

Prethermalization by Random Multipolar Driving on a 78-Qubit Processor

Corresponding Author: Professor Hongzheng Zhao

Any redactions in this file are there to maintain patient confidentiality, the confidentiality of unpublished data, or to remove third-party material.

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Referee #1

(Remarks to the Author)

I have reviewed the manuscript "Prethermalization by Random Multipolar Driving on a 78-Qubit Superconducting Processor" and find it to be an interesting experimental contribution to the field of non-equilibrium quantum dynamics. This work reports the experimental observation of long-lived prethermal phases in many-body systems with tunable heating rates, generated by random multipolar driving on a 78-qubit superconducting processor. The authors characterize these phases by monitoring particle imbalance and subsystem entanglement entropy over 1,000 driving cycles, observing a prethermal plateau and claiming a crossover from area-law to volume-law entanglement scaling. The detailed supplemental information regarding the experimental methods and techniques is particularly commendable, significantly aiding in the reproducibility of the work.

However, several aspects of the manuscript require revision and clarification before I can make a recommendation for publication in Nature. My specific comments are detailed below:

Major Concerns:

* "Beyond Classical" Claim: The claim of "quantum advantage" or "beyond classicality" is not sufficiently supported. The authors do not provide any estimates on the fidelity of their analog quantum simulator with Random Multipolar Driving (RMD). Without such fidelity estimates, it is difficult to reasonably assess the validity of claims regarding the beyond-classical nature of their process. In particular, the significant discrepancy between the prethermal lifetime power-law exponent observed for larger lattice sizes (8x4, 13x4, and 13x8) and the theoretically predicted value raises skepticism about the fidelity of the quantum simulator when utilizing the full system. A quantitative discussion of the simulator's fidelity, including systematic errors, especially for larger system sizes and longer evolution times, is essential to substantiate any "beyond classical" claims.

Specific Revisions Required:

Introduction:

* Terminology: On page 2, in systems with particle number conservation, the term "chemical potential" is meaningless. I recommend replacing it with "site energy" or "site potential."

* Quantum Advantage Claim: At the end of the Introduction, the authors claim that the "experiment potentially demonstrates the quantum advantage." This claim is currently too strong, based solely on the observed failure of GMPS and PEPS to simulate the dynamics. To strengthen this assertion, the authors should either significantly revise this claim to reflect a more cautious interpretation or, ideally, bolster it by studying the limitations of other numerical methods, such as Pauli propagation techniques and variational algorithms such as NQS, which have proven successful in classically simulating other many-body systems operating in a similar regime (i.e. Quantinuum's beyond-classical claim <https://arxiv.org/abs/2503.20870>)

Experimental Setup:

* Decoherence Time vs. Relaxation Time: The authors refer to the time of their qubits as "decoherence time." This should be corrected to "relaxation time." Qubit decoherence time (T_2) takes contributions from both relaxation time (T_1) and dephasing time (T_ϕ).

* Many-Body Decoherence and Driving Cycles: The authors claim that their T_1 times being significantly larger than the characteristic timescale T allows them to experimentally implement more than 1000 driving cycles before notable decoherence occurs. However, T_1 alone is not the correct metric. The authors must also consider the effects of dephasing. The Supplemental Information (SI) indicates that the average T_2 in their device is more than an order of magnitude smaller than their T_1 . While T_2 does not fully represent the many-body decoherence time, the authors need to measure and report an estimate for the many-body decoherence time of their system. This is crucial to convince the reader that they can faithfully implement more than 1000 driving cycles. Furthermore, I would like to see an upper and lower bound estimate on the fidelity of the system as a result of decoherence after 1000 cycles.

* Z Pulse Calibration and High-Frequency Regime: While the method used for calibrating Z pulse distortion effectively captures long time-scale transients ($\tau > 100$ ns), it appears less effective in characterizing and calibrating short time-scale transients. Can the authors comment on how accurately they are able to realize the high-frequency regime ($T \sim 3$ ns)?

* AWG Bandwidth and Pulse Shape: I anticipate that the limited bandwidth of the Arbitrary Waveform Generators (AWGs) in use will further alter the shape of the generated pulse from ideal square pulses to trapezoidal pulses. Can the authors clarify whether this difference influences the physics they study, and if so, how?

Characteristics of Prethermalization by RMD:

* Physical Connection between Lifetimes: This section would greatly benefit from a discussion around the physical connection between the lifetime of the entanglement plateau (τ_S) and the imbalance decay time scale (τ_I), rather than simply quoting the extracted numbers from the experimental data. What are the underlying physical mechanisms that link these two timescales?

Tunable Heating Rate by RMD

* Fig 3a Discrepancy and Error Mechanism: In Figure 3a, the discrepancy between the power-law exponent for larger lattice sizes (8×4 , 13×4 , and 13×8) and the theoretically predicted value is rather large ($\sim 33\%$). The authors state that it is "challenging to obtain a converging scaling exponent, which requires an extremely fast drive and long evolution time." I would like to see a more detailed analysis of the error mechanisms causing this reduction in the power-law exponent. This analysis is crucial to assess the extensibility and reliability of their experimental protocol for larger systems.

* Typo: On page 4, the sentence "where the prethermal lifetime scales exponentially with the increase the driving frequency" is missing "of" after "increase." It should read: "where the prethermal lifetime scales exponentially with the increase of the driving frequency."

Entanglement Dynamics and Scaling:

* Area-Law to Volume-Law Crossover: The measured volume entropy per site is consistently larger than the area entropy per site, as reported in Fig. 4b, even within the prethermal regime. This observation appears to contradict the claim of a crossover in the entanglement structure from area-law to volume-law. The authors need to better justify why they expect area-law entanglement in the prethermal regime prepared using RMD. Furthermore, they should discuss the experimental sources of error that prevent them from observing a proper crossover where the area-law regime is clearly dominant at early times.

Supplemental Information (SI):

* Figure S13b Label: The y-axis label in Figure S13b should be modified to "Interaction duration" or something similar that accurately reflects the quantity being plotted.

Missing Details / Areas for Further Elaboration:

* Fidelity Estimates for Beyond Classicality: As noted above, a quantitative estimate of the simulator's fidelity for RMD experiments, especially for larger system sizes and longer evolution times, is critical to support any "beyond classical" claims. This could involve comparing experimental results to exact numerical simulations (where feasible for smaller systems) or providing theoretical error bounds.

* Impact of Pulse Distortions: A more thorough discussion of the impact of AWG bandwidth limitations and the resulting trapezoidal pulse shapes on the studied physics would strengthen the experimental section. Does this lead to additional heating mechanisms or modify the effective Hamiltonian in a significant way?

* Detailed Error Analysis for Power-Law Exponent: Beyond simply stating the challenge, a more in-depth analysis of the specific error mechanisms (e.g., control errors, decoherence, other systematic errors) that contribute to the reduction in the power-law exponent for larger lattice sizes would be highly valuable. This could include a breakdown of error sources and their estimated contributions.

Physical Intuition for Entanglement Crossover: While the authors discuss the observed entanglement dynamics, providing more physical intuition or a theoretical model for why area-law entanglement is expected in the prethermal regime under RMD, and what specific experimental imperfections might obscure this, would be beneficial.

I look forward to reviewing a revised version of the manuscript that addresses these points comprehensively.

Referee #2

(Remarks to the Author)

The manuscript by Liu et al. presents experimental data from a system of 78 superconducting qubits subject to a random drive. The device is operated in a hybrid analog-digital mode: single-qubit gates are used for state preparation and readout while the dynamics are governed by a time-dependent Hamiltonian. The paper investigates in particular the heating caused by the drive. The authors show that the heating rate can be controlled by imposing a "multipolar" constraint, which had been predicted in theoretical works by some of the authors. The experimental results corroborate theoretical predictions, although significant deviations are observed for the largest systems under consideration.

The authors further claim that the dynamics are difficult to simulate using classical tensor network techniques. The claim is based on the failure to reproduce the experimental data using two particular tensor network algorithms and on the experimentally measured growth of the entanglement entropy. However, I remain unconvinced that the dynamics cannot be simulated with more advanced tensor network algorithms or using Heisenberg picture methods. For the specific Hamiltonian parameters and initial state, it is plausible that the dynamics can be separated into nearly localized dynamics with slow growth of quantum correlations and unstructured heating, which may be amenable to simplifying approximations. Since the key features of the heating appear to be largely independent of system size, it may also be possible to extrapolate numerical results from smaller systems.

Another shortcoming of the work is the lack of an in-depth discussion of the role of noise. The single-qubit decoherence time is on the same scale as the duration of the experiment, indicating that a significant fraction of qubits will have decohered. The authors make no attempt at mitigating this noise and do not comment on its potential contribution to heating. The increasing discrepancy between theory and experiment for larger systems in Fig. 3a suggests that noise may indeed play an important role. It should further be noted that noise can simplify the classical simulation of the system or, at least, reduce the accuracy requirements.

The manuscript is well written and the results are clearly presented. Although the work illustrates curious features of heating under random multipolar driving, its broader implications are limited in my view. For applications in quantum simulation, one would typically aim to avoid the type of heating studied in this work and instead focus on the dynamics that occur up until the prethermal plateau. I include additional detailed comments and questions to the authors below.

1. To what extent are the authors confident that Eq. (1) is an accurate description of their system? Recent work by Google on a similar superconducting platform showed that the couplers can introduce additional terms in the Hamiltonian [Nature 638, 79 (2025)]. Are these terms negligible in the authors' setup?
2. On page 3, the authors mention the decoherence time of the qubits but then quote the value of the relaxation time (T_1). This is misleading, especially because the T_2 time reported in the supplementary materials is smaller than T_1 by more than an order of magnitude. The decoherence time is such that a significant fraction qubits are expected to have decohered on the time scale of the experiment.
3. Do the authors have a (potentially speculative) explanation for the discrepancy between theory and experiment in Fig. 3a?
4. It would be interesting to see a comparison between the random multipolar drive and a standard Trotter decomposition or the evolution under the time-independent average Hamiltonian. The latter would allow the separation of heating induced by the drive from intrinsic heating due to hardware noise.
5. The authors extract the heating time scale by fitting an exponential function to the measured entanglement entropy and spin imbalance. However, the details on data selection are unclear and there is no discussion of error bars / quality of the fit. I recommend that the authors add an extended discussion that addresses these issues.
6. Why do the heating time scales extracted from the entropy and the imbalance differ significantly?
7. Supplementary figures 14 to 16 suggest that the heating slows down in some cases at late times, whereas in supplementary figure 18, the heating appears to speed up. What is the cause for this behavior?
8. Eq. (1) may be confusing to some readers as it omits the hardcore constraint on the bosons. In my view, it would be more natural to represent the model using spins as an XY model.
9. On page 2, the authors write that by measuring the entanglement entropy, their work goes "beyond established experimental results on driven systems where typically only the evolution of local observables is accessible". I do not think that this is a fair characterization since the largest subsystems considered in the manuscript consist of four neighboring qubits. The quantum state tomography used to measure the entanglement only requires the measurement of four-body

operators, which can still be considered local. The authors do not demonstrate any fundamentally new measurement capabilities that are unavailable in similar platforms.

10. The impact of the manuscript could be significantly enhanced by a thorough analysis of the prethermal regime. To what state does the system prethermalize? Does this depend on the order of the drive? How long does the prethermal plateau persist?

Referee #3

(Remarks to the Author)

The authors use a large superconducting quantum processor to investigate prethermalization dynamics under random multipolar driving. Contrary to periodic driving, the mechanism preventing immediate heating and inducing prethermal plateaux under multipolar driving is less understood to date. In their experiments, the authors access rather long times, allowing them to identify long-lived prethermal plateaux under high-frequency driving and their findings support a previously proposed scaling law for the dependence of the heating rate on the multipole order of the drive. Moreover, the authors extract subsystem entanglement entropies, which reveal some anisotropy in the entanglement generation. By careful comparison to two suited tensor network methods, the authors illustrate, that the experimentally accessible system sizes and time scales are hard (if not impossible) to address with classical simulation methods.

The work is original and addresses a very timely topic of non-equilibrium quantum many-body physics. On the one hand, the experiments make a significant contribution to the investigation of multipolar driving protocols with quantitative results for heating rates of large systems, that are in agreement with previous theoretical derivations. Multipolar driving is of high interest, as it may allow for the engineering of new non-equilibrium states of matter, given the control capabilities of quantum processors and quantum simulators as demonstrated in this work. On the other hand, this work is a rather convincing demonstration of the capabilities of superconducting qubits as analog quantum simulator, that yields physical insight at or beyond the boundaries of classical simulation methods.

The conclusions drawn by the authors are well supported by the data shown. Also the comments on potential quantum advantage are appropriate and supported by the reasonable effort invested in tensor network simulations. The presentation is clear, including the abstract, that summarizes the manuscript in a suited manner. The introduction outlines and motivates the general topic of the manuscript appropriately.

However, as the experiment probes physics beyond otherwise accessible regimes, I believe, that the results should be supported by all possible sanity checks. In particular, the authors take the very good approach to assess the accuracy of tensor network simulations by monitoring the conservation of particle number. I have not found analogous data for the quantum simulation in the manuscript or the supplemental material. I believe, that it should be easy to analyze the time-dependence of the total particle number using existing data. On the other hand, Fig. 2 shows a comparison to numerical data for a system of 8 qubits. The next larger system studied experimentally has 32 qubits. I believe, that this should still be feasible by state vector simulations, especially as particle number conservation can be exploited. In my eyes, a demonstration of particle number conservation would significantly strengthen the overall results and I propose, that the authors include such an analysis in a revised version of the manuscript.

Moreover, as the quantitative results for the prethermal lifetimes are a central part of the work, there should be more details about the fitting procedure. It would be beneficial to include a description of the procedure in the supplemental material to convince the reader, that the results do not depend on fine tuning of fitting parameters.

Besides that, I have some further detailed comments, that I list below.

Overall, the manuscript reports high quality research, reporting new insights into the timely topic of manipulating quantum many-body systems far from equilibrium and demonstrating the potential of quantum simulation on superconducting qubits. I believe the manuscript may be suited for publication in Nature upon revision in view of the comments provided.

- 1) The notation in Eqs. (2-4) differs from Eq. (1). Would it be possible to use identical notation across Eqs. (1-4)?
- 2) Reference [83] cited to explain the approach for quantum state tomography (QST) itself refers to another reference for the details of the employed QST. Would it be possible to mention in one sentence what the approach is based on, as there exists a zoo of possible techniques, or to cite an original reference?
- 3) Why is only a single-qubit subsystem considered in Fig. 2a? Wouldn't a larger subsystem be a more demanding check, as for example the page value won't simply be $\log(N)$ due to particle number conservation?
- 4) Statistical error bars should be included in Fig. 3c/d.
- 5) In arguing why plain MPS won't work for simulations it is argued, that swapping would be a major issue. Couldn't that bottleneck simply be avoided by propagating the state via the time-dependent variational principle instead of TEBD?
- 6) It is not mentioned, that 12 subsystem sizes were investigated for the entanglement scaling. Could a list with sketches of those be included in the supplement?
- 7) A unit is missing on the δt axis of supplementary Fig. 1.

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)

After reviewing the revisions made to the manuscript "Prethermalization by Random Multipolar Driving on a 78-Qubit Superconducting Processor", I have the following comments:

* Clarity of Supplementary Material References: The current references to the supplementary material lack specificity. To enhance the paper's readability and allow readers to easily navigate and understand the supporting information, authors must explicitly identify the relevant sections within the supplementary material for each reference in the main text. This will guide readers directly to the pertinent details, improving the overall coherence and logical flow of the arguments presented.

* Error Mitigation Scheme Description: The authors' current reference to a "standard error mitigation scheme" in the supplementary materials is problematic. Error mitigation techniques are highly nuanced and depend heavily on the specific unitary operations under consideration. There is no universally standardized approach that can be assumed. Therefore, the authors are urged to eliminate this generalized language. Instead, they must provide a clear, detailed, and comprehensive description of their specific error mitigation scheme. This description should include a thorough justification for its validity in the context of their experiment, explaining its principles, implementation, and expected efficacy. This detailed explanation is crucial for the reproducibility and critical evaluation of their results.

Appreciated Adjustments and Incorporations:

* Language Adjustment on Simulation Difficulty: The reviewer commends the authors for adjusting the language surrounding the difficulty of simulating the experiment described in the paper. This revision demonstrates an appropriate and more nuanced understanding of the computational challenges involved, contributing to the paper's overall credibility and accuracy.

* Additional Data on Entanglement Crossover: The provision of additional data by the authors, particularly concerning the crossover from area- to volume-law entanglement, is highly appreciated. This data is critical for substantiating key claims made within the manuscript. Specifically, Figure R1-7 from the authors' response is deemed essential and should be seamlessly integrated into Figure 4 of the main text. Its inclusion in the primary narrative is vital for a complete and robust presentation of the findings.

General Comments and Publication Suitability:

* Author's Responsiveness and Manuscript Improvement: The reviewer wishes to express sincere appreciation for the meticulous effort demonstrated by the authors in addressing all previous comments. This diligence has undeniably led to a significant improvement in the quality and clarity of the manuscript.

* Suitability for Nature Physics: While the experiments presented are acknowledged as both impressive and novel in their execution, the conclusions drawn within the manuscript may not possess the broad, paradigm-shifting impact typically expected for publication in Nature. The work, while scientifically sound and demonstrating significant experimental prowess, appears to be more aligned with the scope and readership of Nature Physics.

Referee #2

(Remarks to the Author)

I thank the authors for their detailed response and the careful revision of the manuscript, which has adequately addressed my technical concerns. Even though the authors have weakened their claims of achieving "quantum advantage", I still have reservations about the framing of their work in terms of computational hardness.

