
Supplementary information

Prethermalization by random multipolar driving on a 78-qubit processor

In the format provided by the
authors and unedited

Supplementary Material for: Prethermalization by Random Multipolar Driving on a 78-Qubit Processor

Zheng-He Liu,^{1,2,*} Yu Liu,^{1,2,*} Gui-Han Liang,^{1,2,*} Cheng-Lin Deng,^{1,2,*} Keyang Chen,^{2,3} Yun-Hao Shi,¹ Tian-Ming Li,^{1,2} Lv Zhang,^{1,2} Bing-Jie Chen,^{1,2} Cai-Ping Fang,^{1,2} Da'er Feng,^{1,2} Xu-Yang Gu,^{1,2} Yang He,^{1,2} Kaixuan Huang,⁴ Hao Li,⁴ Hao-Tian Liu,^{1,2} Li Li,^{1,2} Zheng-Yang Mei,^{1,2} Zhen-Yu Peng,^{1,2} Jia-Cheng Song,^{1,2} Ming-Chuan Wang,^{1,2} Shuai-Li Wang,^{1,2} Ziting Wang,⁴ Yongxi Xiao,^{1,2} Minke Xu,^{1,2} Yue-Shan Xu,⁴ Yu Yan,^{1,5} Yi-Han Yu,^{1,2} Wei-Ping Yuan,^{1,2} Jia-Chi Zhang,^{1,2} Jun-Jie Zhao,^{1,2} Kui Zhao,⁴ Si-Yun Zhou,^{1,2} Zheng-An Wang,⁴ Xiaohui Song,¹ Ye Tian,¹ Florian Mintert,^{6,7} Johannes Knolle,^{6,8,9} Roderich Moessner,¹⁰ Yu-Ran Zhang,¹¹ Pan Zhang,^{2,3} Zhongcheng Xiang,^{1,2,12,†} Dongning Zheng,^{1,2,12,13} Kai Xu,^{1,2,4,12,13,‡} Hongzheng Zhao,^{14,§} and Heng Fan^{1,2,4,12,13,¶}

¹Beijing National Laboratory for Condensed Matter Physics,

Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China

²School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

³CAS Key Laboratory for Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China

⁴Beijing Key Laboratory of Fault-Tolerant Quantum Computing,

Beijing Academy of Quantum Information Sciences, Beijing 100193, China

⁵School of Physics, Northwest University, Xian 710127, China

⁶Blackett Laboratory, Imperial College London, London SW7 2AZ, United Kingdom

⁷Helmholtz-Zentrum Dresden-Rossendorf, Bautzner Landstraße 400, 01328 Dresden, Germany

⁸Department of Physics TQM, Technische Universität München, James-Frank-Straße 1, D-85748 Garching, Germany

⁹Munich Center for Quantum Science and Technology (MCQST), 80799 Munich, Germany

¹⁰Max-Planck-Institut für Physik komplexer Systeme, Nöthnitzer Straße 38, 01187 Dresden, Germany

¹¹School of Physics and Optoelectronics, South China University of Technology, Guangzhou 510640, China

¹²Hefei National Laboratory, Hefei 230088, China

¹³Songshan Lake Materials Laboratory, Dongguan, Guangdong 523808, China

¹⁴State Key Laboratory of Artificial Microstructure and Mesoscopic Physics,
School of Physics, Peking University, Beijing 100871, China

CONTENTS

S1. Prethermalization and numerical simulation	2
A. Prethermalization in Floquet systems	2
B. Prethermalization by random multipolar driving	3
C. Numerical results	4
1. Trotter-Suzuki decomposition	4
2. Project entangled pair states	5
3. Grouped matrix product states	6
4. Pauli propagation	9
5. Neural quantum states	11
S2. Experimental setup	15
S3. Device fabrication and performance	16
A. Device fabrication	16
B. Device performance	17
S4. Device control and calibration	19
A. Pulse calibration	19
B. Timing by Floquet engineering	21

* These authors contributed equally to this work.

† zcxiaang@iphy.ac.cn

‡ kaixu@iphy.ac.cn

§ hzhao@pku.edu.cn

¶ hfan@iphy.ac.cn

S5. Impact of noise and benchmark of simulator fidelity	22
A. Effects of decoherence	22
B. Effects of experimental spatio-temporal perturbations	23
C. Quantum simulator performance benchmarking	26
1. Fidelity of quantum simulator	26
2. Conservation of particle number and error mitigation	27
S6. Fitting details and precision	29
S7. Supplementary data	32
A. Dynamics of imbalance in systems with different sizes	32
B. Dynamics of entropy in systems with different sizes	37
C. Area law to volume law crossover	38
D. Subsystems sizes for entropy scaling	40
E. Estimated error in local observables in GMPS simulation	40
F. Waveform inspection	41
References	42

S1. PRETHERMALIZATION AND NUMERICAL SIMULATION

A. Prethermalization in Floquet systems

Periodically driven (Floquet) quantum systems can engineer novel far-from-equilibrium phases of matter absent in static systems [1–21]. These systems can be analysed within the framework of Floquet theory. Consider the Hamiltonian $\hat{H}$ with temporal periodicity, $\hat{H}(t) = \hat{H}(t + T)$, the Floquet theorem states that the time evolution operator can be written as

$$\hat{U}(t) = \hat{P}(t)e^{-i\hat{H}_F t}, \quad (1)$$

where $\hat{H}_F$ is the time-independent Floquet Hamiltonian, and $\hat{P}(t)$ is the stroboscopic operator with period T . In the high-frequency regime, the dynamics can be approximated by the Hamiltonian $\hat{H}_F$, known as the effective Hamiltonian. In general, an exact closed-form expression for $\hat{H}_F$ is difficult to obtain, and a common scheme is to use the Floquet-Magnus expansion (FME), which is a perturbative expansion in the driving frequency

$$\hat{H}_F = \sum_{n=0}^{\infty} \frac{1}{\Omega^n} \hat{H}_F^{(n)}, \quad (2)$$

with the driving frequency $\Omega = 2\pi/T$. The lowest-order effective Hamiltonian is the time average of system Hamiltonian $\hat{H}$ over a period T

$$\hat{H}_F^{(0)} = \frac{1}{T} \int_0^T dt \hat{H}(t). \quad (3)$$

Unlike static systems that conserve energy, Floquet systems typically absorb energy from the drive [8, 22]. The nature of equilibrium states under such heating becomes clearer through the analysis of FME convergence. When the FME converges, the effective Hamiltonian $\hat{H}_F$ can be effectively represented by a summation of local operators. In this case, the system is expected to thermalize to a steady state that, at the local level, becomes effectively indistinguishable from the microcanonical ensemble of a static system described by $\hat{H}_F$. However, it is widely believed that the FME diverges in generic many-body interacting systems, indicating the absence of a quasi-local $\hat{H}_F$. For a non-integrable Floquet system, heating leads to the eventual thermalization where the system reaches the featureless infinite-temperature state.

Theoretical studies on the long-time behaviors of Floquet systems reveal that the FME may converge to the time-averaged Hamiltonian when driving periods are sufficiently short, such that the driving frequency Ω is much larger than the local energy scale of the many-body system. In this regime, Floquet prethermalization appears before the onset of noticeable heating, giving rise to an exponentially long-lived prethermal regime governed by the prethermal Hamiltonian

$$\hat{H}_{\text{pre}} = \sum_{n=0}^N \frac{1}{\Omega^n} \hat{H}_F^{(n)}, \quad (4)$$

obtained by truncating FME at the order N . In contrast to using many-body localization (MBL) to suppress heating [23], Floquet prethermalization is generally applicable to clean systems without the need of strong spatial disorder.

The exponential suppression of the heating rate can be justified by noticing the residual part of the Floquet Hamiltonian satisfying $\|\hat{H}_F - \hat{H}_{\text{pre}}\| < e^{-\mathcal{O}(1/\Omega)}$, with $\|\cdot\|$ denoting the norm of an operator. Higher-order processes lead to exponentially small corrections that become noticeable after an exponentially long time scale [24]. Alternatively, one can also use a simple linear response theory—the Fermi’s golden rule (FGR) [25, 26]—to understand this exponential suppression. We start with a generic Floquet system

$$\hat{H} = \hat{H}_0 + g(t)\hat{K}, \quad (5)$$

where $g(t)$ is taken the Fourier series form $g(t) = \sum_{m>0} g_m \sin(m\Omega t)$, $\hat{H}_0$ and $\hat{K}$ are time-independent operators. For Fourier mode m , the heating rate is given by

$$\Gamma_m(t) = \frac{\dot{E}_m(t)}{E_\infty - E(t)}, \quad (6)$$

where E_∞ denotes the energy at infinite temperature with $E_\infty = 0$ for our model. The transition rate of the mode $\dot{E}_m(t)$ can be obtained from the FGR

$$\dot{E}_m(t) = 2\pi g_m^2 \sum_{i,f} |\langle E_f^0 | \hat{K} | E_i^0 \rangle|^2 (E_f^0 - E_i^0) P_i^0(t) \cdot \delta(E_f^0 - E_i^0 \pm m\Omega), \quad (7)$$

with E_i^0, E_f^0 being eigenstates of $\hat{H}_0$, and $P_i^0(t) = \langle E_i^0 | \rho(t) | E_i^0 \rangle$. Combining all Fourier modes, the heating rate of the system is

$$\Gamma(t) = \sum_{m>0} \Gamma_m(t), \quad (8)$$

with the m -mode rate

$$\Gamma_m = g_m^2 A e^{-m\Omega/\epsilon}, \quad (9)$$

where A and ϵ are Ω -independent constants.

B. Prethermalization by random multipolar driving

Non-periodically driven systems have attracted notable research interest during recent years [27–35], which go beyond the constraint imposed by the strict temporal periodicity. We focus on a family of random but structured driving protocol, random multipolar driving (RMD), characterized by the multipolar correlation n . This protocol involves two elementary evolution operators $\hat{U}_+$ and $\hat{U}_-$, generated by two Hamiltonians $\hat{H}_\pm$. A random selection of $U_\pm$ leads to 0-RMD sequence and normally this quickly destabilizes the system and induces heating. The heating rate can be significantly suppressed by imposing a dipolar structure into the random driving sequence, such that the elementary building blocks now read $\hat{U}_+\hat{U}_-$ and $\hat{U}_-\hat{U}_+$. 1-RMD protocol then contains a random sequence of these two building blocks. Similarly, the n -th multipole can be recursively constructed by anti-aligning two $(n-1)$ -th order operators, and in the $n \rightarrow \infty$ limit it converges to the self-similar Thue-Morse driving [36]. A detailed discussion on the properties of RMD sequences is presented in [26] and here we elaborate on how to estimate the heating rate.

A purely random sequence exhibits a flat and continuous distribution of different frequencies in the Fourier space. Similarly, n -RMD sequence also exhibits a continuous spectrum. However, the overall envelope of the distribution is modified to a power law x^n [26], due to the presence of the multipolar correlation in time. Therefore, to estimate the heating rate one can first approximate the time-dependent RMD signal as

$$g(\tau)\hat{K} = \int dx g_x \sin(x\Omega\tau)\hat{K}, \quad (10)$$

where $g_x = x^n$ captures the suppressed spectrum distribution. According to the linear response theory, each frequency mode x contributes to the heating rate separately, following Supplementary Eq. (9) as in Floquet systems. Therefore, to obtain the entire heating rates from all frequencies, one can simply perform an integral

$$\Gamma = \int_0^\infty dx g_x^2 A e^{-x\Omega/\epsilon} \propto \Omega^{-2n-1} (2n)!, \quad (11)$$

where we use the Gamma function, $F(n) = \int_0^\infty dx x^{n-1} e^{-x} = (n-1)!$. Its inverse defines the heating time scale, $1/\Gamma \propto \Omega^{2n+1}$, which has been observed in our experiments. In fact, a generalized Floquet-Magnus expansion can be employed to investigate higher order processes. It goes beyond this simple linear response theory and still leads to the same prediction on the heating rates, see details in [36].

C. Numerical results

We adopt two advanced tensor network algorithms, Projected Entangled Pair States (PEPS) and grouped Matrix Product States (GMPS), to simulate the dynamics generated by RMD protocols. Tensor networks are powerful methods to simulate quantum many-body physics and quantum circuits, providing highly efficient classical computational frameworks. Notably, they have demonstrated the capacity to surpass Google's quantum supremacy claims [37, 38]. Although tensor network algorithms are limited to simulating dynamics within constrained Hilbert spaces, they allow us to extract critical features from the numerical results. In this section, we initially present the Trotter-Suzuki decomposition, an effective method for approximately factoring time-evolution operators into products of local terms, thus making them suitable for tensor network simulations. Subsequently, we introduce these two methods, respectively, and compare their performance with our experimental observations. These numerical simulations will support our experimental observation, including T^{2n+1} scaling in heating rates, oscillatory dynamics in entanglement entropy during the prethermal regime, as well as the rapid entanglement growth induced by heating.

1. Trotter-Suzuki decomposition

The time-dependent evolution operators are supported over the entire Hilbert space, applying directly to a quantum state is nearly infeasible. To address this challenge, we introduce the second-order Trotter-Suzuki decomposition [39]. This method approximates the short-time evolution operator $e^{-i\hat{H}\delta t}$ as a product of operators with finite support, thereby facilitating practical computation and implementation. Specifically, we split the Hamiltonian as

$$\hat{H} = \hat{H}^1 + \hat{H}^2 + \dots + \hat{H}^m, \quad (12)$$

where $\hat{H}^i$ is a sum of local operators $\hat{H}^i = \sum_{x,y} \hat{h}_{x,y}^i$. The second-order Trotter-Suzuki decomposition approximates the time-evolution operators as

$$e^{-i\hat{H}\delta t} = \prod_{k=m}^1 e^{-i\hat{H}^k\delta t/2} \prod_{k=1}^m e^{-i\hat{H}^k\delta t/2} + \mathcal{O}(\delta t^3). \quad (13)$$

The selection of $\hat{H}^k$ guarantees that each term $\hat{h}_{x,y}^k$ commutes pairwise. This commutativity implies that each operator on the right-hand side of Supplementary Eq. (13) can be further decomposed into a product of local operators.

$$e^{-i\sum_{x,y} \hat{h}_{x,y}^k \delta t/2} = \prod_{x,y} e^{-i\hat{h}_{x,y}^k \delta t/2}. \quad (14)$$

We express our two-dimensional Hamiltonian as a sum of five distinct components.

$$\hat{H}_\pm^1 = \sum_{x,y} (1 \pm \delta h) h_0 [1 - (-1)^y] \hat{a}_{x,y}^\dagger \hat{a}_{x,y}, \quad (15)$$

$$\hat{H}^2 = \sum_{x \text{ is odd}} \sum_y J \hat{a}_{x+1,y}^\dagger \hat{a}_{x,y} + \text{H.c.}, \quad (16)$$

$$\hat{H}^3 = \sum_{x \text{ is even}} \sum_y J \hat{a}_{x+1,y}^\dagger \hat{a}_{x,y} + \text{H.c.}, \quad (17)$$

$$\hat{H}^4 = \sum_x \sum_{y \text{ is odd}} J \hat{a}_{x,y+1}^\dagger \hat{a}_{x,y} + \text{H.c.}, \quad (18)$$

$$\hat{H}^5 = \sum_x \sum_{y \text{ is even}} J \hat{a}_{x,y+1}^\dagger \hat{a}_{x,y} + \text{H.c.} \quad (19)$$

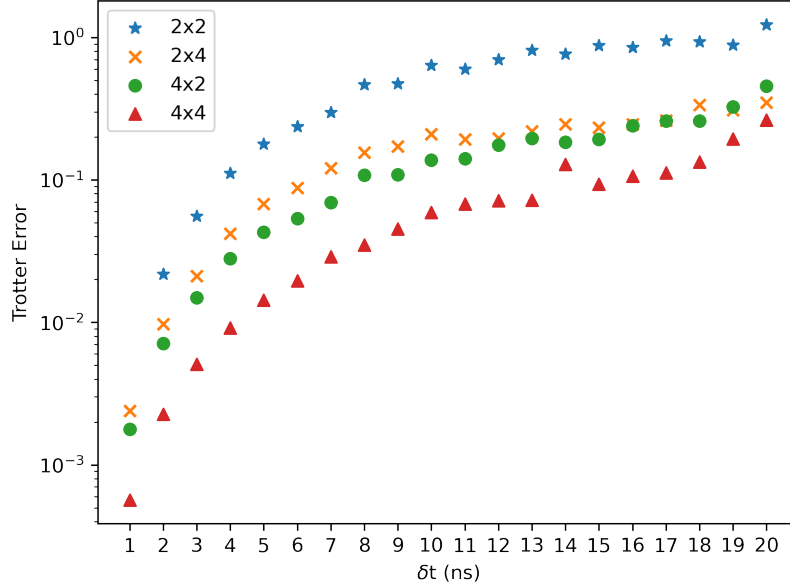
Here, $\hat{H}^1$ is the sum of single-site operators while $\hat{H}^2$ to $\hat{H}^5$ are the sum of two-site operators.

For extended time intervals t , we introduce timestamps $\delta t, 2\delta t, 3\delta t \dots$ and $T, 2T, 3T \dots$, and label them sequentially as $t_1, t_2, \dots$ in a chronological order, with $t_0 = 0$ and $t_N = t$. This construction guarantees that each time interval $t_{i+1} - t_i$ is less than δt and ensures that the Hamiltonian remains time-independent.

We use a raw tensor to estimate the Trotter errors, defined as the maximum discrepancy between the imbalance computed using the exact Hamiltonian and that obtained from an approximate Hamiltonian,

$$\text{Trotter Error} = \|\mathcal{I}_{\text{raw}} - \mathcal{I}_{\text{exa}}\|_{\infty}. \quad (20)$$

Here $\mathcal{I}_{\text{raw}}$ is calculated using raw tensors in conjunction with the Trotter-Suzuki decomposition. We compute the Trotter errors exactly for the parameters $T = 20$ ns and $n = 2$ on $2 \times 2, 2 \times 4, 4 \times 2$ and 4×4 lattices. The results are depicted in Supplementary Fig. 1. In our subsequent simulations, we select $\delta t = 3$ ns, such that the Trotter errors remain negligible.



Supplementary Fig. 1. **Estimation of Trotter errors.** We use Supplementary Eq. (20) to evaluate the Trotter error with different time steps on $2 \times 2, 2 \times 4, 4 \times 2$ and 4×4 lattices. The driving period is set to $T = 20$ ns and $n = 2$.

2. Project entangled pair states

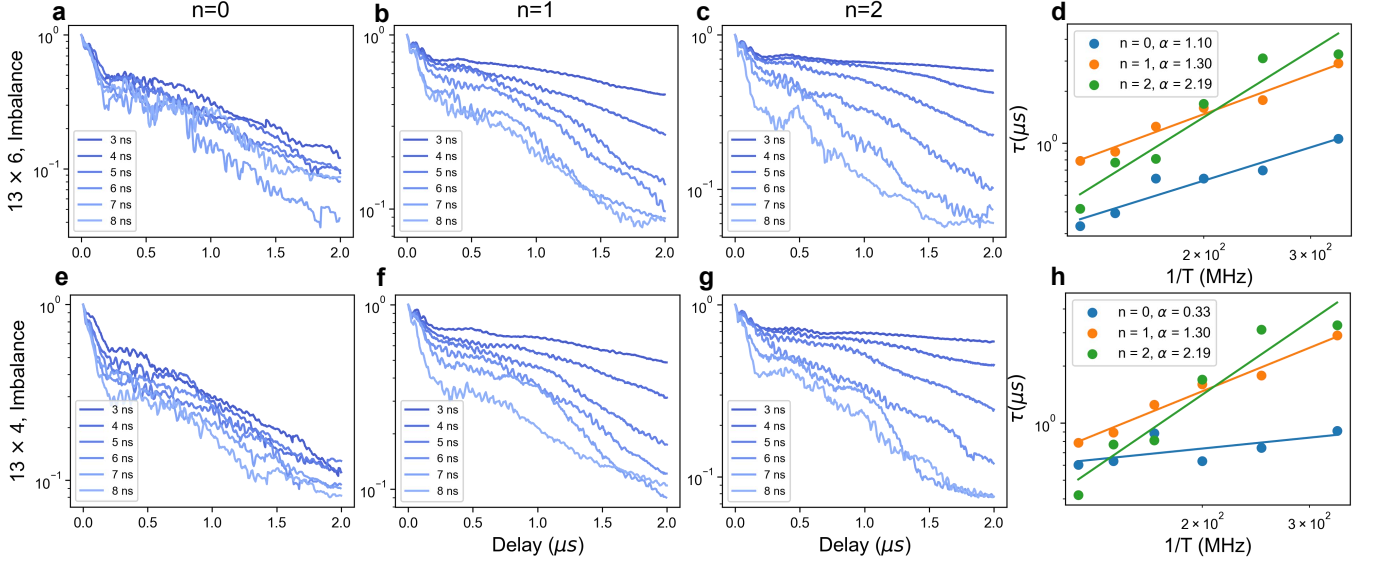
PEPS presents a way to parameterize many-body wave functions in a tensor network whose structure matches the two-dimensional underlying geometry of size $L_x \times L_y$. This approach is scalable in 2D and expressive, making it a potent tool for investigating the physics of strongly correlated systems [40, 41] and for simulating quantum circuits [42].

We adopt the same decomposition strategy as Supplementary Eqs. (15–19). Choose the basis $|0\rangle \leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ for the unoccupied state and $|1\rangle \leftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ for the occupied one. The factorized time-evolution operators in this basis are represented as

$$e^{-i\delta t h \hat{a}_i^\dagger \hat{a}_i} = \begin{bmatrix} 1 & 0 \\ 0 & e^{-i\delta t h} \end{bmatrix}. \quad (21)$$

$$e^{-i\delta t J (\hat{a}_i^\dagger \hat{a}_j + \hat{a}_j^\dagger \hat{a}_i)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\delta t J) & -i \sin(\delta t J) & 0 \\ 0 & -i \sin(\delta t J) & \cos(\delta t J) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (22)$$

Single-site operators are 2×2 matrices and thus can be directly applied to the single physical bonds. The two-site operators are 4×4 full-rank matrices, which will increase the bond dimensions. We apply the two-site time evolution operators using simple update methods [43, 44], simply applying the operators on local tensors without accounting for extra environments, and then



Supplementary Fig. 2. **Dynamics of the imbalance for system sizes 13×6 and 13×4 using PEPS.** The maximum bond dimensions are $\chi = 4$. **a–c**, Size 13×6 , $n = 0, 1, 2$ respectively, imbalance dynamics for T ranging from 3 ns to 8 ns. **d**, We use a linear fitting to determine the power-law scaling, $\tau \propto T^{-\alpha}$, in the log-log scale. **e–h**, Size 13×4 , same as **a–d**. All results are computed on Intel Xeon Platinum 8268 CPUs.

truncate to retain the largest bond dimension χ . This approach ensures that the computational cost is managed while maintaining the essential features of the system's states.

