npj | quantum information

Article

Published in partnership with The University of New South Wales

https://doi.org/10.1038/s41534-024-00892-z

Extending the computational reach of a superconducting qutrit processor

Noah Goss[1,2,6] ✉, Samuele Ferracin[3,6], Akel Hashim[1], Arnaud Carignan-Dugas[3], John Mark Kreikebaum[2,4,5], Ravi K. Naik[2], David I. Santiago[1,2] & Irfan Siddiqi[1,2,4]

Quantum computing with qudits is an emerging approach that exploits a larger, more connected computational space, providing advantages for many applications, including quantum simulation and quantum error correction. Nonetheless, qudits are typically afflicted by more complex errors and suffer greater noise sensitivity which renders their scaling difficult. In this work, we introduce techniques to tailor arbitrary qudit Markovian noise to stochastic Weyl–Heisenberg channels and mitigate noise that commutes with our Clifford and universal two-qudit gate in generic qudit circuits. We experimentally demonstrate these methods on a superconducting transmon qutrit processor, and benchmark their effectiveness for multipartite qutrit entanglement and random circuit sampling, obtaining up to 3× improvement in our results. To the best of our knowledge, this constitutes the first-ever error mitigation experiment performed on qutrits. Our work shows that despite the intrinsic complexity of manipulating higher-dimensional quantum systems, noise tailoring and error mitigation can significantly extend the computational reach of today's qudit processors.

In the noisy intermediate-scale quantum (NISQ) era[1], the ability to effectively couple many quantum two-level systems (qubits) together has led to experimental demonstrations of certain tasks that challenge the limits of current classical capabilities[2–5]. The computational power of these near-term devices can potentially be further boosted by leveraging the innate multi-level structure to encode quantum information in the larger and more connected Hilbert space of d-level systems (qudits)[6–11]. Coherent control of higher-level quantum systems has been demonstrated in superconducting circuits[12–18], trapped ions[19,20], and in photonic circuits[21,22].

The simplest and most immediately experimentally viable member of the qudit family is the quantum three-level system or qutrit. Qutrits can yield specific advantages in quantum simulations, where they are a natural platform for studying spin-1 physics[23,24] and robust and resource efficient for simulating high-energy phenomena[12,25]. Additional applications include improvements in quantum cryptography[26,27], communication[28], compactly synthesizing multi-qubit gates[29–34], and improving qubit readout[35–37]. Eventually, qutrits are expected to provide significant advantages for quantum error correction via improved error thresholds[38–41], errors tailored to erasure[42,43], enhanced fault tolerance[44,45], and compact encodings of both logical qutrits[46] and logical qubits[47,48].

While fault tolerance remains the ultimate goal, alternative efforts in qubit devices to mitigate and extrapolate expectation values beyond the noise present in the system have garnered interest lately due to the lack of an increase in hardware requirements and overall feasibility[49–64]. Notably, recent works have demonstrated that error mitigation protocols can prove effective for large-scale qubit experiments, allowing an exciting pathway to high-fidelity results in the near term for problems of interest[51–53,57,59,61–66]. Recent investigations have also explored the compatibility of noise tailoring, error mitigation, and quantum error correction[67,68]. However, there is a dearth of similar studies for qudit devices, which are now approaching the maturity to scale to larger experimental sizes but are afflicted by more complicated noise processes[69]. This opens the door to an interesting question: can the computational power of contemporary devices be augmented by enlarging their Hilbert space as qudits, while at the same time retaining the ability to generate noiseless expectation values in this more complex noise environment?