1. In point 1-14 of the response, the authors write that the state fidelity of the GMPS simulation for the 2-RMD drive with period $T=3\text{ns}$ exceeds 80% for the entire duration of the experiment, indicating that the dynamics can be classically simulated in this regime. The authors omit this fact from the main text. Moreover, they do not show the classical simulation for the shorter driving period ($T=4\text{ns}$) in Fig. 3b, which could presumably be simulated with a similarly high fidelity. I suggest that the authors address these points in the main text to clearly delineate the regimes in which they have or have not been able to classically simulate the dynamics.

2. The authors state in the response that the dynamics up to the prethermal plateau can be efficiently simulated (likely because the strong staggered potential effectively localizes the system). This should be highlighted in the main text since previous works discussing the computational hardness of prethermalization have focused on this regime (e.g., Ref. [Yang2023] mentioned in response 1-2).

3. Concerning the classical simulation in harder regimes, the authors write in the caption of Fig. 3 that the simulation of the longer driving period ($T=7\text{ns}$) "can only capture the early time evolution". This statement may be too strong given that the GMPS simulation still agrees qualitatively and could be used to extract estimates of the exponents α_I and α_S , potentially on par with the systematic deviations observed in the experimental data. The pertinent question is therefore whether the deviations of the classical numerics from the exact dynamics are significantly larger than the deviations of the experimental data from the exact dynamics. The authors have not addressed this conclusively in either the manuscript or in the response to the referees.

In summary, the manuscript has been greatly improved by the revisions, although further revisions should be made to address the above points.

Referee #3

(Remarks to the Author)

In preparing the revised version of the manuscript, the authors have addressed most of the points I raised in a satisfactory manner.

The claims about quantum advantage have been tuned down and I find them appropriate. The authors declare the physical setting "challenging for numerical methods", which is definitely true. Classical simulation may be possible, but it will require the effort of top experts for the respective method and possibly substantial computational resources. In the appendix it is stated that "These resource requirements far exceed reasonable computational limits." upon giving some estimate of compute times. Since what is reasonable depends very much on perspective, I would rather leave such judgement to the reader, but that's a question of style.

Apart from that, I have a few more questions about the reply and revised version of the manuscript.

1) I am wondering whether the error mitigation scheme of Ref. [80] is really justified in the present case. The authors claim, that particle loss is due to excitation to higher energy levels, i.e. leakage out of the computational subspace. But to my understanding this is not a depolarizing noise channel, which the error mitigation scheme of Ref. [80] is designed for. How can I reconcile this?

2) About the new data from different tensor network methods (Suppl. Fig. 7): I am surprised by the discrepancy, although at least TDVP and GMPS look more or less converged with bond dimension:

- All tensor network architectures treat the two spatial dimensions differently. Is it mentioned anywhere, how they are oriented compared to the initial density wave state? I am wondering, whether this might bias the obtained imbalance.
- Slow convergence with bond dimension may be a consequence of slowly decaying singular values / entanglement spectra. Has it been checked, whether this is the reason why TDVP or GMPS are both seemingly converged?

3) I would also like to remark, that the new data from neural quantum state simulations are clearly not state of the art. It has been pointed out early-on in [Schmitt and Heyl, PRL (2020)], that non-analytic activation functions (like ReLU used in the present work) are not suited when solving the time-dependent Schrödinger equation. Moreover, the mentioned sample size is one order of magnitude smaller than usually required for accurate simulations and adaptive integration schemes often lead to time steps varying over orders of magnitude, meaning that a fixed time step as used in the present work may require orders of magnitude more compute time. It is also mentioned, that some instability at early times immediately renders the simulation inaccurate. Therefore, the presented results are rather meaningless.

Version 2:

Reviewer comments:

Referee #2

(Remarks to the Author)

I thank the authors for their detailed response and careful revisions. All of my concerns have been addressed and I am happy to recommend the manuscript for publication.

Referee #3

(Remarks to the Author)

The revised manuscript and supplemental material is clearly improved in view of my previous comments. However, I am still not convinced, that leakage should be modeled as a depolarizing noise channel.

The authors give two arguments to explain why leakage is depolarizing noise:

1) "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."

Fig. 3(a) in [30] shows dynamics of an imbalance analogous to the present work. The imbalance decays to zero, which is consistent with approaching a maximally mixed state for the individual qubits, which in turn is consistent with depolarizing noise. But it does not prove, that it is a depolarizing noise channel.

2) "We note that a similar error mitigation method has been used in Google's experiment, where random unitaries are implemented [55]"

This fact does not really help my understanding.

My understanding is the following:

Assume two rates γ_0 and γ_1 for leakage from the logical states $|0\rangle$ and $|1\rangle$ out of the computational subspace. Then a leakage error channel for the computational subspace of a single qubit should look something like

$$\rho \rightarrow (1-\gamma_0)|0\rangle\langle 0|\rho|0\rangle\langle 0| + (1-\gamma_1)|1\rangle\langle 1|\rho|1\rangle\langle 1| + (\text{offd. terms})$$

This means that leakage is not trace preserving. Therefore, I don't see how depolarizing noise, which preserves the trace, can be used to model this process.

In view of point 1) above: The decreasing trace of such a leakage channel has the same signature of vanishing imbalance in the long time limit.

I can imagine, that one can actually show, that the employed error mitigation scheme is also suited to address trace-reduction. But if that is what happens, that should be clarified.

If this line of reasoning is wrong, I would appreciate, if the authors could include a brief derivation how leakage becomes depolarizing noise or include a reference, where this is shown. If their point holds, I imagine there is a simple explanation and that could be very helpful for readers.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Report of Referee#1- 2025-05-11012

(1-1) Referee #1 commented

I have reviewed the manuscript "Prethermalization by Random Multipolar Driving on a 78-Qubit Superconducting Processor" and find it to be an interesting experimental contribution to the field of non-equilibrium quantum dynamics. This work reports the experimental observation of long-lived prethermal phases in many-body systems with tunable heating rates, generated by random multipolar driving on a 78-qubit superconducting processor. The authors characterize these phases by monitoring particle imbalance and subsystem entanglement entropy over 1,000 driving cycles, observing a prethermal plateau and claiming a crossover from area-law to volume-law entanglement scaling. The detailed supplemental information regarding the experimental methods and techniques is particularly commendable, significantly aiding in the reproducibility of the work.

However, several aspects of the manuscript require revision and clarification before I can make a recommendation for publication in Nature. My specific comments are detailed below:

Our response:

We would like to thank Referee #1 for reviewing our manuscript, the accurate summary, and the recognition of the novelty in our manuscript. We will respond to comments in the report point by point as follows.

(1-2) Referee #1 commented

Major Concerns:

* "Beyond Classical" Claim: The claim of "quantum advantage" or "beyond classicality" is not sufficiently supported. The authors do not provide any estimates on the fidelity of their analog quantum simulator with Random Multipolar Driving (RMD). Without such fidelity estimates, it is difficult to reasonably assess the validity of claims regarding the beyond-classical nature of their process. In particular, the significant discrepancy between the prethermal lifetime power-law exponent observed for larger lattice sizes (8x4, 13x4, and 13x8) and the theoretically predicted value raises skepticism about the fidelity of the quantum simulator when utilizing the full system. A quantitative discussion of the simulator's fidelity, including systematic errors, especially for larger system sizes and longer evolution times, is essential to substantiate any "beyond classical" claims.

Our response:

We investigate prethermalization and heating in quantum many-body systems using a large-scale quantum processor that remains stable over long timescales. This far-from-equilibrium phenomenon is known to be hard to simulate efficiently with classical numerical methods, see Ref. [Yang2023]. We appreciate that any claim of "quantum advantage" or "beyond classicality" is difficult to prove conclusively. In fact, perhaps all published experimental results so far on intermediate-scale quantum processors can be reproduced, either completely or approximately, using classical algorithms with sufficient computational resources and incorporating the noise/imperfections of the simulator. Therefore, in the revised manuscript, we would like to be more precise and make a careful statement when benchmarking our experimental observations with classical simulations.

In particular, we have systematically compared the numerical efficiency of several algorithms in simulating our system, including TEBD, TDVP, GMPS, Pauli propagation, and NQS. We provide convincing evidence to show that, by incorporating the quasi-1D lattice geometry of the system, the highly optimized GMPS - while being the most efficient among those considered - still fails to accurately simulate the many-body dynamics at long times.

In addition, following the referee's suggestions, we further discuss the system's fidelity and analyze potential sources of error to improve the account of our experimental observations. 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.

Details are provided in the following sections. In particular, Response (1-4) discusses different numerical algorithms; in Responses (1-6) and (1-14), we estimate the system's fidelity; Response (1-8) discusses different error sources that can appear in our experiment.

With this supporting evidence, we hope the referee agrees that the experimentally observed prethermalization with controllable heating rates is novel and challenging to simulate classically with current computational resources. However, whether or not our experiment amounts to a true regime beyond classical simulatability for our current system sizes and times is of secondary nature for our central claim of realizing a novel non-equilibrium phenomenon. We have also revised our statement regarding "quantum advantage" in the updated manuscript.

REFERENCE

[Yang2023] Y. Yang *et al*, *Simulating Prethermalization Using Near-Term Quantum Computers*, PRX Quantum **4**, 030320 (2023).

(1-3) Referee #1 commented

Specific Revisions Required:

Introduction:

* Terminology: On page 2, in systems with particle number conservation, the term "chemical potential" is meaningless. I recommend replacing it with "site energy" or "site potential."

Our response:

In the revised manuscript, we have replaced the "chemical potential" with "site potential".

(1-4) Referee #1 commented

* Quantum Advantage Claim: At the end of the Introduction, the authors claim that the "experiment potentially demonstrates the quantum advantage." This claim is currently too strong, based solely on the observed failure of GMPS and PEPS to simulate the dynamics. To strengthen this assertion, the authors should either significantly revise this claim to reflect a more cautious interpretation or, ideally, bolster it by studying the limitations of other numerical methods, such as Pauli propagation techniques and variational algorithms such as NQS, which have proven successful in classically simulating other many-body systems operating in a similar regime (i.e., Quantinuum's beyond-classical claim <https://arxiv.org/abs/2503.20870>)

Our response:

As suggested by Referee #1, in the revised manuscript, we further implement the Pauli propagation and neural quantum states variational Monte Carlo algorithms to simulate our systems. Computational details can be found in the new Supplementary Materials (SM), and our results indicate that GMPS is still the optimal algorithm for the current simulation task.

In general, the system becomes challenging to simulate after the initial thermalization period, and the algorithmic fidelity in the simulation is difficult to quantify and control. Specifically, for the Pauli propagation method, as we use more Pauli strings to increase precision, the algorithm fails to converge with the current computational resources. This suggests that the Heisenberg picture evolution still leads to notable computational complexity. Importantly, for the long-time evolution, algorithms based on the Heisenberg picture exhibit an intrinsic disadvantage since operators must be evolved backwards in time. Consequently, calculating the expectation value of an observable at any given time requires recalculating the entire evolution from $t = 0$ ns to that specific

time. This leads to an additional $\mathcal{O}(t^2)$ time complexity compared to other numerical methods using the Schrödinger picture.

Regarding Quantinuum's recent paper arxiv: 2503.20870, the authors also use several classical methods to simulate the non-equilibrium evolution generated by the time-dependent Trotterization protocol. It is worth noting that they only numerically simulate around 10 Trotter steps for 56-qubit systems. This indeed requires substantially fewer computational resources compared to our experiment, where we implement 1,000 driving cycles, or Trotter steps, for the 78-qubit system. In addition, we note that gate fidelities accumulate errors in their time evolution, such that error mitigation is needed to reconstruct their signal.

We appreciate that any claim of "quantum advantage" or "beyond classicality" is difficult to prove conclusively. Therefore, we would like to make a more precise and careful statement when benchmarking our experimental observations with classical simulations in the revised manuscript. See details in the revised manuscript.

(1-5) Referee #1 commented

Experimental Setup:

* Decoherence Time vs. Relaxation Time: The authors refer to the time of their qubits as "decoherence time." This should be corrected to "relaxation time." Qubit decoherence time (T_2) takes contributions from both relaxation time (T_1) and dephasing time (T_ϕ).

Our response:

We have replaced the "decoherence time" as "relaxation time".

(1-6) Referee #1 commented

* Many-Body Decoherence and Driving Cycles: The authors claim that their T_1 times being significantly larger than the characteristic timescale T allows them to experimentally implement more than 1,000 driving cycles before notable decoherence occurs. However, T_1 alone is not the correct metric. The authors must also consider the effects of dephasing. The Supplemental Information (SI) indicates that the average T_2^* in their device is more than an order of magnitude smaller than their T_1 . While T_2^* does not fully represent the many-body decoherence time, the authors need to measure and report an estimate for the many-body decoherence time of their system. This is crucial to convince the reader that they can faithfully implement more than 1,000 driving cycles. Furthermore, I would like to see an upper and lower bound estimate on the fidelity of the system as a result of decoherence after 1,000 cycles.

Our response:

In our experiment, we meticulously tracked the entire heating process across over 1,000 driving cycles. The evolution time of the key results presented in the main text (e.g., Fig. 2 and Fig. 4) extends up to $2 \mu s$. As further detailed in the Supplementary Materials, we demonstrate that the average relaxation time T_1 is $26.4 \mu s$, which is much longer than the evolution time, while the Ramsey dephasing time T_2^* exhibits a mean value of $1.2 \mu s$. However, it is crucial to note that, as established in prior studies (and as the Referee noted), the effective dephasing time in a many-body interacting system typically exceeds the individually measured Ramsey dephasing time T_2^* [Dewes2012, Song2019, Guo2021, Liu2025]. This happens because the eigenenergies of many-body systems depend weakly on each qubit flux [Xu2018] and the collective dynamical process also provides a coherence protection [Li2018].

To explicitly demonstrate this, here 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

$$\dot{\rho} = \frac{1}{i\hbar} [H, \rho] + \sum_{m,j} \left(L_{m,j} \rho L_{m,j}^\dagger - \frac{1}{2} \{L_{m,j}^\dagger L_{m,j}, \rho\} \right)$$

where ρ denotes the density matrix of the system, and $L_{m,j}$ 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

$$L_{1,j} = \sqrt{\gamma_{1,j}} \sigma_j^- \text{ with } \gamma_{1,j} = 1/T_{1,j}; \text{ for the dephasing effects, we use } L_{\phi,j} = \sqrt{\gamma_{\phi,j}} \sigma_j^z \text{ with } \gamma_{\phi,j} = 1/T_{2,j} - \gamma_{1,j}/2 \text{ [Krantz2019, Guo2021].}$$

As shown in Fig. R1-1, we compare experimental and numerical results for increasing dephasing time, using a driving period $T = 4 \text{ ns}$ 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 [Yan2019, Shi2023]

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

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 Fig.

R1-1B, the fidelity achieves 0.92 over $2 \mu\text{s}$. Similar results can be obtained for 78-qubit systems, as shown in the response for the Comment (1-14).

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.

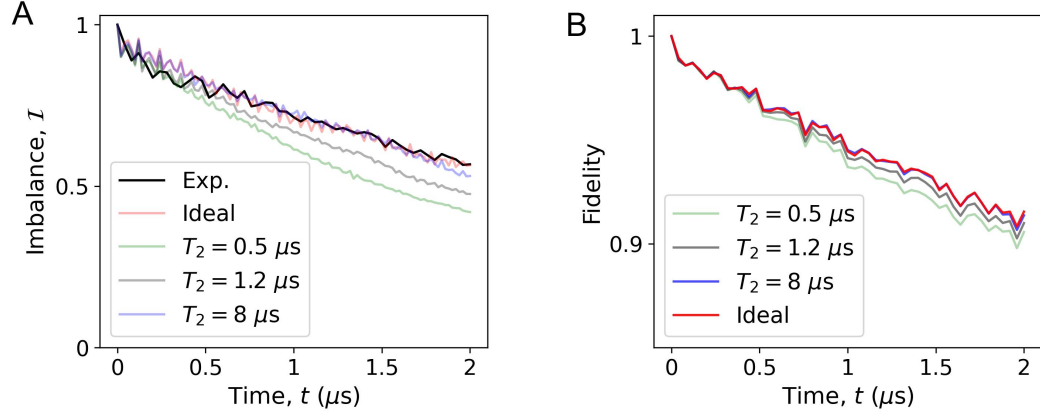


Fig. R1-1: Prethermal dynamics of an 8-qubit system under RMD. (A) Comparison between experimental and numerical results for the spin imbalance across varying dephasing times. (B) Fidelity of the particle number distribution for experimental and simulated data at different dephasing times.

REFERENCES

[Dewes2012] A. Dewes *et al*, *Characterization of a two-transmon processor with individual single-shot qubit readout*, Phys. Rev. Lett. **108**, 057002 (2012).

[Li2018] X. Li *et al*, *Perfect quantum state transfer in a superconducting qubit chain with parametrically tunable couplings*, Phys. Rev. Appl. **10**, 054009 (2018).

[Xu2018] K. Xu *et al*, *Emulating many-body localization with a superconducting quantum processor*, Phys. Rev. Lett. **120**, 050507 (2018).

[Krantz2019] P. Krantz *et al*, *A quantum engineer's guide to superconducting qubits*, Appl. Phys. Rev. **6**, 021318 (2019).

[Yan2019] Z. Yan *et al*, *Strongly correlated quantum walks with a 12-qubit superconducting processor*, Science **364**, 753 (2019).

[Song2019] C. Song *et al*, *Generation of multicomponent atomic Schrodinger cat states of up to 20 qubits*, Science **365**, 574 (2019).