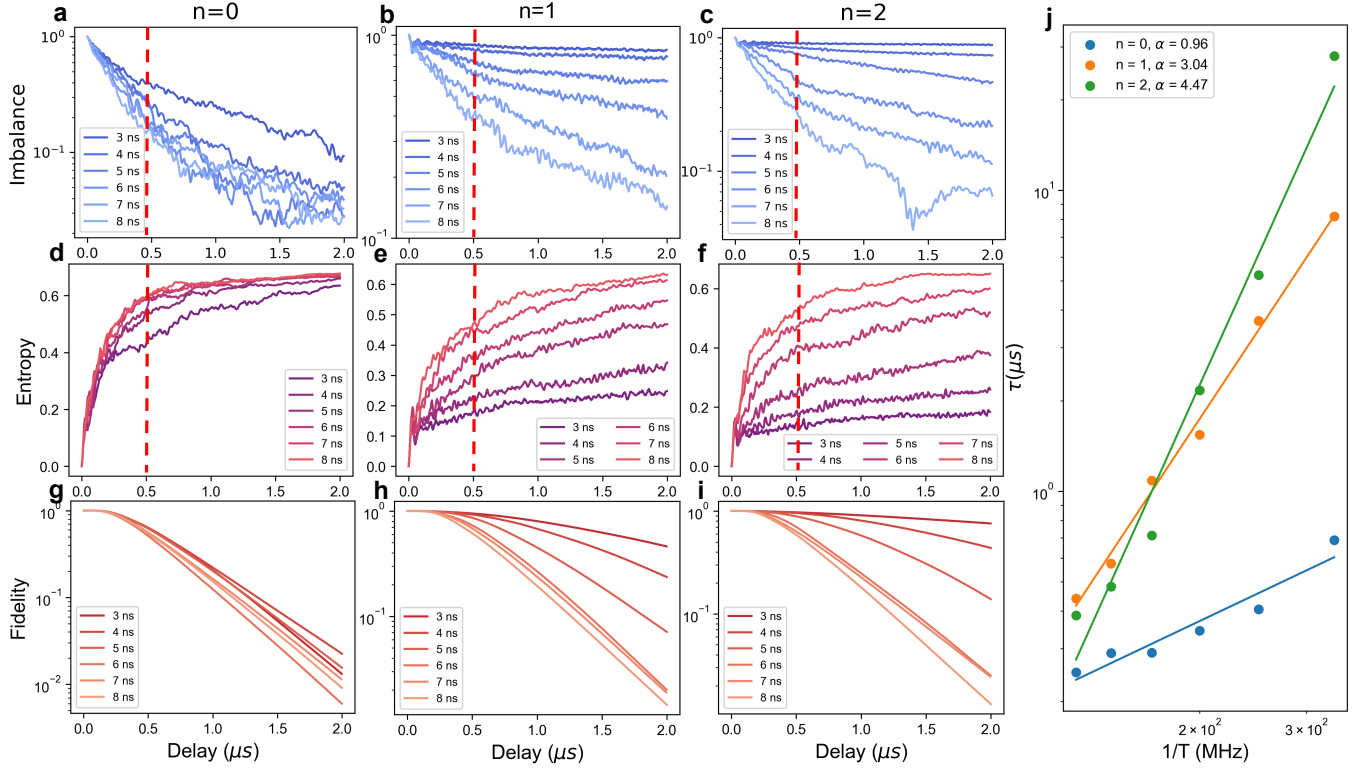
For generic PEPS, computing the expectation value of a single-site observable is computationally expensive due to the fact that fully contracting a PEPS is a $\#P$ -hard problem. To address this, we employ the standard approximate method known as Time-Evolving Block Decimation (TEBD), which was originally designed to simulate the time evolution of matrix product states [45, 46], and can also be utilized to contract two-dimensional PEPS. The computational complexity of this procedure scales as $\mathcal{O}(\chi^{12})$, making it the most resource-intensive step. The dynamics of imbalance from PEPS of $\chi = 4$ are shown in Supplementary Fig. 2. Each curve is obtained by averaging over ten different random realizations. Although PEPS captures the signal of prethermalization over a very short time scale, the systems experience accelerated thermalization during extended dynamic processes, manifesting as rapid decays of the imbalance. Furthermore, the observed scaling law of the thermalization lifetime is significantly shorter than theoretical predictions. Accurately modeling longer dynamics would incur a substantially higher computational cost compared to experiments.

3. Grouped matrix product states

Given the inherent limitations of PEPS methods, particularly the lack of a canonical form, we employ MPS. MPS are renowned for their efficiency and high accuracy in representing one-dimensional gapped quantum phases [45]. With the availability of a canonical form, the expectation values of local operators can be computed by contracting only the local tensors, rather than the entire tensor networks.

The numerical results of the dynamics of the 13×6 system are shown in Supplementary Fig. 3. All computations are performed on Intel Xeon Phi 7285 CPUs and it takes about ten hours to complete one trajectory simulation. We compute the imbalance (Supplementary Fig. 3a–c), the entropy of the upper leftmost site (Supplementary Fig. 3d–f), and the fidelity (Supplementary Fig. 3g–i). Each curve is obtained by averaging over ten trajectories. The results show that the imbalance exhibits exponential decay at short times ($t \lesssim 0.5$ ns, shown by the red dashed lines) and subsequently slows down. This time scale is also consistent with the deceleration as observed in the entropy growth. The short-time imbalance data are used to estimate the decay rate, as shown in Supplementary Fig. 3j, which exhibits the expected algebraic scaling T^{2n+1} approximately.

The decay of the estimated fidelity is also consistent with the heating process. During the prethermal regime, the system remains close to the density wave ordered state and the growth of the entanglement is slow. Therefore, MPS exhibits a high accuracy in numerical simulations. However, heating destabilizes the density wave order and at the same time, fidelity quickly drops. In particular, for $T = 7$ ns and $n = 2$ (corresponding to the orange experimental data in Fig. 3), as shown in Supplementary Fig. 3i, a notable decay of the estimated fidelity is visible already at early times and it drops below 0.05 when $t \approx 1.5$ μ s. This highlights that the non-equilibrium dynamics of a driven quantum system quickly become numerically intractable, especially for



Supplementary Fig. 3. **Grouped MPS results for a 13×6 system.** The maximum bond dimension is set to $\chi = 96$, and we average ten random trajectories for each parameter being studied. **a–c**, Imbalance dynamics for various T , ranging from 3 ns to 8 ns, and $n = 0, 1, 2$, respectively. **d–f**, Single-site entropy for various T ranging from 3 ns to 8 ns, and $n = 0, 1, 2$ respectively. The red dashed line approximately marks the time where noticeable errors appear. **g–i**, The corresponding fidelity of the MPS simulation. **j**, We employ the short-time imbalance dynamics ($t < 0.5 \mu\text{s}$) to determine the lifetime τ . It exhibits a power-law dependence $\tau \propto (1/T)^\alpha$. The slopes are $\alpha(n=0) = 0.96$, $\alpha(n=1) = 3.04$, $\alpha(n=2) = 4.47$, close to the theoretical prediction, $2n+1$.

2D systems like ours.

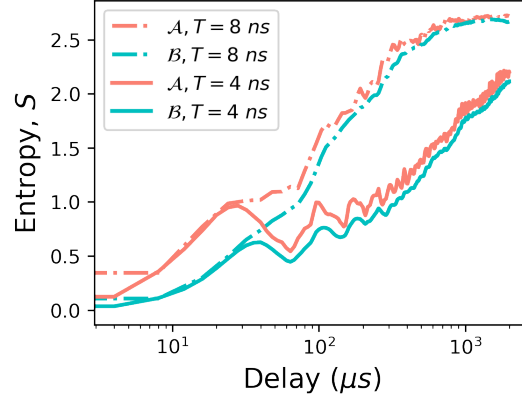
Utilizing the canonical form we can conveniently calculate the von Neumann entropy of a subsystem

$$S = -\text{Tr}(\rho \log \rho), \quad (23)$$

with the reduced density matrix ρ . For instance, when calculating the von Neumann entropy of the subsystem $\mathcal{B}$ as depicted in Fig. 4a in the main text, all of the four qubits in $\mathcal{B}$ are in one column and grouped into the fifth MPS tensor from left to right. We begin by moving the canonical site to the fifth tensor. This arrangement ensures that when tracing out the tensors 1~4 and 6~13 tensors we obtain two identity matrices separately. Then we reshape the fifth tensor to a tensor of dimensions $(\chi, 2, 2, 2, 2, 2, \chi)$ and modify it to a $16 \times 4\chi^2$ matrix C . Here, the bold numbers signify the physical dimensions of the subsystem $\mathcal{B}$. The reduced density matrix is $\rho_{\mathcal{B}} = CC^\dagger$ and von Neumann entropy can be calculated straightforwardly. For the subsystem $\mathcal{A}$, computation of the reduced density matrix is more complicated. We still benefit from the canonical form by tracing out the 1~4 and 9~13 tensors, but need to contract the 5~8 tensors explicitly to get $\rho_{\mathcal{A}}$. For $\chi = 96$, the evaluation of $S_{\mathcal{A}}$ takes approximately 6.9 seconds, whereas the evaluation of $S_{\mathcal{B}}$ only takes about 0.54 seconds on a personal computer.

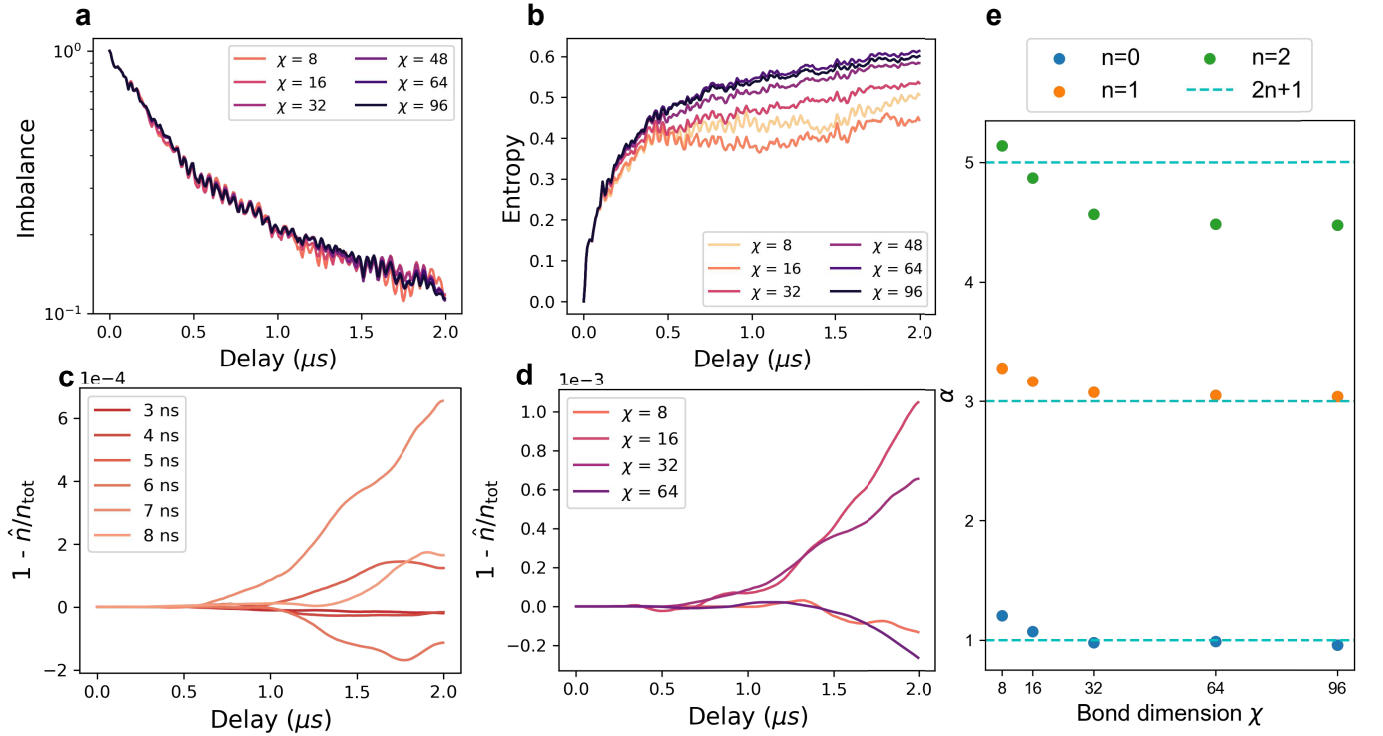
For simplicity, we first analyse the dynamics of the entanglement entropy of the top leftmost single site, as shown in Supplementary Fig. 3d–f. Entanglement entropy grows at different rates for different multipolar order and the driving period, T . Crucially, for multipolar order $n \geq 1$, the increase of entanglement entropy can be substantially suppressed by using a fast drive, highlighting the importance of the temporal correlation in stabilizing the system. However, heating is still inevitable and leads to notable entanglement generation at later times, significantly increasing the computational resources in simulating the time evolution.

We further analyze the entanglement dynamics of subsystems $\mathcal{A}$ and $\mathcal{B}$, as illustrated in Supplementary Fig. 4. For a rapid drive with $T = 4$ ns, our numerical simulation successfully reproduces the oscillation behavior of entanglement within the prethermal regime. It is worth noting that, the subsystem $\mathcal{A}$ exhibits a more pronounced oscillation amplitude than the subsystem $\mathcal{B}$, indicating a non-uniform entanglement distribution in this prethermal plateau. This behavior has also been experimentally detected, as shown in Fig. 4a in the main text. However, due to the presence of noise and decoherence, the oscillatory dynamics



Supplementary Fig. 4. **Entanglement dynamics for different subsystems.** Subsystems $\mathcal{A}$ and $\mathcal{B}$ both exhibit oscillatory behaviors when $T = 4$ ns. In the experiment, however, we only observe the pronounced oscillatory dynamics in subsystem $\mathcal{A}$. For subsystem $\mathcal{B}$, due to the presence of noise and decoherence, as shown in Fig. 4a, its entanglement entropy quickly saturates at a prethermal plateau in real experiments. We use 1-RMD drives.

in subsystem $\mathcal{B}$ becomes invisible experimentally. Also, for a slower drive (dashed dotted line in Supplementary Fig. 4), the system heats up swiftly without exhibiting the prethermal plateau.



Supplementary Fig. 5. **Grouped MPS with varying maximum bond dimensions.** **a–b**, Dynamics of the imbalance and the single-site entropy for the 13×6 lattice, with driving period $T = 7$ ns and 2-RMD protocol. **c**, Dynamics of the total particle number for the 2-RMD drive with different periods. The relative errors are at scale 10^{-4} and the particle numbers are well preserved. **d**, Dynamics of the total particle number for different maximum bond dimensions. The driving period is $T = 7$ ns and we study the 2-RMD driving. **e**, Fitted scaling exponents α using the imbalance dynamics. Different bond dimensions χ are used. Dots represent the fitted values, while the dashed lines correspond to $2n + 1$.

Moreover, we investigate the performance of grouped MPS with different maximum bond dimensions. In Supplementary Fig. 5a–b, the congruence of these imbalance and entropy is sustained only for a short time, beyond which discrepancies appear at later times. Varying values of χ also affects the largest reachable entanglement entropy. Even with $\chi = 96$ the entropy is still

far from the Page value at long times, suggesting the computational challenge in simulating the heating process. We calculate the scaling exponents by extracting the prethermal lifetime for different T based on the imbalance decay and present the results in Supplementary Fig. 5e. For $n = 0$ and 1 the exponents converge to the theoretically prediction $\alpha = 2n + 1$ (dashed lines) when $\chi \geq 32$. However, for $n = 2$, the exponent converges to around 4.5, which is still smaller than 5. It happens because for large n , typically one needs an extremely long time scale to extract the scaling exponent, as shown in [26], which is far beyond the reach of MPS simulations. In fact, similar behavior also appears in our experiments. As shown in Fig. 3a, due to the presence of noise and decoherence, for large system sizes the scaling exponent for $n = 2$ is around 3.6.

We also check the conservation of total particle number

$$\hat{n} = \sum_{x,y} \hat{a}_{x,y}^\dagger \hat{a}_{x,y}, \quad (24)$$

during the MPS simulation. It remains almost conserved at the value $n_{\text{tot}} = L_x L_y / 2$ throughout the driven dynamics. This quantity serves as a critical criterion for the performance of our algorithms. Our analysis reveals that the relative errors in particle numbers are around 10^{-4} (Supplementary Fig. 5c), indicating that convergence is robustly achieved. Furthermore, this conservation is consistently maintained across different values of χ (Supplementary Fig. 5d).

To further demonstrate that GMPS is the most suitable algorithm for the underlying lattice geometric of this system. We provide a comparison with two other one-dimensional MPS-based algorithms: TEBD and the two-site Time-Dependent Variational Principle (TDVP). We force the lattice coordinates to one-dimensional sites through the map $(x, y) \rightarrow y + xL_y$. We compute a single trajectory for the $n = 1, T = 8$ ns case. We use ITensors.jl package [47] to implement TEBD and TDVP algorithm. Supplementary Fig. 6 shows the imbalance and single-site entanglement entropy dynamics in different lattice sizes (L_x, L_y) with bond dimensions are fixed at $\chi = 64$.

In a 4×2 lattice (Supplementary Fig. 6a–b), no truncation is needed and all simulation methods lead to the same exact evolution. For large system sizes (Supplementary Fig. 6c–h), the TDVP simulation result is similar to GMPS, while TEBD shows a clear discrepancy from $t \approx 0.3 \mu\text{s}$. Such discrepancy appears mainly due to the overhead from the extra SWAP operations in TEBD. Supplementary Fig. 7 depicts the dynamics with system size 13×6 using different bond dimensions. At very early times, all methods quickly achieve converged results. However, as heating happens we observe slow convergence with bond dimension. Such a behavior can be attributed to the slow decay of the singular values spectra. As shown in Supplementary Fig. 8, we monitor the 64 largest singular values of the middle tensor in GMPS. At very early times ($t = 40$ ns and 88 ns), the system is close to a product state during the prethermal regime, and hence the spectrum is dominated by a few large singular values. In such a case, a slight increase in the bond dimension leads to the converged simulation results. However, the spectrum becomes flat as time evolves, especially when the system leaves the prethermal plateau and heats towards the infinite temperature, e.g., Supplementary Fig. 8c,d with $t = 200$ ns and 600 ns. Therefore, substantially more singular values must be included to accurately reconstruct the wavefunction, leading to a slow convergence in the simulation results. In fact, this behavior generally happens during the heating process in ergodic systems, irrespective of the underlying driving protocol and Hamiltonian parameters. Therefore, further increasing the computational resource only marginally improves the simulation accuracy.

Supplementary Fig. 9 shows the CPU time per Trotter step. The TDVP exhibits significantly lower computational efficiency compared to the Trotter-Suzuki decomposition based algorithms. Moreover, our GMPS implementation is 30% faster than TEBD. Based on this evidence, our GMPS approach demonstrates the highest efficiency and accuracy for this problem.

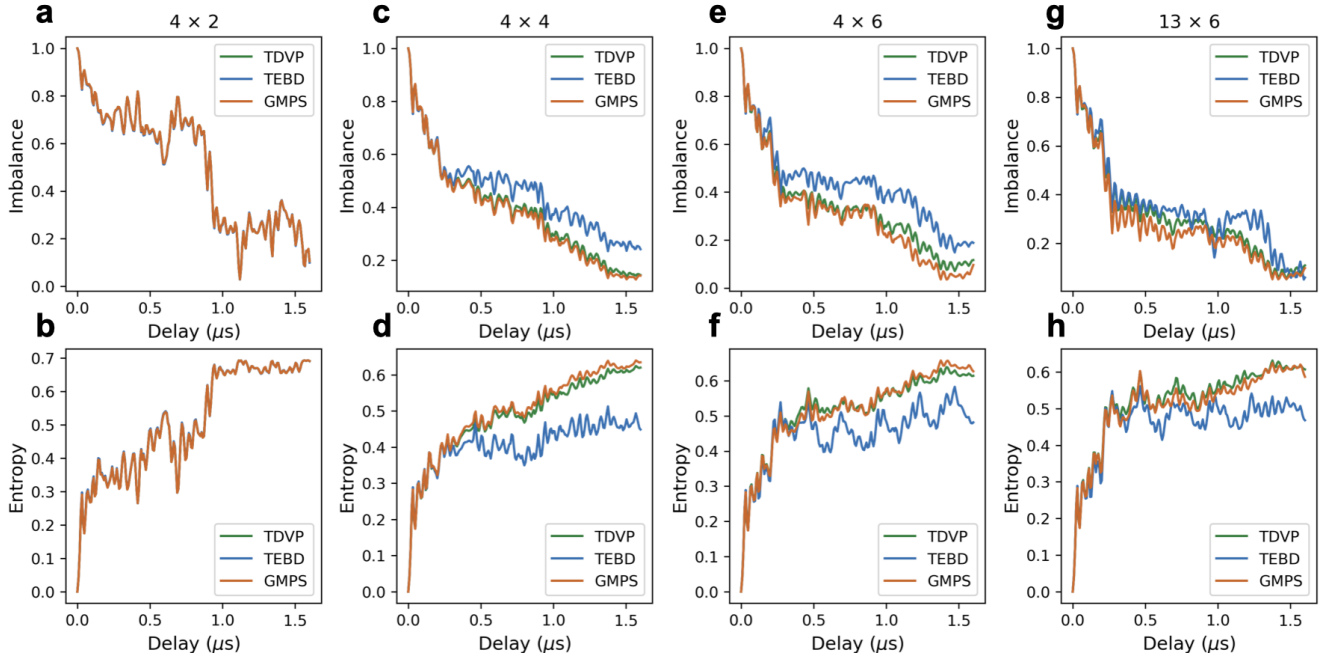
4. Pauli propagation

It has been shown that a near-Clifford quantum circuit can be efficiently simulated using the Heisenberg picture in quantum mechanics, even the state is highly entangled [42, 48]. In the Heisenberg's framework, the expectation values of observables are calculated by evolving the operators themselves, rather than the quantum states. The evolution of the expectation value of the imbalance operator is given by:

$$\mathcal{I}(t) = \langle \psi_0 | \hat{\mathcal{I}}(t) | \psi_0 \rangle = \langle \psi_0 | \hat{U}_1^\dagger(\delta t) \cdots \hat{U}_{t/\delta t}^\dagger(\delta t) \hat{\mathcal{I}}(0) \hat{U}_{t/\delta t}(\delta t) \cdots \hat{U}_1(\delta t) | \psi_0 \rangle. \quad (25)$$

To perform this simulation, we employ the Pauli Propagation algorithm, which is a method for evolving operators in the Heisenberg picture. The central idea of this algorithm is to represent the time-evolving imbalance operator $\hat{\mathcal{I}}(t)$ as a linear combination of Pauli strings:

$$\hat{\mathcal{I}}(t) = \sum_P a_P \hat{P}. \quad (26)$$



Supplementary Fig. 6. **Comparison between GMPS, TEBD, and TDVP.** **a,c,e,g.** Dynamics of the imbalance for different lattice sizes. **b,d,f,h.** Dynamics of the single-site entanglement entropy for different lattice sizes. TEBD shows discrepancy at early time and the simulation results of TDVP are almost the same as GMPS. The bond dimension of MPS is fixed as 64.

In this expression, a_P is a complex coefficient, and $\hat{P}$ is a Pauli string. We first map the Hamiltonian of the system onto a spin-1/2 model by defining the creation and annihilation operators $\hat{a}^\dagger$ and $\hat{a}$ in terms of Pauli matrices $\{\hat{I}, \hat{X}, \hat{Y}, \hat{Z}\}$:

$$\hat{a}^\dagger = \frac{1}{2}(\hat{X} + i\hat{Y}), \quad \hat{a} = \frac{1}{2}(\hat{X} - i\hat{Y}). \quad (27)$$

This transformation yields the XY Hamiltonian, up to a constant term:

$$\hat{H}_{XY} = \sum_i \frac{h_i}{2} \hat{Z}_i + \frac{J}{2} \sum_{\langle i,j \rangle} (\hat{X}_i \hat{X}_j + \hat{Y}_i \hat{Y}_j) + \text{const.} \quad (28)$$

During the time evolution, each Pauli string subjected to a Pauli rotation will either remain unchanged or branch into two new Pauli strings. The rules governing this branching are derived from the Baker-Campbell-Hausdorff formula:

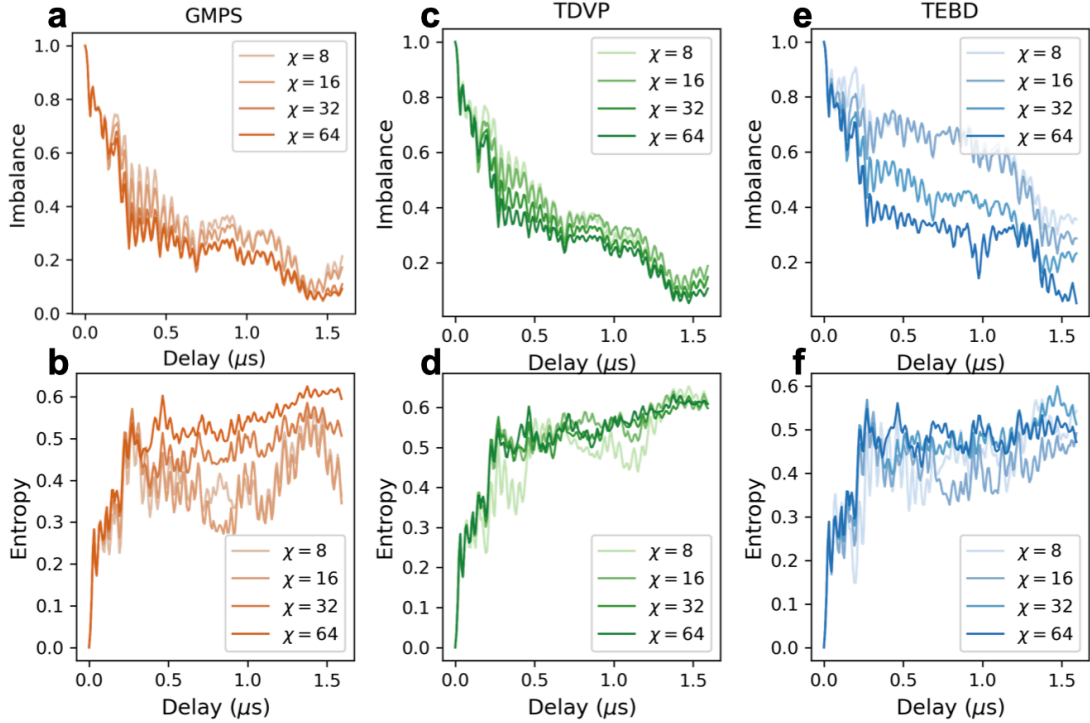
$$e^{i\theta\hat{\sigma}/2} \hat{P} e^{-i\theta\hat{\sigma}/2} = \begin{cases} \hat{P}, & \text{if } [\hat{P}, \hat{\sigma}] = 0 \\ \cos(\theta) + i \sin(\theta) \hat{\sigma} \hat{P}, & \text{else.} \end{cases} \quad (29)$$

It's worth to note that in the Pauli propagation scheme, conserved operators commute with the Hamiltonian, so these conservation laws always hold. At $t = 0$, the imbalance operator is given by

$$\hat{\mathcal{I}}(0) = \frac{1}{L_x L_y} \sum_x \left(\sum_{y \text{ is odd}} \hat{Z}_{x,y} - \sum_{y \text{ is even}} \hat{Z}_{x,y} \right) = \frac{2}{L_x L_y} \sum_x \sum_{y \text{ is odd}} \hat{Z}_{x,y}. \quad (30)$$

Here we use the particle number conservation law $\sum_{x,y} \hat{Z}_{x,y} = 0$ to eliminate half of the terms in the expression for $\hat{\mathcal{I}}(0)$.

