Scalable error mitigation for noisy quantum circuits produces competitive expectation values

Received: 2 October 2021
Accepted: 9 December 2022
Published online: 6 February 2023

Youngseok Kim, Christopher J. Wood, Theodore J. Yoder, Seth T. Merkel, Jay M. Gambetta, Kristan Temme & Abhinav Kandala

Noise in existing quantum processors only enables an approximation to ideal quantum computation. However, for the computation of expectation values, these approximations can be improved by error mitigation. This has been experimentally demonstrated in small systems but the scaling of these methods to larger circuit volumes remains unknown. Here we demonstrate the utility of zero-noise extrapolation for practically relevant quantum circuits using up to 26 qubits, circuit depths of 120 and 1,080 CNOT gates. We study the scaling of the method for canonical examples of product states and entangling Clifford circuits of increasing size, and extend it to simulating the quench dynamics of two-dimensional Ising spin lattices with varying couplings. These experiments reveal that the accuracy of physically relevant observables after error mitigation substantially exceeds previously expected values. Furthermore, we show that the efficacy of error mitigation is greatly enhanced by additional error suppression techniques and native gate decomposition that reduce the circuit time. By combining these methods, the accuracy of our quantum simulation surpasses the classical approximations obtained from an established tensor network method. These results establish the potential of a useful quantum advantage using noisy, digital quantum processors.

Decoherence and unitary gate errors currently limit the volume and fidelity of quantum circuits. Although this can be remedied with the advent of quantum error correction, a fully fault-tolerant hardware architecture is not immediately accessible. Although existing quantum processors have achieved a size that pushes the boundaries of classical simulability, it is important to ask if such noisy machines can perform useful computations. In this context, recent theoretical work has proven that even noisy shallow-depth quantum circuits can outperform their noiseless classical counterparts. Furthermore, recently proposed error mitigation techniques present a path to obtain accurate expectation values even on noisy quantum computers. The general operating principle of these techniques is to reconstruct a noise-free estimate of an expectation value, from multiple noisy experiments. These techniques are particularly attractive for the near term because they typically require no additional overhead for the most valuable resources—number of qubits and circuit depth. Although error mitigation only improves the estimation of expectation values, such computations represent a sizable portion of near-term quantum algorithms, for instance, the estimation of molecular energies in variational quantum eigensolvers, or kernel estimation in machine learning. This has motivated a number of theoretical proposals and experimental demonstrations of error mitigation, which can be broadly categorized into general-purpose and problem-specific techniques.

A widely adopted general-purpose error mitigation technique is zero-noise extrapolation. Here expectation values are measured for

IBM Quantum, IBM T.J. Watson Research Center, Yorktown Heights, NY, USA. e-mail: youngseok.kim1@ibm.com; akandala@us.ibm.com

Nature Physics | Volume 19 | May 2023 | 752–759

752
circuits run at varying noise levels to extrapolate to its zero-noise limit. The noise levels in the circuit are typically varied either by stretching the control pulses in time or by the insertion of noisy identity-equivalent operations. Recent extensions also involve learning the noise amplification factors with near-Clifford training circuits. Initial small-scale experiments with just four qubits of a fixed-frequency superconducting transmon processor and stretching of pulses in entangling cross-resonance gates demonstrated the promise of the technique, enabling variational ansatz with an increased circuit depth to achieve ground-state energies with far improved accuracy for molecular and interacting spin Hamiltonians. However, the hardware platform has witnessed tremendous improvements since this initial demonstration. The size of the largest quantum processors based on this architecture now lies at 65 qubits (ref. 7), reported transmon coherence times are exceeding several hundreds of microseconds and cross-resonance gate fidelities are approaching 99.9% with sub-100-ns gate times, as well as steady progress in holistic metrics. This raises a tantalizing question of how much these hardware improvements could extend the reach of zero-noise extrapolation to perform accurate quantum computation at a larger scale, into the realm of quantum advantage. In this context, once we approach a scale where the general performance of an error mitigation technique cannot be compared with exact numerics, it is particularly important to understand the error bounds, as discussed elsewhere. For zero-noise extrapolation, an upper bound on the error of the zero-noise estimate after nth-order Richardson extrapolation is \( O(NT\lambda)^n \), where \( N \) is the number of qubits, \( \lambda \) is a parameter that describes the strength of noise (for example, \( \lambda \approx \lambda_0 \) for purely amplitude-damping noise with energy relaxation time \( T_\lambda \)) and \( T \) is the total evolution time. Therefore, the performance of zero-noise extrapolation is expected to be limited by \( NT\lambda \) being a small number. This would also imply a fairly restrictive path to scaling—hardware noise rates would need to scale with increasing circuit volumes.