[Guo2021] Q. Guo *et al*, *Observation of energy-resolved many-body*

localization, Nat. Phys. **17**, 234 (2021).

[Shi2023] Y.-H. Shi *et al*, *Quantum simulation of Hawking radiation and curved spacetime with a superconducting on-chip black hole*, Nat. Commun. **14**, 3263 (2023).

[Liu2025] Y. Liu *et al*, *Interplay between disorder and topology in Thouless pumping on a superconducting quantum processor*, Nat. Commun. **16**, 108 (2025).

(1-7) Referee #1 commented

* Z Pulse Calibration and High-Frequency Regime: While the method used for calibrating Z pulse distortion effectively captures long time-scale transients ($\tau > 100$ ns), it appears less effective in characterizing and calibrating short-time-scale transients. Can the authors comment on how accurately they can realize the high-frequency regime ($T \sim 3$ ns)?

Our response:

We thank Referee #1 for this comment. To meet the high-frequency requirements of random multipolar driving (RMD), we utilize a dual-channel digital-to-analog converter (DAC) operating at 2 Gs/s with 16-bit vertical resolution controlled by a field-programmable gate array (FPGA) chip. This configuration, according to Nyquist's sampling theorem, enables us to generate microwave signals with frequencies up to about 1 GHz. Additionally, as discussed in the Supplementary Materials, by measuring the step responses of the Z pulse, we have carefully corrected the Z pulse distortion for both qubits and couplers, due to the finite bandwidth of various electrical components, such as cables, filters, and resonators [Krantz2019, Yan2019, Kwon2021, Xiang2023, Li2025]. Although capturing longer transients (~ 100 ns) may appear more effective, our approach can also correct distortions at the 10 ns scale. In Fig. R1-2, we show the experimental "PulseShape" measurements, which clearly capture the 10 ns scale signal distortion.

To further calibrate the accuracy of the period in the high-frequency driving regime, we experimentally evaluate the performance of fast Floquet driving in a two-qubit system. We pump a photon to one qubit at their idle points, then tune the qubits to the operating points with a frequency difference $\Delta\omega$ between them, and simultaneously apply fast Floquet driving with a period of T . We monitor the population evolution versus time for different T . (Here, the definition of T is consistent with that in RMD, and a total Floquet period is $2T$.) For a given T which meets the condition $2\Delta\omega \cdot T = 2\pi$, energy swap occurs between these two qubits with an effective coupling strength of $g \cdot J_1(2AT)$, where A is the driving amplitude, g is the coupling strength, and $J_n(\cdot)$ is the n -th order Bessel function [Wu2018, Shi2023]. In Fig. R1-3, we present the

experimental results of the driving period T , spanning from 3 to 6 ns, where population swapping between qubits at the corresponding period indicates the nanosecond-scale accuracy in the GHz-frequency domain.

REFERENCES

[Wu2018] Y. Wu *et al*, *An efficient and compact switch for quantum circuits*, *npj Quantum Inf.* **4**, 50 (2018).

[Krantz2019] P. Krantz *et al*, *A quantum engineer's guide to superconducting qubits*, *Appl. Phys. Rev.* **6**, 021318 (2019).

[Yan2019] Z. Yan *et al*, *Strongly correlated quantum walks with a 12-qubit superconducting processor*, *Science* **364**, 753 (2019).

[Kwon2021] S. Kwon *et al*, *Gate-based superconducting quantum computing*, *J. Appl. Phys.* **129**, 041102 (2021)

[Shi2023] Y.-H. Shi *et al*, *Quantum simulation of topological zero modes on a 41-qubit superconducting processor*, *Phys. Rev. Lett.* **131**, 080401 (2023).

[Xiang2023] Z. Xiang *et al*, *Simulating Chern insulators on a superconducting quantum processor*, *Nat. Commun.* **14**, 5433 (2023).

[Li2025] T.-M. Li. *et al*, *High-precision pulse calibration of tunable couplers for high-fidelity two-qubit gates in superconducting quantum processors*, *Phys. Rev. Appl.* **23**, 024059 (2025).

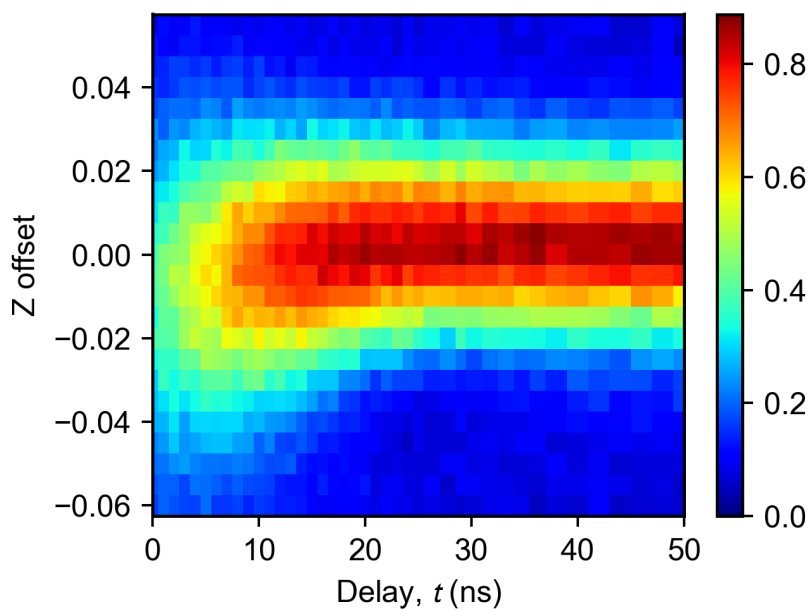


Fig. R1-2: Transient Z pulse after correction between 0~50 ns. The distortion in short time is well mitigated.

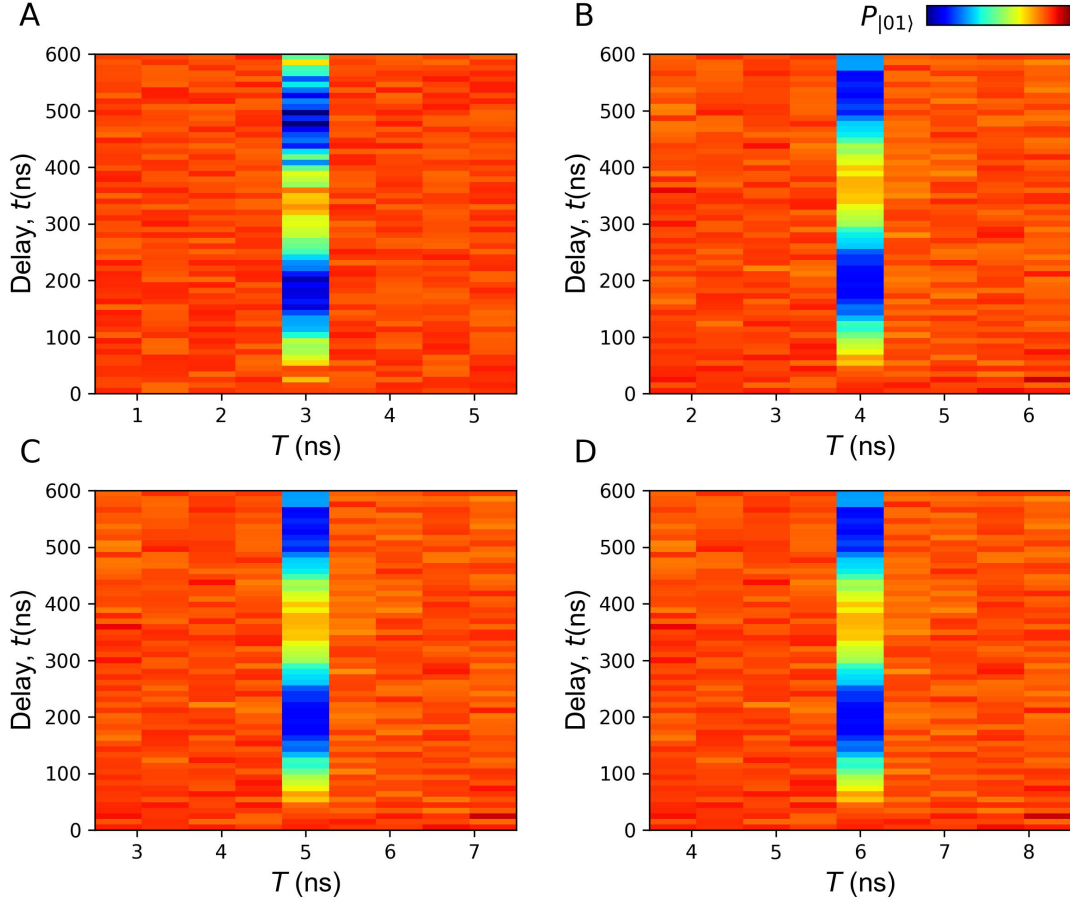


Fig. R1-3: Experimental demonstration of Floquet-driven population swapping between qubits. (A-D) show results for driving periods ranging from 3 to 6 ns.

(1-8) Referee #1 commented

* AWG Bandwidth and Pulse Shape: I anticipate that the limited bandwidth of the Arbitrary Waveform Generators (AWGs) in use will further alter the shape of the generated pulse form from ideal square pulses to trapezoidal pulses. Can the authors clarify whether this difference influences the physics they study, and if so, how?

Our response:

The experimental implementation of a single unitary operator U_+ or U_- is realized by selectively tuning the 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 spatiotemporal perturbations.
2. The characteristic power law scaling $T^{-(2n+1)}$ of the lifetime can be influenced by temporal noise, leading to a reduction of the scaling exponent as experimentally observed for $n = 2$.

Our response here also addresses the questions (1-10, 1-15, 1-16).

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.

At longer times, heating occurs at a rate that exhibits the scaling law T^{2n+1} . This originates from the power law suppression at low frequencies (T^n , see Fig. R1-4, panel A) in the spectrum of the driving protocol [Zhao2021], which remains unchanged for trapezoidal pulses (see panel B).

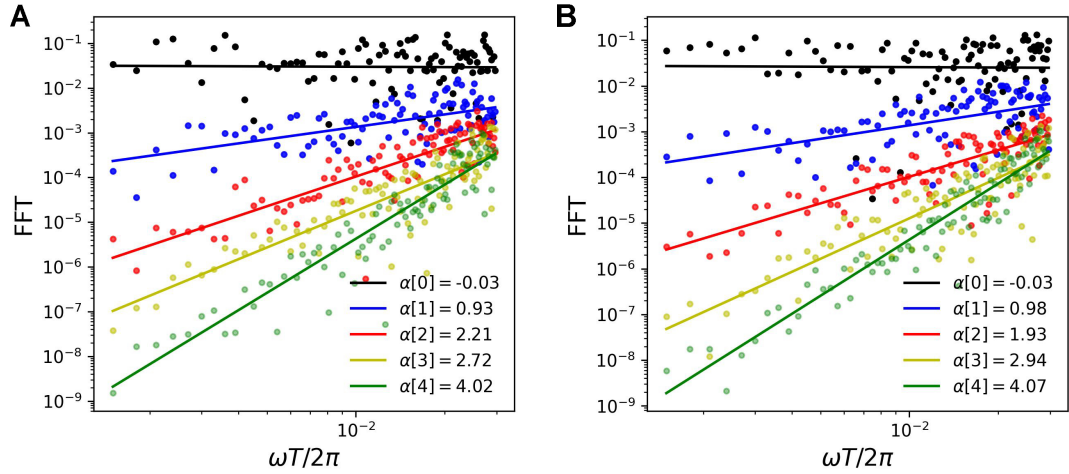


Fig. R1-4: Fourier spectrum of random multipolar sequences for different n when shapes of the pulses are chosen to be rectangle (A) or trapezoidal (B) with rising edge 0.5 ns . The driving period is $T = 3 \text{ ns}$, and the results are averaged over 30 different samples.

As numerically verified in [Zhao2021] and rigorously proved in [Mori 2021], 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 the figure below, where we simulate an 8-qubit system, using multipolar operators U_+ and U_- generated by a trapezoidal onsite potential with 0.5 ns rising and falling edges. 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 a converged scaling exponent $2n + 1$. This is evident in Fig. R1-5: 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 timescales (around 1,000 driving cycles) are typically required. In our previous theoretical study [Zhao2021], at least 10^6 driving cycles were needed to verify the scaling exponent for $n = 3$, which is far beyond the current experimental capabilities.

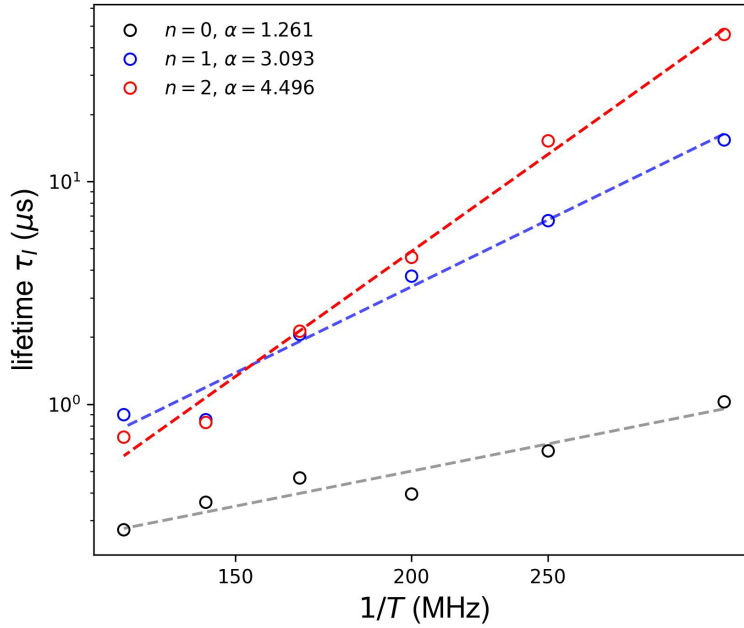


Fig. R1-5: Prethermal lifetime of the RMD system with trapezoidal pulse. The characteristic scaling still appears in the high-frequency regime.

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 qubit frequencies and random error of coupling strengths in such a large system [Yan2019]. 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 operators, 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 [Mori2021], 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 time scale, 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. As shown in Fig. R1-6, 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 clearly shortened in the high-frequency regime. This is sharply different from the time-independent perturbation as shown in Fig. R1-5, 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.

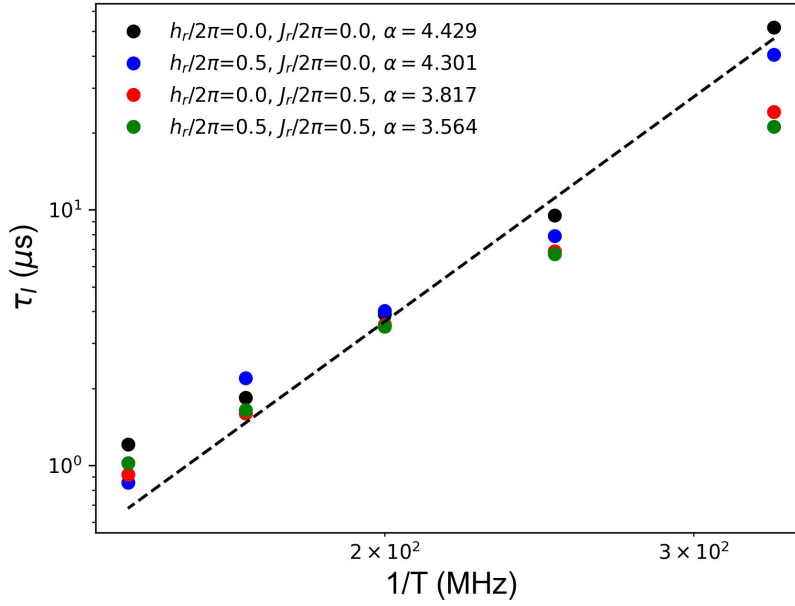


Fig. R1-6: Prethermal lifetime of the RMD 8-qubit system with temporal disorder. The dashed line represents the exact $2n + 1$ power law.

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 may potentially be 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 particle loss (see our response to Comment 3-2) and other forms of decoherence. A complete characterization of all possible error sources goes beyond the scope of the current work.

REFERENCES

[Krantz2019] P. Krantz *et al*, *A quantum engineer's guide to superconducting qubits*, Appl. Phys. Rev. **6**, 021318 (2019).

[Yan2019] Z. Yan *et al*, *Strongly correlated quantum walks with a 12-qubit superconducting processor*, Science **364**, 753 (2019).

[Zhao2021] H. Zhao *et al*, *Random Multipolar Driving: Tunably Slow Heating through Spectral Engineering*, Phys. Rev. Lett. **126**, 040601 (2021).

[Mori2021] T. Mori *et al*, *Rigorous Bounds on the Heating Rate in the Thue-Morse Quasiperiodically and Randomly Driven Quantum Many-Body Systems*, Phys. Rev. Lett. **127**, 050602 (2021).

(1-9) Referee #1 commented

Characteristics of Prethermalization by RMD:

* Physical Connection between Lifetimes: This section would greatly benefit from a discussion around the physical connection between the lifetime of the entanglement plateau (τ_S) and the imbalance decay time scale (τ_I), rather than simply quoting the extracted numbers from the experimental data. What are the underlying physical mechanisms that link these two timescales?