In this work, we attempt to answer this question and show that error mitigation can be utilized to significantly extend the existing resources of a superconducting qutrit processor. More explicitly, we introduce two techniques for tailoring and mitigating noise in qutrit circuits: randomized compiling (RC)[68,70] and noiseless output extrapolation (NOX)[66]. We then explore the performance of these techniques in a variety of multi-qutrit experiments using fixed frequency transmons. In particular, we study state tomography of a three-qutrit Greenberger–Horne–Zeilinger (GHZ) state, as

[1]Quantum Nanoelectronics Laboratory, Department of Physics, University of California at Berkeley, Berkeley, CA, 94720, USA. [2]Applied Mathematics and Computational Research Division, Lawrence Berkeley National Lab, Berkeley, CA, 94720, USA. [3]Keysight Technologies Canada, Kanata, ON, K2K 2W5, Canada. [4]Materials Sciences Division, Lawrence Berkeley National Lab, Berkeley, CA, 94720, USA. [5]Present address: Google Quantum AI, Mountain View, CA, 94043, USA. [6]These authors contributed equally: Noah Goss, Samuele Ferracin. ✉e-mail: noahgoss@berkeley.edu

well as random circuit sampling (RCS) with two and three qutrits. We find that in all cases, despite the more complex noise environment, our results benefit greatly from these protocols, achieving up to a 3 times improvement in fidelity. Our work is the first to experimentally demonstrate that error mitigation can be effectively implemented on qutrit platforms, and it paves the way to scaling near-term, large-scale qutrit computations.

Results

Computing, twirling, and mitigating with qutrits

Computing. With qutrits, information is encoded in the three energy levels that correspond to the computational basis states $|0\rangle$, $|1\rangle$, and $|2\rangle$. As for qubits, gates are unitary operators relative to the computational basis. An important set of qutrit gates is the Weyl operators, which are a generalization of the of the qubit Pauli group in a higher dimensional Hilbert space. The action of the Weyl gates is specified by the operators $W_{p,q} = \omega^{-pq/2} Z^p X^q$, where X and Z are defined by their action on the basis state $|n\rangle$ as $X|n\rangle = |n \oplus 1\rangle$ and $Z|n\rangle = \omega^n |n\rangle$ and $\omega = e^{2i\pi/3}$. Like the Pauli gates for qubit systems, the Weyl operators form a unitary 1-design[71] and are normalized by the qutrit Clifford group. That is, by definition[72], given a gate G, we have that a product of local Weyl gates $\otimes_i W_{p_i,q_i}$ conjugated by G remains a product of Weyl gates, e.g.,

$$G(\otimes_i W_{p_i,q_i})G^{-1} = \otimes_i W_{p_i',q_i'}, \tag{1}$$

if and only if G is part of the qutrit Clifford group.

Natural generalizations of two-qubit gates can also be constructed. For example, the two-qutrit analog of the controlled-Z (CZ) gate is defined as,

$$\mathrm{CZ} = \sum_{n=0}^{n=2} |n\rangle\langle n| \otimes Z^n. \tag{2}$$

Both CZ and its inverse $\mathrm{CZ}^\dagger$ are Clifford and universal (when combined with arbitrary SU(3) rotations) entangling gates that can be performed on our system[14]. Critically, the natural adoption of twirling and mitigation methods that have been developed for qubits is only possible through the ability to implement two-qutrit Clifford gates (see Fig. 1).

Twirling. Being a unitary 1-design, Weyl operators can be used to *twirl* noise[68,71,73,74]—that is, to transform arbitrary Markovian noise processes into stochastic channels of the form

$$\mathcal{W}(\rho) = \sum_{\bar{p},\bar{q}}^{4^n} \mathrm{prob}(W_{\bar{p},\bar{q}}) W_{\bar{p},\bar{q}}\, \rho\, W_{\bar{p},\bar{q}}^\dagger, \tag{3}$$

where ρ is an n-qutrit state, $W_{\bar{p},\bar{q}} = \otimes_{i=1}^n W_{q_{k_i},p_{k_i}}$ is a tensor product of one-qutrit Weyl operators, and $\mathrm{prob}(W_{\bar{p},\bar{q}})$ is the probability that a Weyl error $W_{\bar{p},\bar{q}}$ occurs. Twirling noise processes can significantly improve the performance of noisy devices[68]. Indeed, while coherent errors can accumulate quadratically in the number of noisy gates, stochastic channels accumulate linearly and dramatically lower worst-case error rates[75].