We implement our simulation using the PauliPropagation.jl package [49]. The space and time complexity of the simulation scales proportionally with the number of Pauli strings. According to the propagation rule in Eq. (29), each Pauli string splits into two distinct Pauli operators during a single update step, inducing exponential growth in the number of terms. To mitigate this computational burden, a reduction in the number of Pauli strings is essential. Leveraging the y -reflection symmetry $(x, y) \leftrightarrow (L_x - x, y)$ inherent to our system, we merge the symmetry-related Pauli strings into a single representative operator. This symmetry exploitation reduces the Pauli string count by nearly a half. Further reduction is achieved through coefficient truncation and weight truncation strategies. The former is to discard the Pauli strings with $|a_P|$ less than a given tolerance threshold δ during the simulation and the latter is to discard the Pauli strings with weight exceeding a prescribed value w .



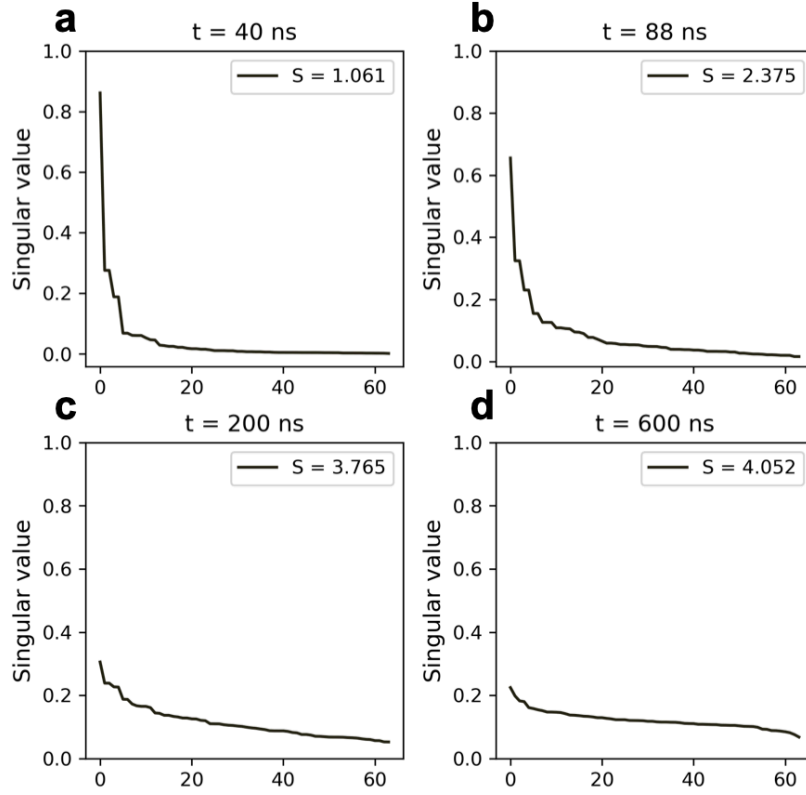
Supplementary Fig. 7. **GMPS, TEBD, and TDVP with different bond dimensions.** At very early times, all methods quickly achieve converged results. However, as heating happens we observe slow convergence with bond dimension. Such a behavior can be attributed to the slow decay of the singular values spectra.

We simulate the dynamics of the 13×6 system with truncation threshold $\delta = 10^{-7}$ and $w = 5$. Results are presented in Supplementary Fig. 10. Supplementary Fig. 10a–c display the time evolution of imbalance under different driving levels, with extracted lifetimes τ summarized in Supplementary Fig. 10g. The observed scaling exponents fall below the theoretically predicted $2n + 1$, particularly for the $n = 0$ case. Supplementary Fig. 10d–f quantify the number of Pauli strings (M) required to approximate $\hat{U}$. We observe that the number of Pauli strings reaches the maximum values near $t = 0.5 \mu\text{s}$, indicating that after initial thermalization, the system is challenging to classical simulate.

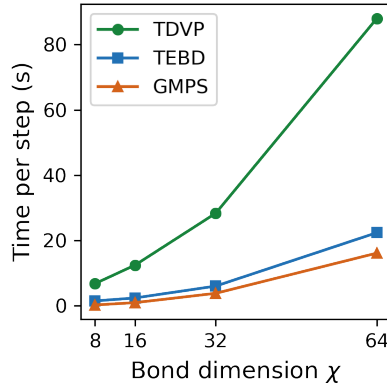
We further quantify the computational resources required for Pauli propagation. The time and space complexities are determined by the maximum numbers of Pauli strings $M_{\max}$ during the simulations. In Supplementary Fig. 11a, shows results where $\delta = 10^{-7}$ is held constant while w is varied and Supplementary Fig. 11b presents the data with $w = 5$ fixed while δ is adjusted. A critical limitation arises with weight truncation, we may discard Pauli strings with highly significant coefficients and risks uncontrolled error accumulation. Notably, Supplementary Fig. 11a indicates that the maximum number of Pauli strings grows exponentially with w , without yet exhibiting convergence. Supplementary Fig. 11c demonstrates the CPU time per Trotter step (2 layers of single Pauli rotation gates and 8 layers of two-qubit Pauli rotation gates), revealing a computational complexity scaling approximately as $\mathcal{O}(M_{\max}^{1.75})$. Furthermore, Heisenberg-picture simulations employ a backward-in-time propagation scheme. Evaluating an observable like $\mathcal{I}(t)$ requires recomputing the full evolution path from $t = 0$ to t . Consequently, if a Schrödinger-picture simulation requires N update steps, the Heisenberg-picture analogue necessitates $\mathcal{O}(N^2)$ operations. This intrinsic scaling renders long-time dynamics, such as those explored in this work, prohibitively expensive to simulate within the Heisenberg picture framework.

5. Neural quantum states

Neural networks exhibit high expressivity and strong generalization capabilities, achieving significant success across diverse domains. In quantum many-body physics, they have been demonstrated to function effectively as variational ansatz for approximating ground states and simulating quantum real-time dynamics [50–52]. We represent the quantum state by a group convolutional neural network (GCNN) that respects the symmetries of the lattice. The symmetries include the reflection sym-



Supplementary Fig. 8. **The largest 64 singular values in GMPS.** We record the singular values spectrum for the middle tensor at different times. **a,b** At early times, the spectrum is dominated by a few large singular values and GMPS quickly converges by slightly increasing the bond dimension. **c,d** At later times when heating happens, the spectrum becomes flat, leading to the slow convergence with bond dimension.

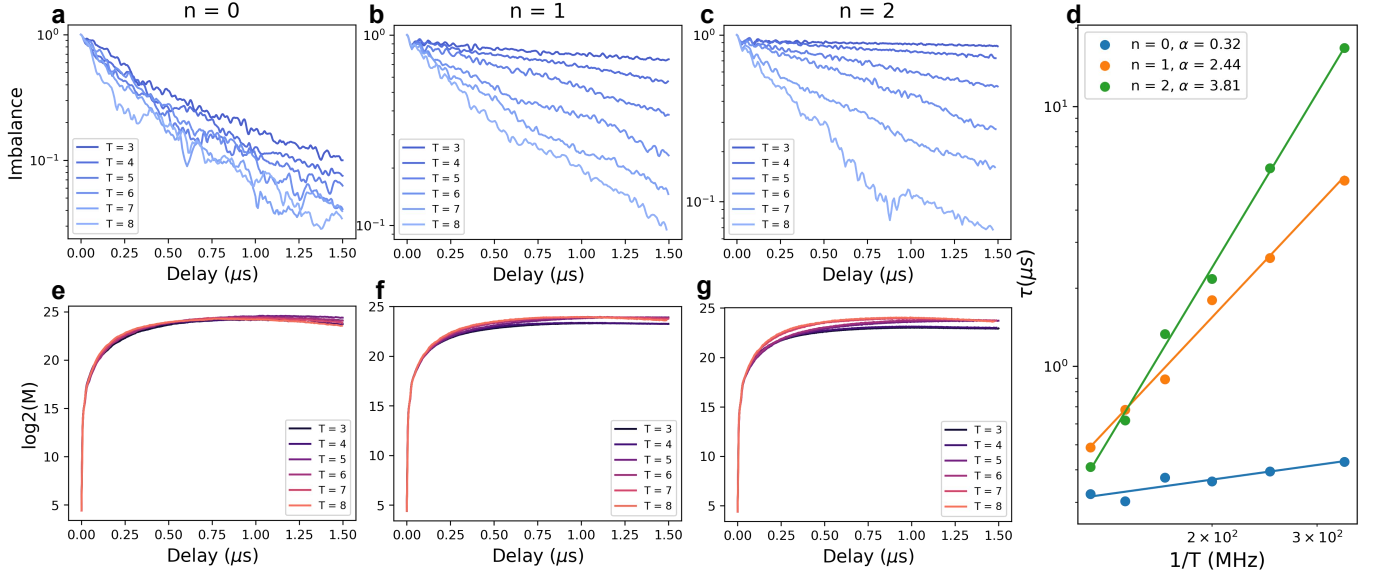


Supplementary Fig. 9. **CPU time per Trotter step varies with bond dimension.** MPS exhibits the highest efficiency among the update schemes.

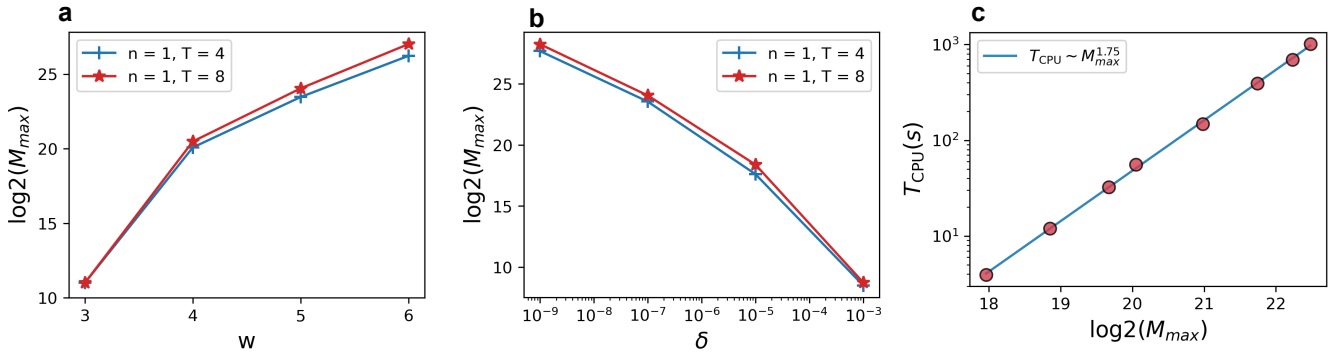
metry along the x -axis for each row. The activation function is the polynomial expansion of $\log \cosh(z)$.

$$\sigma(z) = \frac{z^2}{2} - \frac{z^4}{12} + \frac{z^6}{45}. \quad (31)$$

The use of symmetries allows us to reduce the number of parameters and efficiently capture the structure of the quantum state. The time evolution is simulated using the time-dependent variational principle with a adaptive Heun integrator. For each time segment to update, gradients are estimated using Monte Carlo samples. Our neural quantum network framework is implemented within the NetKet package [53], and all computations are performed on NVIDIA GeForce RTX 5090 GPUs.

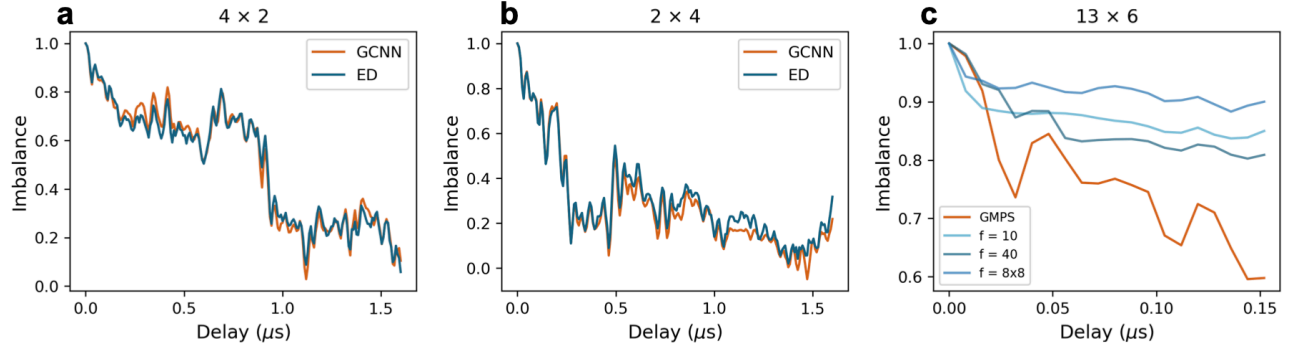


Supplementary Fig. 10. **Pauli propagation results a 13×6 lattice.** We use $\delta = 10^{-7}$, $w = 5$ for numerical simulation and perform average over ten random trajectories. **a–c**, Dynamics of the imbalance for various T , ranging from 3 ns to 8 ns, and $n = 0, 1, 2$. **d**, We employ the short-time imbalance dynamics ($t < 0.5$ ns) to determine the lifetime τ . It exhibits a power-law dependence $\tau \propto (1/T)^\alpha$. The slopes are $\alpha(n=0) = 0.32$, $\alpha(n=1) = 2.44$, $\alpha(n=2) = 3.81$, deviating from the theoretical prediction, $2n + 1$. **e–g**, The number of Pauli strings M grows in time and reaches the maximum values at $t \approx 0.5 \mu\text{s}$, suggesting that the numerical simulation becomes challenging after a short time.



Supplementary Fig. 11. **Estimation of the computational resources.** The maximum number of Pauli strings M_{max} obtained in a 13 qubit system with different parameters. In **a**, $\delta = 10^{-7}$ is fixed and in **b**, $w = 5$ is fixed. **c**, The CPU time versus the M_{max} per Trotter step, which shows a power-law dependence $T_{CPU} \propto M_{max}^{1.75}$.

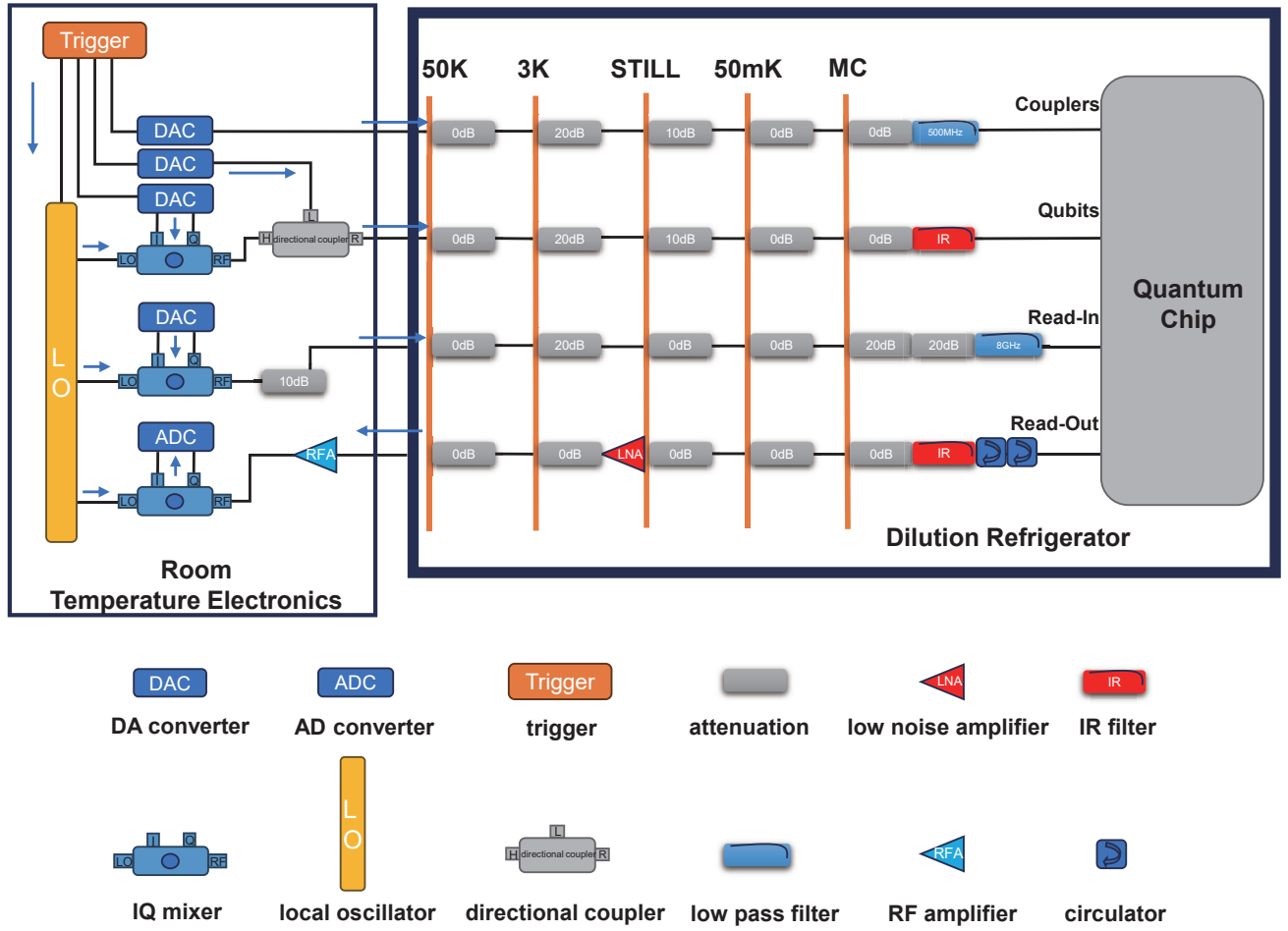
We present the results in Supplementary Fig. 12. We first benchmark our GCNN against the exact solution for small systems (4QE2 and 2QE4). We use a single layer with 10 features in the GCNN algorithm and 1,024 samples to update the network. The Hilbert spaces are of dimension $2^8 = 256$, so the samples are able to explore the whole spaces. As shown in Supplementary Fig. 12a and b, GCNN achieves high accuracy throughout the entire simulation. When enlarging the system size to 13QE6, we consider different neural network structures, including a single-layer network with 10 and 40 features, and a two-layer network with 8 features per layer, using 10,000 samples to update the networks. In Supplementary Fig. 12c, we compare the GCNN results with those from GMPS. The fidelity of GMPS within this entire time window, 0–0.2 μs , is almost 1 and hence the GMPS result (red in Supplementary Fig. 12e) is very close to the exact solution. We find that the single-layer GCNN with 40 features has the best performance. However, it only agrees with GMPS at a short timescale around 0.02 μs , corresponding to the first 3 driving cycles. Increasing the number of features or layers of the neural network can suppress the variance in gradient estimation and may further improve the simulation accuracy, but this goes beyond our current computational capabilities. In a recent work [54], GCNN has been used with 40 features to simulate dynamics in a 56-qubit transverse field Ising model, requiring approximately one week to complete 19 Trotter steps. Our system contains 78 qubits and requires long-time thermalization monitoring. Under



Supplementary Fig. 12. **Comparison between GCNN and GMPS simulation of the imbalance dynamics.** We use 4×2 , 2×4 , 13×6 lattice geometries in panels a, b and c, respectively. For the largest system size, the single-layer GCNN with 40 features has the best performance. However, it only agrees with GMPS at a short timescale around $0.02 \mu\text{s}$, corresponding to the first 3 driving cycles. Increasing the number of features or layers of the neural network can suppress the variance in gradient estimation and may further improve the simulation accuracy, but this quickly increases the computational resource.

identical settings, simulating one trajectory would demand roughly one year on a single GPU. These resource requirements far exceed reasonable computational limits.

In summary, we perform extensive numerical simulations via tensor network techniques, Pauli propagation and neural quantum states. GMPS exhibits the best performance among the others. GMPS is capable of capturing the early time dynamics, based on these data we confirm the T^{2n+1} in the suppressed heating rates. However, the appearance of heating accelerates the growth of the entanglement entropy and GMPS method requires a significant amount of computational resources in emulating our 78-qubit experiments over 1,000 driving cycles. Therefore, our experiment demonstrates non-equilibrium many-body dynamics in driven systems that is challenging to simulate classically.



Supplementary Fig. 13. Room temperature and cryogenic wiring.

S2. EXPERIMENTAL SETUP

Our experiments are performed on a superconducting quantum processor, named *Chuang-tzu 2.0*, comprises 78 qubits arranged in a square lattice configuration with 6 rows and 13 columns. Each pair of nearest-neighbor qubits is interconnected by an adjustable coupler, resulting in a total of 137 couplers within the processor. In Supplementary Fig. 13, we show the room temperature and cryogenic wiring layout. The cryogenic environment, maintained at approximately 12 mK, is achieved using a BlueFors dilution refrigerator (model XLD1000). A carefully designed arrangement of attenuators and filters has been implemented to effectively suppress thermal noise, ensuring precise control of qubits and couplers.

To meet the high-frequency requirements of RMD, we utilize a dual-channel digital-to-analog converter (DAC) operating at 2 Gs/s with 16-bit vertical resolution and dual-channel analog-to-digital converter (ADC) at 1 Gs/s with 14-bit vertical resolution, controlled by a field-programmable gate array (FPGA) chip. This system, functioning as the Z control for both qubits and couplers, is capable of generating nanosecond-scale pulses, ensuring the precision and speed necessary for advanced quantum operations. Moreover, when combined with a microwave source acting as the local oscillator (LO), the two DAC channels can modulate the in-phase and quadrature components of an IQ mixer, enabling frequency up-conversion to generate signals in the GHz range. These signals are critical for qubit XY control and dispersive readout operations. We also perform zero calibration of each IQ mixer to reduce intrinsic and mirror leakages. The custom DAC/ADC devices are synchronized using a trigger signal and a 250 MHz reference clock. Communication between the host computer and the devices is achieved through a high-speed local area Ethernet network, leveraging fiber-optic connections with a bandwidth of 10,000 Mbps.