Our response:

After a short transient, the system evolves to the prethermal state governed by the averaged Hamiltonian H_{eff} . According to the eigenstate thermalization hypothesis (ETH), the expectation value of a local observable $\langle O \rangle$, e.g., the particle number imbalance, is close to $\text{Tr}[\rho_{\text{pre}} O]$, with the Gibbs ensemble $\rho \sim e^{-\beta H_{\text{eff}}}$, where β denotes the inverse local temperature during the prethermal regime. In fact, this ensemble also determines the entanglement entropy of a subsystem at the temperature β . When notable heating occurs, the system's energy starts to increase and $\beta(t)$ explicitly depends on time, and drops towards 0 approaching infinite temperature. Crucially, at a fixed time t , both the entanglement entropy and the expectation value of a local observable should correspond to the same local temperature, see systematic discussions in Refs [Machado2019, Fleckenstein2021]. Therefore, to quantify how fast heating happens, both the lifetime of the entanglement plateau (τ_S) and the imbalance decay time scale (τ_I) are suitable measures, which can be estimated by fitting numerically the time evolution as detailed in our response to Comment (3-4).

Note that the precise value of the prethermal lifetime certainly depends on the precise value of $\langle O \rangle$ and the entanglement entropy, which relies on the microscopic details of the system and also the fitting procedure. However, crucially, both time scales exhibit the same universal scaling law $T^{-(2n+1)}$ in the high-frequency regime, as we have experimentally observed.

REFERENCES

[Machado2019] F. Machado *et al*, *Exponentially slow heating in short and long-range interacting Floquet systems*, Phys. Rev. Research **1**, 033202 (2019)

[Fleckenstein2021] C. Fleckenstein and M. Bukov, *Prethermalization and thermalization in periodically driven many-body systems away from the high-frequency limit*, Phys. Rev. B **103**, L140302 (2021).

(1-10) Referee #1 commented

Tunable Heating Rate by RMD

* Fig 3a Discrepancy and Error Mechanism: In Figure 3a, the discrepancy between the power-law exponent for larger lattice sizes (8x4, 13x4, and 13x8) and the theoretically predicted value is rather large (~33%). The authors state that it is "challenging to obtain a converging scaling exponent, which requires an extremely fast drive and long evolution time." I would like to see a more detailed analysis of the error mechanisms causing this reduction in the power-law exponent. This analysis is crucial to assess the extensibility and reliability of their experimental protocol for larger systems.

Our response:

The deviation from the theoretical prediction can be attributed to the accumulation of weak temporal noise. In particular, its effect does not scale up with the system size. Therefore, our driving protocol can be extended to larger systems where the novel prethermal phenomenon should also appear. Please see details in our Response (1-8).

(1-11) Referee #1 commented

* Typo: On page 4, the sentence "where the prethermal lifetime scales exponentially with the increase the driving frequency" is missing "of" after "increase." It should read: "where the prethermal lifetime scales exponentially with the increase of the driving frequency."

Our response:

We thank Referee #1 for the comment. In the revised manuscript, we have addressed this typo.

(1-12) Referee #1 commented

Entanglement Dynamics and Scaling:

* Area-Law to Volume-Law Crossover: The measured volume entropy per site is consistently larger than the area entropy per site, as reported in Fig. 4b, even within the prethermal regime. This observation appears to contradict the claim of a crossover in the entanglement structure from area-law to volume-law. The authors need to better justify why they expect area-law entanglement in the prethermal regime prepared using RMD. Furthermore, they should discuss the experimental sources of error that prevent them from observing a proper crossover where the area-law regime is clearly dominant at early times.

Our response:

In our experimental setup, the initial state is prepared as a density wave ordered product state, which should have zero entanglement entropy if prepared perfectly. 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.

Under the drive, the many-body system indeed quickly develops long-range correlations throughout the entire system, leading to 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 law for generic non-integrable systems. Indeed, as noted by the referee, the ratio s_V/s_A already reaches a value of order 1, suggesting that volume-law entangled states have been generated. Considering the fact that the initial state is a product state, the crossover from area-law to volume-law scaling should occur during the short transient before the prethermal regime. We confirm this below in Fig. R1-7 where the crossover of the value of s_A and s_V is visible at early times, around 5~10 ns.

We now clarify our statement regarding the entanglement growth in the revised manuscript.

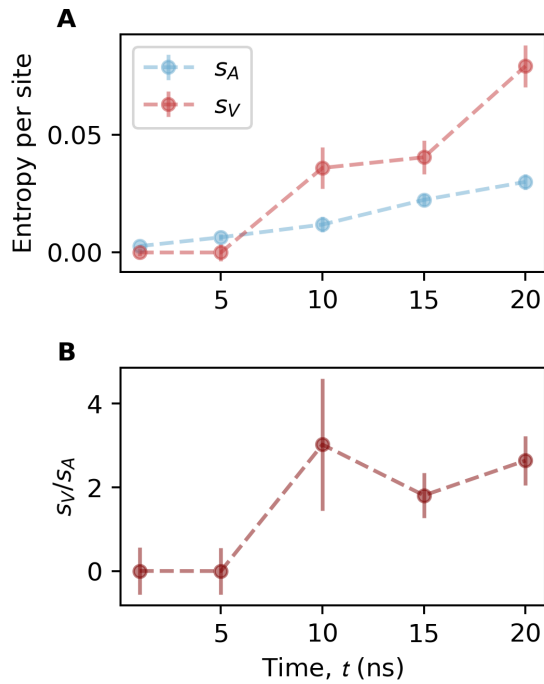


Fig. R1-7: Fitted values of s_A and s_V at early time to distinguish different contributions to the entanglement entropy. The crossover from area-law to volume-law occurs around 5~10 ns.

(1-13) Referee #1 commented

Supplemental Information (SI):

* Figure S13b Label: The y-axis label in Figure S13b should be modified to "Interaction duration" or something similar that accurately reflects the quantity being plotted.

Our response:

We thank Referee #1 for this suggestion. In response, we have updated the y-axis label in the corresponding figure (now labeled as S17b due to the addition of supplementary figures) in the revised Supplemental Materials.

(1-14) Referee #1 commented

Missing Details / Areas for Further Elaboration:

* Fidelity Estimates for Beyond Classcality: As noted above, a quantitative estimate of the simulator's fidelity for RMD experiments, especially for larger system sizes and longer evolution times, is critical to support any "beyond classical" claims. This could involve comparing experimental results to exact numerical simulations (where feasible for smaller systems) or providing theoretical error bounds.

Our response:

As stated in Comment (1-2), we appreciate that any claim of "quantum advantage" or "beyond classicality" is difficult to prove conclusively. In addition, determining the state fidelity of our analog simulator is very challenging. Therefore, we use occupation distribution fidelity (see definition in Response 1-6) as a measure to benchmark our device and show that 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 Fig. R1-1B, 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 $n = 2, T = 3 \text{ ns}$, 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 Fig. R1-8 panel B, for $n = 2, T = 3 \text{ ns}$ (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 and smaller multipolar order, heating happens faster, and the accuracy of GMPS drops accordingly, which also induces 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 experimentally observed non-equilibrium phenomenon. Importantly, our findings highlight that the quantum processor is particularly useful for the study of the many-body heating process, where numerical simulation faces a formidable challenge due to rapid entanglement growth. We now include the discussion in the revised manuscript and the SM.

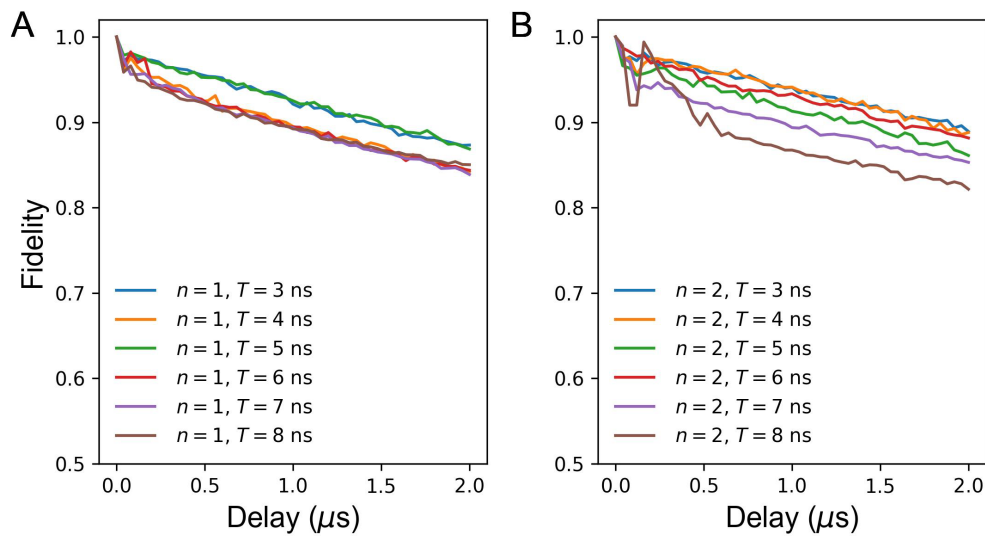


Fig. R1-8: The fidelity for the 78-qubit experimental data and the numerical results from GPMS, when $n = 1$ (A) and $n = 2$ (B), and the driving period ranging 3 ns to 8 ns.

(1-15) Referee #1 commented

* Impact of Pulse Distortions: A more thorough discussion of the impact of AWG bandwidth limitations and the resulting trapezoidal pulse shapes on the studied physics would strengthen the experimental section. Does this lead to additional heating mechanisms or modify the effective Hamiltonian in a significant way?

Our response:

The effective Hamiltonian only gets weakly modified by static perturbations induced by, e.g., trapezoidal deformation of the pulse shape. More crucially, time-dependent perturbation can introduce additional heating, which may induce a deviation between the experimentally observed scaling exponent and its theoretical prediction. See detailed discussion in the Response (1-8).

(1-16) Referee #1 commented

*Detailed Error Analysis for Power-Law Exponent: Beyond simply stating the challenge, a more in-depth analysis of the specific error mechanisms (e.g., control errors, decoherence, other systematic errors) that contribute to the reduction in the power-law exponent for larger lattice sizes would be highly valuable. This could include a breakdown of error sources and their estimated contributions. Physical Intuition for Entanglement Crossover: While the authors discuss the observed entanglement dynamics, providing more physical intuition or a theoretical model for why area-law entanglement is expected in the prethermal regime under RMD, and what specific experimental imperfections might obscure this, would be beneficial.

Our response:

Please see our [Response \(1-6\)](#), [\(1-8\)](#) and [\(1-12\)](#).

(1-17) Referee #1 commented

I look forward to reviewing a revised version of the manuscript that addresses these points comprehensively.

Our response:

We sincerely thank Referee #1 for their valuable feedback and insightful suggestions, which have significantly improved our manuscript. We believe the revised version now meets *Nature's* rigorous publication standards.

Report of Referee#2- 2025-05-11012

(2-1) Referee #2 commented

The manuscript by Liu et al. presents experimental data from a system of 78 superconducting qubits subject to a random drive. The device is operated in a hybrid analog-digital mode: single-qubit gates are used for state preparation and readout while the dynamics are governed by a time-dependent Hamiltonian. The paper investigates in particular the heating caused by the drive. The authors show that the heating rate can be controlled by imposing a "multipolar" constraint, which had been predicted in theoretical works by some of the authors. The experimental results corroborate theoretical predictions, although significant deviations are observed for the largest systems under consideration.

The authors further claim that the dynamics are difficult to simulate using classical tensor network techniques. The claim is based on the failure to reproduce the experimental data using two particular tensor network algorithms and on the experimentally measured growth of the entanglement entropy. However, I remain unconvinced that the dynamics cannot be simulated with more advanced tensor network algorithms or using Heisenberg picture methods. For the specific Hamiltonian parameters and initial state, it is plausible that the dynamics can be separated into nearly localized dynamics with slow growth of quantum correlations and unstructured heating, which may be amenable to simplifying approximations. Since the key features of the heating appear to be largely independent of system size, it may also be possible to extrapolate numerical results from smaller systems.

Another shortcoming of the work is the lack of an in-depth discussion of the role of noise. The single-qubit decoherence time is on the same scale as the duration of the experiment, indicating that a significant fraction of qubits will have decohered. The authors make no attempt at mitigating this noise and do not comment on its potential contribution to heating. The increasing discrepancy between theory and experiment for larger systems in Fig. 3a suggests that noise may indeed play an important role. It should further be noted that noise can simplify the classical simulation of the system or, at least, reduce the accuracy requirements.

Our Response:

We would like to thank Referee #2 for reviewing our manuscript and the accurate summary.

Prethermalization and heating in quantum many-body systems are known to be hard to simulate efficiently with classical numerical methods, see Ref. [Yang2023]. We appreciate that any claim of "quantum advantage" or "beyond classicality" is difficult to prove conclusively. In fact, perhaps all published

experimental results so far on intermediate-scale quantum processors can be reproduced, either completely or approximately, using classical algorithms with sufficient computational resources and incorporating noise/imperfections of the simulator. Therefore, in the revised manuscript, we would like to be more precise and make careful statements when benchmarking our experimental observations with classical simulations.

In the previous version of our manuscript, we systematically compared the numerical efficiency of several algorithms in simulating our system, including TEBD, TDVP and GMPS. We now further use Pauli propagation and neural quantum states variational Monte Carlo algorithms to simulate our system. We provide convincing evidence to show that, by incorporating the quasi-1D lattice geometry of the system, the highly optimized GMPS is the most efficient option; however, it still fails to simulate the many-body dynamics accurately at long times. In addition, we present an in-depth discussion of the possible experimental imperfections and the error mitigation methods used in this work. See details below, in Response (1-8) and (3-2), and in the revised manuscript.

With this evidence, we hope Referee #2 agrees that the experimentally observed prethermalization with controllable heating rates is novel and challenging to simulate classically with current computational resources. We also agree with Referee #2 that there might be better numerical algorithms to simulate our experiment classically. However, whether or not our experiment amounts to a true regime beyond classical simulatability for our current system sizes and times is of secondary nature for our central claim of realizing a novel non-equilibrium phenomenon.

REFERENCE

[Yang2023] Y. Yang *et al*, *Simulating Prethermalization Using Near-Term Quantum Computers*, PRX Quantum **4**, 030320 (2023).

(2-2) Referee #2 commented

The manuscript is well written and the results are clearly presented. Although the work illustrates curious features of heating under random multipolar driving, its broader implications are limited in my view. For applications in quantum simulation, one would typically aim to avoid the type of heating studied in this work and instead focus on the dynamics that occur up until the prethermal plateau. I include additional detailed comments and questions to the authors below.

Our Response:

We agree with Referee #2 on the statement "For applications in quantum simulation, one would typically aim to avoid the type of heating...". However, it is worth noting that, before developing strategies that may be practically useful

for quantum simulation, we have to first gain an in-depth understanding of unwanted heating in generic many-body systems. This indeed poses a fundamental challenge in numerical simulation, as the interplay between the explicit time-dependent control and many-body interaction notably increases the complexity in simulating the non-equilibrium heating phenomenon via classical means.

For periodically driven systems with the discrete time-translation symmetry (DTTS), heating suppression via prethermalization has been theoretically investigated extensively in the past decade, and has also been experimentally observed recently. Yet, it remains largely unexplored of heating that occurs in generic time-dependent systems in the absence of DTTS. RMD protocol interpolates between the fully random time-dependence and highly structured periodic protocols with a controllable degree of temporal correlations. Realization of these protocols on quantum processors significantly advances our understanding of heating and provides crucial insights into stabilizing modern quantum simulators with non-periodic drives.

In addition, importantly, studying heating is also of great importance in addressing fundamental questions in non-equilibrium statistical physics. Specifically, RMD protocols pave the way for realizing non-equilibrium phases of matter with novel forms of spatio-temporal order, e.g., time rondeau order [Moon2024], which cannot be realized with conventional periodic drives.

Therefore, we believe our work will attract broad interest from different research communities, including not only quantum simulation and quantum computation, but also non-equilibrium phases of matter in many-body systems.

REFERENCE

[Moon2024] L. Moon *et al*, *Experimental Observation of a Time Rondeau Crystal: Temporal Disorder in Spatiotemporal Order*, arXiv: 2404.05620.

(2-3) Referee #2 commented

1. To what extent are the authors confident that Eq. (1) is an accurate description of their system? Recent work by Google on a similar superconducting platform showed that the couplers can introduce additional terms in the Hamiltonian [Nature 638, 79 (2025)]. Are these terms negligible in the authors' setup?

Our Response:

The authors in Google's recent work [Andersen2025] presented a more comprehensive modelling of their device, for instance, by including the density–density interaction and next-nearest-neighbor (NNN) hopping processes. Those terms in general scale as $\mathcal{O}(J^2/\eta)$, substantially weaker than the energy scales appearing in Eq. 1 in our main text. Concretely, in our

experiments, the coupling strength J between qubits is tuned to be 2 MHz, which is significantly smaller than the anharmonicity η of approximately -200 MHz. This leads to the substantial separation of energy scales $|\eta/J| \sim 100$, ensuring that the system is well captured by the 2D hard-core Bose-Hubbard model (or equivalently the 2D XY model) [Yan2019, Gong2021]. Furthermore, the coupler frequencies are operated at ~ 6.3 GHz, while the qubits are tuned at ~ 4.2 GHz, resulting in a detuning exceeding 2 GHz. This large frequency difference, which is orders of magnitude larger than the qubit-coupler coupling strength (~ 70 MHz), firmly places the system in the dispersive regime, thereby justifying the theoretical model employed in our analysis [Yan2018, Liang2023, Shi2023].

In practice, we have also deliberately engineered the qubit energy level to suppress undesired NNN couplings and ZZ interactions induced by higher energy levels [ZhaoP2021]. In addition, for this 78-qubit quantum processor, we implement an in-house developed coupler architecture, under which the distance between neighboring qubits can be increased considerably [Liang2023], thus the NNN hopping terms are further suppressed.