In this work, we experimentally study the performance of zero-noise extrapolation for circuits employing up to 26 transmon qubits of a fixed-frequency quantum processor, for a maximum circuit depth of 120 with a maximum CNOT count of 1,080 gates (Fig. 1). This represents a substantial increase in circuit volume from initial demonstrations (four qubits and circuit depth of 6), which would have required an -130 times improvement in coherence, based on the bounds discussed above. However, our experiments demonstrate that the realistic bounds for several observables of interest are far more accessible, evidenced by the efficacy of mitigation in the absence of any such dramatic hardware improvement. We initially present a couple of canonical examples—the \( T_1 \) decay of multiqubit product states and 21-qubit Greenberger–Horne–Zeilinger (GHZ) states. Even for a fixed number of qubits \( N \), the errors in extrapolation display circuit-specific dependencies on the weight and locality of mitigated observables. We then extend our experiments to non-Clifford circuits, studying the time dynamics of a 26 spin Ising Hamiltonian for varying system parameters. We show that the performance of zero-noise extrapolation can be significantly enhanced by suppressing \( \lambda \) with additional error suppression strategies and reducing the circuit evolution time \( T \) with more efficient gate decomposition and compiling (Fig. 1). With these improvements, we show that we can already begin to achieve an accuracy on error-mitigated observables, measured off a noisy quantum processor, which is able to surpass the accuracy of an established tensor network method in the simulation of dynamics.

We perform our experiments on a superconducting processor with 27 transmon qubits and fixed connectivity in the heavy-\( \Delta \) geometry. The connectivity is favourable towards frequency imprecision and crosstalk of fixed-frequency superconducting architectures, as well as enabling the implementation of quantum error-correcting codes. The control and readout of the qubits are performed with microwave pulses, with entangling operations based off the cross-resonance effect, with a native ZX interaction. Additional details of this device are provided in Supplementary Information.

First, we study the performance of zero-noise extrapolation for trivial, multiqubit product states prepared by simultaneous \( T_1 \) decay circuits on an initial \( |\psi\rangle = |0\rangle^\otimes N \) state (Fig. 2a, inset). As shown in the original proposal, the noise can be amplified by stretching the circuit’s time evolution, under the assumption of time-invariant noise. Here this amounts to simply stretching the wait times in the \( T_1 \) decay circuits. To reduce the effect of coherence fluctuations, all the stretched circuits in this work are averaged together, as discussed previously. Unless otherwise specified, all the experiments discussed in the text employ 100,000 shots for every stretch factor, and the reported expectation values are corrected for measurement error using a tensor product of individual qubit readout calibration matrices. Figure 2a depicts the decay of \( \langle Z \rangle \) for the considered stretch factors \( c = 1.0, 1.5 \) and 2.0, as well as the mitigated observable constructed by linear extrapolation. The mitigated observable shows far superior accuracy over the raw data, even for circuit times going up to 100 \( \mu s \), comparable to the \( T_1 \) value of the measured qubit. On a line of five qubits, we then compare the accuracy of increasing weight-Z observables after linear extrapolation, going up to \( \langle ZZZZZ \rangle \), for varying delay times. As shown in Fig. 2b, the lower-weight observables display superior accuracy, for the same state-preparation circuit/number of qubits.

Next we study the performance of zero-noise extrapolation for low-weight observables for large entangled states. We prepare a GHZ state by sequentially applying CNOTs for 21 qubits along the longest one-dimensional chain in the layout shown in Fig. 2d. There are two sets of observables we examine: the local observables \( Z \) and the non-local observables \( Z_1 \) for \( j = 1, 2, \ldots N = 20 \), all of which have ideal expectations of +1. The CNOT gates are constructed from the calibrated cross-resonance sequences, and all the pulses in the circuit are extended by the considered stretch factors \( c = 1.0, 1.3 \) and 1.6. The circuit structure leads to long idling times, and we see that dynamical decoupling sequences (Fig. 1) suppress qubit dephasing and enhance the quality of error mitigation. Figure 2c,d illustrates the unmitigated and mitigated expectation values obtained from linear extrapolation. Due to the circuit structure, unmitigated values of these observables suffer \( j \)-dependent error rates. The non-local observables are especially exposed to more sources of errors for increasing \( j \). For example, consider a simple error model where \( X \) errors occur after both two-qubit and one-qubit gates (including idle locations) in the circuit, but with different probabilities \( p_0 \) and \( p_j \). We also make the simplifying but largely realistic assumption that qubits in the ground state (that is, before any gate is applied) do not suffer errors, except for some initialization \( X \) error at rate \( p_0 \). A specific observable is sensitive to errors only at a subset of locations. Generally, if \( X \) error at any of the \( E_i \) two-qubit locations, \( E_i \) one-qubit locations or \( E_i \) initializations flips observable \( O \), the probability that its expectation is -1 is