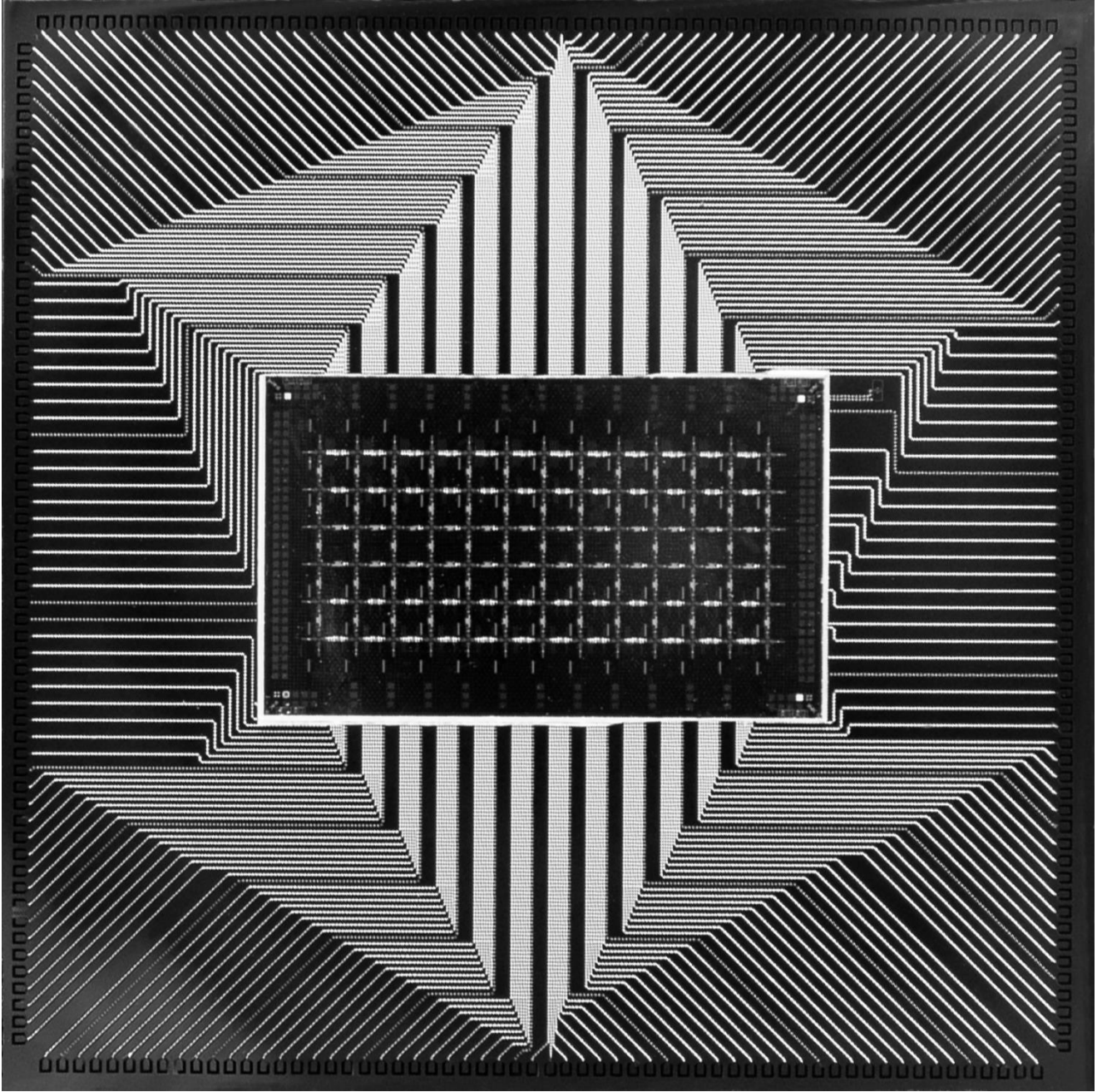
S3. DEVICE FABRICATION AND PERFORMANCE

A. Device fabrication

The optical micrograph of the processor is displayed in Supplementary Fig. 14. It is noteworthy that the qubit connectivity in our processor, while maintaining a square lattice structure, differs slightly from processors such as Sycamore [55], which feature a square lattice with zig-zag edges. We employ a tunable coupler with a capacitively connecting pad (TCCP) architecture [56], which facilitates a $1,200\text{ }\mu\text{m}$ spacing between adjacent qubits, thereby ensuring sufficient wiring space. The qubits and couplers are fabricated on the qubit layer chip, whereas the control lines, readout lines, and readout cavities are integrated into the wiring layer chip. These two chips are interconnected using flip-chip bonding technology. The readout cavities for the 6 qubits in each column are multiplexed onto a single readout line and are capacitively coupled to the qubits through interfacial capacitance. The control lines are similarly coupled to the qubits and couplers, enabling excitation and biasing via interfacial capacitance and mutual inductance.

The superconducting qubit processor utilized in the experiment was fabricated using an aluminum film deposited on a sapphire substrate, combined with a flip-chip bonding process. The detailed fabrication procedure is outlined as follows:

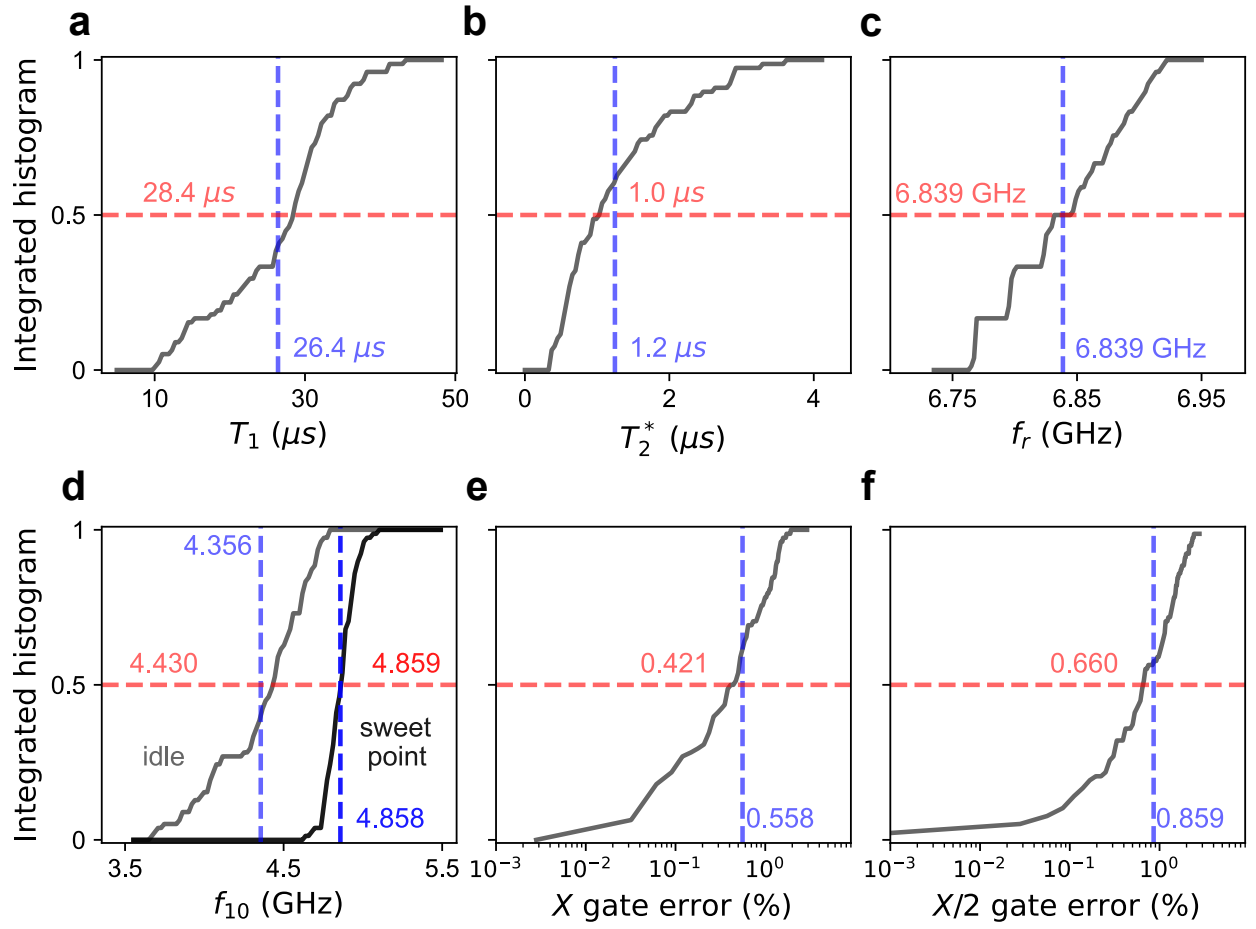
1. Two two-side polished sapphire wafers are annealed by electron beam evaporation equipment. Then 100 nm aluminum is evaporated, followed by in-situ oxidation to generate a dense aluminum oxide. This dense aluminum oxide can both protect the aluminum film and make subsequent etching processes more uniform.
2. Apply SPR-955 photoresist to the wafer surface. Use laser direct writing to expose patterns of the qubit layer and wiring layer on two wafers respectively. Develop with Tetramethylammonium hydroxide solution and etch with heated one. Finally, use the N-Methylpyrrolidone to remove the photoresist to obtain two the qubit wafer and wiring wafer.
3. We need to prepare niobium on the wafers as a transition layer between aluminum and indium, as well as for the alignment mark of the electron beam exposure. Spin-coat LOR-5A and SPR-955 photoresists sequentially, then expose and develop to obtain patterns with undercut structures. Use magnetron sputtering to in-situ argon ion cleaning to remove the aluminum surface aluminum oxide layer and then deposit 100 nm niobium. Finally, remove the photoresist and lift-off the excess niobium film from the surface of the wafers.
4. Prepare Josephson junctions on the qubit wafers. Spin-coat MAA-EL9 and PMMA-A5 bilayer resist sequentially and then deposit 10 nm aluminum for conductivity. Perform multiple exposures with different doses using electron beam lithography, and then remove the surface 10 nm aluminum, and develop using MIBK solution diluted with isopropanol. Employ in-situ argon ion cleaning with electron beam evaporation to remove the surface aluminum oxide layer, followed by double-angle evaporation with 65 nm and 100 nm aluminum respectively. Introduce in-situ oxidation between the two evaporation steps to form the insulating region in the Josephson junction's two-layer superconductors. By controlling the oxidation time, the Josephson junction's resistance and Josephson energy can be managed. Finally, remove the photoresist and lift-off the excess aluminum film, and its resistance is tested using a probe station to ensure it meets the design specifications.
5. Both the qubit wafer and wiring layer wafer need to prepare air-bridges to achieve better ground sharing and reduce various parasitic modes. Start by spin-coating SPR-220 photoresist, exposing and developing the pillar patterns for the air-bridges. Then perform reflow to create arched air-bridges surfaces. Use electron beam evaporation to deposit 500 nm aluminum after in-situ argon ion cleaning, followed by in-situ oxidation. Spin-coat S-1813 photoresist, expose and develop the areas outside the air-bridges, and then etch using heated aluminum etchant type-D. Clean with pure oxygen using reactive ion etching and then remove the resist to obtain the air-bridges.
6. Prepare indium bumps for mechanically and superconductively connecting the two chips. Spin-coat AZ-4620 photoresist for the first exposure to create the undercut region. Wait for a period to allow the photoresist to rehydrate and degas, then spin-coat S-1813 photoresist for the second exposure in the indium bump area. After development, a structure with a deep undercut can be obtained. Use an indium thermal evaporation equipment to deposit $7\text{ }\mu\text{m}$ of indium after in-situ argon ion cleaning. Clean with a reactive ion etching using a mix of argon and oxygen, and remove the photoresist and lift-off the excess indium film to obtain the indium bumps.
7. Finally, we will dice the wafers into chips and proceed with the packaging. Spin-coat AZ-4620 to protect the wafer, then use a dicing saw to cut the wafer into chips for the qubit layer ($11\times 18\text{ mm}^2$) and wiring layer ($35\times 35\text{ mm}^2$). After removing the photoresist, the two chips are bonded using flip-chip bonding equipment. Lastly, the chips are packaged into sample boxes using an ultrasonic wire bonder. After verifying that all connections from the sample box to the chip are functioning correctly using a multimeter, the chip can undergo low-temperature testing.



Supplementary Fig. 14. Optical micrograph of the processor *Chuang-tzu 2.0*.

B. Device performance

In Supplementary Fig. 15, we display the basic performance of qubits, including mean energy relaxation time $\bar{T}_1$, Ramsey dephasing time at idle frequency T_2^* , qubit readout frequency f_r , idle frequency f_{10} , and single-qubit gate error. The flux crosstalk between neighbouring qubits is below 10^{-3} . The hopping strength between neighboring qubits can be precisely adjusted using couplers, with a tunable range spanning from -10 MHz to 3 MHz. This flexibility allows for straightforward achievement of the target value, $J/2\pi = 2$ MHz. Notably, the average anharmonicity is $U/2\pi = -200$ MHz, with a ratio $|U/J| \approx 100 \gg 1$, our processor can be regarded as a hard-core bosonic system [57]. All measured qubit probabilities are corrected to mitigate the readout errors.



Supplementary Fig. 15. **Device performance of the processor *Chuang-tzu 2.0*.** **a**, Distribution of the mean energy relaxation time $\overline{T}_1$. **b**, Distribution of the dephasing time $\overline{T}_2^*$ measured by Ramsey experiments at idle points. **c**, Distribution of the readout frequency f_r . **d**, Distribution of qubit frequency of sweet points (the darker curve) and idle points (the lighter curve). **e**, Distribution of the X gate error. **f**, Distribution of the $X/2$ gate error. The single-qubit gates are characterized by quantum process tomography (QPT). In these panels, the blue vertical dashed lines indicate the mean values, while the red horizontal dashed lines represent the median values.

S4. DEVICE CONTROL AND CALIBRATION

A. Pulse calibration

When signals propagate from the DAC device to the qubits and couplers, the pulse shape undergoes distortion arising from the finite bandwidth limitations of various electrical components. The distortion of the Z pulses can induce an unintended frequency drift prior to the readout process. In our experiments, we implement precise correction protocol to effectively compensate for the distortion of Z pulses.

Typically, the measurement circuit can be modeled as a linear time-invariant (LTI) system, where the output is fully determined by its impulse response $h(t)$:

$$y(t) = S[x(t)] = \int_{-\infty}^{+\infty} d\tau x(\tau)h(t - \tau) = x(t) * h(t), \quad (32)$$

where $f * g$ denotes the convolution between f and g , $h(t)$ represents the system's output when subjected to a Dirac delta function $\delta(t)$ as the input

$$h(t) = S[\delta(t)]. \quad (33)$$

Alternatively, the unit step response of the Z pulse during propagation can be measured, i.e., $s(t) = S[\theta(t)]$, with the unit step function $\theta(t)$, offering a more practical and convenient implementation, and $h(t)$ is given by

$$h(t) = \dot{s}(t). \quad (34)$$

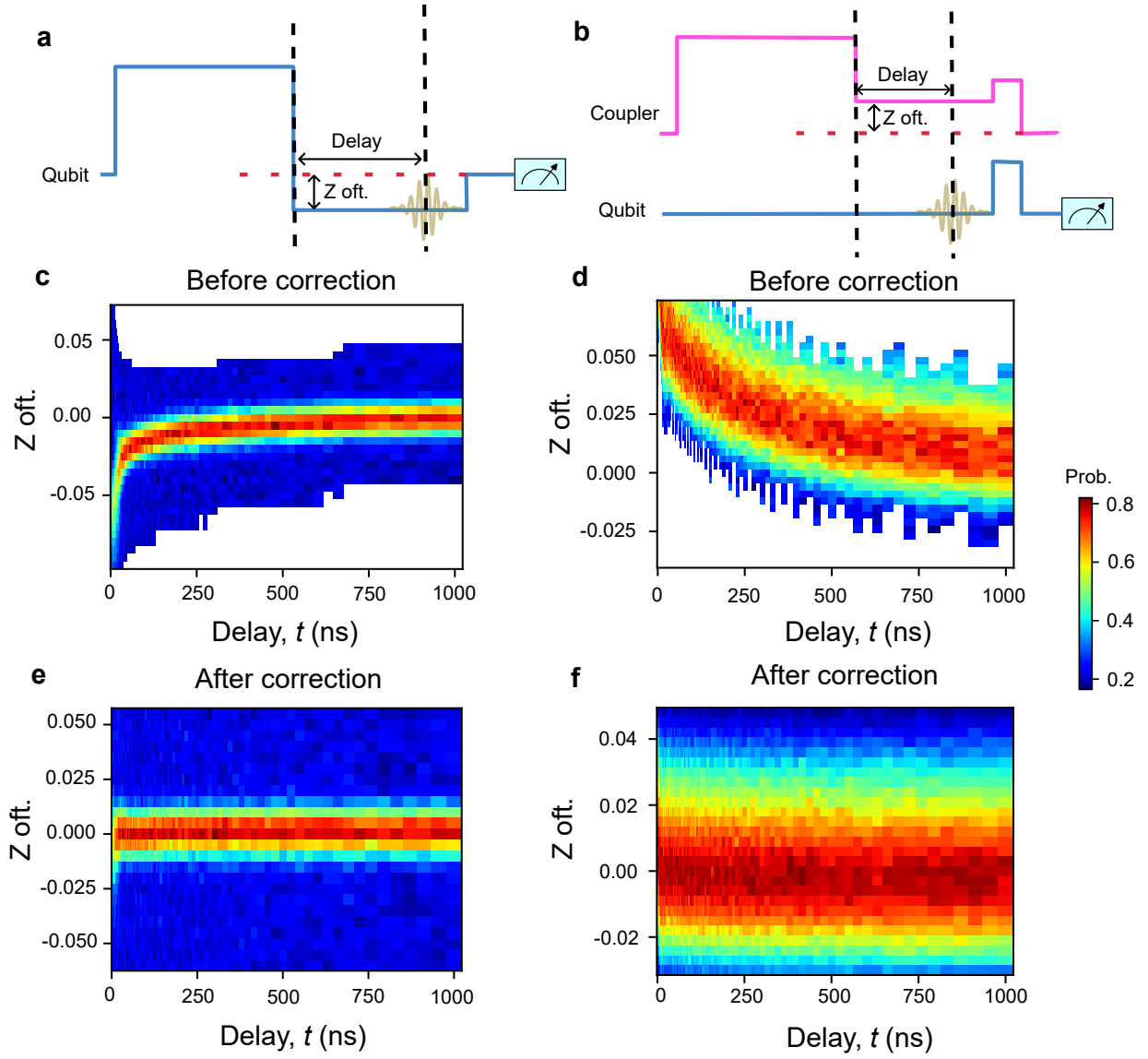
In frequency domain, we have Fourier transformations $Y(j\omega) = \mathcal{F}[y(t)]$, $H(j\omega) = \mathcal{F}[h(t)]$ and $X(j\omega) = \mathcal{F}[x(t)]$, such that Supplementary Eq. (32) is transformed as

$$Y(j\omega) = H(j\omega)X(j\omega), \quad (35)$$

with $H(j\omega) \approx 1$. Thus, the target signals can be effectively reconstructed through deconvolution with the inverse of the frequency response. According to Supplementary Eq. (34), the key procedure for distortion correction is to obtain the unit step response $s(t)$. The pulse sequences for correcting the Z pulse distortion in qubits and couplers [58] is shown in Supplementary Fig. 16a and b, respectively.

To correct the distortion in a qubit, a long pulse with duration about 3 μs and a fixed amplitude is applied to the Z channel, generating a step signal. After a specific delay, a π -pulse is applied to the qubit, which can only excite the qubit to the $|1\rangle$ state when its frequency matches that of the π -pulse. The probability of the qubit being in the $|1\rangle$ state is then measured, as illustrated in Supplementary Fig. 16c. By determining the peak population of the $|1\rangle$ -state for each delay, the Z offset data can be fitted using an exponential decay function. By employing the deconvolution method, we generate a corrected waveform that is subsequently transmitted to the arbitrary waveform generator (AWG). As a result, the qubit experiences a significantly flattened Z pulse, as demonstrated in Supplementary Fig. 16e.

In addition, the Z distortion in the coupler can be effectively corrected by exploiting its interaction with the neighboring qubit. This is achieved by first applying a step signal to the coupler, followed by a π -pulse to the neighboring qubit after a specific delay. The qubit and coupler are then biased into resonance to enable efficient energy exchange by a fine-tuned swapping duration and amplitude. Following this swapping process, the probability of the $|1\rangle$ state is measured. Similarly, the population of the $|1\rangle$ state is maximized only when the coupler frequency precisely matches the qubit frequency. As demonstrated in Supplementary Fig. 16d and f, the peak probability of the $|1\rangle$ state can be fitted to accurately characterize and correct the Z distortion in the couplers.



Supplementary Fig. 16. **Calibration of the Z pulse distortion for qubits and couplers.** **a**, Schematic of the experimental pulse sequence for correcting Z distortion of qubits. **b**, Schematic of the experimental pulse sequence for correcting Z distortion of couplers. **c**, **d**, The $|1\rangle$ state probability before correction for qubits and couplers, respectively. **e**, **f**, The $|1\rangle$ state probability after correction for qubits and couplers, respectively.

B. Timing by Floquet engineering

Another challenge in achieving precise RMD control lies in the timing misalignment between adjacent qubits, when applying rapid pulse with the short period less than 5 ns.

Here, we employ Floquet engineering [59] to synchronize the timing sequence of the flux bias pulses between qubit-qubit pairs. As shown in Supplementary Fig. 17a, the Floquet sequence $\{\hat{U}_- \hat{U}_+ \hat{U}_- \hat{U}_+ \hat{U}_- \hat{U}_+ \cdots\}$, with the same characteristic time T , is simultaneously applied to two adjacent qubits. The coupling strength is dynamically modulated according to the degree of timing misalignment between the qubits. To mimic the the modulation dynamics, we use a sinusoidal waveform ($j = 1, 2$)

$$\omega_j(t) = \omega_c + A \sin(\mu t + \varphi_j), \quad (36)$$

where $A = \delta h \cdot h$ is the driving amplitude, $\mu = \pi/T$ is the driving frequency, and φ_j characterizes the timing offset. The Hamiltonian reads

$$\hat{H} = -\frac{1}{2}\omega_1(t)\hat{\sigma}_1^z - \frac{1}{2}\omega_2(t)\hat{\sigma}_2^z + g(\hat{\sigma}_1^+ \hat{\sigma}_2^- + \text{H.c.}), \quad (37)$$

with Pauli matrices $\hat{\sigma}_{x,y,z}$, $\sigma_{\pm} = (\hat{\sigma}_x \pm i\hat{\sigma}_y)/2$, and the direct coupling strength g . By applying the unitary transformation

$$\hat{U} = \exp \left\{ i \int_0^t dt [\omega_1(t) + \omega_2(t)] \right\} = \exp \left\{ -\frac{i}{2} \sum_j \left[\omega_c t + \frac{A}{\mu} \cos(\mu t + \varphi_j) \right] \hat{\sigma}_j^z \right\}, \quad (38)$$

the effective Hamiltonian can be calculated as

$$\hat{H}' = \hat{U} \hat{H} \hat{U}^\dagger + i \partial_t \hat{U} \hat{U}^\dagger \quad (39a)$$

$$= g \exp \left\{ i \frac{A}{\mu} [\cos(\mu t + \varphi_1) - \cos(\mu t + \varphi_2)] \right\} \hat{\sigma}_1^+ \hat{\sigma}_2^- + \text{H.c.} \quad (39b)$$

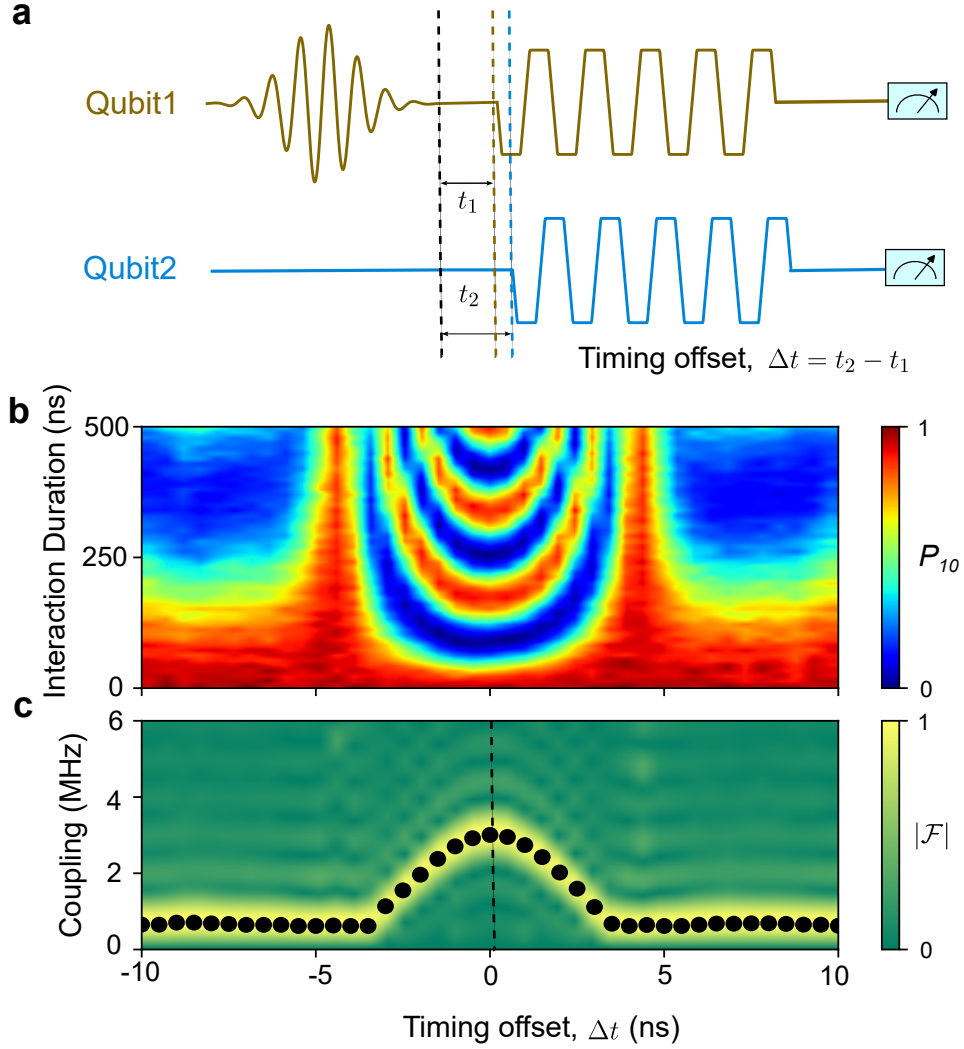
$$= g \exp \left\{ 2i \frac{A}{\mu} \sin \left(\frac{\Delta\varphi}{2} \right) \sin \left(\mu t + \frac{\varphi_1 + \varphi_2}{2} \right) \right\} \hat{\sigma}_1^+ \hat{\sigma}_2^- + \text{H.c.} \quad (39c)$$

$$= g \sum_{n=-\infty}^{\infty} J_n \left(2 \frac{A}{\mu} \sin \left(\frac{\Delta\varphi}{2} \right) \right) e^{in\theta} \hat{\sigma}_1^+ \hat{\sigma}_2^- + \text{H.c.}, \quad (39d)$$

where $\Delta\varphi = \varphi_2 - \varphi_1$ is the phase difference, $\theta = \mu t + (\varphi_1 + \varphi_2)/2$, $J_n(\cdot)$ is the n th-order Bessel function, and the last line is from Jacobi-Anger series $e^{ix \sin \theta} = \sum_n J_n(x) e^{in\theta}$. Under rotating-wave approximation (RWA), when the condition $\mu \gg g$ is satisfied, only the lowest frequency term $n = 0$ remains, leading to the effective coupling strength

$$g^{\text{eff}} = g J_0 \left(2 \frac{A}{\mu} \sin \left(\frac{\Delta\varphi}{2} \right) \right). \quad (40)$$

The effective coupling strength reaches its maximum value g , if and only if the phase difference satisfies $\Delta\varphi = 2m\pi$ with an arbitrary integer m . This relationship enables us to achieve precise timing alignment by systematically characterizing the coupling strength as a function of the phase difference $\Delta\varphi$ or, equivalently, the timing offset Δt in the Floquet sequence implementation. A typical result is shown in Supplementary Fig. 17b and c.