To further evaluate the effectiveness of the model, we numerically calculate the dynamics driven by RMD in various systems: (i) a 4-qubit array (2×2 lattice) including 2 levels; (ii) a system containing all 4 qubits and 4 couplers including 2 levels; (iii) a system containing all 4 qubits and 4 couplers including 3 levels, based our experimental parameters. As illustrated in Fig. R2-1, we plot the imbalance as a function of evolution time across different configurations. The results demonstrate that the hard-core approximation maintains high accuracy, further validating its applicability in our experimental framework.

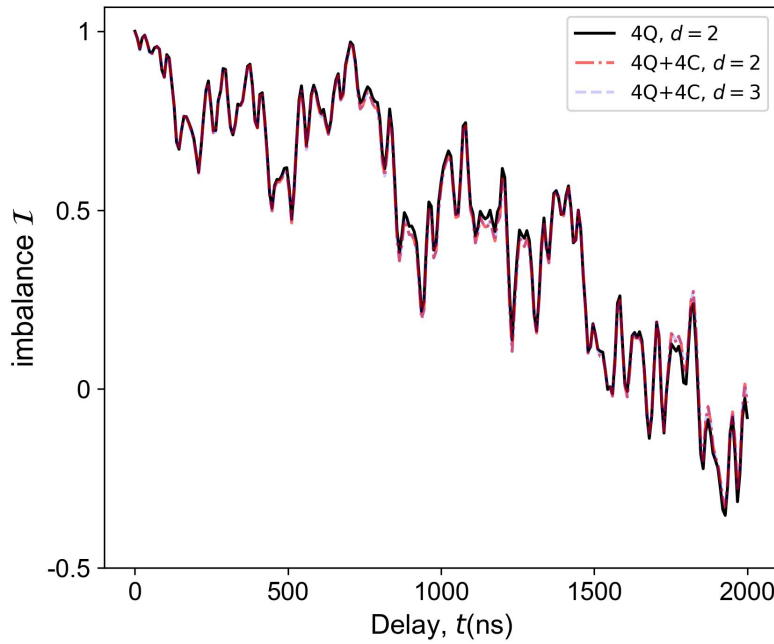


Fig. R2-1: Numerical results of the dynamics in different systems for a single RMD realization, with the multipolar order $n = 1$ and $T = 8 \text{ ns}$.

Finally, we also stress that the presence of NNN coupling plays no critical role in studying RMD systems. Neither the prethermal phenomenon nor the heating rate scaling depends on the specific form of the Hamiltonian. Please see our in-depth discussion in the Response (2-12).

In the revised manuscript, we have added a higher-order term in Eq. (1) and included the discussion above.

REFERENCES

[Yan2018] F. Yan *et al*, *Tunable coupling scheme for implementing high-fidelity two-qubit gates*, Phys. Rev. Appl. **10**, 054062 (2018).

[Yan2019] Z. Yan *et al*, *Strongly correlated quantum walks with a 12-qubit superconducting processor*, Science **364**, 753 (2019).

[Gong2021] M. Gong *et al*, *Quantum walks on a programmable two-dimensional 62-qubit superconducting processor*, Science **372**, 948 (2021).

[ZhaoP2021] P. Zhao *et al*, *Suppression of static ZZ interaction in an all-transmon quantum processor*, Phys. Rev. Appl. **16**, 024037 (2021)

[Liang2023] G.-H. Liang *et al*, *Tunable-coupling architectures with capacitively connecting pads for large-scale superconducting multiqubit processors*, Phys. Rev. Appl. **20**, 044028 (2023).

[Shi2023] Y.-H. Shi *et al*, *Quantum simulation of Hawking radiation and curved spacetime with a superconducting on-chip black hole*, Nat. Commun. **14**, 3263 (2023).

[Andersen2025] T. Andersen *et al*. *Thermalization and criticality on an analogue–digital quantum simulator*, Nature **638**, 79 (2025).

(2-4) Referee #2 commented

2. On page 3, the authors mention the decoherence time of the qubits but then quote the value of the relaxation time (T_1). This is misleading, especially because the T_2 time reported in the supplementary materials is smaller than T_1 by more than an order of magnitude. The decoherence time is such that a significant fraction qubits are expected to have decohered on the time scale of the experiment.

Our Response:

In our experiment, the evolution time of the key results presented in the main text extends up to $2\ \mu s$. As further detailed in the Supplementary Materials, we demonstrate that the average relaxation time T_1 is $26.4\ \mu s$, which is much longer than the evolution time, while the Ramsey dephasing time 2_*^T exhibits a mean value of $1.2\ \mu s$. However, it is crucial to note that, as established in prior studies, the effective dephasing time in a many-body interacting system typically exceeds the individually measured Ramsey dephasing time 2_*^T [Dewes2012, Song2019, Guo2021, Liu2025]. This happens because the eigenenergies of many-body systems depend weakly on each qubit flux [Xu2018] and the collective dynamical process also provides a coherence protection [Li2018].

To explicitly demonstrate this, here 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

$$\dot{\rho} = \frac{1}{i\hbar} [H, \rho] + \sum_{m,j} \left(L_{m,j} \rho L_{m,j}^\dagger - \frac{1}{2} \{L_{m,j}^\dagger L_{m,j}, \rho\} \right)$$

where ρ denotes the density matrix of the system, and $L_{m,j}$ is the Lindblad operators for the j -th qubit with $m \in \{1, \phi\}$, capturing the dissipation induced the coupling to the environment. For the energy relaxation effects, we use

$L_{1,j} = \sqrt{\gamma_{1,j}} \sigma_j^-$ with $\gamma_{1,j} = 1/T_{1,j}$; for the dephasing effects, we use $L_{1,j} =$

$\sqrt{\gamma_{\phi,j}} \sigma_j^z$ with $\gamma_{\phi,j} = 1/T_{2,j} - \gamma_{1,j}/2$ [Krantz2019, Guo2021].

As shown in Fig. R2-2, we compare experimental and numerical results for increasing dephasing time, using a 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 $T_2 = 1.2\ \mu s$. To further evaluate the effects of decoherence, we employ the occupation distribution fidelity defined as [Yan2019, Shi2023]

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

where N is the occupied particle number, $p_i(t)$ and $q_i(t)$ are the measured and theoretical probability distributions at a given time t , respectively. As shown in Fig. R2-2B, the fidelity achieves 0.92 over $2\ \mu s$. Similar results can also be obtained for 78-qubit systems, as shown in our Response (1-14).

In summary, the effective many-body dephasing time is much longer than the individually measured Ramsey dephasing time, exceeding the time scale

required for the observation of the novel prethermalization and its characteristic lifetime scaling.

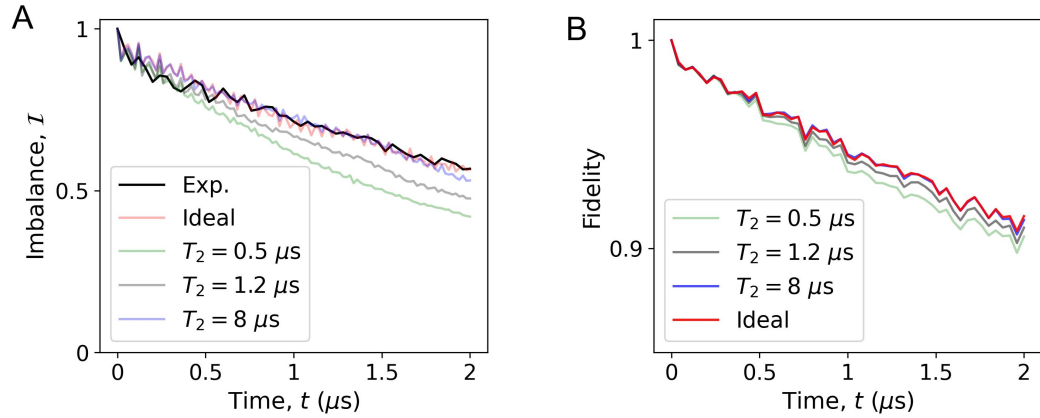


Fig. R2-2: Prethermal dynamics of an 8-qubit system under RMD. (A) Comparison between experimental and numerical results for the spin imbalance across varying dephasing times. (B) Fidelity of the occupation distribution for experimental and simulated data at different dephasing times.

REFERENCES

[Dewes2012] A. Dewes *et al*, Characterization of a two-transmon processor with individual single-shot qubit readout, *Phys. Rev. Lett.* **108**, 057002 (2012).

[Xu2018] K. Xu *et al*, *Emulating many-body localization with a superconducting quantum processor*, *Phys. Rev. Lett.* **120**, 050507 (2018).

[Krantz2019] P. Krantz *et al*, *A quantum engineer's guide to superconducting qubits*, *Appl. Phys. Rev.* **6**, 021318 (2019).

[Song2019] C. Song *et al*, *Generation of multicomponent atomic Schrödinger cat states up to 20 qubits*, *Science* **365**, 574 (2019).

[Guo2021] Q. Guo *et al*, *Observation of energy-resolved many-body localization*, *Nat. Phys.* **17**, 234 (2021).

[Liu2025] Y. Liu *et al*, *Interplay between disorder and topology in Thouless pumping on a superconducting quantum processor*, *Nat. Commun.* **16**, 108 (2025).

(2-5) Referee #2 commented

3. Do the authors have a (potentially speculative) explanation for the discrepancy between theory and experiment in Fig. 3a?

Our Response:

The characteristic power law scaling for RMD protocols is robust to static perturbations, induced by, e.g., the trapezoidal deformation of the pulse shape. The discrepancy between the experiment and the theory can potentially be attributed to the accumulated effects of weak temporal noise in the drive. Please see systematic discussions of different perturbation effects in our Response (1-8).

(2-6) Referee #2 commented

4. It would be interesting to see a comparison between the random multipolar drive and a standard Trotter decomposition or the evolution under the time-independent average Hamiltonian. The latter would allow the separation of heating induced by the drive from intrinsic heating due to hardware noise.

Our Response:

As shown in Fig. 3 in the main text, the agreement between the experimentally observed scaling exponent for multipolar orders $n = 0$ and 1 suggests that the drive-induced heating is dominating the evolution. For $n = 2$, the prethermal phenomenon persists for a long timescale, where the noise effect may accumulate in time and become visible at later times. See systematic discussions of different perturbation effects in our response to Comment (1-8).

(2-7) Referee #2 commented

5. The authors extract the heating time scale by fitting an exponential function to the measured entanglement entropy and spin imbalance. However, the details on data selection are unclear and there is no discussion of error bars / quality of the fit. I recommend that the authors add an extended discussion that addresses these issues.

Our Response:

We thank Referee #2 for this comment. The details of fitting can now be found in Section S6 in the revised Supplementary Materials and also in the Response (3-4) to Referee #3.

(2-8) Referee #2 commented

6. Why do the heating time scales extracted from the entropy and the imbalance differ significantly?

Our Response:

After a short transient, the system evolves to the prethermal state governed by the averaged Hamiltonian H_{eff} . According to the eigenstate thermalization hypothesis (ETH), the expectation value of a local observable $\langle O \rangle$, e.g, the particle number imbalance, is close to $\text{Tr}[\rho_{\text{pre}} O]$, with the Gibbs ensemble $\rho \sim e^{-\beta H_{\text{eff}}}$, where β denotes the inverse local temperature during the prethermal regime. In fact, this ensemble also determines the entanglement entropy of a subsystem at the temperature β . When notable heating occurs, the system's energy starts to increase and $\beta(t)$ explicitly depends on time. Crucially, at a fixed time t , both the entanglement entropy and the expectation value of a local observable should correspond to the same local temperature; see detailed discussions in Refs. [Machado2019, Fleckenstein2021]. Therefore, to quantify how fast heating happens, both the lifetime of the entanglement plateau (τ_S) and the imbalance decay time scale (τ_I) are suitable measures.

Both time scales obey the universal scaling law $C \cdot T^{-(2n+1)}$ with a prefactor C in the high-frequency regime. Observing this power-law scaling with the tunable exponent is precisely the key achievement in our experiment. However, the prefactor C depends on the precise value of $\langle O \rangle$, the entanglement entropy, and also on the fitting procedure. Therefore, in general, τ_S and τ_I should have different values.

REFERENCES

[Machado2019] F. Machado *et al*, *Exponentially slow heating in short and long-range interacting Floquet systems*, Phys. Rev. Research **1**, 033202 (2019).

[Fleckenstein2021] C. Fleckenstein and M. Bukov, *Prethermalization and thermalization in periodically driven many-body systems away from the high-frequency limit*, Phys. Rev. B **103**, L140302 (2021).

(2-9) Referee #2 commented

7. Supplementary figures 14 to 16 suggest that the heating slows down in some cases at late times, whereas in supplementary figure 18, the heating appears to speed up. What is the cause for this behavior?

Our Response:

In these figures, we present our experimental results of imbalance using 10 different realizations of RMD protocols. For randomly driven many-body systems, the fluctuation among different realizations barely affects the time

evolution. However, at a late stage of the time evolution, their effects can appear and potentially lead to the behavior mentioned by the referee. The repetition times have now been declared in the revised manuscript to improve the description of our experimental observation.

(2-10) Referee #2 commented

8. Eq. (1) may be confusing to some readers as it omits the hardcore constraint on the bosons. In my view, it would be more natural to represent the model using spins as an XY model.

Our Response:

We thank Review #2 for this comment. Now we explicitly clarify that our system is equivalent to the XY model after Eq. (1).

(2-11) Referee #2 commented

9. On page 2, the authors write that by measuring the entanglement entropy, their work goes "beyond established experimental results on driven systems where typically only the evolution of local observables is accessible". I do not think that this is a fair characterization since the largest subsystems considered in the manuscript consist of four neighboring qubits. The quantum state tomography used to measure the entanglement only requires the measurement of four-body operators, which can still be considered local. The authors do not demonstrate any fundamentally new measurement capabilities that are unavailable in similar platforms.

Our Response:

We thank Referee #2 for this comment. In our experiments, we systematically measure the entanglement entropy dynamics for subsystems of increasing size, up to 4 qubits with different spatial geometries. Our in-depth analysis of the entanglement entropy is sufficient to distinguish the different contributions from the area and the volume to the entanglement entropy. Importantly, we clearly observe the crossover from area-law to volume-law scaling of the entanglement during the initial transient evolution, and the long-lived entanglement plateau followed by the late-stage rapid growth of entanglement entropy.

To the best of our knowledge, this is the first experimental observation of all the striking features of entanglement growth during the entire heating process using 1,000 driving cycles. This has been made possible not only because of the efficient tomography techniques, but, crucially, also due to the long-time stability of our quantum processor. In comparison, other quantum

simulators either lack an efficient entanglement entropy measurement scheme—e.g., ultracold atoms [Rubio-Abadal2020] and NV centers [Peng2021]—or the capability of performing stable time-dependent control of the device to simulate a driven system [Karamlou2024].

Increasing the subsystem size is exponentially costly in experiments, and developing fundamentally new measurement schemes is definitely worthwhile for future work. However, it goes beyond the scope of the current study. We now clarify our statement regarding the entanglement observation; see details in the revised manuscript.

REFERENCES

[Rubio-Abadal2020] A. Rubio-Abadal *et al*, *Floquet Prethermalization in a Bose-Hubbard System*, Phys. Rev. X 10, 021044 (2020).

[Peng2021] Pai Peng *et al*, *Floquet prethermalization in dipolar spin chains*, Nat. Phys. 17, 444 (2021).

[Karamlou2024] A. Karamlou *et al*, *Probing entanglement in a 2D hard-core Bose-Hubbard lattice*, Nature **629**, 561 (2024).

(2-12) Referee #2 commented

10. The impact of the manuscript could be significantly enhanced by a thorough analysis of the prethermal regime. To what state does the system prethermalize? Does this depend on the order of the drive? How long does the prethermal plateau persist?

Our Response:

Prethermalization is a ubiquitous phenomenon for generic quantum many-body systems driven by the RMD protocol in the high-frequency regime, as shown in [Zhao2021, Mori2021]. It generally appears in different settings, irrespective of the specific microscopic details of the Hamiltonian and the choice of the initial states.

After a short transient, the system evolves to the prethermal state governed by the averaged Hamiltonian H_{eff} . The expectation value of a local observable $\langle O \rangle$, e.g., the imbalance of particle number, will deviate from its initial value and evolve towards its prethermal value $\langle O \rangle_{\text{pre}}$. For generic non-integrable systems and according to the eigenstate thermalization hypothesis (ETH), $\langle O \rangle_{\text{pre}}$ follows the thermal prediction as $\text{Tr}[\rho_{\text{pre}} O]$, with the Gibbs ensemble $\rho \sim e^{-\beta H_{\text{eff}}}$ where β denotes the inverse local temperature during the prethermal regime. Therefore, as long as $\langle O \rangle_{\text{pre}}$ does not vanish, prethermalization appears in the $\langle O \rangle_{\text{pre}}$ and persists until the onset of heating.

In addition, the algebraically suppressed heating rate is also a universal feature of RMD systems. As shown in [Zhao2021, Mori2021], the characteristic scaling $T^{-(2n+1)}$, with the multipolar order n , of the prethermal lifetime should appear in the high-frequency regime for generic many-body Hamiltonian systems with local interactions. In the current experiment, for concreteness, we use specific initial states and Hamiltonian parameters to clearly demonstrate this long-lived prethermal phenomenon. Now we have included these discussions in the revised manuscript.

Therefore, the experimental observation of the long-lived prethermal phenomenon with tunable heating rates in RMD systems is sufficiently generic. It significantly enriches our understanding of paths towards quantum thermalization and provides crucial insights into stabilizing modern quantum simulators with non-periodic time-dependent protocols. Our experiment will further motivate the active research of non-equilibrium physics in the future. In particular, pursuing an in-depth understanding of the stability of many-body localization and topological phases of matter in RMD systems, as well as identifying different prethermalization mechanisms for broader classes of non-periodic drives, remains an open and interesting direction.