\[
P_{\text{err}} = \frac{1}{2}(1 - (1 - 2p_j)^{E_i}(1 - 2p_j)^{E_i}(1 - 2p_j)^{E_i}),
\]

and thus, \( \langle O \rangle = (1 - 2p_j)^{E_i}(1 - 2p_j)^{E_i}(1 - 2p_j)^{E_i} \). A local observable \( Z_1 \) is sensitive to errors on qubits \( j - 1 \) and \( j \) after the CNOT from \( j - 1 \) to \( j \) and one of the two initializations, implying \( E_0 = 1 \), \( E_1 + E_2 = 2(N - j) \) and \( E_2 = 3 \) for \( j < N - 1 \) and \( E_1 = 2 \) for \( j = N - 1 \). For a non-local observable \( Z_1 \), \( E_1 + E_2 = 2(N - 1) \) is constant with \( j \), but \( E_2 = 2 + j \) and \( E_6 = j \). Assuming \( p_0 \), \( p_j \) and small \( N \), we, therefore, expect a near-linear increase in the local expectations \( \langle Z_1, Z_1 \rangle \) with \( n \) and an exponential decrease in the non-local expectations \( \langle Z_i, Z_i \rangle \). These trends are present in the experimental data. Consequently, error-mitigated observables follow the same trends. This suggests that one should pay careful attention to the circuit structure when applying zero-noise extrapolation, and choose to mitigate observables that are less sensitive to noise, such as local, low-weight observables. Although this might appear to be restrictive, we note that
the performance of zero-noise extrapolation, which are not incorporated in the appropriately scaling the duration and amplitude of the microwave pulses that compose the gates. A number of additional techniques are employed to enhance the performance of zero-noise extrapolation, which are not incorporated in the bare circuit. The insertion of dynamical decoupling $X_\tau X_\tau$ sequences is used to suppress dephasing during qubit idle times. Pauli twirling averages out the off-diagonal coherent errors in the Pauli basis and is implemented by sandwiching the two-qubit gates with additional single-qubit gates drawn from a set of Pauli operations. The twirling gates can be combined with original single-qubit gates into arbitrary single-qubit rotations $U(\theta, \phi) = R_x(\theta)R_y(\pi/2)R_z(\phi)$ to minimize any additional circuit time. The total evolution time is dominated by two-qubit gates, and can be reduced by native gate decomposition. The standard decomposition of $R_{z\theta}(\theta)$ operations for arbitrary $\theta$ uses two CNOT gates, each of which is constructed on our hardware using cross-resonance pulses to tune fully entangling $R_{z\theta}(\pi/2)$ operations. Alternately, the circuit time can be substantially reduced by constructing $R_{z\theta}(\theta)$ gates from partially entangling cross-resonance pulses that implement $R_{z\theta}(\theta)$.

several physical Hamiltonians of interest are indeed described by such observables. This benefit is also highlighted in noisy numerics extending to larger system sizes—the accuracy of several observables can be substantially improved by zero-noise extrapolation for 100-qubit GHZ circuits for realistic error rates (Supplementary Section IV).

Finally, we discuss the performance of zero-noise extrapolation on short-depth non-Clifford circuits. As an example, we study the quench dynamics of a transverse-field Ising model. The problem entails studying the evolution of a system under the following Hamiltonian:

$$H = -J \sum_{\langle i, j \rangle} Z_i Z_j + h \sum_i X_i,$$

where $J$ is the strength of the exchange coupling between the nearest-neighbour spins with indices $\langle i, j \rangle$ and $h$ is the transverse-magnetic field. The study of quantum quenches with interacting spin-1/2 systems provides a rich playground for explorations of fundamental questions in condensed-matter and statistical physics\textsuperscript{34}, and have already been extensively explored on analogue quantum simulators\textsuperscript{35-37}. These models are also becoming a particularly attractive application of existing noisy, digital quantum computers\textsuperscript{38-41} with their increased control and addressability. The spins can be naturally encoded in the physical qubits, nearest-neighbour interactions are accessible with the local connectivity of qubits and relevant quantities can be measured by local low-weight observables. Although this Hamiltonian can be fully diagonalized by a spin-to-fermion mapping in one dimension and even analytical solutions are known for translationally invariant chains, no such procedure is known for the two-dimensional (2D) model. This makes the 2D Ising model an attractive testbed for the pursuit of quantum advantage on near-term quantum processors\textsuperscript{42}.