In our experiments, we twirl the noise using the RC protocol from ref. 70 generalized to qutrits. To illustrate this procedure, let us consider a "target" circuit of the type in Fig. 2a, i.e., a circuit that alternates between cycles of single-qutrit gates (represented by Completely Positive Trace-Preserving, or CPTP, maps $\mathcal{U}_j$) and cycles of multi-qutrit gates (represented by CPTP maps $\mathcal{H}_j$), implementing the operation

$$\mathcal{C} = \mathcal{U}_{m+1}\mathcal{H}_m\mathcal{U}_m \cdots \mathcal{H}_1\mathcal{U}_1. \tag{4}$$

Expressing a noisy implementation of $\mathcal{U}_j$ (respectively $\mathcal{H}_j$) as $\mathcal{D}_j\mathcal{U}_j$ (respectively $\mathcal{F}_j\mathcal{H}_j$), a noisy implementation of this circuit performs the operation

$$\widetilde{\mathcal{C}} = \mathcal{D}_{m+1}\mathcal{U}_{m+1}\mathcal{E}_m\mathcal{H}_m\mathcal{U}_m \cdots \mathcal{E}_1\mathcal{H}_1\mathcal{U}_1, \tag{5}$$

where $\mathcal{E}_j = \mathcal{F}_j\mathcal{H}_j\mathcal{D}_j\mathcal{H}_j^{-1}$ is the combined noise of $\mathcal{U}_j$ and $\mathcal{H}_j$. Analogous to the original protocol, to perform RC, we recompile randomly chosen Weyl gates and their inverses into the cycles of one-qubit gates (Fig. 2b, c). Implementing the circuit $N > 1$ times with different choices of random Weyl gates and averaging over the various implementations is equivalent to implementing a circuit

$$\widetilde{\mathcal{C}}_{\mathrm{RC}} = \mathcal{U}_{m+1}\mathcal{W}_m\mathcal{H}_m\mathcal{U}_m \cdots \mathcal{W}_1\mathcal{H}_1\mathcal{U}_1 \tag{6}$$

afflicted by stochastic noise processes $\mathcal{W}_j$ of the type in Eq. (3), up to statistical fluctuations that decrease exponentially with N. Specifically, denoting by $\widetilde{\mathcal{C}}_1, \ldots, \widetilde{\mathcal{C}}_N$ the N circuits with random Weyl operators, Hoeffding's inequality[76] ensures that for every state ρ, observable O, and