REFERENCES

[Zhao2021] H. Zhao *et al*, *Random Multipolar Driving: Tunably Slow Heating through Spectral Engineering*, Phys. Rev. Lett. **126**, 040601 (2021).

[Mori2021] T. Mori *et al*, *Rigorous Bounds on the Heating Rate in Thue-Morse Quasiperiodically and Randomly Driven Quantum Many-Body Systems*, Phys. Rev. Lett. **127**, 050602 (2021)

Report of Referee#3- 2025-05-11012

(3-1) Referee #3 commented

The authors use a large superconducting quantum processor to investigate prethermalization dynamics under random multipolar driving. Contrary to periodic driving, the mechanism preventing immediate heating and inducing prethermal plateaux under multipolar driving is less understood to date. In their experiments, the authors access rather long times, allowing them to identify long-lived prethermal plateaux under high-frequency driving and their findings support a previously proposed scaling law for the dependence of the heating rate on the multipole order of the drive. Moreover, the authors extract subsystem entanglement entropies, which reveal some anisotropy in the entanglement generation. By careful comparison to two suited tensor network methods, the authors illustrate, that the experimentally accessible system sizes and time scales are hard (if not impossible) to address with classical simulation methods.

The work is original and addresses a very timely topic of non-equilibrium quantum many-body physics. On the one hand, the experiments make a significant contribution to the investigation of multipolar driving protocols with quantitative results for heating rates of large systems, that are in agreement with previous theoretical derivations. Multipolar driving is of high interest, as it may allow for the engineering of new non-equilibrium states of matter, given the control capabilities of quantum processors and quantum simulators as demonstrated in this work. On the other hand, this work is a rather convincing demonstration of the capabilities of superconducting qubits as analog quantum simulator, that yields physical insight at or beyond the boundaries of classical simulation methods.

The conclusions drawn by the authors are well supported by the data shown. Also the comments on potential quantum advantage are appropriate and supported by the reasonable effort invested in tensor network simulations. The presentation is clear, including the abstract, that summarizes the manuscript in a suited manner. The introduction outlines and motivates the general topic of the manuscript appropriately.

Our Response:

We thank Referee #3 for the careful review of our manuscript, the recognition of the novelty of our results, and the comments that helped to improve our manuscript. We are enclosing a new version of our manuscript revised according to all comments.

(3-2) Referee #3 commented

However, as the experiment probes physics beyond otherwise accessible regimes, I believe, that the results should be supported by all possible sanity checks. In particular, the authors take the very good approach to assess the accuracy of tensor network simulations by monitoring the conservation of particle number. I have not found analogous data for the quantum simulation in the manuscript or the supplemental material. I believe, that it should be easy to analyze the time-dependence of the total particle number using existing data.....In my eyes, a demonstration of particle number conservation would significantly strengthen the overall results and I propose, that the authors include such an analysis in a revised version of the manuscript.

Our Response:

We agree that a quantitative assessment of the conservation law will significantly strengthen our analysis, and have therefore incorporated it into the revised manuscript. Our results suggest that, due to the presence of depolarizing error, the conservation law is weakly broken, but overall it is well preserved throughout the long-time evolution. In addition, by properly performing an error mitigation method, the effect induced by the particle loss is further suppressed.

Concretely, we first calculate the particle number of an 8-qubit system after calibrating the readout error. The result (black dashed curve) is shown in Fig. R3-1 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 (excitations to higher energy levels in superconducting systems), the system gradually loses particles, but it still retains around 80% of the initial filling after 1,000 driving cycles. In general, such an error weakly speeds up the heating induced by the RMD protocol. However, if we assume that the depolarizing error is uniform in space and has a sufficiently weak strength, one can easily correct its contribution using simple error mitigation schemes, e.g. the one proposed in Ref. [Joseph2021]. According to Eqs. (4) and (8) in Ref. [Joseph2021], this can be achieved by using the experimentally measured particle number N , instead of the initial particle number, as the normalization factor in Eq. (5) in the main text. As shown in Fig. R3-1, 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 Fig. R3-2. 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 $0.2 \mu s$. 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 $2 \mu s$. The error mitigation further suppresses the heating process.

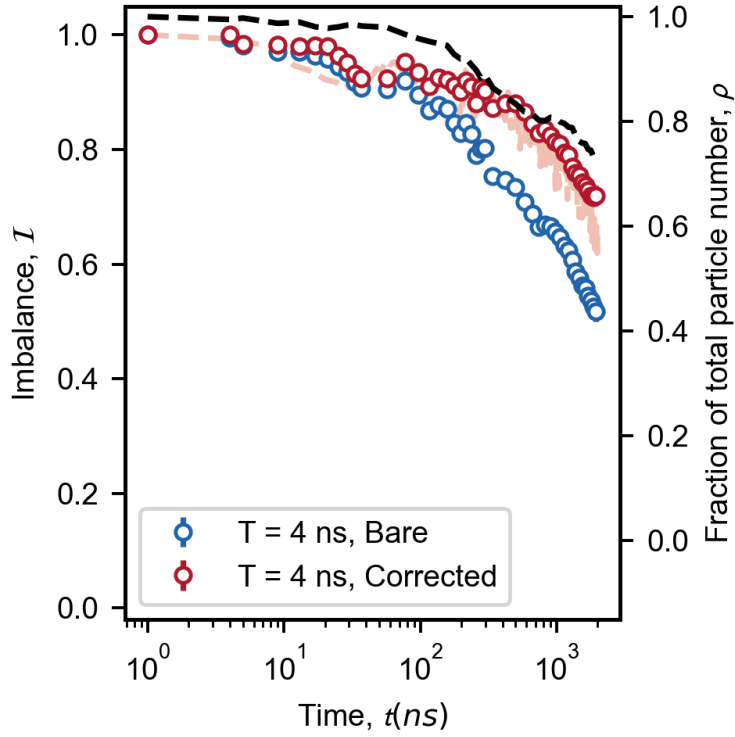


Fig. R3-1: The imbalance and the fraction of the particle number (the black dashed curve) in 8 qubits. A standard error mitigation scheme is used to correct the contribution from the particle loss. The corrected results agree well with the exact numerical simulation (the orange dashed curve).

In Fig. R3-2B, we present the results for a smaller multipolar order and slower 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 μ 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 [Zhao2022], 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

0.2~1.5 μs to further reduce the influence of the particle loss. See details in our response to Comment (3-4).

With this evidence, we confirm that the particle number conservation is well preserved in our experiment. Effects induced by particle loss can be further suppressed using a standard error mitigation scheme. We now include these discussions in the revised manuscript to strengthen our experimental observation of the novel non-equilibrium phenomenon in RMD systems.

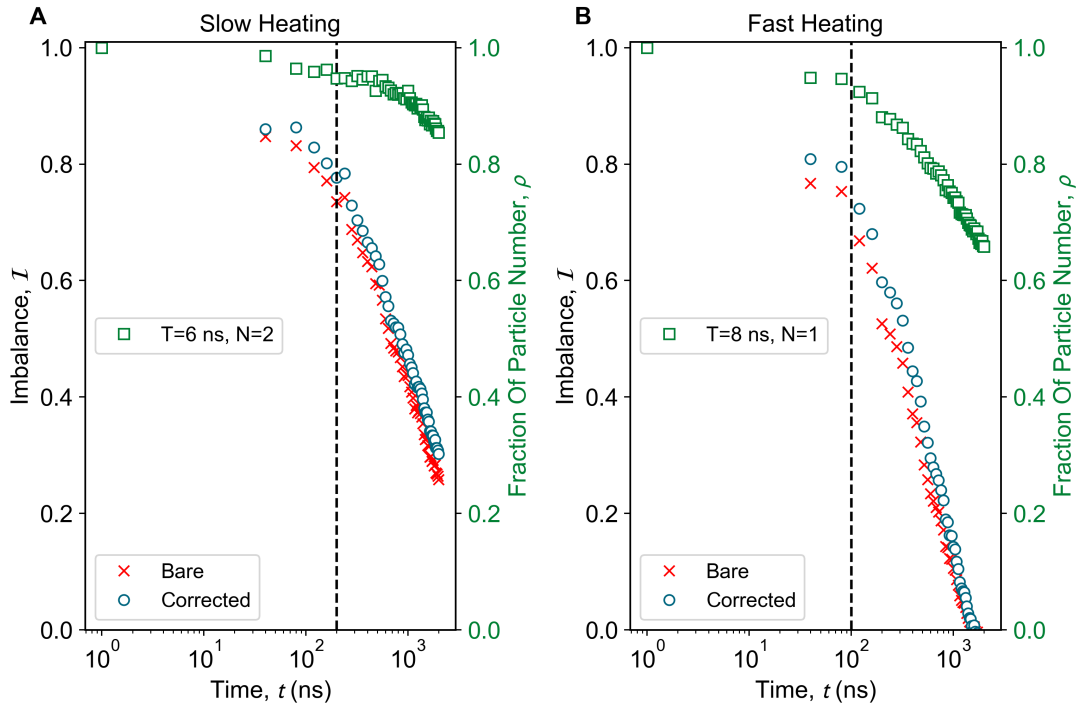


Fig. R3-2: The imbalance and the fraction of the particle number in 78 qubits. A standard error mitigation scheme is used to correct the contribution from the particle loss.

REFERENCES

[Joseph2021] J. Vovrosh *et al*, *Simple mitigation of global depolarizing errors in quantum simulations*, Phys. Rev. E **104**, 035309 (2021).

[Zhao2022] H. Zhao *et al*, *Suppression of Interband Heating for Random Driving*, Phys. Rev. Lett. **129**, 120605 (2022).

(3-3) Referee #3 commented

... On the other hand, Fig. 2 shows a comparison to numerical data for a system of 8 qubits. The next larger system studied experimentally has 32 qubits. I believe, that this should still be feasible by state vector simulations, especially as particle number conservation can be exploited ...

Our response:

The Hilbert space dimension for a 32-qubit system with the particle number conservation law still remains large (about 2^{29}). Exact numerical simulation is still challenging with the current computational resources and only approximate methods, e.g. GMPS, can be implemented. See our response to Comment (1-14) for details where we benchmark the 78-qubit system.

(3-4) Referee #3 commented

Moreover, as the quantitative results for the prethermal lifetimes are a central part of the work, there should be more details about the fitting procedure. It would be beneficial to include a description of the procedure in the supplemental material to convince the reader, that the results do not depend on fine tuning of fitting parameters.

Our response:

In the "Characteristics of Prethermalization by RMD" section of the main text, we have briefly demonstrated the fitting method used to extract the prethermal lifetime. We provide more details here and also in the revised Supplementary Materials, section S.6.

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

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

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,

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

As shown in Fig. R3-3, 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.3 c and d, where the error bar 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 Figs. 26-32, the resulting fitted functions agree well with the experimentally observed time evolution of the imbalance and the entanglement entropy.

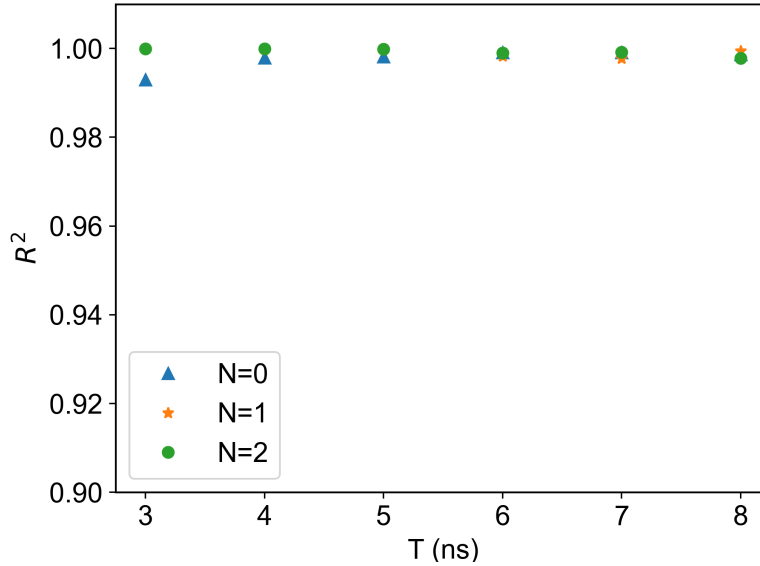


Fig. R3-3: R^2 of different parameter setups in 78 qubits.

(3-5) Referee #3 commented

Besides that, I have some further detailed comments, that I list below.

Overall, the manuscript reports high quality research, reporting new insights into the timely topic of manipulating quantum many-body systems far from equilibrium and demonstrating the potential of quantum simulation on superconducting qubits. I believe the manuscript may be suited for publication in Nature upon revision in view of the comments provided.

Our response:

We thank the constructive suggestions and useful comments from Referee #3, which substantially improve the quality and impact of our manuscript. We have carefully revised the manuscript according to all comments from Referee #3. We believe that the current version of our manuscript with these revisions has now met the high standards of *Nature*.

(3-6) Referee #3 commented

1) The notation in Eqs. (2-4) differs from Eq. (1). Would it be possible to use identical notation across Eqs. (1-4)?

Our response:

In the revised manuscript, the indices in Eq. (1) have been modified from (i, j) to (x, y) to align with Eqs. (2-4), reflecting the two-dimensional square lattice configuration of our superconducting quantum chip.

(3-7) Referee #3 commented

2) Reference [83] cited to explain the approach for quantum state tomography (QST) itself refers to another reference for the details of the employed QST. Would it be possible to mention in one sentence what the approach is based on, as there exists a zoo of possible techniques, or to cite an original reference?

Our response:

In the revised manuscript, we have clarified that "This (QST) can be achieved by performing statistically complete measurement sets, typically implemented through projective measurements along the Pauli eigenbases."

(3-8) Referee #3 commented

3) Why is only a single-qubit subsystem considered in Fig. 2a? Wouldn't a larger subsystem be a more demanding check, as for example the page value won't simply be $\log(N)$ due to particle number conservation?

Our response:

We thank the Referee for this comment. In Fig. 2a, we demonstrate the prethermal dynamics of subsystem entanglement entropy under RMD. Our main goals are threefold: (1) demonstrating frequency-dependent heating rates under different driving conditions ($T = 4\text{ ns}, 8\text{ ns}$); (2) experimentally observing the prethermal plateau; and (3) validating our experimental precision against numerical simulations in an 8-qubit small-scale system. Crucially, these phenomena are fully captured by analyzing entanglement entropy within a single-qubit subsystem.

In general, the particle number conservation modifies the entanglement scaling. However, for the hardcore boson system at half filling as in our experiment, the leading order contribution to the entanglement entropy equals the page value [Vidmar2017]. In supplementary Fig. 32, we have presented additional data of 2-qubit subsystems in the 78-qubit device. After the long-time evolution, the entanglement entropy falls slightly below the theoretical $2 \ln 2$ expectation due to decoherence.

REFERENCE

[Vidmar2017] Lev Vidmar and Marcos Rigol, *Entanglement Entropy of Eigenstates of Quantum Chaotic Hamiltonians*, Phys. Rev. Lett. **119**, 220603 (2017).

(3-9) Referee #3 commented

4) Statistical error bars should be included in Fig. 3c/d.

Our response:

We have added error bars to Fig. 3c/d in the revised manuscript.

(3-10) Referee #3 commented

5) In arguing why plain MPS won't work for simulations it is argued, that swapping would be a major issue. Couldn't that bottleneck simply be avoided by propagating the state via the time-dependent variational principle instead of TEBD?

Our response:

The time-dependent variational principle (TDVP) algorithm is a variational method for simulating the time evolution of matrix product states (MPS). Within the TDVP framework, the Hamiltonian is typically represented as a matrix product operator (MPO) rather than employing a Trotter-Suzuki decomposition. Therefore, we can avoid the explicit introduction of SWAP operators present in some Trotter-based methods. However, the TDVP algorithm introduces significant computational overhead from two primary sources:

1. Local Effective Hamiltonian Construction: For each site (or pair of sites), TDVP requires the computation of a local effective Hamiltonian H_{eff} via an MPS-MPO contraction. For Hamiltonians exhibiting long-range interactions, the corresponding MPO representation can acquire a large bond dimension, rendering the MPS-MPO contraction computationally expensive.

2. Time Evolution of Local Tensors: TDVP needs to apply the matrix exponential $\exp(-iH_{\text{eff}} dt)$ to the local tensors, the dimension of H_{eff} scales as χ^2 (where χ is the MPS bond dimension), making this step costly even employing Krylov subspace techniques.

We systematically compare the performance of TEBD (with SWAP operators), two-site TDVP, and GMPS in the revised SM. We compute a single trajectory for the $n = 1, T = 8 ns$ case. In our study of a 13×6 lattice system, the Hamiltonian MPO possesses a bond dimension of 14. The results are shown in Fig. R3-4. We use *ITensorMPS.jl* package to implement TEBD and

TDVP algorithms. The left figure shows the CPU time per Trotter step. Our results indicate that, for a fixed bond dimension, TDVP exhibits significantly lower computational efficiency compared to the Trotter-Suzuki decomposition-based algorithms. Moreover, our GMPS implementation is 30% faster than TEBD. The right figure shows the imbalance dynamics, where colors progressing from light to dark correspond to bond dimensions of 8, 16, 32, 64, respectively. Both TEBD and TDVP fail to capture the long-time dynamics and thermalize quickly. TEBD performs even less accurately due to the overhead from SWAP operations.