Supplementary Fig. 17. **Synchronization of the flux bias pulses between qubit-qubit pairs by Floquet engineering.** **a**, Experimental pulse sequences for calibrating the timing between the flux bias pulses of a qubit-qubit pair. The qubit is excited to the state $|1\rangle$ using a π pulse, followed by the measurement of the coupling strength as a function of the timing offset. **b**, The experimentally measured $|10\rangle$ probability, of the qubit versus the timing offset Δt . **c**, Fourier spectrum of the data presented in **b**. The characteristic time T is chosen as 5 ns.

S5. IMPACT OF NOISE AND BENCHMARK OF SIMULATOR FIDELITY

A. Effects of decoherence

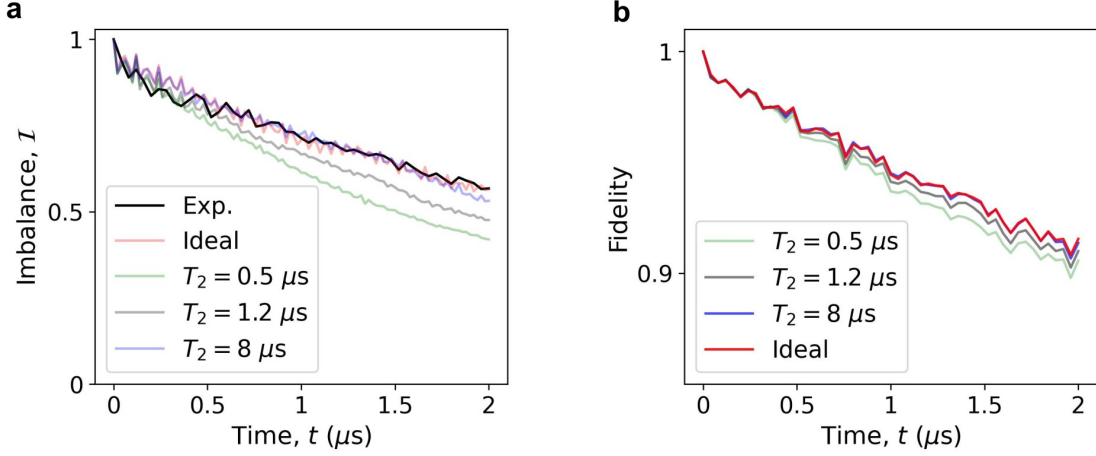
The decoherence error consists of two channels: the energy relaxation and the dephasing. On the one hand, the energy relaxation is an irreversible source. Within our evolution time, the energy relaxation is negligible due to the long average T_1 of $26.4 \mu\text{s}$. On the other hand, for many-body interacting systems, the effective dephasing time in a many-body interacting system typically exceeds the individually measured Ramsey dephasing time T_2^* as reported in [60, 61]. This happens because the eigenenergies of many-body systems depend weakly on each qubit flux [62] and the collective dynamical process also provides a coherence protection [63].

To show this, we numerically simulate the dynamics of an 8-qubit system driven by RMD governed by the Lindblad Master Equation to examine the effects of decoherence. The Lindblad Master Equation has a form of

$$\dot{\rho} = \frac{1}{i\hbar}[\hat{H}, \rho] + \sum_{m,j} (\hat{L}_{m,j} \rho \hat{L}_{m,j}^\dagger - \frac{1}{2} \{ \hat{L}_{m,j}^\dagger \hat{L}_{m,j}, \rho \}), \quad (41)$$

where ρ denotes the density matrix of the system, and $\hat{L}_{m,j}$ is the Lindblad operators for the j -th qubit with $m \in \{1, \phi\}$,

capturing the dissipation induced by the coupling to the environment. For the energy relaxation effects, we use $\hat{L}_{1,j} = \sqrt{\gamma_{1,j}}\hat{\sigma}_j^-$ with $\gamma_{1,j} = 1/T_{1,j}$; for the dephasing effects, we use $\hat{L}_{\phi,j} = \sqrt{\gamma_{\phi,j}/2}\hat{\sigma}_j^z$ with $\gamma_{\phi,j} = 1/T_{2,j} - \gamma_{1,j}/2$ [60, 64].



Supplementary Fig. 18. **Prethermal dynamics of an 8-qubit system under RMD.** (a) Comparison between experimental and numerical results for the spin imbalance across varying dephasing times. (b) Fidelity of the particle number distribution for experimental and simulated data at different dephasing times.

In Supplementary Fig. 18a, we compare experimental and numerical results for increasing dephasing time, using a RMD driving period $T = 4$ ns and the order $n = 1$. Notably, the experimentally observed imbalance decays significantly more slowly than the one obtained for the Ramsey dephasing time $T_2^* = 1.2 \mu s$, indicating that the actual dephasing time for our interacting many-body system is much longer. We find that selecting a T_2^* value of $8 \mu s$ yields numerical results that match our experimental data well. To further evaluate the effects of decoherence, we employ the occupation distribution fidelity defined as [57, 65]

$$F(t) = \frac{1}{N} \sum_i \sqrt{p_i(t)q_i(t)}, \quad (42)$$

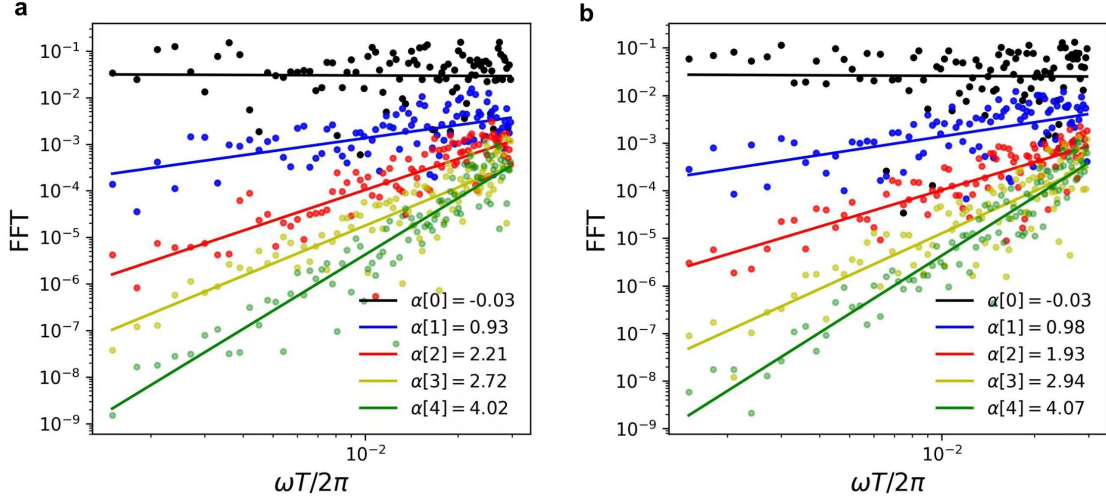
where N is the occupied particle number, $p_i(t)$ and $q_i(t)$ are the measured and theoretical probability distributions, respectively. For the data shown in Supplementary Fig. 18b, the fidelity achieves 0.92 over $2 \mu s$. Similar results can be obtained for 78-qubit systems, as shown in the Supplementary Fig. 22. In summary, the effective many-body dephasing time is much longer than the Ramsey dephasing time, exceeding the time scale required for the observation of the novel prethermalization and its characteristic lifetime scaling.

B. Effects of experimental spatio-temporal perturbations

The experimental implementation of a single unitary operator $\hat{U}_+$ or $\hat{U}_-$ is realized by selectively tuning qubit frequency to $\omega_c + \delta h \cdot h_0$ or $\omega_c - \delta h \cdot h_0$, while we simultaneously maintain a fixed coupling strength through precise control of the coupler biases. Due to the finite sampling rate of our DAC, both the rising and falling edges of the control pulses are inherently limited by the minimum duration of 0.5 ns, resulting in trapezoidal pulse shapes. This certainly introduces perturbations, which can be random in both space and time, to the unitary operators being realized in practice. We systematically analyze these perturbation effects and conclude that:

1. The appearance of prethermalization is robust to spatio-temporal perturbations;
2. The characteristic power law scaling T^{-2n-1} of the lifetime can be influenced by the temporal noise, leading to a reduction of the scaling exponent as experimentally observed for $n = 2$.

For simplicity, we first assume that perturbations are static in time. For RMD protocols using high-frequency drives, at early times, the dynamics is dominated by the averaged Hamiltonian H_{eff} and evolves towards the corresponding prethermal states. Because of the large potential energy difference h_0 in H_p (Eq. (3) in the main text), the system preserves a notable particle number imbalance in the prethermal stage. This phenomenon is robust as long as the perturbation strength is much smaller than h_0 . The perturbation induced by the difference between the trapezoidal and the square pulse only weakly perturbs H_{eff} , inducing negligible variation in the prethermal phenomenon.



Supplementary Fig. 19. **Fourier spectrum of random multipolar sequences for different n** The shapes of the pulses are chosen to be rectangular (a) or trapezoidal (b) with the rising edge 0.5 ns. The driving period is $T = 3$ ns, and the results are averaged over 30 different samples.

At longer times, heating occurs at a rate which exhibits the scaling law T^{2n+1} . This originates from the power law suppression at low frequencies (T^n , see Supplementary Fig. 19a) in the spectrum of the driving protocol [26], which remains unchanged for trapezoidal pulses (see Supplementary Fig. 19b).

As numerically verified in [26] and rigorously proved in [36], the resulting scaling law of the heating rate is universal and does not depend on the specific form of the Hamiltonian, as long as the system is locally interacting. Therefore, any time-independent local perturbations, including the trapezoidal pulses, do not severely affect the characteristic behavior of RMD systems. We have numerically verified this in Supplementary Fig. 20, where we simulate an 8-qubit system, using multipolar operators U_+ and U_- generated by trapezoidal onsite potentials with 0.5 ns rising and falling edges. Clearly, the scaling T^{2n+1} appears in the high-frequency regime, the same as the square pulse.

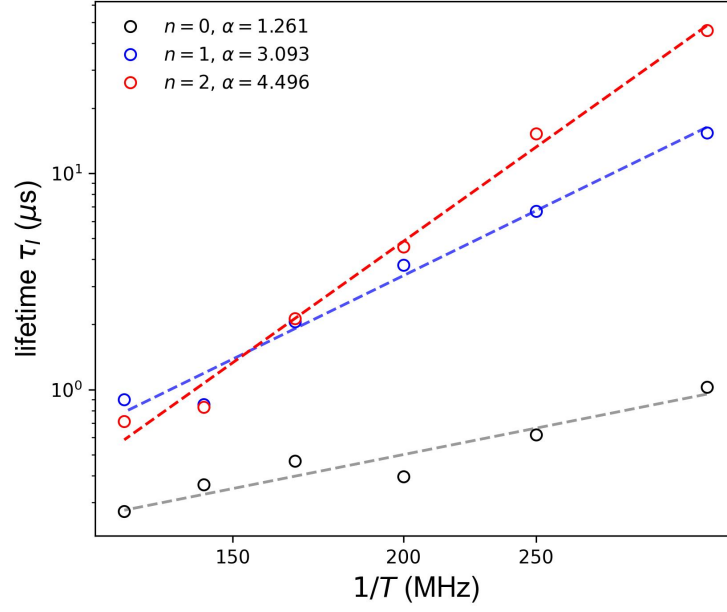
We would also like to point out that, for a larger multipolar order, it is generally necessary to use a higher driving frequency and a longer simulation time to obtain the converging scaling exponent $2n + 1$. This is evident in Supplementary Fig. 20: the prethermal lifetimes for $n = 0$ and $n = 1$ are well separated even within the intermediate frequency regime. In contrast, for $n = 2$, higher driving frequencies and longer time scales (around 1,000 driving cycles) are typically required. In the previous theoretical study [26], at least 10^6 driving cycles were needed to verify the scaling exponent for $n = 3$, which is far beyond the current experimental capabilities. Now we discuss time-dependent perturbations that can occur during the random pulse realization. In fact, it remains an outstanding challenge to precisely determine their concrete form in the experiment. For instance, noise can potentially originate from the temporal variations of the pulse shape generated from the AWG, imperfect calibration of the pulse distortion, unbalanced qubits' frequencies, and random error of coupling strengths in such a large system [57]. They potentially induce temporal perturbation in qubits' frequencies and coupling strengths during rapid modulation of RMD. To numerically demonstrate their effect, when implementing the multipolar time evolution operator, we explicitly introduce spatio-temporal disorder to the onsite potential and the coupling strength, where the perturbation strength is quantified by h_r and J_r . In numerical simulations, for simplicity, we assume that the perturbation strength is uniform in space. According to a simple Fermi's golden rule (FGR) argument [36], such white noise opens a new heating channel and its effect only becomes noticeable at the time scale $t_{\text{noise}} \sim \mathcal{O}(1/h_r^2)$ or $\mathcal{O}(1/J_r^2)$. Note, this FGR estimation remains valid even in the thermodynamic limit, i.e., their effect usually does not scale up with the system size. After this timescale, the noise dominates over other physical processes and the system eventually becomes featureless. Hence, t_{noise} naturally restricts the time window where we can investigate the RMD protocol and observe the predicted scaling law.

For the multipolar order $n = 0$ and 1, the system generally heats rapidly, and t_{noise} is irrelevant to the RMD scaling. However, for $n = 2$, the prethermal phenomenon persists for a long time so that the noise effect accumulates and may appear at late times. To model this, we introduce spatio-temporal noise by modifying the elementary Hamiltonian as

$$\hat{H}'_{\pm} = \hat{H}'_{\kappa} + (1 \pm \delta h) \hat{H}'_p, \quad (43)$$

where the on-site term is given by

$$\hat{H}'_p = \sum_{x,y} \{h_0[1 - (-1)^y] + V_{x,y}\} \hat{a}_{x,y}^{\dagger} \hat{a}_{x,y}, \quad (44)$$



Supplementary Fig. 20. **Prethermal lifetime of the RMD system with trapezoidal pulses.** The characteristic scaling still appears in the high-frequency regime.

with $V_{x,y}$ uniformly distributed in $[-h_r, h_r]$, and the hopping term takes the form

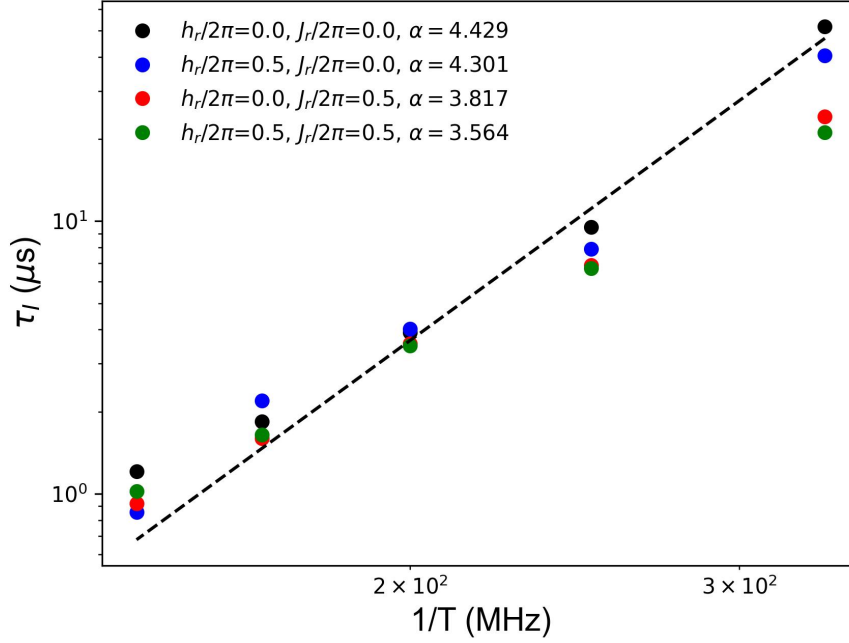
$$\hat{H}'_{\kappa} = \sum_{x,y} \{ [J + W_{(x,x+1),y}] \hat{a}_{x,y}^{\dagger} \hat{a}_{x+1,y} + [J + W_{x,(y,y+1)}] \hat{a}_{x,y}^{\dagger} \hat{a}_{x,y+1} + \text{H.c.} \}, \quad (45)$$

with W uniformly distributed in $[-J_r, J_r]$. The sampling of the spatial disorder will be regenerated for the multipolar operators at different times. As shown in Supplementary Fig. 21, we plot the prethermal lifetime obtained for different noise strengths. For the noise-free simulation (black dots), the fitted scaling exponent is close to the expected value of 5 (black dashed line). In the presence of temporal perturbations, the prethermal lifetimes are shortened in the high-frequency regime. This is sharply different from the time-independent perturbation as shown in Supplementary Fig. 20, where the scaling exponent remains unchanged. Consequently, the shortened lifetime translates into a reduction of the scaling exponent α .

Note that, in practice, the noise strength also exhibits spatial fluctuations because the control over different qubits can differ slightly. However, spatial fluctuations are averaged out for a sufficiently large system size. Indeed, as shown in Fig. 3a, the fitted scaling exponent is approximately the same for large system sizes, 8×4 , 13×4 , 13×6 . Therefore, our driving protocol can be extended to larger system sizes where the novel prethermal phenomenon should also appear.

These numerical results suggest that our experiment is robust to static perturbations, highlighting the universal nature of the power-law suppression in the heating rate of RMD systems. The experimentally observed reduction of the scaling exponent can be potentially attributed to the accumulated effects of the weak temporal noise during the implementation of the pulse.

Of course, our results may also be influenced by other experimental imperfections, such as the particle loss (see the next subsection) and other forms of decoherence. A complete characterization of all possible error sources goes beyond the scope of the current work.



Supplementary Fig. 21. **Prethermal lifetime of the RMD 8-qubit system with temporal disorder.** The dashed line represents the exact $2n + 1$ power law.

C. Quantum simulator performance benchmarking

We further discuss the system's fidelity and analyse potential error resources to strengthen the reliability of our experimental results. Unlike digital devices, where gate infidelities can be quantified relatively straightforwardly, estimating quantum state fidelity is more challenging in analog simulators. Instead, we assess particle number conservation in our 78-qubit system and compute the occupation distribution fidelity to demonstrate that our experiment remains sufficiently stable against imperfections throughout the entire time evolution.

1. Fidelity of quantum simulator

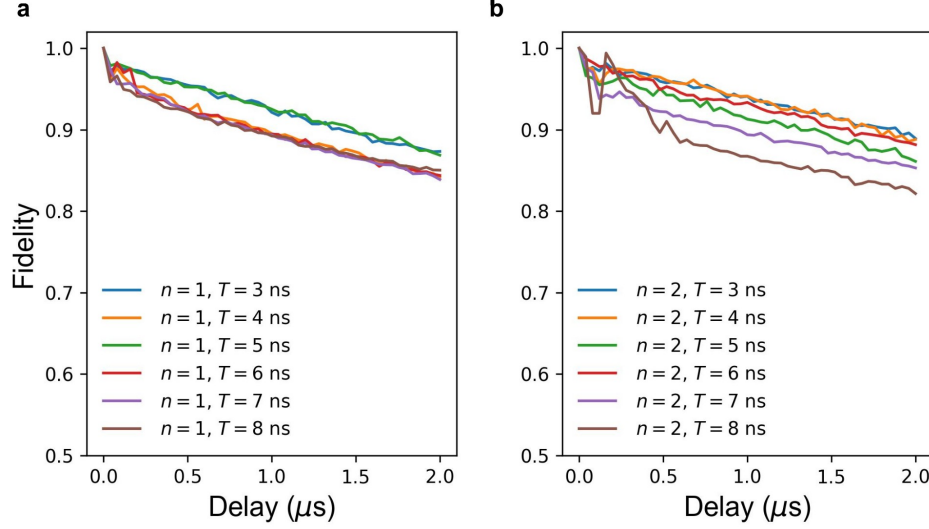
We use occupation distribution fidelity (see definition in Supplementary Eq. (42)) as a measure to benchmark our device and show our experiment remains sufficiently robust during the long-time evolution.

For a small system size, we can simulate the system accurately, and as shown in Supplementary Fig. 18, the occupation distribution fidelity remains above 90% throughout the entire evolution. For 78-qubit systems, however, the numerical simulation is challenging for generic driving parameters, due to the rapid entanglement growth. Yet, before the onset of heating, the long-lived prethermal regime only generates a low amount of entanglement entropy, such that our GMPS simulation remains sufficiently accurate. Therefore, we can calculate the occupation distribution fidelity to faithfully benchmark the 78-qubit experiment in the high-frequency regime even at long times.

Concretely, for $n = 2$, $T = 3$ ns, we estimate that the GMPS simulation with bond dimension $\chi = 48$ maintains more than 80% quantum state fidelity of the many-body wavefunction at time $2 \mu\text{s}$. Note, different from the occupation distribution fidelity, quantum state fidelity generally decays exponentially for larger system sizes. A value of 80% state fidelity for the 78-qubit system indeed suggests that the numerical simulation is almost exact throughout the entire time evolution. We use this set of driving parameters to benchmark our experiment. As shown in Supplementary Fig. 22b, for $n = 2$, $T = 3$ ns (blue), the distribution fidelity remains above 90% at $2 \mu\text{s}$, confirming that our 78-qubit system is sufficiently stable after 1,000 driving cycles. For slower driving (panel a) and also with a smaller multipolar order, heating happens faster, and the accuracy for GMPS drops accordingly, which also induces the decay in the occupation distribution fidelity when we compare experimental and numerical results.

These observations confirm the robustness of our 78-qubit quantum processor, strengthening the reliability of the experi-

mentally observed non-equilibrium phenomenon. Importantly, our finding highlights that the quantum processor is particularly useful for the study of the many-body heating process, where the numerical simulation faces a formidable challenge due to the rapid entanglement growth.



Supplementary Fig. 22. **Occupation distribution fidelity for the 78-qubit experimental data.** For $n = 2, T = 3$ ns, GMPS is sufficiently accurate which we use to benchmark our quantum processor. The distribution fidelity remains above 90% at $2 \mu\text{s}$, confirming that our 78-qubit system is sufficiently stable after 1,000 driving cycles. However, for slow driving and $n = 1$, the accuracy of GMPS drops down, which also leads to the decay in the occupation fidelity. We use $n = 1$ (a) and $n = 2$ (b) and the driving period ranging 3 ns to 8 ns.