Here we study the quench dynamics of a 2D spin lattice that follows the connectivity of our processor, excluding a single inferior qubit. Starting in the ground state of the Hamiltonian at zero field ($h = 0$) with the qubits initialized in $|\psi_0\rangle = |0^{26}\rangle$, the time evolution of the system is studied once the transverse field is suddenly turned on at time $\tau = 0$.

$$|\psi(\tau)\rangle = e^{-iH\tau} |\psi_0\rangle$$

\textsuperscript{2}
We implement the dynamics on our hardware by a first-order Trotter decomposition of the time evolution.

\[ e^{-iHt} = e^{-iH_1t}e^{-iH_2t} \approx \prod_i e^{-iH_1 \delta t}e^{-iH_2 \delta t}, \]

where \( H_1 = -\sum_i \Delta Z_i \) and \( H_2 = \hbar \sum X_i \). The maximum evolution time \( T \) is discretized into time steps \( \delta t = T/n \), with \( n \) Trotter layers, and

\[ e^{-iH_1 \delta t} = \prod_k \exp \left( i \delta t \Delta Z_k \right), \]

\[ e^{-iH_2 \delta t} = \prod_x \exp \left( -i \hbar \delta t X_x \right). \]

Unless otherwise specified, the experiments described here fix \( \hbar = 1 \), \( T = 10 \), \( \delta t = 0.5 \) and study the dynamics for varying \( J \). The evolution can now be implemented via a combination of single-qubit rotations and ZZ two-qubit gates between pairs of connected qubits on the lattice. We study the dynamics for up to 26 spins on a 2D lattice and \( n = 20 \) Trotter steps; Fig. 3a,b shows that each Trotter layer implements three blocks of parallelized two-qubit gates on the next-nearest neighbours for a maximum circuit depth of 120.

Although a comparison to the ideal result—obtained by exact numerical simulation—is possible for sizes up to 26 spins, a mere increase in system size by a factor of 2–3 renders this unfeasible. At this larger scale, only approximate classical methods are available. It is important to point out here that the quantum experiment itself only provides an approximation to the idealized simulation due to hardware noise. To establish the validity of approximate quantum computation with error mitigation, it is, therefore, prudent to compare with established approximate classical methods that scale efficiently.

Simulating the time dynamics of general quantum systems is bounded-error quantum polynomial time-complete—i.e., the class of decision problems solvable in polynomial time by a quantum computer. No classical approximation method is, therefore, expected to produce a universally accurate approximation. However, a wide range

**Fig. 2** Scaling of zero-noise extrapolation for product states and entangling Clifford circuits. a, Error-mitigated (red) and unmitigated decay of \( \langle Z \rangle \) for a standard \( T \) decay circuit. The unmitigated experiments are performed for three different stretch factors, namely, \( c = 1.0, 1.5 \) and \( 2.0 \), by stretching the \( T \) delay times (denoted as \( \text{Id} \) in the inset of b). The mitigated curve is obtained by a linear extrapolation of the observables measured from the stretched experiments. The inset depicts the device connectivity and highlights the five qubits employed for the experiments shown in this figure. b, Error-mitigated decay of increasing weight \( Z \) observables, from the simultaneous measurement of \( \langle Z_i^a \rangle \) for a standard \( T \) decay circuit. The unmitigated experiments are performed for three different stretch factors, namely, \( c = 1.0, \ 1.3 \) and 1.6. The inset in d depicts the GHZ state-preparation circuit. The qubit distance for the \( x \) axis in d is defined by the difference in the qubit index for non-local observables. All the experiments shown in this figure have ideal expectation values with a magnitude of 1. The mean value of the error-mitigated results in a–d are computed from 50 numerical experiments by bootstrapping the experimental distributions and standard deviations are provided as an error bar.
of approximate classical simulation methods have been developed that provide good approximations in various limiting cases. Starting from a limiting case where the approximation works well, we want to provide a comparison with quantum approximation and see for which parameter values the experimental expectation values provide a higher accuracy. Here we focus on a comparison with tensor network methods since our experiment considers a locally interacting spin Hamiltonian. A quasi-one-dimensional method such as matrix product states scales exponentially in the shorter lattice length to preserve locality. However, it is still a practical method for the lattice considered in this experiment and can be made to closely track the exact numerical evolution with a bond dimension of only $512$. Since we consider the dynamics of 2D systems and will eventually be interested in scaling to larger lattices, we compare the experimental results with the standard projected entangled pair state (PEPS) method. These states are the current prime candidate to capture the ground-state properties of gapped 2D spin systems.
This method can also be used to approximate the dynamics of weakly entangled systems, but it is generally expected that the PEPS simulation will become inaccurate when strong entanglement is produced. Here we consider the original PEPS time evolution algorithm with bond dimensions $D = 4$. The algorithm scales as $O(D^N N^2)$ in time and $O(D^3 N)$ in memory, and can become numerically unstable for larger bond dimensions. At each Trotter step, we estimate the average magnetization of the interacting spin system by a measurement of the weight-1 observables averaged over 26 qubits—$(X, Y)$ and $(Z)$. We quantify and compare with the experimental error in terms of the normalized Euclidean distance between the magnetization vectors obtained from PEPS and the ideal vector. Figure 3c–e shows that PEPS exhibits excellent agreement with the exact numerics over the entire range of Trotterized evolution when the $R_{\delta t}(\theta)$ gates are weakly entangling for $J = 0.1$, but displays increasingly poorer accuracy with increasing $J$ and Trotter steps.