Based on this evidence, our GMPS approach demonstrates the highest efficiency and accuracy for this problem.

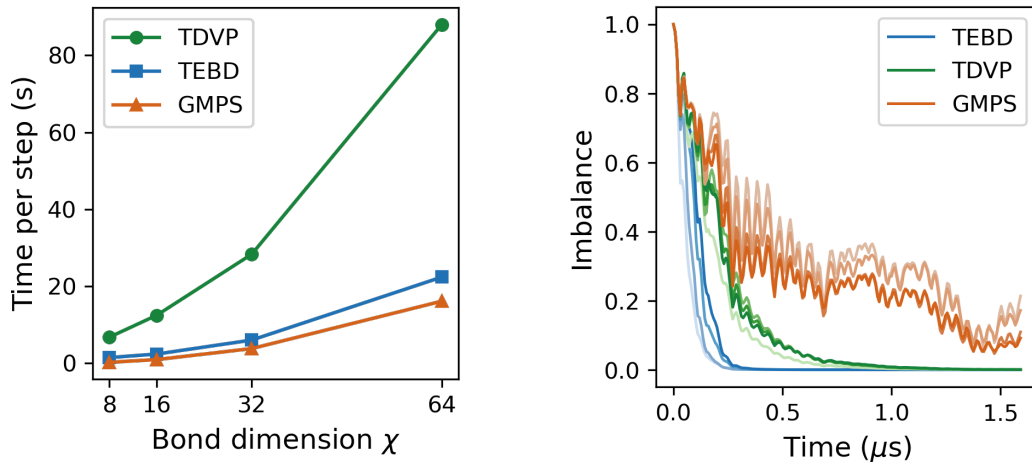


Fig. R3-4: Comparison of different numerical algorithms.

(3-11) Referee #3 commented

6) It is not mentioned, that 12 subsystem sizes were investigated for the entanglement scaling. Could a list with sketches of those be included in the supplement?

Our response:

We thank Referee #3 for this suggestion. In the revised Supplementary Materials and the figure below, we have included 12 subsystems used for the entanglement scaling, along with an accompanying schematic diagram illustrating their spatial distribution across the device architecture. The configurations are listed in increasing order by volume (1 to 4). Labels $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ correspond to Fig. 4a, with index values ranging from 0 to 3.

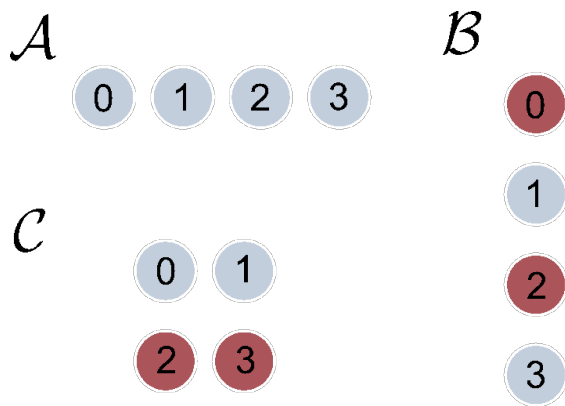


Fig. R3-5: Schematic of configurations with indices.

	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}

Tab. R3-1: Subsystem configurations for entanglement entropy calculation.

(3-12) Referee #3 commented

7) A unit is missing on the δt axis of supplementary Fig. 1.

Our response:

We have addressed this typo in the revised Supplementary Materials.

Referees' comments:

(1-1) Reviewer #1 commented

After reviewing the revisions made to the manuscript "Prethermalization by Random Multipolar Driving on a 78-Qubit Superconducting Processor", I have the following comments:

* Clarity of Supplementary Material References: The current references to the supplementary material lack specificity. To enhance the paper's readability and allow readers to easily navigate and understand the supporting information, authors must explicitly identify the relevant sections within the supplementary material for each reference in the main text. This will guide readers directly to

the pertinent details, improving the overall coherence and logical flow of the arguments presented.

Our response:

We would like to thank Reviewer #1 for the suggestion. We have explicitly cited the specific sections of the Supplementary Materials (SM) in the revised main text.

(1-2) Reviewer #1 commented

* Error Mitigation Scheme Description: The authors' current reference to a "standard error mitigation scheme" in the supplementary materials is problematic. Error mitigation techniques are highly nuanced and depend heavily on the specific unitary operations under consideration. There is no universally standardized approach that can be assumed. Therefore, the authors are urged to eliminate this generalized language. Instead, they must provide a clear, detailed, and comprehensive description of their specific error mitigation scheme. This description should include a thorough justification for its validity in the context of their experiment, explaining its principles, implementation, and expected efficacy. This detailed explanation is crucial for the reproducibility and critical evaluation of their results.

Our response:

We thank Reviewer #1 for this suggestion, and here we would like to provide a comprehensive description of the error mitigation scheme. At long times, our devices suffer from particle loss, due to the excitations to higher energy levels. By treating this effectively as depolarizing errors, we performed an error mitigation scheme proposed by some of us in [Joseph2021], which notably improves the quality of our results.

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 excitations, is implemented. In our previous work[Zhao2022], 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 [Zhao2022], 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 [Google2019], which essentially heats up the system towards the infinite temperature.

Concretely, for qubit systems, a single qubit depolarizing error channel can be modeled via [Nielsen2010, Google2019, Joseph2021]

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

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 [Joseph2021]

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

where p_i is the error on the i -th qubit, $I_i/2$ 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

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

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

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

where $d > 2$ is the Hilbert space dimension of a qudit on a given site. 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. The expectation value of an operator $\hat{O}$ can now be calculated as

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

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

$$\langle I \rangle_{noisy} = (\mathbf{1} - p_{tot})\langle I \rangle_{ideal}$$

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 $\hat{N}_i$ acting on the i -th site is the tensor product of the 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

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

Since $d > 2$, we know $(2/d)^n \ll 1$ after a large number of driving cycles n , leading to

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

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

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

As shown in Supplementary Fig. 25, the notable improvement in the quality of results on the 8-qubit devices further confirms the validity of this method. In the revised SM Sec. S5. C. 2, we have now provided details of this error mitigation scheme.

REFERENCE

[Google2019] F. Arute *et al*, *Quantum supremacy using a programmable superconducting processor*, Nature **574**, 505-510 (2019), see Section IV of Supplementary Information for the consideration using depolarizing channel for random circuits.

[Zhao2022] H. Zhao *et al*, *Suppression of Interband Heating for Random Driving*, Phys. Rev. Lett. **129**, 120605 (2022).

[Nielsen2010] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).

[Joseph2021] Joseph Vovrosh *et al*, *Simple mitigation of global depolarizing errors in quantum simulations*, Phys. Rev. E **104**, 035309 (2021).

(1-3) Reviewer #1 commented

Appreciated Adjustments and Incorporations:

* Language Adjustment on Simulation Difficulty: The reviewer commends the authors for adjusting the language surrounding the difficulty of simulating the experiment described in the paper. This revision demonstrates an appropriate and more nuanced understanding of the computational challenges involved, contributing to the paper's overall credibility and accuracy.

Our response:

We thank Reviewer #1 for their suggestions, which significantly improve the overall credibility and accuracy of our manuscript.

(1-4) Reviewer #1 commented

* Additional Data on Entanglement Crossover: The provision of additional data by the authors, particularly concerning the crossover from area- to volume-law

entanglement, is highly appreciated. This data is critical for substantiating key claims made within the manuscript. Specifically, Figure R1-7 from the authors' response is deemed essential and should be seamlessly integrated into Figure 4 of the main text. Its inclusion in the primary narrative is vital for a complete and robust presentation of the findings.

Our response:

We have now followed this suggestion and revised the presentation of Figure 4.

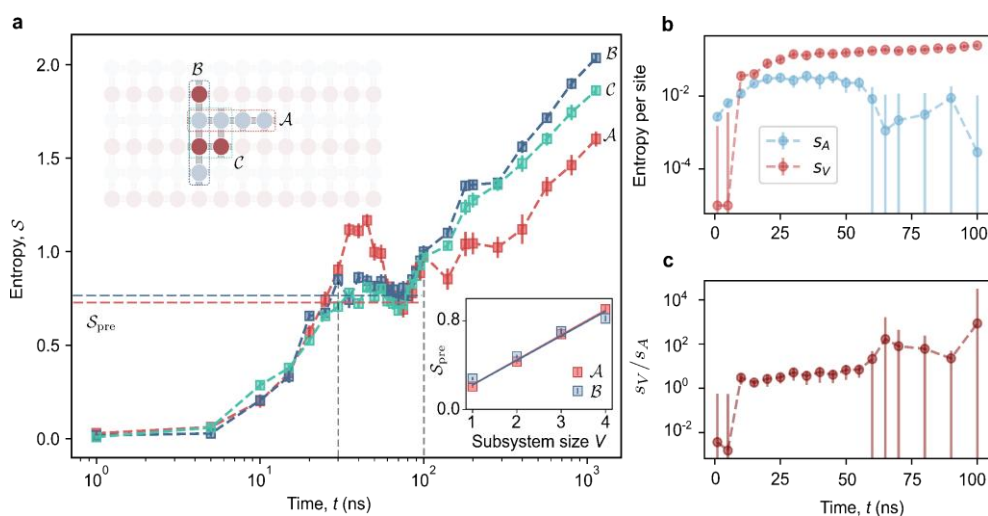


Fig. R1-1 Revised Figure 4. The early 20 ns behaviour is integrated in panels b and c.

(1-5) Reviewer #1 commented

General Comments and Publication Suitability:

* Author's Responsiveness and Manuscript Improvement: The reviewer wishes to express sincere appreciation for the meticulous effort demonstrated by the authors in addressing all previous comments. This diligence has undeniably led to a significant improvement in the quality and clarity of the manuscript.

Our response:

We thank Reviewer #1 for this comment.

(1-6) Reviewer #1 commented

* Suitability for Nature Physics: While the experiments presented are

acknowledged as both impressive and novel in their execution, the conclusions drawn within the manuscript may not possess the broad, paradigm-shifting impact typically expected for publication in Nature. The work, while scientifically sound and demonstrating significant experimental prowess, appears to be more aligned with the scope and readership of Nature Physics.

Our response:

We thank Reviewer #1 for their positive evaluation of our experiment. We would like to emphasize that beyond the purely technological/methodological advances underpinning the experimental results, our work has a broad and significant impact across multiple research fields, which together we believe fully meet the highly selective publication criteria of Nature.

Indeed, our advanced 78-qubit superconducting chip, featuring a high-performance design and precise time-dependent control, represents a state-of-the-art quantum processor. Our large-scale, long-time analog quantum simulation explores regimes that pose formidable challenges for classical computation, thereby offering fundamental and wide-ranging implications for quantum simulation and computation.

This has enabled the experimental observation of exotic non-equilibrium phases of matter, a topic of intense current interest. We have found tunable, universal scaling laws in randomly driven systems. Together with the development of our classical simulation algorithms, all of this is of great significance in non-equilibrium many-body physics, quantum control, and computational science.

We therefore believe that our timely findings meet the high standards for publication in Nature.

Referees' comments:

(2-1) Reviewer #2 commented

I thank the authors for their detailed response and the careful revision of the manuscript, which has adequately addressed my technical concerns. Even though the authors have weakened their claims of achieving "quantum advantage", I still have reservations about the framing of their work in terms of

computational hardness.

1. In point 1-14 of the response, the authors write that the state fidelity of the GMPS simulation for the 2-RMD drive with period $T=3\text{ns}$ exceeds 80% for the entire duration of the experiment, indicating that the dynamics can be classically simulated in this regime. The authors omit this fact from the main text. Moreover, they do not show the classical simulation for the shorter driving period ($T=4\text{ns}$) in Fig. 3b, which could presumably be simulated with a similarly high fidelity. I suggest that the authors address these points in the main text to clearly delineate the regimes in which they have or have not been able to classically simulate the dynamics.

Our response:

During the prethermal regime, as the staggered potential severely limits the entanglement growth, GMPS can faithfully simulate the evolution with high accuracy. We therefore use these data to benchmark our quantum processor. However, classical simulation becomes particularly challenging when heating happens, as we have already stated in the manuscript, *"Therefore, our quantum processor allows us to simulate the entire heating dynamics towards the highly-entangled infinite temperature states that is extremely costly to simulate classically"*.

Following the Reviewer #2's suggestion, the GMPS simulation of high-frequency driving ($T=4\text{ ns}$) has now been added in Fig. 3b; also see figure below. This simulation at early times is sufficiently accurate where the state fidelity exceeds 0.8 before $1\text{ }\mu\text{s}$ (see Supplementary Fig. 4). 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 [Machado2019, Mallayya2019]. 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 $2 \mu s$, capable of simulating the heating process in a regime where the classical algorithm faces formidable challenges.

For the slow driving ($T=7$ ns), heating occurs at early times and the entanglement entropy grows much faster than the high-frequency driving ($T = 4$ ns). The GMPS becomes even less efficient and a significant discrepancy between the numerical simulation and the experimental results is visible even at early times. It happens because the GMPS state fidelity already drops to 0.62 at $0.5 \mu s$, and hence the evolution of the local observables is no longer reliable.

In conclusion, both the late time evolution in the high-frequency regime, as well as the intermediate and late time evolution for slow drivings, cannot be efficiently simulated using classical methods due to the rapid entanglement growth. We have now clarified this in the main text and also in the revised SM Sec. S5. C. 1.

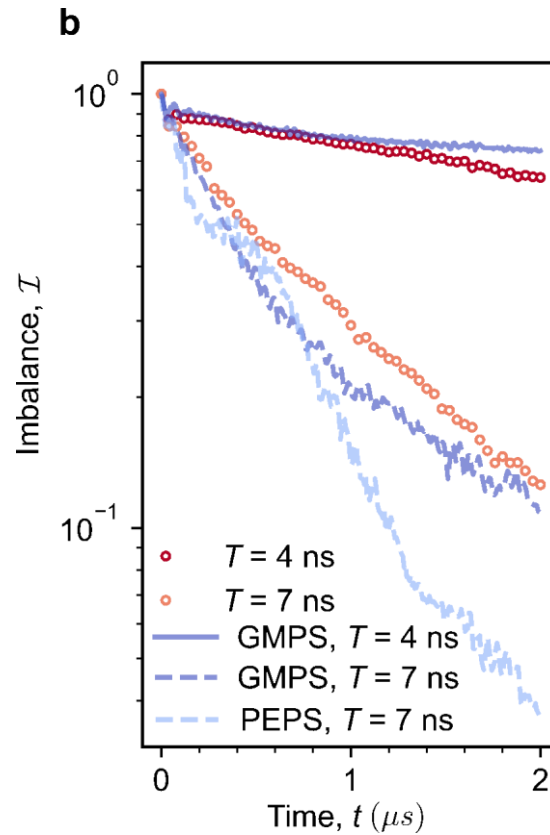


Fig. R2-1 Revised Figure 3b.

[Machado2019] F. Machado *et al*, *Exponentially slow heating in short and*

long-range interacting Floquet systems, Phys. Rev. Research **1**, 033202 (2019) .

[Mallayya2019] Krishnanand Mallayya and Marcos Rigol, *Heating Rates in Periodically Driven Strongly Interacting Quantum Many-Body Systems*, Phys. Rev. Lett. **123**, 240603 (2019).

(2-2) Reviewer #2 commented

2. The authors state in the response that the dynamics up to the prethermal plateau can be efficiently simulated (likely because the strong staggered potential effectively localizes the system). This should be highlighted in the main text since previous works discussing the computational hardness of prethermalization have focused on this regime (e.g., Ref. [Yang2023] mentioned in response 1-2).

Our response:

We agree with Review #2 that the strong staggered potential constrains the entanglement growth and hence GMPS can simulate the prethermal regime with high accuracy. However, for other Hamiltonian parameters, fast entanglement growth can occur during the prethermal regime, and the classical simulation can be computationally costly.

We have now commented on the effects induced by the staggered potential in the revised manuscript.

(2-3) Reviewer #2 commented

3. Concerning the classical simulation in harder regimes, the authors write in the caption of Fig. 3 that the simulation of the longer driving period ($T=7\text{ns}$) "can only capture the early time evolution". This statement may be too strong given that the GMPS simulation still agrees qualitatively and could be used to extract estimates of the exponents α_I and α_S , potentially on par with the systematic deviations observed in the experimental data. The pertinent question is therefore whether the deviations of the classical numerics from the exact dynamics are significantly larger than the deviations of the experimental data from the exact dynamics. The authors have not addressed this conclusively in either the manuscript or in the response to

the referees.

In summary, the manuscript has been greatly improved by the revisions, although further revisions should be made to address the above points.

Our response:

Reviewer #2 indeed raised a fundamental question: how to benchmark quantum simulation results without the knowledge of exact solutions. For digital quantum simulation, one standard method is to use cross entropy benchmarking (XEB) [Google2019] to estimate the circuit fidelity. In contrast, there is no such general measure for analog quantum simulation where the dynamics is not generated by discrete quantum gates but rather continuous time Hamiltonian evolution.

In our work, in the high-frequency regime where the numerical simulation is sufficiently accurate, we have systematically analyzed particle number conservation and the particle distribution fidelity. All of these benchmarks convincingly support the conclusion that our quantum processor remains stable and reliable throughout the long-time evolution. Therefore, we are confident that our experiment has provided reliable simulation results even in regimes where numerical simulations face challenges. In the figure below, we show the systematic error in local observables, estimated based on the state infidelity in the GMPS simulation, see details in [Nielsen2010], 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.