Concretely, the simulation in Fig. 3b at early times is sufficiently accurate where the state fidelity exceeds 0.8 before $1 \mu\text{s}$ (see Supplementary Fig. 3). Within this time window, it accurately matches our experimental results, confirming the exceptional stability of our quantum processor. However, at later times, notable heating occurs together with noticeable entanglement entropy growth. This significantly increases the complexity of our GMPS simulation and the state fidelity decreases drastically to 0.45 at later times. It is worth highlighting that, both the early time GMPS results and our experiments suggest a universal exponential decay in time of the particle number imbalance. This is indeed expected during the heating process in the high-frequency regime with a constant heating rate; see systematic studies in [66, 67]. At later times, our experimental observation still remains consistent with this exponential decay, whereas the GMPS results clearly deviate from this behavior. Therefore, this comparison provides strong evidence that our quantum processor is more reliable throughout the entire time evolution up to $2 \mu\text{s}$, capable of simulating the heating process in a regime where the classical algorithm faces formidable challenges.

2. Conservation of particle number and error mitigation

For superconducting devices, restricting to a qubit system is always an approximation. The existence of high-energy levels in the system can induce leakage out of the operating subspace, leading to particle loss in our Bose-Hubbard system. In the spin representation, this corresponds to enlarging the original qubit system to the qudit system with the local Hilbert space dimension $d > 2$. This effect can become visible at the late-stage evolution of a quantum simulation task, especially when an external random drive, which can trigger high-energy excitation, is implemented. In our previous work [30], we numerically studied a multi-band Bose-Hubbard model, where the lowest band corresponds to the target Bose-Hubbard system driven by the RMD protocol, and the higher bands mimic the existence of high-energy levels. As clearly shown by the blue data in Fig. 3(a) in [30], particle loss generally accelerates the heating process in the lowest-band system towards a maximally mixed state for a local subsystem. For the particle number imbalance being studied, this results in a faster decay of its expectation value towards zero. This particle-loss induced effect can be effectively captured by depolarizing errors, which replace a quantum state with a maximally mixed state with some probability. We note that a similar error mitigation method has been used in Google's experiment, where random unitaries are implemented [55], which essentially heats up the system towards the infinite temperature.

Concretely, for qubit systems, a n -single qubit depolarizing error channel can be modeled via [55, 68, 69]

$$\mathcal{E}^{\otimes n}(\rho) = (1-p)^n \rho + \sum_{\alpha \in [x,y,z]} \sum_{j=1}^n (1-p)^{n-1} \frac{p}{3} \hat{\sigma}_{\alpha}^j \rho \hat{\sigma}_{\alpha}^j + \dots \quad (46)$$

where $\mathcal{E}$ is the error channel, ρ is the exact density matrix after the target evolution, p is the probability of an error occurring for each qubit, which is assumed to be equal for each qubit, and for each Pauli error, and ... indicates higher-order terms corresponding to errors on multiple qubits. One important feature of this error model is that the depolarizing error commutes with any unitary operator. Thus, for a many-qubit system the error on the i -th qubit with purely depolarizing errors is [69]

$$\mathcal{E}^i(\rho) = (1 - p_i)\rho + p \text{Tr}_i[\rho] \otimes \frac{I_i}{2}, \quad (47)$$

where p_i is the error on the i -th qubit, $\frac{I_i}{2}$ is the maximally mixed state for the i -th qubit. Instead of dealing with each qubit error, we can assume that there is a total error probability p_{tot} , so we have an effective depolarizing channel on the entire state

$$\rho_{noisy} = (1 - p_{tot})\rho_{ideal} + p_{tot} \frac{I^{\otimes n}}{2^n}. \quad (48)$$

Note, in principle one can directly relate p_{tot} to the local p_i , but this is unnecessary as we can directly measure the total error. To incorporate the leakage out of the computational space, one can generalize this expression to a qudit system

$$\rho_{noisy} = (1 - p_{tot})\rho_{ideal} + p_{tot} \frac{I^{\otimes n}}{d^n}, \quad (49)$$

where $d > 2$ is the Hilbert space dimension of a qudit on a given site. Note, theoretically, leakage out of the two-level computational subspace reduces the trace of the density matrix when projected onto the two-level subspace. However, when dealing with the entire d -level local Hilbert space, the dynamics is indeed trace-preserving. Our key assumption here is that drive-induced heating effectively generates chaotic evolution in the d -level system, which eventually evolves toward a maximally mixed state. One can thus develop the following depolarizing error mitigation procedure in this expanded space to account for leakage. We stress that even though this may not be a good approximation for a single driving cycle, it becomes a reasonable approximation for the error channel of many-body systems after many driving cycles and the ensemble average over different RMD realizations.

According to Eq. (49), the expectation value of an operator $\hat{O}$ can now be calculated as

$$\langle \hat{O} \rangle_{noisy} = (1 - p_{tot})\langle \hat{O} \rangle_{ideal} + p_{tot} \frac{\text{Tr}(\hat{O})}{d^n}. \quad (50)$$

For the imbalance operator $\hat{I}$, we have $\text{Tr}(\hat{I}) = 0$, leading to

$$\langle \hat{I} \rangle_{noisy} = (1 - p_{tot})\langle \hat{I} \rangle_{ideal}. \quad (51)$$

Crucially, the total error probability can be determined directly on our device by measuring the total particle number $\hat{N}$, whose conservation law is now weakly broken due to the particle loss. Note that we only get non-trivial contributions to these operators when the state remains in the computational qubit subspace. Hence, for an n -qubit system, the particle number operator acting on the i -th site is the tensor product of the $\hat{N}_i$ particle number matrix $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ on the i -th site and identity operators on the remaining sites in the qubit subspace, leading to $\text{Tr}(\hat{N}_i) = 2^{n-1}$. So the trace of the total particle number operator is $\text{Tr}(\hat{N}) = 2^{n-1}n$. Therefore, we can get

$$\langle \hat{N} \rangle_{noisy} = (1 - p_{tot})\langle \hat{N} \rangle_{ideal} + p_{tot} \left(\frac{2}{d} \right)^n \frac{n}{2}. \quad (52)$$

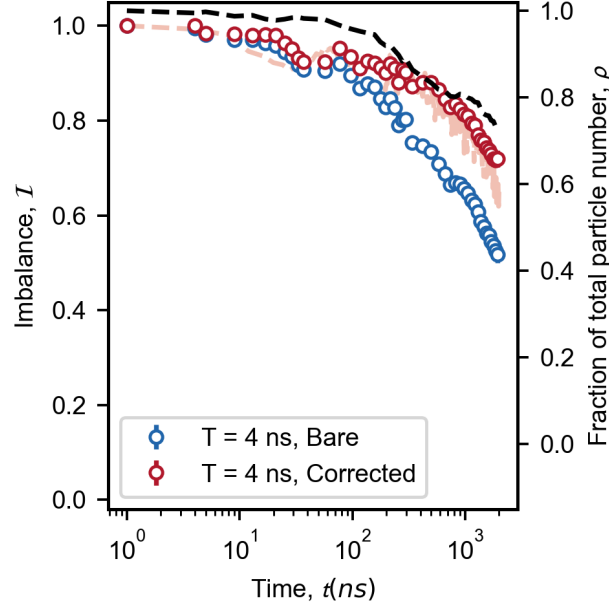
Since $d > 2$, we know $(2/d)^n \ll 1$ for large n , leading to

$$\langle \hat{N} \rangle_{noisy} = (1 - p_{tot})\langle \hat{N} \rangle_{ideal}. \quad (53)$$

Therefore, by measuring the remaining particle number in the system, one can extract $1 - p_{tot}$ at each time, and hence $\langle \hat{I} \rangle_{ideal}$ is obtained

$$\langle \hat{I} \rangle_{ideal} = \langle \hat{I} \rangle_{noisy} / (1 - p_{tot}) = \frac{\langle \hat{N} \rangle_{ideal}}{\langle \hat{N} \rangle_{noisy}} \times \langle \hat{I} \rangle_{noisy}. \quad (54)$$

To show the validity of this method, we first calculate the particle number of the 8-qubit system after calibrating the readout error. The result (black dashed curve) is shown in Supplementary Fig. 23 below. For our initial density-wave state at the half-filling, the particle number is 4 for the 8-qubit system. Due to the presence of depolarizing error, the system gradually loses



Supplementary Fig. 23. **Imbalance and fraction of the particle number (the black dashed curve) in 8 qubits.** The depolarizing error mitigation scheme is used to correct the contribution from the particle loss. The corrected results agree well with the exact numerical simulation (the orange dashed curve).

particles, but it still retains around 80% of the initial filling after 1,000 driving cycles. This effect can be mitigated using the aforementioned error mitigation scheme. As shown in Supplementary Fig. 23, the corrected results agree well with the exact numerical simulation.

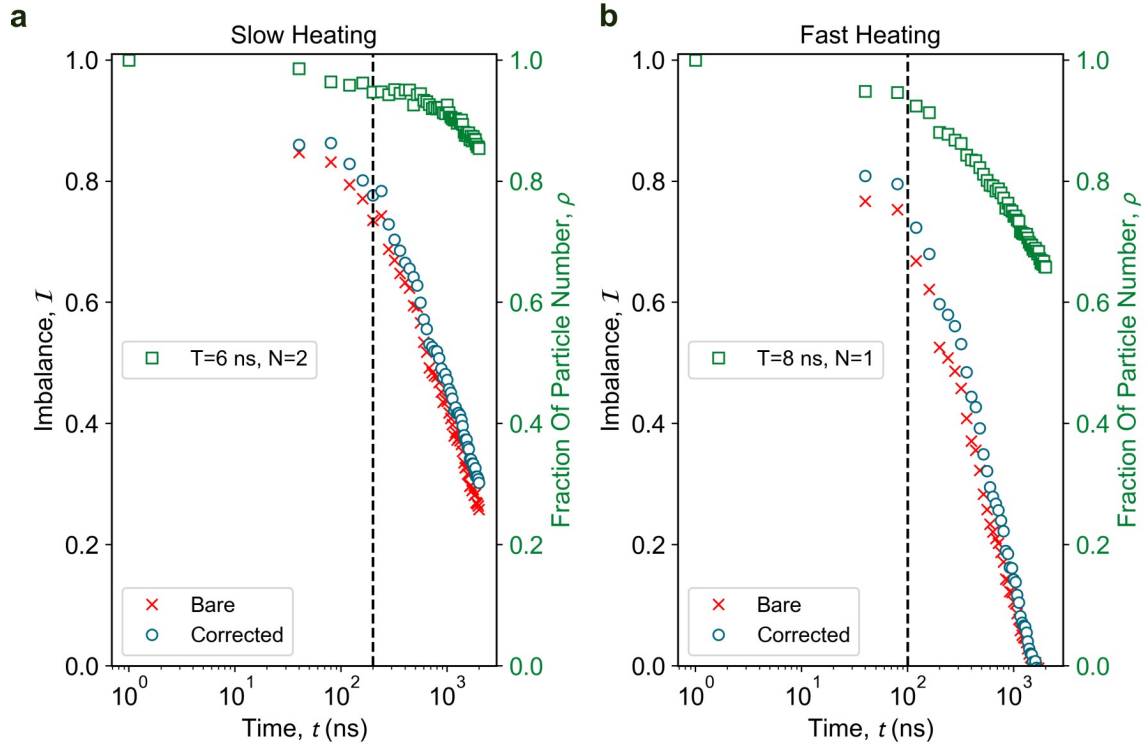
We perform a similar comparison in the 78-qubit system with two representative sets of driving parameters in Supplementary Fig. 24. In panel **a**, in the high-frequency regime and with a large multipolar order, the system exhibits a long-lived prethermal plateau, and notable heating occurs after the black dashed line around $0.2 \mu\text{s}$. During the prethermal regime, the conservation law is exceptionally well preserved, with the fraction of total particle number above 95%, which drops to 85% after the entire time evolution at $2 \mu\text{s}$. The error mitigation further suppresses the heating process.

In panel **b**, we present the results for a smaller multipolar order and faster driving, where RMD-induced heating occurs swiftly. Similarly, the prethermal regime retains around 95% of the particle number, but more particle loss appears at the very late stage of the experiment, dropping to 70% around $1 \mu\text{s}$. The comparison between panels **a** and **b** potentially suggests that quantum thermalization within the low energy levels accelerates the particle loss to higher energy levels. However, it is worth highlighting that the accelerated particle loss indeed only weakly affects the imbalance decay induced by the RMD protocol. This becomes manifest by comparing the bare (red) with the corrected data (blue), where error mitigation only introduces marginal changes, especially at late times where the system already heats up to the infinite temperature. The key reason for this observation lies in the fact that the driving-induced heating rate is much faster than the particle loss rate of the system. This has been systematically explored in our previous theoretical work [30], where we have explicitly studied high-energy excitations by numerically simulating the two-band Bose-Hubbard systems. A clear separation between these two rates protects the experimental observation of the prethermalization and its characteristic lifetime scaling. In addition, when we extract prethermal lifetime, we only use the data within $0.2\text{--}1.5 \mu\text{s}$ to further reduce the influence of the particle loss. See details in S6.

With these evidences, we confirm that the particle number conservation is well preserved in our experiment. Effects induced by particle loss can be further suppressed using the depolarizing error mitigation scheme.

S6. FITTING DETAILS AND PRECISION

For a single random drive sequence, we extract the lifetime τ_I from the decay of imbalance according to the exponential function $I \sim e^{-\frac{t}{\tau_I}}$. Equivalently, the lifetime can be determined by performing linear regression on $\ln(I)$ versus the evolution time t . Similarly, the lifetime τ_S obtained from the entropy dynamics is fitted according to the function: $S = A - Be^{-at}$. We focus on the time domain $0.2\text{--}1.5 \mu\text{s}$ since for longer times, particle loss may influence the accuracy of the fitting. These fitting



Supplementary Fig. 24. **Imbalance and fraction of the particle number in 78 qubits.** Depolarizing error mitigation scheme is used to correct the contribution from the particle loss.

parameters can be obtained by optimization, and we restrict the optimization parameters within the interval

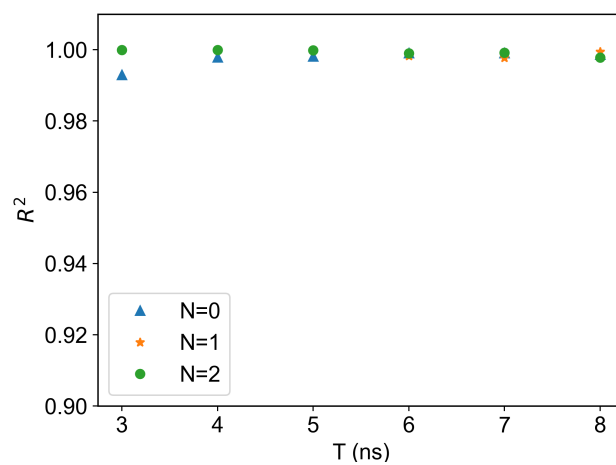
$$\begin{cases} A & \in (0.9 \ln 2, 1.99 \ln 2) \\ a & \in (0, \infty) \\ B & \in (0, 1.99 \ln 2) \end{cases}, \quad (55)$$

with initial values $(A, a, B) = (1.6 \ln 2, 1, \ln 2)$. The lifetime can be obtained accordingly, $\tau_S = 1/a$. The coefficient of determination (R^2) can be used to assess the fitting accuracy,

$$R^2 = 1 - \frac{\sum_i^n (y_{i,\text{exp}} - y_{i,\text{fit}})^2}{\sum_i^n (y_{i,\text{exp}} - y_{\text{avg}})^2}. \quad (56)$$

As shown in Supplementary Fig. 25, for different driving parameters, the value of R^2 exceeds 0.99, demonstrating exceptional fitting accuracy in our experiment.

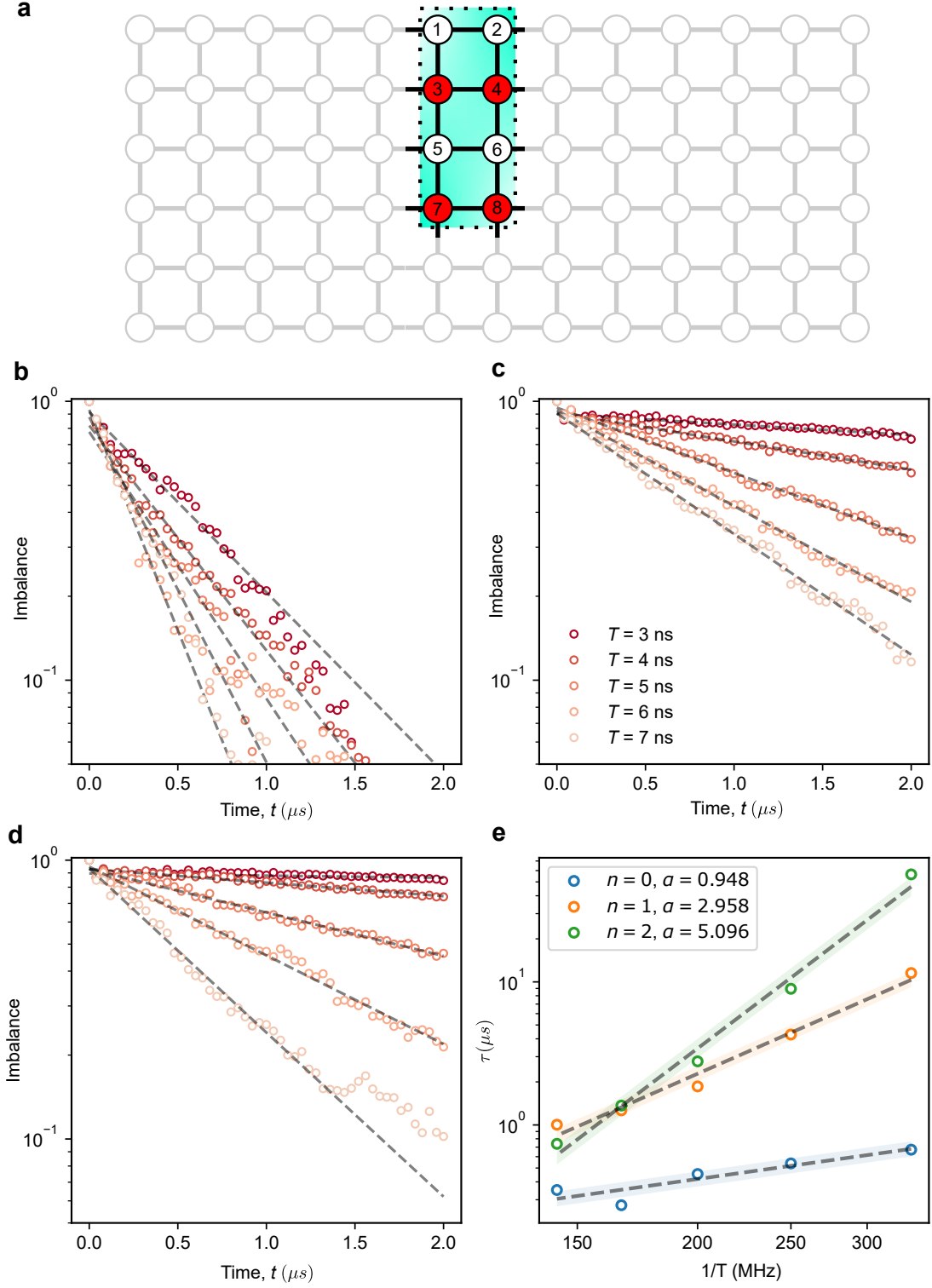
We have performed 10 independent random realizations and obtained the averaged lifetimes as shown in Fig. 3c and d, where the errorbar corresponds to their standard deviation. The averaged lifetime should not depend on the fine-tuning of the fitting parameters. In fact, as we have already shown in Supplementary Figs. 26–32, the resulting fitted functions agree well with the experimentally observed time evolution of the imbalance and the entanglement entropy.



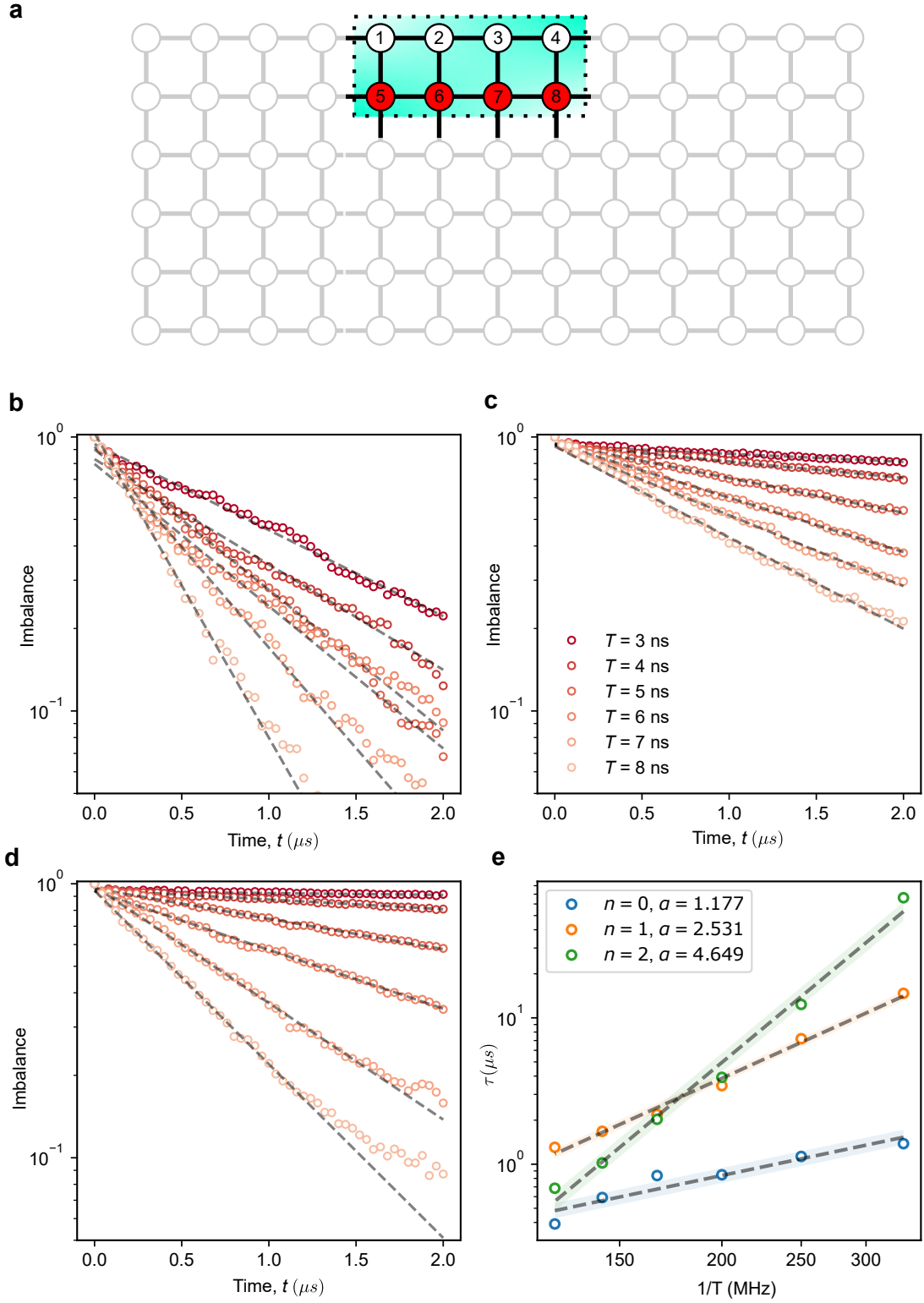
Supplementary Fig. 25. R^2 of different parameter setups in 78 qubits.