With our digital approach, the time evolution for different coupling parameters can be experimentally studied by simply varying the angle of single-qubit rotations $R_y(2\theta)$ and two-qubit gates $R_{\delta t}(-2\theta)$.

Arbitrary $R_{\delta t}(\theta)$ gates can be implemented on our device by using two CNOT gates and additional single-qubit rotations, where the CNOT gates are themselves constructed from cross-resonance-driven $R_{\delta t}(\pi/2)$ interactions that are native to our hardware. This enables us to compare the performance of zero-noise extrapolation on circuits with the same structure and evolution time, and systematically increasing the amount of entanglement with increasing $J$. Error mitigation is ultimately limited by the noise in the circuit, and to extend its reach, we adopt additional error reduction strategies such as dynamical decoupling, previously discussed for GHZ circuits. Furthermore, an important source of crosstalk in fixed-frequency architectures is the ZZ interaction, which can be an important limitation to circuit fidelity and particularly problematic for circuits simulating Ising Hamiltonians. In this context, Pauli twirling is an attractive approach to average out the off-diagonal coherent errors of the circuit in the Pauli basis and improve the circuit fidelity. It also enables the suppression of coherent errors in gate rescaling that could otherwise lead to unphysical extrapolations. Here each two-qubit gate is sandwiched between twirling gates and inverting gates such that the same unitary is implemented; for two-qubit Clifford operations such as CNOT gates, the twirling gates can be randomly sampled from a set of tensor products of all the single-qubit Pauli gates. By incorporating these error suppression strategies with linear extrapolation, we estimate the magnetization and the corresponding error for $J = 0.1$, 0.3, and 0.5236 (Fig. 3c–e, respectively). Interestingly, the accuracy of error-mitigated magnetization does not reveal a strong dependence on the degree of entanglement in experiment. As a consequence, our error-mitigated experiment provides an increasingly competitive approximation to the magnetization over PEPS for increasing $J$. Therefore, we focus on the case of $J = 0.5236$, which specifically requires the implementation of $R_{\delta t}(\pi/6)$ gates for our chosen Hamiltonian and Trotterization parameters.

With CNOT decomposition, our longest circuit times exceed 80 $\mu$s, limiting the accuracy of mitigated observables. For $J = 0.5236$, the performance of our experiment can be further enhanced by employing a more time-effective gate decomposition that uses shorter cross-resonance pulses to calibrate partially entangling $R_{\delta t}(\pi/6)$ gates. Details of the calibration and benchmarking are discussed in Supplementary Information. With this decomposition, we achieve an almost four times reduction in circuit time. These considerations and strategies highlight the benefit of pulse-level stretching for error amplification and the associated freedom over choice of stretch factors. For large-volume circuits, error amplification via the insertion of identity-equivalent gates can lead to circuit times beyond the coherence budget. Additionally, we show that a Pauli twirling strategy can also be employed with the non-Clifford fractional $R_{\delta t}(\theta)$ gates, although with a limited set of twirling gates. Even though the partial twirling set and limited number of twirling instances may not guarantee the Pauli-ness of the noise channel, we nevertheless observe clear improvements in the quality of the observable (Supplementary Section II-B).