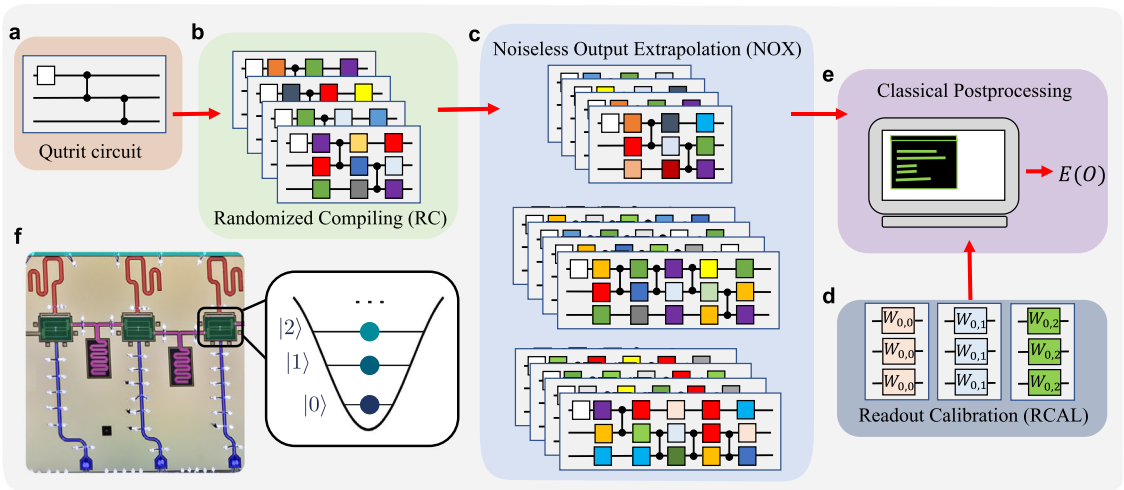

Fig. 1 | Schematic of error mitigation in qutrit circuits. a An arbitrary qutrit circuit is **b** randomly compiled into many logically equivalent copies, effectively twirling the noise of the multi-qutrit gate cycles into stochastic channels. **c** The noise present in the multi-qutrit gates in the circuit is amplified by insertions of the identity. **d** The readout assignment errors are efficiently characterized by measuring the single qutrit confusion matrices. **e** Via classical post-processing, noiseless expectation values are achieved from the combination of RC, NOX, and RCAL. **f** A false-colored micrograph of the system of three capacitively coupled, fixed-frequency transmon qutrits used in the experiment.

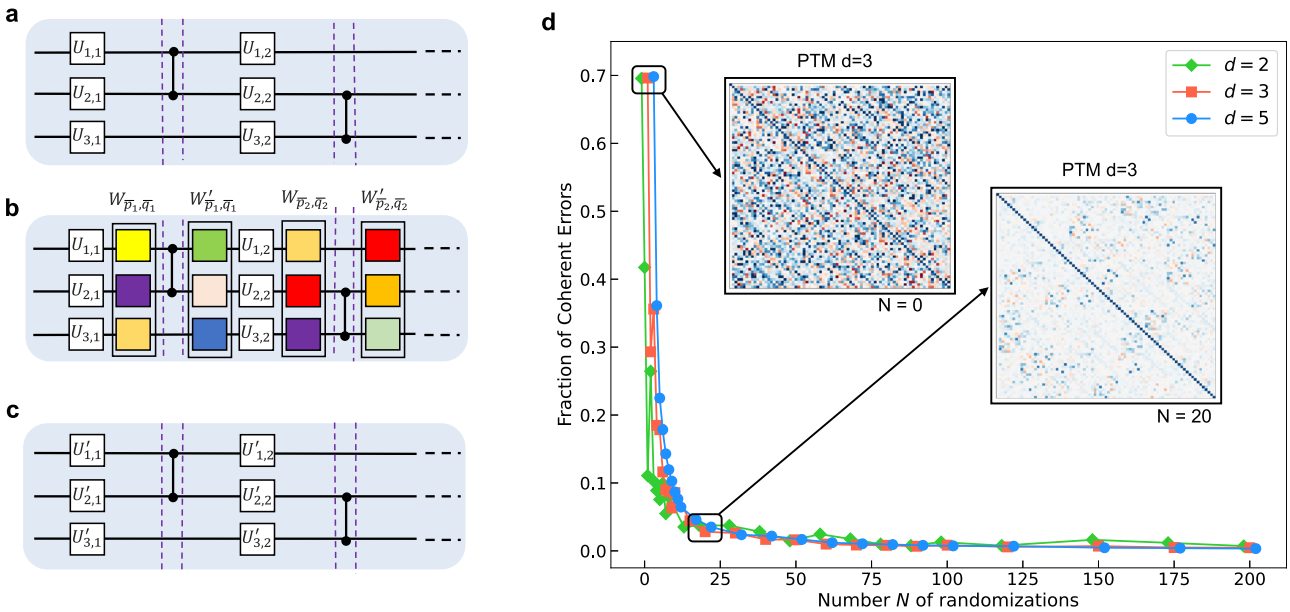

Fig. 2 | Randomized compiling (RC) for qudit circuits. a The input circuit, which alternates cycles of one- and multi-qudit gates. **b** Random Weyl gates ($W_{\vec{p}_j,\vec{q}_j}$) and their inverses ($W'_{\vec{p}_j,\vec{q}_j} = W^{\dagger}_{\vec{p}_j,\vec{q}_j} H_j W_{\vec{p}_j,\vec{q}_j}$) are added before and after every cycle of multi-qudit gates H_j. **c** Before executing the circuit, the extra gates are recompiled into the native one-qudit gates. In this way, the returned circuit has the same depth as the input one. **d** Numerical study of the fraction of coherent errors under RC randomizations in randomly generated two-qudit Pauli Transfer Matrices (PTM) in $d \in \{2, 3, 5\}$. The formula used to quantify the fraction of coherent errors is derived in the "Methods". All PTMs are chosen to initially have 70% of their errors be coherent. As can be seen, the coherent errors decrease exponentially with N, confirming the validity of Eq. (7). Furthermore, the fraction of coherent errors per depth does not depend on d. This demonstrates that for an equal suppression of coherent errors, RC does not require additional twirls in higher dimensions. The inset PTMs visualize the suppression of off-diagonal terms as a function of N.

