| \tag{S6}$$

Here, $||\cdot||$ denotes the operator and vector norms. Since $|\phi(t)\rangle$ and $|\phi_\perp(t)\rangle$ are normalized ($||\phi(t)|| = ||\phi_\perp(t)|| = 1$), and $\hat{O}$ is bounded with eigenvalues λ_k satisfying $|\lambda_k| \leq 1$ (e.g., Pauli string operators), the operator norm $||\hat{O}||$ is bounded by the largest eigenvalue:

$$||\hat{O}|| \leq 1 \tag{S7}$$

Thus, Eq. (S6) simplifies to:

$$|\langle\phi(t)|\hat{O}|\phi_\perp(t)\rangle| \leq 1 \tag{S8}$$

Similarly, $|\langle\phi_\perp(t)|\hat{O}|\phi(t)\rangle| \leq 1$. By omiting higher orders of Δ for small leakage, we have

$$O_{non-th} - 2\Delta \leq \langle\hat{O}\rangle \leq O_{non-th} + 2\Delta, \tag{S9}$$

namely, non-thermal behavior can be determined via local observable measurements as long as Δ is below some threshold.

In fact, we can derive a significantly tighter bound by leveraging the ETH. Let us expand $|\phi(t)\rangle$ and $|\phi_\perp(t)\rangle$ in the energy eigenbasis:

$$|\phi(t)\rangle = \sum_n d_n e^{-iE_n t}|E_n\rangle, \quad |\phi_\perp(t)\rangle = \sum_\alpha c_\alpha e^{-iE_\alpha t}|E_\alpha\rangle \tag{S10}$$

where $|E_n\rangle$ are non-thermalizing eigenstates, $|E_\alpha\rangle$ are thermalizing eigenstates, and d_n, c_α are their respective coefficients. The cross term then becomes:

$$\langle\phi(t)|\hat{O}|\phi_\perp(t)\rangle = \sum_{n,\alpha} d_n^* c_\alpha e^{i(E_n - E_\alpha)t}\langle E_n|\hat{O}|E_\alpha\rangle \tag{S11}$$

The phase factors $e^{i(E_n-E_\alpha)t}$ oscillate rapidly at late times, causing dephasing and suppression of the sum. By the ETH, the magnitude of off-diagonal matrix elements $\langle E_n|\hat{O}|E_\alpha\rangle$ can be expressed as:

$$|\langle E_n|\hat{O}|E_\alpha\rangle| \sim e^{-S(\bar{E})/2} f(\bar{E},\omega) \tag{S12}$$

Here $\bar{E} = \frac{E_n+E_\alpha}{2}$ represents the average energy, and $\omega = E_n - E_\alpha$ is the energy difference. The term $S(\bar{E})$ corresponds to the thermodynamic entropy evaluated at energy $\bar{E}$, which is proportional to the system size L while $f(\bar{E},\omega)$ is a smooth function of $\bar{E}$ and ω. Consequently, Eq. (S8) is

$$|\langle\phi(t)|\hat{O}|\phi_\perp(t)\rangle| \leq \delta \tag{S13}$$

where δ is an exponentially small parameter with respect to the system size due to strong dephasing and exponential small terms from off-diagonal prediction of ETH. Consequently,

$$(1-\Delta^2)O_{non-th} + \Delta^2 O_{th} - 2\delta\Delta\sqrt{1-\Delta^2} \leq \langle\hat{O}\rangle \leq (1-\Delta^2)O_{non-th} + \Delta^2 O_{th} + 2\delta\Delta\sqrt{1-\Delta^2} \tag{S14}$$

As δ is much smaller than Δ for a large system, we omit $\delta\Delta$ term while keep the quadratic Δ term instead. We thus have:

$$\langle\hat{O}\rangle - O_{th} \sim (1-\Delta^2)(O_{non-th} - O_{th}) \tag{S15}$$

This relation demonstrates that if O_{th} deviates from O_{non-th}, we can unambiguously distinguish between thermal and non-thermal behaviors through the observable expectation value.

[1] P. W. Anderson, Absence of diffusion in certain random lattices, Physical review **109**, 1492 (1958).
[2] V. Oganesyan and D. A. Huse, Localization of interacting fermions at high temperature, Physical Review B—Condensed Matter and Materials Physics **75**, 155111 (2007).