S7. SUPPLEMENTARY DATA

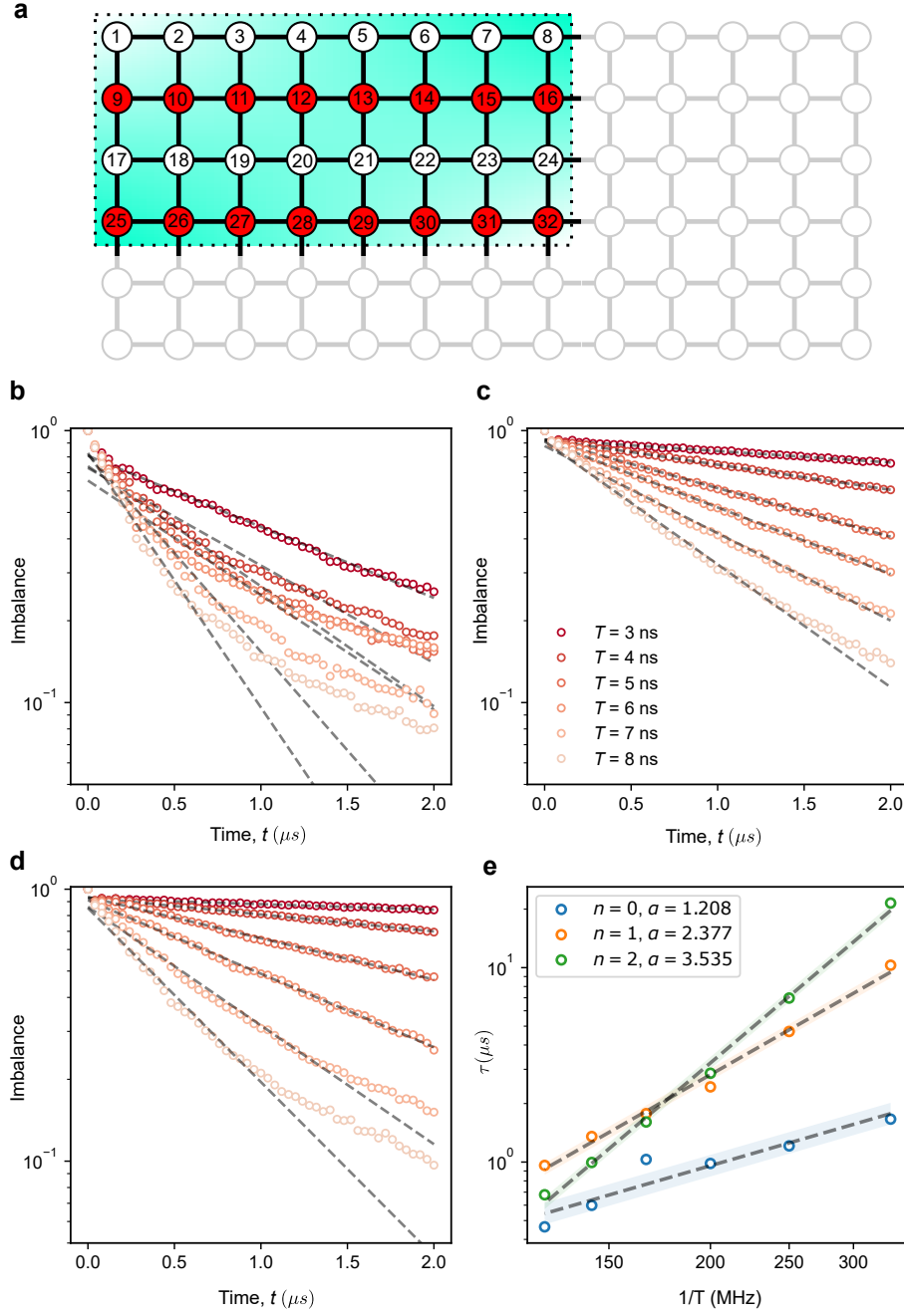
A. Dynamics of imbalance in systems with different sizes



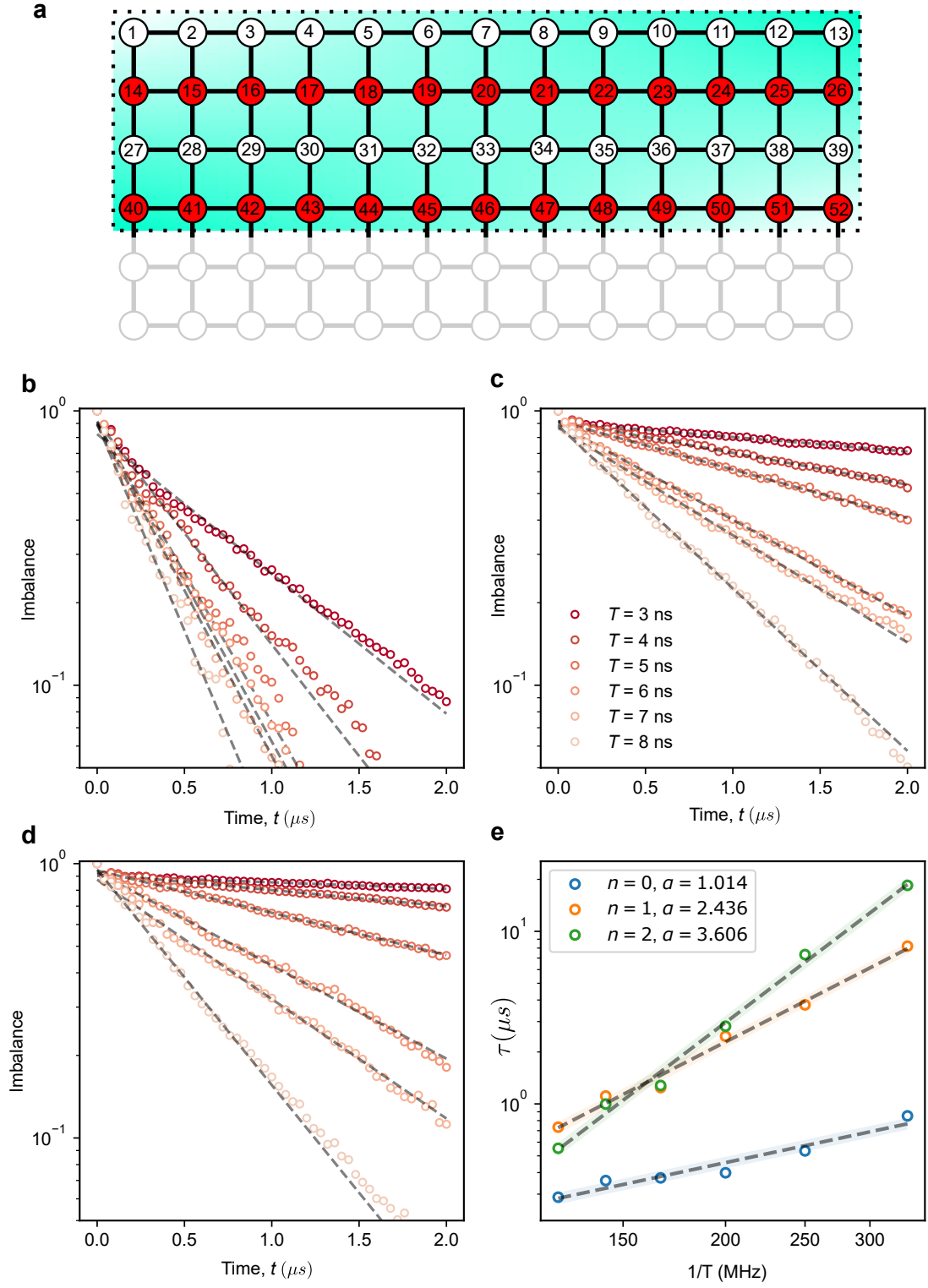
Supplementary Fig. 26. **Dynamics of imbalance in the system of size 2×4 .** **a**, Spatial configuration of the 2×4 system within the processor. **b–d**, Time evolution of the imbalance for driving periods T ranging from 3 ns to 7 ns, corresponding to $n = 0, n = 1, n = 2$, individually. Markers represent experimental data, while dashed lines denote fitted curves used to extract the decay rate $-1/\tau_I$, where τ_I is the characteristic decay time. **e**, Power-law scaling of τ_I , presented on a log-log scale. The scaling exponents, $\alpha(n = 0) = 0.948$, $\alpha(n = 1) = 2.958$, $\alpha(n = 2) = 5.096$, follow the theoretical prediction, $\alpha = 2n + 1$. The results are averaged over 10 independent RMD sequences.



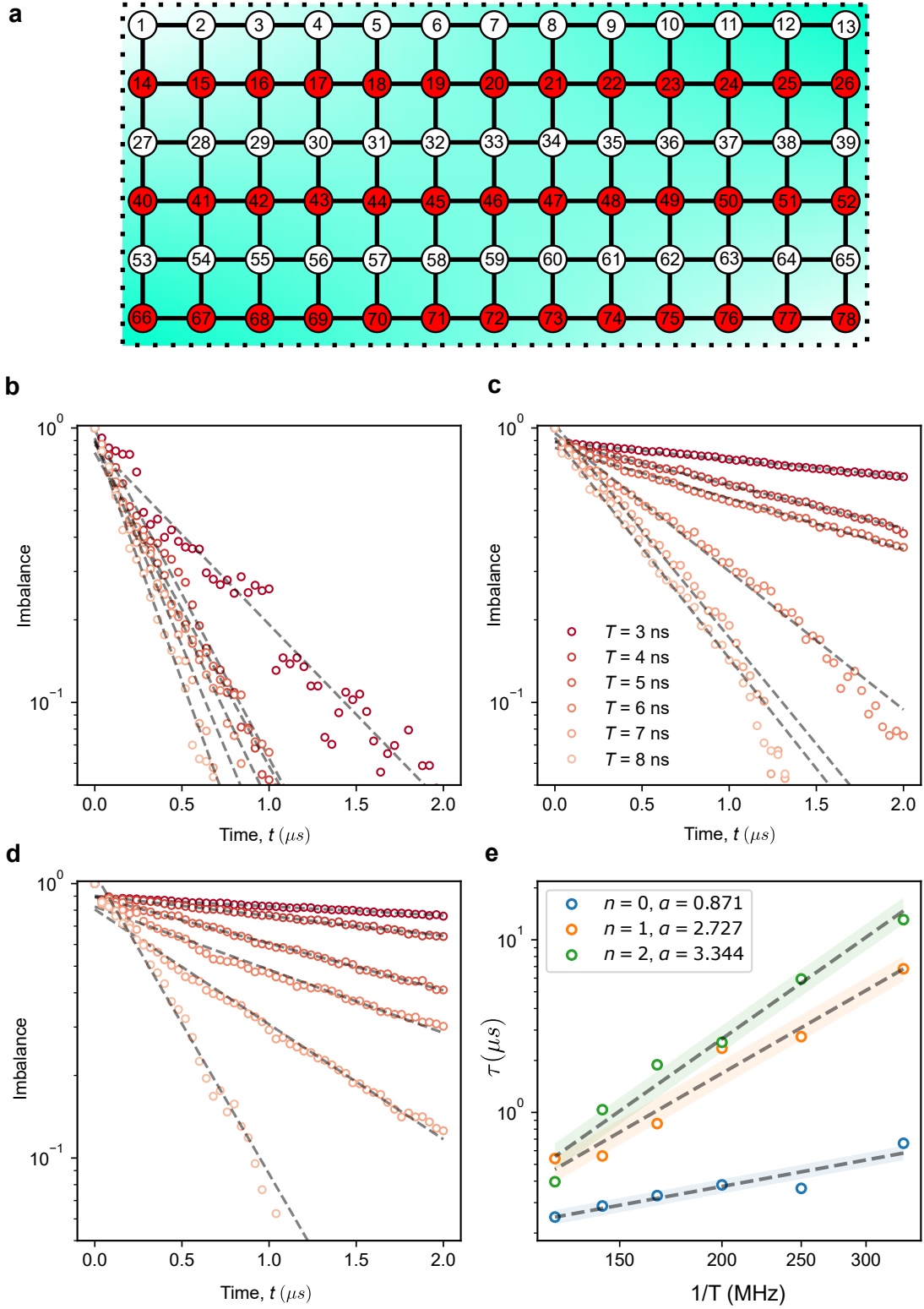
Supplementary Fig. 27. **Dynamics of imbalance in the system of size 4×2 .** **a**, Spatial configuration of the 4×2 system within the processor. **b–d**, Time evolution of the imbalance for driving periods T ranging from 3 ns to 8 ns, corresponding to $n = 0, n = 1, n = 2$, individually. Markers represent experimental data, while dashed lines denote fitted curves used to extract the decay rate $-1/\tau_I$, where τ_I is the characteristic decay time. **e**, Power-law scaling of τ_I in a log-log scale, with extracted exponents $\alpha(n = 0) = 1.177, \alpha(n = 1) = 2.531, \alpha(n = 2) = 4.649$. The results are averaged over 10 independent RMD sequences.



Supplementary Fig. 28. **Dynamics of imbalance in the system of size 8×4 .** **a**, Spatial configuration of the 8×4 system within the processor. **b–d**, Time evolution of the imbalance for driving periods T ranging from 3 ns to 8 ns, corresponding to $n = 0$, $n = 1$, $n = 2$, individually. **e**, Power-law scaling of the imbalance decay time, presented in a log-log scale, with extracted exponents $\alpha(n = 0) = 1.208$, $\alpha(n = 1) = 2.377$, $\alpha(n = 2) = 3.535$. The results are averaged over 10 independent RMD sequences.

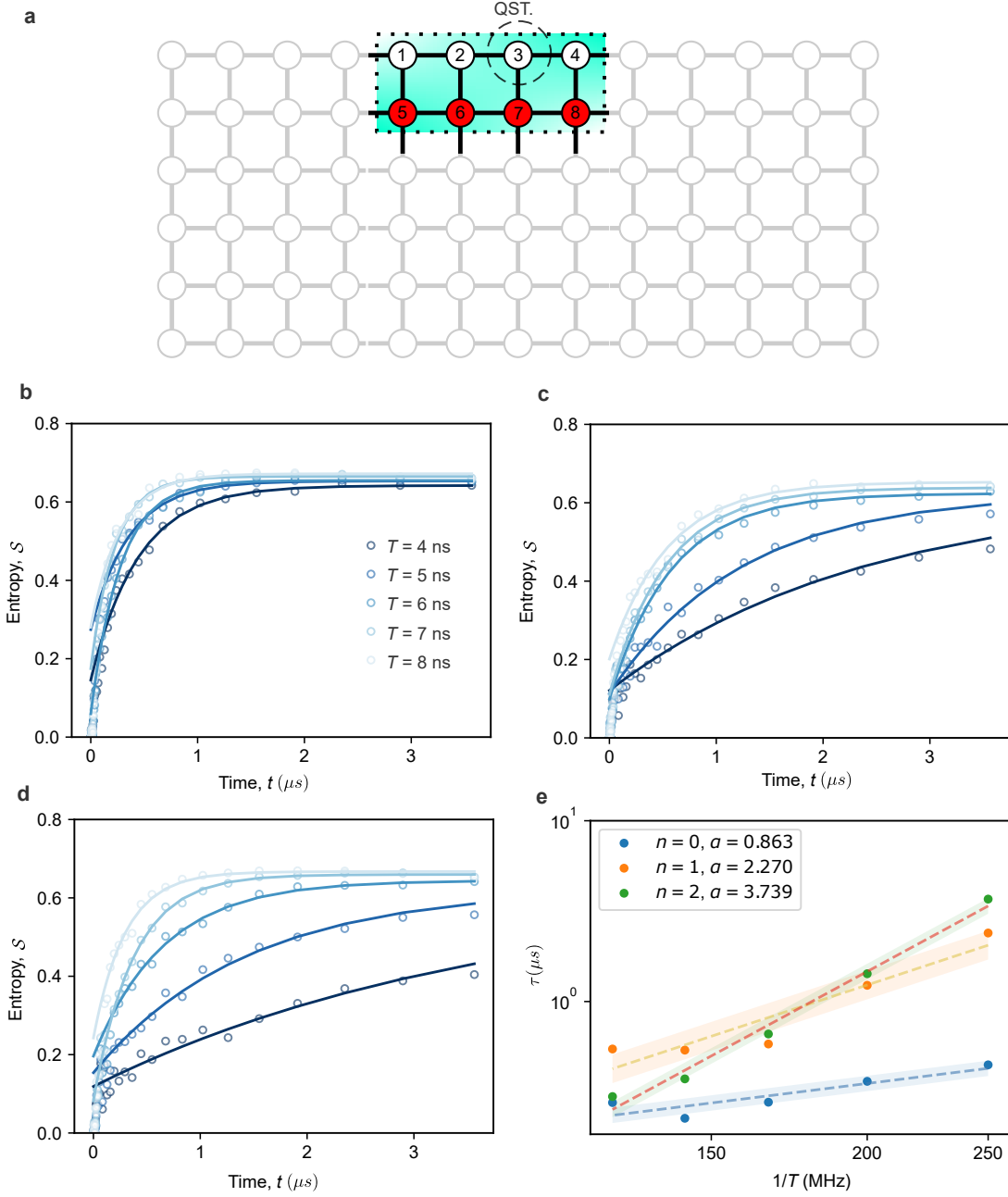


Supplementary Fig. 29. **Dynamics of imbalance in the system of size 13×4 .** **a**, Spatial configuration of the 13×4 system within the processor. **b–d**, Time evolution of the imbalance for driving periods T ranging from 3 ns to 8 ns, corresponding to $n = 0$, $n = 1$, $n = 2$, individually. Markers represent experimental data, while dashed lines denote fitted curves used to extract the decay rate $-1/\tau_I$, where τ_I is the characteristic decay time. **e**, Power-law scaling of τ_I , presented on a log-log scale, with extracted exponents $\alpha(n = 0) = 1.014$, $\alpha(n = 1) = 2.436$, $\alpha(n = 2) = 3.606$. The results are averaged over 10 independent RMD sequences.

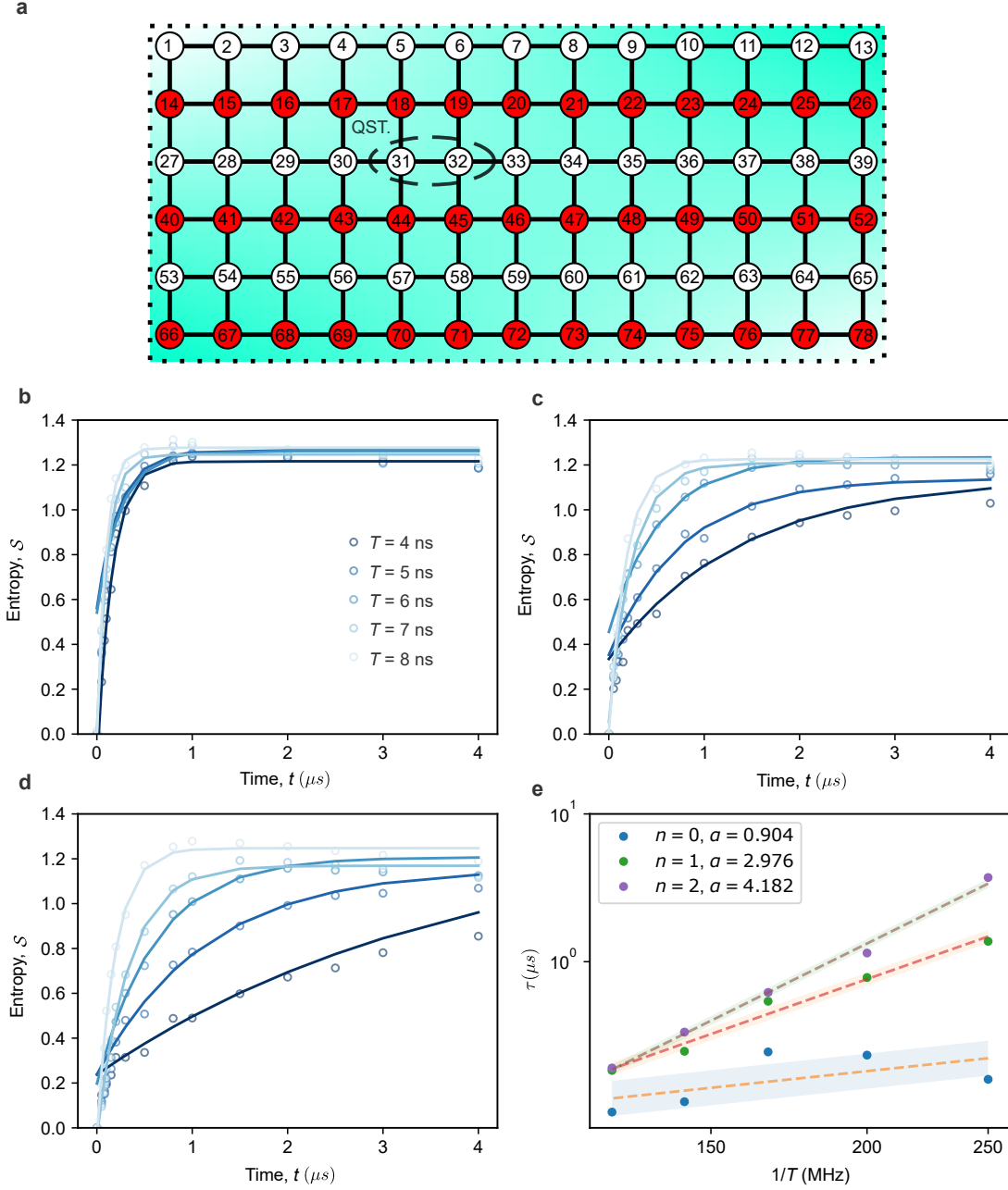


Supplementary Fig. 30. **Dynamics of imbalance in the system of size 13×6 .** **a**, Spatial configuration of the 13×6 system within the processor. **b–d**, Time evolution of the imbalance for driving periods T ranging from 3 ns to 8 ns, corresponding to $n = 0$, $n = 1$, $n = 2$, individually. Markers represent experimental data, while dashed lines denote fitted curves used to extract the decay rate $-1/\tau_I$, where τ_I is the characteristic decay time. **e**, Power-law scaling of τ_I , presented on a log-log scale, with extracted exponents $\alpha(n = 0) = 0.871$, $\alpha(n = 1) = 2.727$, $\alpha(n = 2) = 3.344$. The results are averaged over 10 independent RMD sequences.

B. Dynamics of entropy in systems with different sizes



Supplementary Fig. 31. **Dynamics of entropy in the system of size 4×2 .** **a**, Spatial configuration of the 4×2 system and the subsystem for performing quantum state tomography within the processor. **b–d**, Time evolution of the subsystem entanglement entropy for driving periods T ranging from 4 ns to 8 ns, corresponding to $n = 0$, $n = 1$, $n = 2$, individually. **e**, Power-law scaling of the prethermal lifetime presented on a log-log scale, with extracted exponents $\alpha(n=0) = 0.863$, $\alpha(n=1) = 2.270$, $\alpha(n=2) = 3.739$. The curves in **b–d** are the fitted results using the form $S \sim S_M(1 - e^{-t/\tau_S})$. The results are averaged over 10 independent RMD sequences.

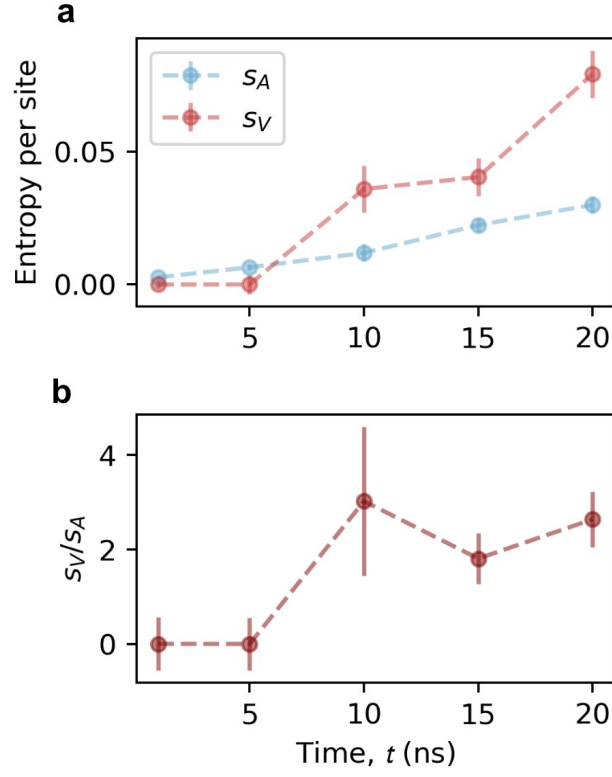


Supplementary Fig. 32. **Dynamics of entropy in the system of size 13×6 .** **a**, Spatial configuration of the 13×6 system and the subsystem for performing quantum state tomography (QST) within the processor. **b–d**, Time evolution of the subsystem entanglement entropy for driving periods T ranging from 4 ns to 8 ns, corresponding to $n = 0$, $n = 1$, $n = 2$, individually. **e**, Power-law scaling of the prethermal lifetime in a log-log scale, with extracted exponents $\alpha(n = 0) = 0.904$, $\alpha(n = 1) = 2.976$, $\alpha(n = 2) = 4.182$. The curves in **b–d** are the fitted results using the form $S \sim S_M(1 - e^{-t/\tau_S})$. The results are averaged over 10 independent RMD sequences.