positive number $\epsilon < 1$, we have

$$\begin{aligned} &\mathrm{prob}\left(|E(O|\widetilde{\mathcal{C}}_{\mathrm{RC}}) - \tfrac{1}{N}\sum_{k=1}^{N} E(O|\widetilde{\mathcal{C}}_k)| < \epsilon\right) \\ &\geq 1 - \exp\left(-2\epsilon^2 N\right), \end{aligned} \tag{7}$$

where $E(O|\mathcal{C}) = \mathrm{Tr}\left[O\mathcal{C}(\rho)\right]$ is the expectation value of O at the end of a circuit $\mathcal{C}$. Importantly, the r.h.s. of the above inequality does not depend on the dimension of the Hilbert space. Hence, achieving the desired approximation level requires implementing the same number of circuits for qutrits as for qubits (Fig. 2d).

Mitigating. While implementing RC alone leads to significant performance gains, employing it in tandem with other error-mitigation protocols leads to even larger gains. In our experiments, we utilize RC in combination with NOX, a protocol designed to mitigate stochastic errors that commute with cycles of gates. In its simplest version, NOX requires implementing the target circuit $\widetilde{\mathcal{C}}_{\mathrm{RC}}$ alongside several copies of it. Each copy implements the same computation as $\widetilde{\mathcal{C}}_{\mathrm{RC}}$, but the noise of one of its cycles is amplified in a controlled way. Specifically, the jth copy implements the operation

$$\widetilde{\mathcal{C}}^{(j)}_{\mathrm{RC}} := \mathcal{U}_{m+1}\mathcal{W}_m\mathcal{H}_m\mathcal{U}_m \cdots \left(\mathcal{W}_j\right)^{\alpha_j}\mathcal{H}_j\mathcal{U}_j \cdots \mathcal{W}_1\mathcal{H}_1\mathcal{U}_1, \tag{8}$$

where $\alpha_j > 1$ is an integer specified by the user. By combining the outputs of the target circuit and those of the various copies, NOX returns the quantity

$$E_{\mathrm{NOX}}(O) = E(O|\widetilde{\mathcal{C}}_{\mathrm{RC}}) - \sum_{j=1}^{m} \frac{E(O|\widetilde{\mathcal{C}}_{\mathrm{RC}}) - E(O|\widetilde{\mathcal{C}}^{(j)}_{\mathrm{RC}})}{\alpha_j - 1}. \tag{9}$$

$E_{\mathrm{NOX}}(O)$ is an estimator of the correct result, $E(O|\mathcal{C})$, and its bias is quadratically smaller than that of $E(O|\widetilde{\mathcal{C}})$[66].

In order to perform noise amplification, in our experiments we utilize Unitary Folding (UF)[54]. That is, given a noisy cycle $\mathcal{WH}$ and a number α such that $\mathcal{H}^{\alpha+1} = \mathcal{H}$, we replace $\mathcal{WH}$ with $(\mathcal{WH})^{\alpha+1}$. In the instance where the error channel $\mathcal{W}$ commutes with the cycle $\mathcal{H}$, i.e., $\mathcal{WH} = \mathcal{HW}$, UF leads to $(\mathcal{WH})^{\alpha+1} = \mathcal{W}^{\alpha+1}\mathcal{H}$, which provides the desired noise amplification. In the more general situation, UF leaves a remainder term:

$$(\mathcal{WH})^{\alpha+1} = \mathcal{W}^{\alpha+1}\mathcal{H} + \mathcal{RH}. \tag{10}$$

In Methods section "Error amplification via unitary folding (UF)", we briefly argue why the term $\mathcal{RH}$ may have little impact on mitigated results in realistic scenarios. This argument is reflected in our results since we observe significant improvements when resorting to UF amplification despite its possible imperfections. That being said, we acknowledge that more sophisticated techniques can be used to amplify noise when $\mathcal{WH} \not\approx \mathcal{HW}$; some methods rely on learning the errors by using cycle-based noise reconstruction techniques[66,77], and some others rely on pulsed-based inverse evolution[78]. We leave the utilization of those methods for future work.