The impact of the combined error suppression strategies is visualized in Fig. 3f by tracking the evolution of magnetization vector. Furthermore, the error-mitigated trajectory displays excellent qualitative agreement with the ideal noiseless path (Fig. 3g), whereas the trajectory obtained from PEPS diverges at increasing Trotter steps. With the optimal gate decomposition, we see that the error-mitigated experiment provides a closer approximation to magnetization than PEPS even for Trotter steps exceeding $n = 6$ for the largest value of $J = 0.5236$ considered in this work (Fig. 3e). However, we also note that the accuracy of larger-weight observables, after error mitigation, is not as competitive (Supplementary Section III). We also compare the magnetization error for an increasing number of qubits (and varying 2D lattices), and map this for increasing numbers of Trotter steps (Fig. 3h). We see that our experiment is closer to the ideal simulation than PEPS for a large part of this phase space, particularly for increasing number of qubits and Trotter steps. These observations of accurate coherent dynamics in error-mitigated expectation values going up to 26 qubit circuits with depth $d = 60$, in conjunction with the scaling of PEPS ($D = 4$) error (Fig. 3i), raise the enticing possibility of studying quench dynamics on noisy, digital and fully programmable quantum processors that are beyond exact diagonalization as well as efficient, classical approximate techniques.

The experiments discussed here highlight several important considerations to access competitive computations from noisy, quantum processors. We demonstrate that zero-noise extrapolation can be well beyond the restrict circuit volumes defined by previous theoretical bounds, to improve the quality of physically relevant observables and show that a number of experimental strategies can be employed to further enhance its performance on a noisy device. Having demonstrated its scalability, future efforts to improve noise scaling will enable higher-order Richardson extrapolation to access an even greater degree of accuracy (Supplementary Sections IV and V), which we expect to be continually supplemented by improvements in coherence and quality of quantum hardware. With existing error rates and linear extrapolation, we already see that the error-mitigated magnetization dynamics of 2D Ising quenches for 26 spins can provide a better approximation than PEPS at $D = 4$ in the limit of increasing entanglement. In this context, the development of 65-qubit processors based on our hardware platform, with comparable error rates and 2D connectivity, makes these devices an attractive platform for the exploration of error-mitigated quantum dynamics of 2D spin systems that are classically intractable.

**Online content**

Any additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41567-022-01914-3.

**References**

1. Shor, P. W. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A* 52, R2493 (1995).

2. Steane, A. M. Error correcting codes in quantum theory. *Phys. Rev. Lett.* 77, 793–797 (1996).

3. Zhang, E. J. et al. High-performance superconducting quantum processors via laser annealing of transmon qubits. *Sci. Adv.* 8, eabf6690 (2022).