C. Area law to volume law crossover

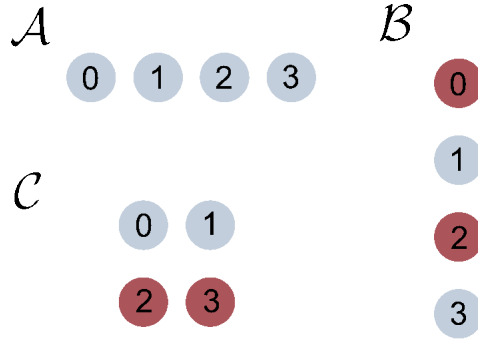
In our experimental setup, the initial state is prepared as a density wave ordered product state, which should have zero entanglement entropy if the initial state preparation is perfect. However, due to experimental noise and readout error, we observe a small but finite value, (0.0283, 0.0175, 0.0084) for subsystems ($\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$) in Fig. 4a, and the area-law scaling is obeyed.

Subject to the drive, the many-body system indeed quickly develops long-range correlations throughout the entire system, leading to the noticeable entanglement growth as shown in Fig. 4a. In fact, according to the eigenstate thermalization hypothesis (ETH), within the prethermal plateau, the system locally equilibrates for which the entanglement entropy should follow a volume



Supplementary Fig. 33. **Fitted values of s_A and s_V at early time.** The comparison clearly distinguishes different contributions to the entanglement entropy. The cross-over from area-law to volume-law occurs around 5–10 ns.

law for generic non-integrable systems. Indeed, the ratio s_V/s_A already reaches a value of order 1 in Fig. 4c, suggesting that volume-law entangled states have been generated. Considering the fact that the initial state is a product state, the cross-over from area-law to volume law scaling should occur during the short transient before the prethermal regime. We confirm this below in Supplementary Fig. 33, where the crossover of the value of s_A and s_V is visible at early times around 5–10 ns.



Supplementary Fig. 34. **Schematic of configurations with indices.**

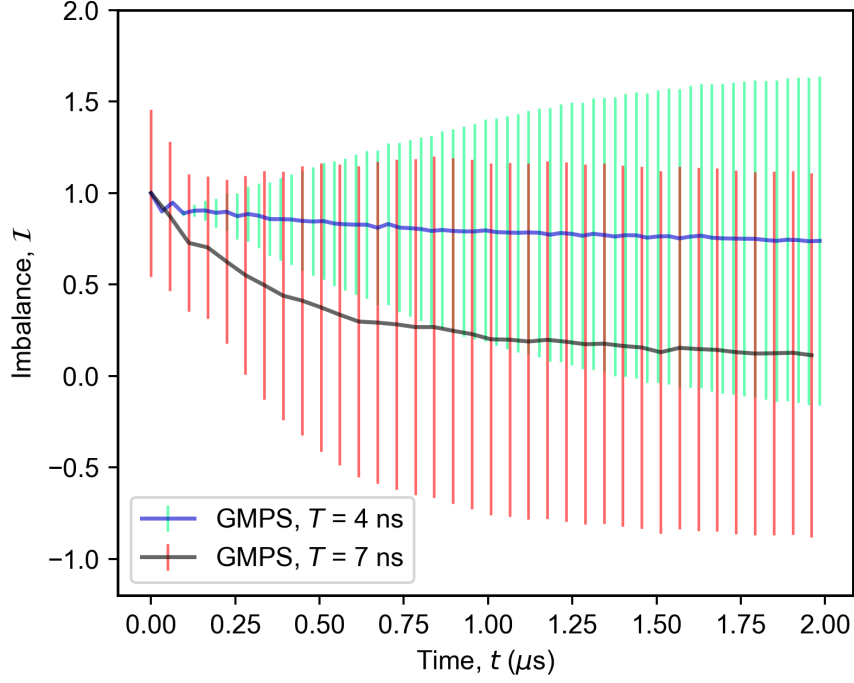
	Indices of $\mathcal{A}$	Indices of $\mathcal{B}$	Indices of $\mathcal{C}$
Volume=1	1	0	0
Volume=2	0, 1	0, 1	0, 1
Volume=3	0, 1, 2	0, 1, 2	0, 1, 2
Volume=4	0, 1, 2, 3	0, 1, 2, 3	0, 1, 2, 3

Supplementary Table I. **Subsystem configurations for entanglement entropy calculation.**

D. Subsystems sizes for entropy scaling

The configurations investigated for the entanglement scaling are listed in increasing order by volume (1 to 4) in Supplementary Fig. 34 and Supplementary Table I. Labels $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ correspond to Fig. 4a, with index values ranging from 0 to 3.

E. Estimated error in local observables in GMPS simulation

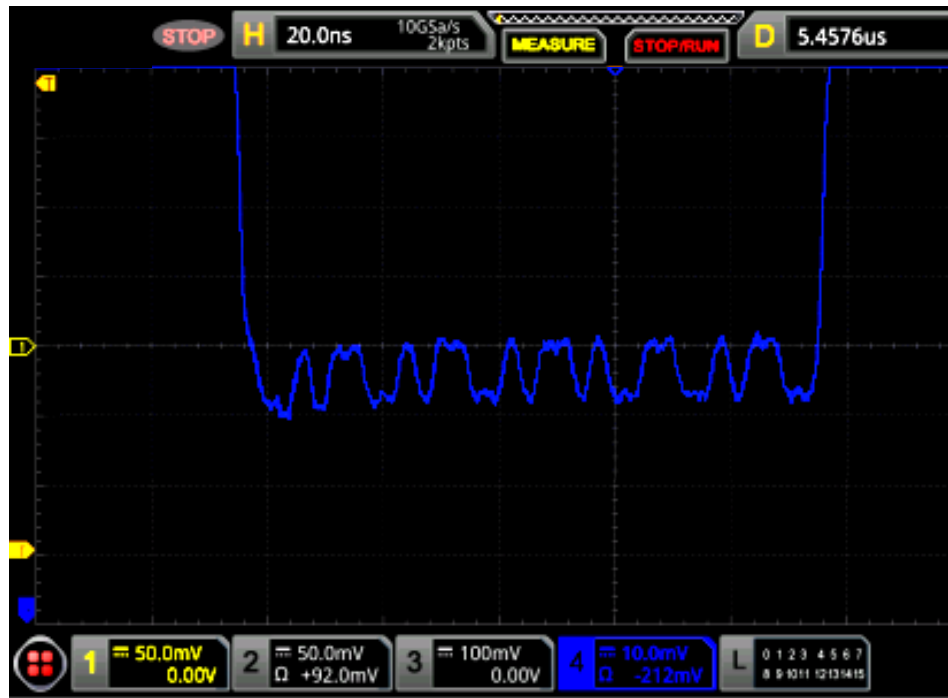


Supplementary Fig. 35. **Estimated error in local observables in GMPS simulation.** We estimate the error bar from the state infidelity $\sqrt{1 - F^2}$. The value of the error bar quickly increases to the order of 1, suggesting that the numerical simulation is unreliable when heating happens.

In Supplementary Fig. 35, we show the systematic error in local observables, estimated based on the state infidelity in the GMPS simulation, using methods provided in Ref. [68] Sec. 9.2.3. Due to the entanglement growth during the heating process, the state fidelity drops over time, and the corresponding error estimation in local observables quickly increases to the order of 1. Therefore, in the absence of a quantitative understanding of the systematic errors involved, extraction of late-time observables is unreliable, even though their behavior might follow a similar trend as our experimental results.

F. Waveform inspection

We have checked the waveform accuracy from oscilloscope measurements in Supplementary Fig. 36.



Supplementary Fig. 36. Waveforms obtained from oscilloscope measurements.

-
- [1] P. Titum, E. Berg, M. S. Rudner, G. Refael, and N. H. Lindner, Anomalous Floquet-Anderson insulator as a nonadiabatic quantized charge pump, *Phys. Rev. X* **6**, 021013 (2016).
 - [2] A. Eckardt, Colloquium: Atomic quantum gases in periodically driven optical lattices, *Rev. Mod. Phys.* **89**, 011004 (2017).
 - [3] S. Choi, J. Choi, R. Landig, G. Kucsko, H. Zhou, J. Isoya, F. Jelezko, S. Onoda, H. Sumiya, V. Khemani, C. von Keyserlingk, N. Y. Yao, E. Demler, and M. D. Lukin, Observation of discrete time-crystalline order in a disordered dipolar many-body system, *Nature* **543**, 221–225 (2017).
 - [4] J. Rovny, R. L. Blum, and S. E. Barrett, Observation of discrete-time-crystal signatures in an ordered dipolar many-body system, *Phys. Rev. Lett.* **120**, 180603 (2018).
 - [5] M. Gring, M. Kuhnert, T. Langen, T. Kitagawa, B. Rauer, M. Schreitl, I. Mazets, D. A. Smith, E. Demler, and J. Schmiedmayer, Relaxation and prethermalization in an isolated quantum system, *Science* **337**, 1318–1322 (2012).
 - [6] K. Mallayya, M. Rigol, and W. De Roeck, Prethermalization and thermalization in isolated quantum systems, *Phys. Rev. X* **9**, 021027 (2019).
 - [7] M. S. Rudner and N. H. Lindner, Band structure engineering and non-equilibrium dynamics in Floquet topological insulators, *Nat. Rev. Phys.* **2**, 229–244 (2020).
 - [8] M. Ueda, Quantum equilibration, thermalization and prethermalization in ultracold atoms, *Nat. Rev. Phys.* **2**, 669–681 (2020).
 - [9] K. Decker, C. Karrasch, J. Eisert, and D. Kennes, Floquet engineering topological many-body localized systems, *Phys. Rev. Lett.* **124**, 190601 (2020).
 - [10] P. Peng, C. Yin, X. Huang, C. Ramanathan, and P. Cappellaro, Floquet prethermalization in dipolar spin chains, *Nat. Phys.* **17**, 444–447 (2021).
 - [11] J. Randall, C. E. Bradley, F. V. van der Gronden, A. Galicia, M. H. Abobeih, M. Markham, D. J. Twitchen, F. Machado, N. Y. Yao, and T. H. Taminiau, Many-bodylocalized discrete time crystal with a programmable spin-based quantum simulator, *Science* **374**, 1474–1478 (2021).
 - [12] X. Zhang, W. Jiang, J. Deng, *et al.*, Digital quantum simulation of Floquet symmetry-protected topological phases, *Nature* **607**, 468–473 (2022).
 - [13] J. Naji, R. Jafari, L. Zhou, and A. Langari, Engineering Floquet dynamical quantum phase transitions, *Phys. Rev. B* **106**, 094314 (2022).
 - [14] M. P. Zaletel, M. Lukin, C. Monroe, C. Nayak, F. Wilczek, and N. Y. Yao, Colloquium : Quantum and classical discrete time crystals, *Rev. Mod. Phys.* **95**, 031001 (2023).
 - [15] W. Beatriz, C. Fleckenstein, A. Pillai, E. de Leon Sanchez, A. Akkiraju, J. Diaz Alcala, S. Conti, P. Reshetikhin, E. Druga, M. Bukov, and A. Ajoy, Critical prethermal discrete time crystal created by two-frequency driving, *Nat. Phys.* **19**, 407–413 (2023).
 - [16] G. He, B. Ye, R. Gong, Z. Liu, K. W. Murch, N. Y. Yao, and C. Zu, Quasi-Floquet prethermalization in a disordered dipolar spin ensemble in diamond, *Phys. Rev. Lett.* **131**, 130401 (2023).

- [17] W. W. Ho, T. Mori, D. A. Abanin, and E. G. Dalla Torre, Quantum and classical Floquet prethermalization, *Ann. Phys.* **454**, 169297 (2023).
- [18] K. A. Harkins, C. Selco, C. Bengs, D. Marchiori, L. J. I. Moon, Z.-R. Zhang, A. Yang, A. Singh, E. Druga, Y.-Q. Song, and A. Ajoy, Anomalous extended Floquet prethermal lifetimes and applications to long-time quantum sensing, *arXiv:2410.09028*.
- [19] J. Kwan, P. Segura, Y. Li, S. Kim, A. V. Gorshkov, A. Eckardt, B. Bakkali-Hassani, and M. Greiner, Realization of one-dimensional anyons with arbitrary statistical phase, *Science* **386**, 1055–1060 (2024).
- [20] C. Wang, F.-M. Liu, M.-C. Chen, H. Chen, X.-H. Zhao, C. Ying, Z.-X. Shang, J.-W. Wang, Y.-H. Huo, C.-Z. Peng, X. Zhu, C.-Y. Lu, and J.-W. Pan, Realization of fractional quantum Hall state with interacting photons, *Science* **384**, 579–584 (2024).
- [21] Z. Fu, R. Moessner, H. Zhao, and M. Bukov, Engineering hierarchical symmetries, *Phys. Rev. X* **14**, 041070 (2024).
- [22] L. DAlessio and M. Rigol, Long-time behavior of isolated periodically driven interacting lattice systems, *Phys. Rev. X* **4**, 041048 (2014).
- [23] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, Colloquium : Many-body localization, thermalization, and entanglement, *Rev. Mod. Phys.* **91**, 021001 (2019).
- [24] T. Mori, T. Kuwahara, and K. Saito, Rigorous bound on energy absorption and generic relaxation in periodically driven quantum systems, *Phys. Rev. Lett.* **116**, 120401 (2016).
- [25] D. A. Abanin, W. De Roeck, and F. m. c. Huveneers, Exponentially slow heating in periodically driven many-body systems, *Phys. Rev. Lett.* **115**, 256803 (2015).
- [26] H. Zhao, F. Mintert, R. Moessner, and J. Knolle, Random multipolar driving: Tunably slow heating through spectral engineering, *Phys. Rev. Lett.* **126**, 040601 (2021).
- [27] E. Boyers, P. J. Crowley, A. Chandran, and A. O. Sushkov, Exploring 2D synthetic quantum Hall physics with a quasiperiodically driven qubit, *Phys. Rev. Lett.* **125**, 160505 (2020).
- [28] C. I. Timms, L. M. Sieberer, and M. H. Kolodrubetz, Quantized floquet topology with temporal noise, *Phys. Rev. Lett.* **127**, 270601 (2021).
- [29] D. M. Long, P. J. Crowley, and A. Chandran, Nonadiabatic topological energy pumps with quasiperiodic driving, *Phys. Rev. Lett.* **126**, 106805 (2021).
- [30] H. Zhao, J. Knolle, R. Moessner, and F. Mintert, Suppression of interband heating for random driving, *Phys. Rev. Lett.* **129**, 120605 (2022).
- [31] G. Guarneri, M. T. Mitchison, A. Purkayastha, D. Jaksch, B. Buča, and J. Goold, Time periodicity from randomness in quantum systems, *Phys. Rev. A* **106**, 022209 (2022).
- [32] S. Kumar and S. Choudhury, Prethermalization in aperiodically driven classical spin systems, *Phys. Rev. E* **110**, 064150 (2024).
- [33] Z. Ma, J. Yan, H. Zhao, and L.-Y. Peng, Stable time rondeau crystals in dissipative many-body systems, *Phys. Rev. B* **112**, L140301 (2025).
- [34] J. Wu, C. Liu, D. Bulmash, and W. W. Ho, Geometric quantum drives: Hyperbolically driven quantum systems and beyond, *arXiv:2503.08242*.
- [35] S. Pilatowsky-Cameo, S. Choi, and W. W. Ho, Critically slow hilbert-space ergodicity in quantum morphic drives, *Phys. Rev. Lett.* **135**, 140402 (2025).
- [36] T. Mori, H. Zhao, F. Mintert, J. Knolle, and R. Moessner, Rigorous bounds on the heating rate in Thue-Morse quasiperiodically and randomly driven quantum many-body systems, *Phys. Rev. Lett.* **127**, 050602 (2021).
- [37] F. Pan and P. Zhang, Simulation of quantum circuits using the big-batch tensor network method, *Phys. Rev. Lett.* **128**, 030501 (2022).
- [38] F. Pan, K. Chen, and P. Zhang, Solving the sampling problem of the Sycamore quantum circuits, *Phys. Rev. Lett.* **129**, 090502 (2022).
- [39] S. Paegel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, and C. Hubig, Time-evolution methods for matrix-product states, *Ann. Phys.* **411**, 167998 (2019).
- [40] F. Verstraete and J. I. Cirac, Renormalization algorithms for quantum-many body systems in two and higher dimensions, *arXiv:cond-mat/0407066*.
- [41] F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, Criticality, the Area Law, and the Computational Power of Projected Entangled Pair States, *Phys. Rev. Lett.* **96**, 220601 (2006).
- [42] H.-J. Liao, K. Wang, Z.-S. Zhou, P. Zhang, and T. Xiang, Simulation of IBM's kicked Ising experiment with Projected Entangled Pair Operator, *arXiv:2308.03082*.
- [43] H. C. Jiang, Z. Y. Weng, and T. Xiang, Accurate Determination of Tensor Network State of Quantum Lattice Models in Two Dimensions, *Phys. Rev. Lett.* **101**, 090603 (2008).
- [44] D.-W. Qu, W. Li, and T. Xiang, Thermal tensor network simulations of the Heisenberg model on the Bethe lattice, *Phys. Rev. B* **100**, 125121 (2019).
- [45] G. Vidal, Efficient Simulation of One-Dimensional Quantum Many-Body Systems, *Phys. Rev. Lett.* **93**, 040502 (2004).
- [46] E. M. Stoudenmire and S. R. White, Minimally entangled typical thermal state algorithms, *New J. Phys.* **12**, 055026 (2010).
- [47] M. Fishman, S. R. White, and E. M. Stoudenmire, The ITensor Software Library for Tensor Network Calculations, *SciPost Phys. Codebases*, **4** (2022).
- [48] T. Begušić, J. Gray, and G. K.-L. Chan, Fast and converged classical simulations of evidence for the utility of quantum computing before fault tolerance, *Science Advances* **10**, eadk4321 (2024).
- [49] M. S. Rudolph, T. Jones, Y. Teng, A. Angrisani, and Z. Holmes, Pauli propagation: A computational framework for simulating quantum systems, *arXiv:2505.21606*.
- [50] D. Wu, R. Rossi, F. Vicentini, N. Astrakhantsev, F. Becca, X. Cao, J. Carrasquilla, F. Ferrari, A. Georges, M. Hibat-Allah, M. Imada, A. M. Läuchli, G. Mazzola, A. Mezzacapo, A. Millis, J. R. Moreno, T. Neupert, Y. Nomura, J. Nys, O. Parcollet, R. Pohle, I. Romero, M. Schmid, J. M. Silvester, S. Sorella, L. F. Tocchio, L. Wang, S. R. White, A. Wietek, Q. Yang, Y. Yang, S. Zhang, and G. Carleo, Variational benchmarks for quantum many-body problems, *Science* **386**, 296–301 (2024).
- [51] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, *Science* **355**, 602–606 (2017).

- [52] M. Schmitt and M. Heyl, Quantum many-body dynamics in two dimensions with artificial neural networks, *Phys. Rev. Lett.* **125**, 100503 (2020).
- [53] F. Vicentini, D. Hofmann, A. Szabó, D. Wu, C. Roth, C. Giuliani, G. Pescia, J. Nys, V. Vargas-Calderón, N. Astrakhantsev, and G. Carleo, NetKet 3: Machine Learning Toolbox for Many-Body Quantum Systems, *SciPost Phys. Codebases*, **7** (2022).
- [54] R. Haghshenas, E. Chertkov, M. Mills, W. Kadow, S.-H. Lin, Y.-H. Chen, C. Cade, I. Niesen, T. Begušić, M. S. Rudolph, *et al.*, Digital quantum magnetism at the frontier of classical simulations, [arXiv:2503.20870](https://arxiv.org/abs/2503.20870).
- [55] F. Arute, K. Arya, R. Babbush, *et al.*, Quantum supremacy using a programmable superconducting processor, *Nature* **574**, 505–510 (2019).
- [56] G.-H. Liang, X.-H. Song, C.-L. Deng, X.-Y. Gu, Y. Yan, Z.-Y. Mei, S.-L. Zhao, Y.-Z. Bu, Y.-X. Xiao, Y.-H. Yu, M.-C. Wang, T. Liu, Y.-H. Shi, H. Zhang, X. Li, L. Li, J.-Z. Wang, Y. Tian, S.-P. Zhao, K. Xu, H. Fan, Z.-C. Xiang, and D.-N. Zheng, Tunable-coupling architectures with capacitively connecting pads for large-scale superconducting multiqubit processors, *Phys. Rev. Appl.* **20**, 044028 (2023).
- [57] Z. Yan, Y.-R. Zhang, M. Gong, Y. Wu, Y. Zheng, S. Li, C. Wang, F. Liang, J. Lin, Y. Xu, C. Guo, L. Sun, C.-Z. Peng, K. Xia, H. Deng, H. Rong, J. Q. You, F. Nori, H. Fan, X. Zhu, and J.-W. Pan, Strongly correlated quantum walks with a 12-qubit superconducting processor, *Science* **364**, 753–756 (2019).
- [58] T.-M. Li, J.-C. Zhang, B.-J. Chen, K. Huang, H.-T. Liu, Y.-X. Xiao, C.-L. Deng, G.-H. Liang, C.-T. Chen, Y. Liu, H. Li, Z.-T. Bao, K. Zhao, Y. Xu, L. Li, Y. He, Z.-H. Liu, Y.-H. Yu, S.-Y. Zhou, Y.-J. Liu, X. Song, D. Zheng, Z. Xiang, Y.-H. Shi, K. Xu, and H. Fan, High-precision pulse calibration of tunable couplers for high-fidelity two-qubit gates in superconducting quantum processors, *Phys. Rev. Appl.* **23**, 024059 (2025).
- [59] Y.-H. Shi, Y. Liu, Y.-R. Zhang, Z. Xiang, K. Huang, T. Liu, Y.-Y. Wang, J.-C. Zhang, C.-L. Deng, G.-H. Liang, Z.-Y. Mei, H. Li, T.-M. Li, W.-G. Ma, H.-T. Liu, C.-T. Chen, T. Liu, Y. Tian, X. Song, S. Zhao, K. Xu, D. Zheng, F. Nori, and H. Fan, Quantum simulation of topological zero modes on a 41-qubit superconducting processor, *Phys. Rev. Lett.* **131**, 080401 (2023).
- [60] Q. Guo, C. Cheng, Z.-H. Sun, Z. Song, H. Li, Z. Wang, W. Ren, H. Dong, D. Zheng, Y.-R. Zhang, R. Mondaini, H. Fan, and H. Wang, Observation of energy-resolved many-body localization, *Nat. Phys.* **17**, 234–239 (2021).
- [61] Y. Liu, Y.-R. Zhang, Y.-H. Shi, T. Liu, C. Li, Y.-Y. Wang, H. Li, T. Li, C.-L. Deng, S.-Y. Zhou, T. Liu, J.-C. Zhang, G.-H. Liang, Z.-Y. Mei, W.-G. Ma, H.-T. Liu, Z.-H. Liu, C.-T. Chen, K. Huang, X. Song, S. P. Zhao, Y. Tian, Z. Xiang, D. Zheng, F. Nori, K. Xu, and H. Fan, Interplay between disorder and topology in Thouless pumping on a superconducting quantum processor, *Nat. Commun.* **16**, 108 (2025).
- [62] K. Xu, J.-J. Chen, Y. Zeng, Y.-R. Zhang, C. Song, W. Liu, Q. Guo, P. Zhang, D. Xu, H. Deng, K. Huang, H. Wang, X. Zhu, D. Zheng, and H. Fan, Emulating many-body localization with a superconducting quantum processor, *Phys. Rev. Lett.* **120**, 050507 (2018).
- [63] X. Li, Y. Ma, J. Han, T. Chen, Y. Xu, W. Cai, H. Wang, Y. Song, Z.-Y. Xue, Z.-q. Yin, and L. Sun, Perfect quantum state transfer in a superconducting qubit chain with parametrically tunable couplings, *Phys. Rev. Appl.* **10**, 054009 (2018).
- [64] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, A quantum engineers guide to superconducting qubits, *Appl. Phys. Rev.* **6**, 021318 (2019).
- [65] Y.-H. Shi, R.-Q. Yang, Z. Xiang, Z.-Y. Ge, H. Li, Y.-Y. Wang, K. Huang, Y. Tian, X. Song, D. Zheng, K. Xu, R.-G. Cai, and H. Fan, Quantum simulation of Hawking radiation and curved spacetime with a superconducting on-chip black hole, *Nature Communications* **14**, 3263 (2023).
- [66] F. Machado, G. D. Kahanamoku-Meyer, D. V. Else, C. Nayak, and N. Y. Yao, Exponentially slow heating in short and long-range interacting Floquet systems, *Phys. Rev. Res.* **1**, 033202 (2019).
- [67] K. Mallayya and M. Rigol, Heating rates in periodically driven strongly interacting quantum many-body systems, *Phys. Rev. Lett.* **123**, 240603 (2019).
- [68] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, 2010).
- [69] J. Vovrosh, K. E. Khosla, S. Greenaway, C. Self, M. S. Kim, and J. Knolle, Simple mitigation of global depolarizing errors in quantum simulations, *Phys. Rev. E* **104**, 035309 (2021).