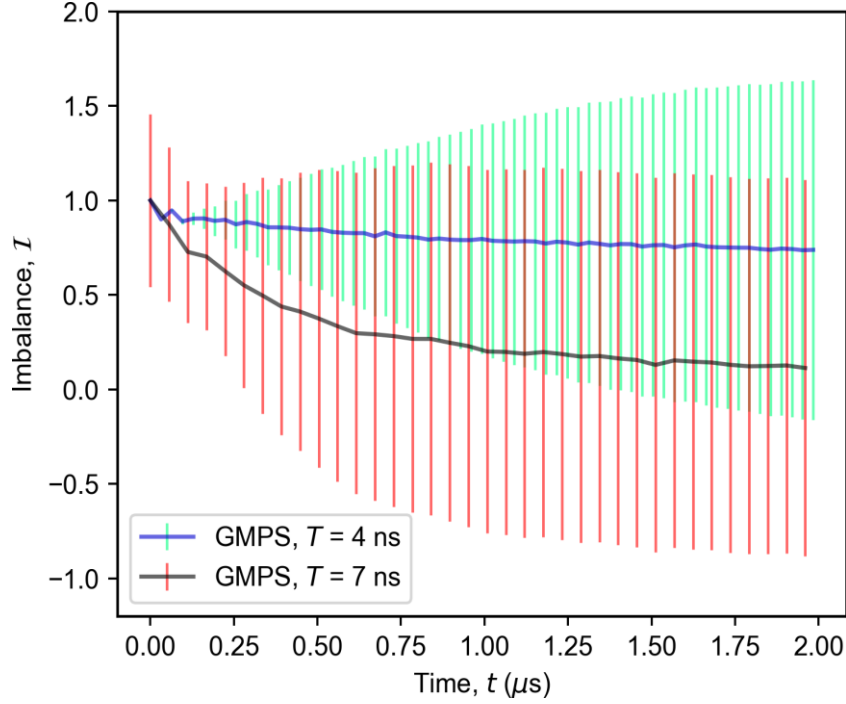


Fig. R2-2. GMPS simulation of local observables, with an increasing error bar estimated from the state infidelity $\sqrt{1 - F^2}$. The value of the error bar quickly increases to the order of 1, suggesting that the numerical simulation is unreliable.

Therefore, we would like to conclude that, in the absence of exact solutions of the heating dynamics, both the late time evolution in the high-frequency regime, as well as the intermediate and late time evolution for slow drivings, cannot be efficiently simulated using classical methods due to the rapid entanglement growth. More generally, the question of how to benchmark analog quantum simulation for generic simulation tasks remains open. Addressing this question is of great importance, which, however, goes beyond the scope of the current work.

We have now added these discussions in the revised SM Sec. S7. E.

[Google2019] F. Arute *et al*, *Quantum supremacy using a programmable superconducting processor*, Nature **574**, 505-510 (2019).

[Nielsen2010] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).

Referees' comments:

(3-1) Reviewer #3 commented

In preparing the revised version of the manuscript, the authors have addressed most of the points I raised in a satisfactory manner.

The claims about quantum advantage have been tuned down and I find them appropriate. The authors declare the physical setting "challenging for numerical methods", which is definitely true. Classical simulation may be possible, but it will require the effort of top experts for the respective method and possibly substantial computational resources. In the appendix it is stated that "These resource requirements far exceed reasonable computational limits." upon giving some estimate of compute times. Since what is reasonable depends very much on perspective, I would rather leave such judgement to the reader, but that's a question of style.

Apart from that, I have a few more questions about the reply and revised version of the manuscript.

(3-2) Reviewer #3 commented

1) I am wondering whether the error mitigation scheme of Ref. [80] is really justified in the present case. The authors claim, that particle loss is due to excitation to higher energy levels, i.e. leakage out of the computational subspace. But to my understanding this is not a depolarizing noise channel, which the error mitigation scheme of Ref. [80] is designed for. How can I reconcile this?

Our response:

At long times, our devices suffer from particle loss, due to the excitations to higher energy levels. This 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 replaces a quantum state with a maximally mixed state with some probability. By treating this effectively as depolarizing errors, we performed an error mitigation scheme proposed by some of us in [Joseph2021], which notably improves the quality of our results.

See details in the response (1-2).

[Joseph2021] Joseph Vovrosh *et al*, *Simple mitigation of global depolarizing errors in quantum simulations*, Phys. Rev. E **104**, 035309 (2021).

(3-3) Reviewer #3 commented

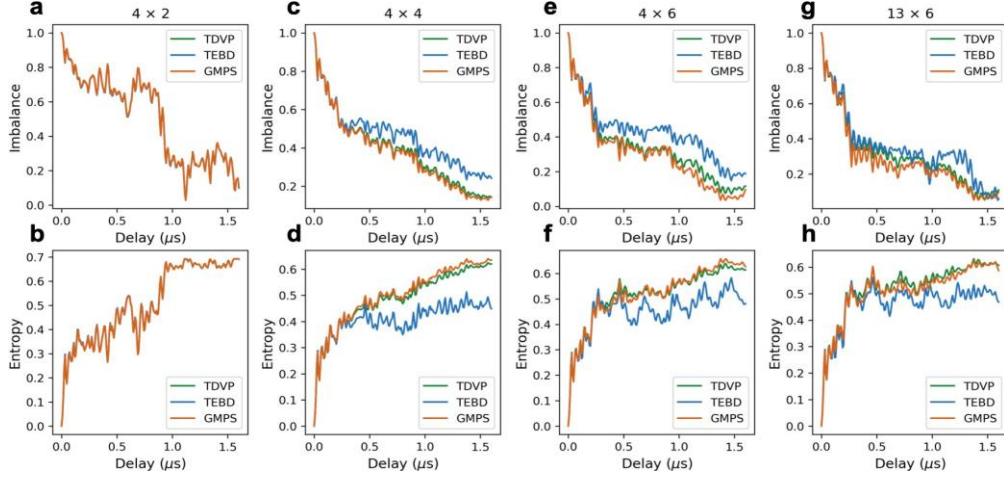
2) About the new data from different tensor network methods (Suppl. Fig. 7): I am surprised by the discrepancy, although at least TDVP and GMPS look more or less converged with bond dimension:

- All tensor network architectures treat the two spatial dimensions differently. Is it mentioned anywhere, how they are oriented compared to the initial density wave state? I am wondering, whether this might bias the obtained imbalance.
- Slow convergence with bond dimension may be a consequence of slowly decaying singular values / entanglement spectra. Has it been checked, whether this is the reason why TDVP or GMPS are both seemingly converged?

Our response:

We thank the reviewer for their comments. The TEBD and TDVP simulations have been refined using the standard iTensor library, with the updated results now presented in the Supplementary Material.

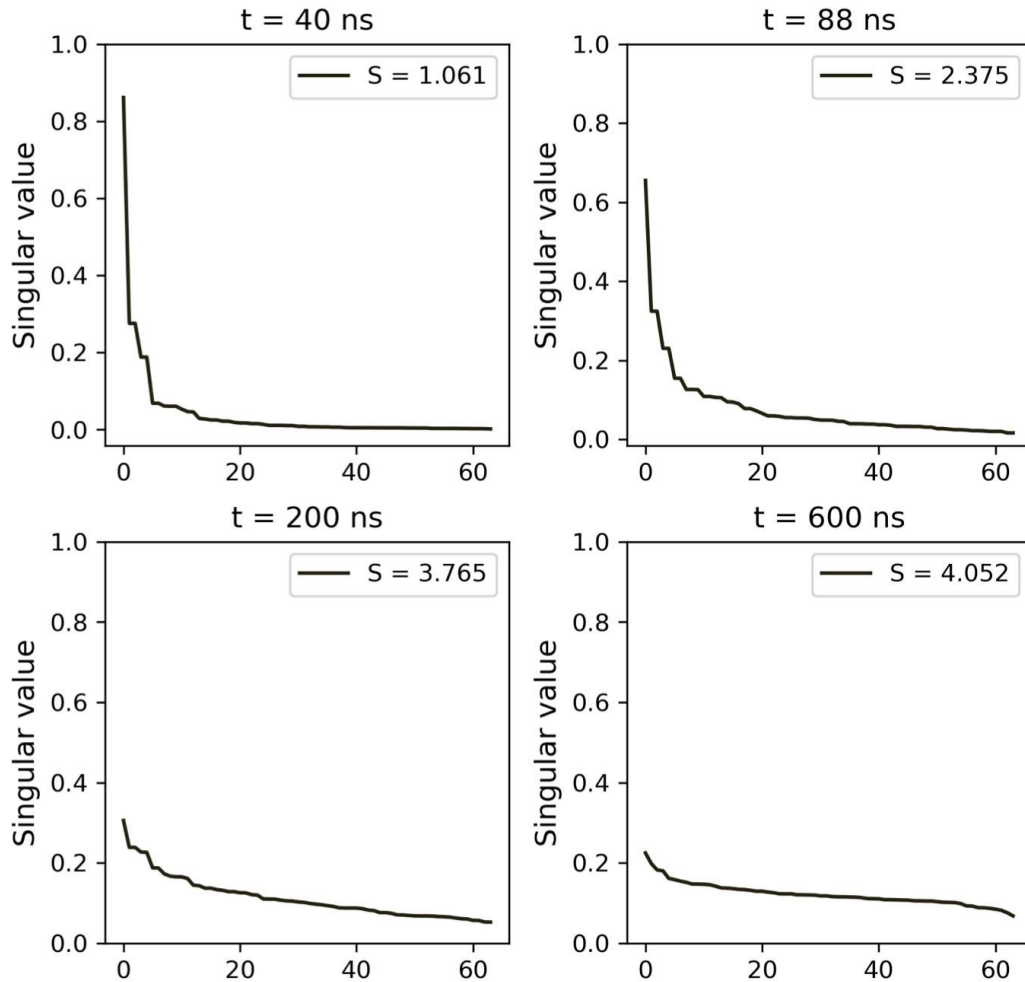
In the figure below, we have presented the simulation results of local observables and single-site entanglement entropy for different system sizes. We use a small system size (panels a and b) to benchmark all three methods, which lead to the same evolution results. For larger system sizes, as shown in panels c to h. Clearly, given a fixed bond dimension, TDVP (green) and GMPS (red) lead to similar simulation outcomes, while TEBD (blue) generates a notable discrepancy at long times. Nevertheless, as shown in Fig. 9 in SM, TDVP is particularly costly in terms of CPU times, in contrast to our GMPS method.



Regarding the first question, our initial state is the same as in the experiment and does not depend on the MPS geometry. The design of the MPS architecture is primarily motivated by the principle of preserving locality. The discrepancy observed in the TEBD results is due to the necessity of introducing extra SWAP operators to handle non-nearest-neighbor interactions, which are not required in the TDVP and GMPS schemes.

To address the second question, we monitor the 64 largest singular values of the middle tensor in GMPS. At early times ($t=40$ ns and 88 ns), the system is close to a product state, 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 starts heating towards the infinite temperature, e.g., the bottom two panels 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.

These observations convincingly support the conclusion that our quantum processor provides crucial insights into quantum dynamics in regimes where classical algorithms face challenges.



(3-4) Reviewer #3 commented

3) I would also like to remark, that the new data from neural quantum state simulations are clearly not state of the art. It has been pointed out early-on in [Schmitt and Heyl, PRL (2020)], that non-analytic activation functions (like ReLU used in the present work) are not suited when solving the time-dependent Schrödinger equation. Moreover, the mentioned sample size is one order of magnitude smaller than usually required for accurate simulations and adaptive integration schemes often lead to time steps varying over orders of magnitude, meaning that a fixed time step as used in the present work may require orders of magnitude more compute time. It is also mentioned, that some instability at early times immediately renders the simulation inaccurate. Therefore, the presented results are rather meaningless.

Our response:

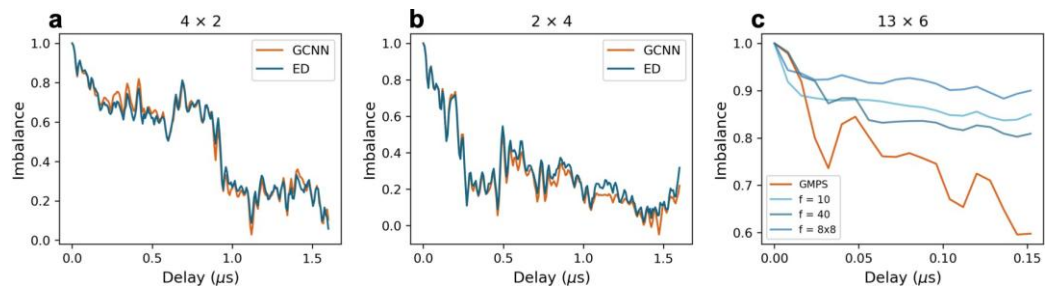
We thank Reviewer #3 for the helpful comments. In the revised manuscript, we have incorporated the polynomial activation function proposed in [Schmitt2020], which indeed improves our GCNN simulation. However, given the current computational resources, we find that GCNN cannot capture the heating dynamics in large system sizes and the highly-optimized GMPS is still the most efficient algorithm.

We first benchmark our GCNN against the exact solution for small systems (4×2 and 2×4 , panels a and b). We use a single layer with 10 features in the GCNN algorithm and 1,024 samples to update the network. As shown in panels a and b, GCNN achieves high accuracy throughout the entire simulation. When enlarging the system size to 13×6 , 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 network. In panel c, we compare the GCNN results with those from GMPS. In this time scale the fidelity of GMPS is almost 1. 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 $0.02\ \mu\text{s}$, corresponding to 3 driving cycles. The deviation quickly increases as time evolves, suggesting that GCNN fails to capture the many-body evolution during the heating process.

Increasing the number of features or layers of the neural network may further improve the simulation accuracy, but this goes beyond our current computational capabilities. We also note that the longest time scale that has been achieved in [Schmitt2020] is around $Jt \sim 10$, with J the local energy scale of the system. This is much shorter than the time scale ($Jt \sim 100-1000$) required to study the long-time heating dynamics. GCNN has also been recently used by Quantinuum to simulate a 56-qubit system classically [Quantinuum2025]. It requires about one week of computation to complete 19 Trotter steps, with results achieving convergence only in the first 3 Trotter steps. So simulating the long-time heating dynamics as performed in our experiment remains impractical at present. Therefore, our systematic comparison of different classical algorithms leads to the conclusion that GMPS is the most efficient option for the current quantum simulation task. However, it still faces formidable challenges in simulating the long-time heating dynamics of our quantum many-body system.

We have also included these discussions in Sec. S1. C. 5 in the revised

[supplementary materials](#).



[Schmitt2020] Markus Schmitt, Markus Heyl, *Quantum Many-Body Dynamics in Two Dimensions with Artificial Neural Networks*, Phys. Rev. Lett. **125**, 100503 (2020).

[Quantinuum2025] Reza Haghshenas, Eli Chertkov, Michael Mills *et al*, *Digital quantum magnetism at the frontier of classical simulations*, arXiv:2503.20870.

Referees' comments:

(2-1) Reviewer #2 commented

I thank the authors for their detailed response and careful revisions. All of my concerns have been addressed and I am happy to recommend the manuscript for publication.

Our response:

[We thank Reviewer #2 for the recommendation.](#)

Referees' comments:

(3-1) Reviewer #3 commented

The revised manuscript and supplemental material is clearly improved in view of my previous comments. However, I am still not convinced, that leakage should be modeled as a depolarizing noise channel.

The authors give two arguments to explain why leakage is depolarizing noise:

1) "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."

Fig. 3(a) in [30] shows dynamics of an imbalance analogous to the present work. The imbalance decays to zero, which is consistent with approaching a maximally mixed state for the individual qubits, which in turn is consistent with depolarizing noise. But it does not prove, that it is a depolarizing noise channel.

2) "We note that a similar error mitigation method has been used in Google's experiment, where random unitaries are implemented [55]"

This fact does not really help my understanding.

My understanding is the following:

Assume two rates γ_0 and γ_1 for leakage from the logical states $|0\rangle$ and $|1\rangle$ out of the computational subspace. Then a leakage error channel for the computational subspace of a single qubit should look something like

$\rho \rightarrow (1-\gamma_0)|0\rangle\langle 0|\rho|0\rangle\langle 0| + (1-\gamma_1)|1\rangle\langle 1|\rho|1\rangle\langle 1| + (\text{offd.})$

terms)

This means that leakage is not trace preserving. Therefore, I don't see how depolarizing noise, which preserves the trace, can be used to model this process.

In view of point 1) above: The decreasing trace of such a leakage channel has the same signature of vanishing imbalance in the long time limit.

I can imagine, that one can actually show, that the employed error mitigation scheme is also suited to address trace-reduction. But if that is what happens, that should be clarified.

If this line of reasoning is wrong, I would appreciate, if the authors could include a brief derivation how leakage becomes depolarizing noise or include a reference, where this is shown. If their point holds, I imagine there is a simple explanation and that could be very helpful for readers.

Our response:

The referee's concern is how to handle leakage from the two-level computational subspace using depolarizing errors. In fact, as detailed in our Supplemental Information (Sec. S5 C2), our treatment is explicitly based on an d -level local Hilbert space, where leakage corresponds to population transfer into the additional high-energy levels. In this enlarged Hilbert space, the evolution remains trace-preserving, and drive-induced heating makes the local reduced density matrix approach a maximally mixed state over all d levels. Our error-mitigation scheme is applied directly in this d -level space, so it naturally incorporates leakage without requiring a trace-preserving map restricted to the computational subspace. Crucially, we have explicitly validated this error-mitigation scheme in the 8-qubit systems, and the resulting time evolution of local observables from the quantum processor precisely follows our numerical exact simulation.

To avoid any possible confusion for readers, we have added the following clarifying sentences to the Supplemental Information in Section S5 C2:

“ 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 hope this fully resolves the referee’s question.