4. Arute, F. et al. Quantum supremacy using a programmable superconducting processor. *Nature* 574, 505–510 (2019).

5. Wu, Y. et al. Strong quantum computational advantage using a superconducting quantum processor. *Phys. Rev. Lett.* 127, 180501 (2021).
6.  Bravyi, S., Gosset, D., Koenig, R. & Tomamichel, M. Quantum advantage with noisy shallow circuits. Nat. Phys. 16, 1040–1045 (2020).
7.  Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. Phys. Rev. Lett. 119, 180509 (2017).
8.  Li, Y. & Benjamin, S. C. Efficient variational quantum simulator incorporating active error minimization. Phys. Rev. X 7, 021050 (2017).
9.  Kandala, A. et al. Error mitigation extends the computational reach of a noisy quantum processor. Nature 567, 491–495 (2019).
10. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. Nat. Commun. 5, 4213 (2014).
11. Kandala, A. et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. Nature 549, 242–246 (2017).
12. Havlíček, V. et al. Supervised learning with quantum-enhanced feature spaces. Nature 567, 209–212 (2019).
13. Schuld, M. & Killoran, N. Quantum machine learning in feature Hilbert spaces. Phys. Rev. Lett. 122, 040504 (2019).
14. Dumitrescu, E. F. et al. Cloud quantum computing of an atomic nucleus. Phys. Rev. Lett. 120, 210501 (2018).
15. Song, C. et al. Quantum computation with universal error mitigation on a superconducting quantum processor. Sci. Adv. 5, eaaw5686 (2019).
16. Zhang, S. et al. Error-mitigated quantum gates exceeding physical fidelities in a trapped-ion system. Nat. Commun. 11, 587 (2020).
17. Koczor, B. Exponential error suppression for near-term quantum devices. Phys. Rev. X 11, 031057 (2021).
18. Huggins, W. J. et al. Virtual distillation for quantum error mitigation. Phys. Rev. X 11, 041036 (2021).
19. McClean, J. R., Kimchi-Schwartz, M. E., Carter, J. & de Jong, W. A. Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. Phys. Rev. A 95, 042308 (2017).
20. Bonet-Monroig, X., Sagastizabal, R., Singh, M. & O’Brien, T. E. Low-cost error mitigation by symmetry verification. Phys. Rev. A 98, 062339 (2018).
21. McArdle, S., Yuan, X. & Benjamin, S. Error-mitigated digital quantum simulation. Phys. Rev. Lett. 122, 180505 (2019).
22. He, A., Nachman, B., de Jong, W. A. & Bauer, C. W. Zero-noise extrapolation for quantum-gate error mitigation with identity insertions. Phys. Rev. A 102, 012426 (2020).
23. Lowe, A. et al. Unified approach to data-driven quantum error mitigation. Phys. Rev. Res. 3, 033098 (2021).
24. Czarnik, P., Arrasmith, A., Coles, P. J. & Cincio, L. Error mitigation with Clifford quantum-circuit data. Quantum 5, 592 (2021).
25. Koch, J. et al. Charge-insensitive qubit design derived from the Cooper pair box. Phys. Rev. A 76, 042319 (2007).
26. Chow, J. M. et al. Simple all-microwave entangling gate for fixed-frequency superconducting qubits. Phys. Rev. Lett. 107, 080502 (2011).
27. Gordon, R. T. et al. Environmental radiation impact on lifetimes and quasiparticle tunneling rates of fixed-frequency transmon qubits. Appl. Phys. Lett. 120, 074002 (2022).
28. Wei, K. X. et al. Hamiltonian engineering with multicolor rives for fast entangling gates and quantum crosstalk cancellation. Phys. Rev. Lett. 129, 060501 (2022).
29. Jurcevic, P. et al. Demonstration of quantum volume 64 on a superconducting quantum computing system. Quantum Sci. Technol. 6, 025020 (2021).
30. Murg, V., Verstraete, F. & Cirac, J. I. Variational study of hard-core bosons in a two-dimensional optical lattice using projected entangled pair states. Phys. Rev. A 75, 033605 (2007).
31. Chamberland, C., Zhu, G., Yoder, T. J., Hertzberg, J. B. & Cross, A. W. Topological and subsystem codes on low-degree graphs with flag qubits. Phys. Rev. X 10, 011022 (2020).
32. Hertzberg, J. B. et al. Laser-annealing Josephson junctions for yielding scaled-up superconducting quantum processors. npj Quantum Inf. 7, 129 (2021).
33. Carroll, M., Rosenblatt, S., Jurcevic, P., Lauer, I. & Kandala, A. Dynamics of superconducting qubit relaxation times. npj Quantum Inf. 8, 132 (2022).
34. Sachdev, S. Quantum phase transitions. Phys. World 12, 33 (1999).
35. Browaeys, A. & Lahaye, T. Many-body physics with individually controlled Rydberg atoms. Nat. Phys. 16, 132–142 (2020).
36. Bernien, H. et al. Probing many-body dynamics on a 51-atom quantum simulator. Nature 551, 579–584 (2017).
37. Zhang, J. et al. Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator. Nature 551, 601–604 (2017).
38. Smith, A., Kim, M. S., Pollmann, F. & Knolle, J. Simulating quantum many-body dynamics on a current digital quantum computer. npj Quantum Inf. 5, 106 (2019).
39. Sopena, A., Gordon, M. H., Sierra, G. & López, E. Simulating quench dynamics on a digital quantum computer with data-driven error mitigation. Quantum Sci. Technol. 6, 045003 (2021).
40. Vovrosh, J. et al. Simple mitigation of global depolarizing errors in quantum simulations. Phys. Rev. E 104, 035309 (2021).
41. Urbanek, M. et al. Mitigating depolarizing noise on quantum computers with noise-estimation circuits. Phys. Rev. Lett. 127, 270502 (2021).
42. Childs, A. M., Maslov, D., Nam, Y., Ross, N. J. & Su, Y. Toward the first quantum simulation with quantum speedup. Proc. Natl Acad. Sci. USA 115, 9456–9461 (2018).
43. Nielsen, M. A. & Chuang, I. L. Quantum Computation and Quantum Information: 10th Anniversary Edition (Cambridge Univ. Press, 2010).
44. Moyal, J. E. Quantum mechanics as a statistical theory. Math. Proc. Camb. Philos. Soc. 45, 99–124 (1949).
45. Keldysh, L. V. et al. Diagram technique for nonequilibrium processes. Sov. Phys. JETP 20, 1018–1026 (1965).
46. White, S. R. & Feiguin, A. E. Real-time evolution using the density matrix renormalization group. Phys. Rev. Lett. 93, 076401 (2004).
47. Orús, R. A practical introduction to tensor networks: matrix product states and projected entangled pair states. Ann. Phys. 349, 117–158 (2014).
48. Vidal, G. Efficient classical simulation of slightly entangled quantum computations. Phys. Rev. Lett. 91, 147902 (2003).
49. Verstraete, F. & Cirac, J. I. Renormalization algorithms for quantum-many body systems in two and higher dimensions. Preprint at https://arxiv.org/abs/cond-mat/0407066 (2004).
50. Verstraete, F., Murg, V. & Cirac, J. I. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. Adv. Phys. 57, 143–224 (2008).
51. Schuch, N., Wolf, M. M., Vollbrecht, K. G. H. & Cirac, J. I. On entropy growth and the hardness of simulating time evolution. New J. Phys. 10, 033032 (2008).
52. Bennett, C. H. et al. Purification of noisy entanglement and faithful teleportation via noisy channels. Phys. Rev. Lett. 76, 722–725 (1996).
53. Knill, E. Fault-tolerant postselected quantum computation: threshold analysis. Preprint at https://arxiv.org/abs/quant-ph/0404104 (2004).
54. Wallman, J. J. & Emerson, J. Noise tailoring for scalable quantum computation via randomized compiling. Phys. Rev. A 94, 052325 (2016).
55. Hashim, A. et al. Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor. Phys. Rev. X 11, 041039 (2021).

56. Earnest, N., Tornow, C. & Egger, D. J. Pulse-efficient circuit transpilation for quantum applications on cross-resonance-based hardware. Phys. Rev. Res. 3, 043088 (2021).

57. Stenger, J. P. T., Bronn, N. T., Egger, D. J. & Pekker, D. Simulating the dynamics of braiding of Majorana zero modes using an IBM quantum computer. Phys. Rev. Res. 3, 033171 (2021).

58. Sundaresan, N. et al. Reducing unitary and spectator errors in cross resonance with optimized rotary echoes. PRX Quantum 1, 020318 (2020).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023
Methods
Partial gate calibration
The fractional $R_{Z\theta}(\theta)$ gates employed for the quench dynamics were constructed from $R_{Z\theta}(\theta)$ sequences. Specifically, the $R_{Z\theta}(\theta = \pi/6)$ gates considered here were calibrated to $R_{Z\theta}(\pi/2)$ by repeating the gate three times. More details are described in Supplementary Information.

Dynamical decoupling
We insert the sequence $\tau/4 - R_x(\pi) - \tau/2 - R_x(-\pi) - \tau/4$ for every idling period of time, where $\tau$ is the idling time minus the length of two $R_x(\pi)$ pulses. The impact of dynamical decoupling on the circuit is further highlighted in Supplementary Information.

Pauli twirling
This is implemented by sandwiching Clifford gates between randomly sampled twirling gates such that the net operation—in the absence of noise—is unchanged. The twirling gates inserted before the Clifford operation are typically sampled from a complete set of Pauli operations, namely, $G = \{I, X, Y, Z\}^\otimes N$ for $N$ qubits, and an appropriate set of Pauli operations are then applied after the Clifford gate for a net equivalent unitary operation. For our quench circuits with the CNOT decomposition of $R_{Z\theta}(\theta)$ operations, each CNOT (Clifford) gate is sandwiched between gates randomly sampled from $G$ and Pauli gates that preserve the net CNOT operation, and the expectation values are then averaged over multiple twirling instances. For each CNOT, the twirling gates are combined with other single-qubit operations in the circuit to avoid any increase in circuit time.

The native gate decomposition for $R_{Z\theta}(\theta)$ substantially shortens the circuit time but involves non-Clifford gates. For this decomposition, we randomly sample twirling gates on the [control, target] pair from subset $G_{Z\theta} \subseteq G$ that commute with $R_{Z\theta}(\theta)$, that is, $G_{Z\theta} = \{I, I, X, X, Y, Y, Z, Z\}$. For every draw, the same set of twirling gates are applied before and after $R_{Z\theta}(\theta)$ to implement the same unitary. The impact of Pauli twirling has been further discussed in Supplementary Information.

Data availability
Source data are available for this paper. All other data that support the plots within this Article and other findings of this study are available from the corresponding author upon reasonable request.

Acknowledgements
We thank I. Lauer for contributions towards two-qubit gate calibration, and D. T. McClure and N. Sundaresan for valuable discussions on the device bring-up.

Author contributions
A.K. and K.T. designed the experiments. Y.K. and A.K. performed the experiments. C.J.W. and T.J.Y. ran the noisy simulations. S.T.M. provided the code for implementing Pauli twirling. K.T. ran the tensor network simulations. Y.K., C.J.W., T.J.Y., J.M.G., K.T. and A.K. analysed the data and wrote the paper.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41567-022-01914-3.

Correspondence and requests for materials should be addressed to Youngseok Kim or Abhinav Kandala.

Peer review information Nature Physics thanks Xiao Yuan, Jinzhao Sun, Yukun Zhang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.