In addition to RC and NOX, we employ readout calibration (RCAL) to mitigate measurement noise[79]. RCAL requires implementing three simple circuits (the first with a $W_{0,0}$ gate on every qutrit, the second with a $W_{0,1}$ gate, and the third one with a $W_{0,2}$ gate) to estimate the probabilities of state-dependent readout errors—i.e., the probabilities that an output $s \in \{0, 1, 2\}$ is incorrectly reported as $s' \neq s$. It then uses this information to efficiently suppress readout errors by inverting confusion matrices.

Experiment

We test our ability to effectively twirl and extrapolate noiseless results in two sets of multi-qutrit experiments. In the first experiment, we use state tomography to reconstruct a three-qutrit GHZ state. In the second experiment, we perform random circuit sampling (RCS) for two and three qutrits at a variety of depths. In both experiments, we find significant improvements in our results via the combination of RC and NOX.

Our experimental device consists of fixed-frequency transmons, with fixed coupling mediated by coplanar waveguide resonators. Single-qutrit gates are performed via Rabi oscillations and virtual Z gates in two-level subspaces of the qutrit, and two-qutrit gates are performed via a tunable cross-Kerr entangling interaction. More information on how we perform single and two-qutrit gates can be found in refs. 12,13 and ref. 14, respectively and additional device characterization is available in Supplementary Note 1.

Multipartite qutrit entanglement. Experimental demonstrations of multipartite entanglement have been instrumental in demonstrating that local realism can be violated by quantum mechanics[80]. So far, studies of multipartite entanglement have mostly focused on coupled qubits, and experimental demonstrations of multipartite qutrit entanglement have only been performed in a few cases[12,81]. In this work, we generate a maximally entangled state on a $D = 27$ dimensional Hilbert space using only three transmon qutrits. Specifically, we realize the qutrit GHZ state $|\Psi\rangle_{\mathrm{GHZ}} = \frac{1}{\sqrt{3}}(|000\rangle + |111\rangle + |222\rangle)$.

To characterize this highly-entangled state, we perform tomography on the full three-qutrit Hilbert space[82], requiring a total of 729 circuits. We measure an informationally-complete set of projections to experimentally reconstruct the density matrix for $\rho = |\Psi\rangle\langle\Psi|_{\mathrm{GHZ}}$, finding a state fidelity of $\mathcal{F} = \mathrm{Tr}\left(\sqrt{\sigma}\rho\sqrt{\sigma}\right) = 0.818$, where σ is the ideal density matrix. Even in the unmitigated case, our work represents, to the best of our knowledge, the highest fidelity exploration of multipartite qutrit entanglement to date.

Next, we perform the same experiment using RC and NOX, requiring a total of 43,740 ($729 \times 20 \times 3$) circuits, where we use 20 logically-equivalent compilations of the bare circuit and for each of the identity insertions for the two different two-qutrit gate cycles. We find a mitigated state fidelity of $\mathcal{F} = 0.951$, resulting in a greater than 3x reduction in infidelity compared to the unmitigated case. The circuits and the experimentally reconstructed density matrices can be found in Fig. 3. Finally, we purify the reconstructed state measured with RC (no NOX)[83] by numerically finding the nearest density matrix that is idempotent ($\rho^2 = \rho$)[84], which improved the state fidelity from 0.912 to 0.998 (see Supplementary Note 4). These results demonstrate the power of RC to tailor coherent errors to purely stochastic channels. A summary of the results of all the state tomography experiments and mitigation strategies can be found in Table 1.

Table 1 | A summary of the results of the three-qutrit state tomography experiment employing the various mitigation strategies outlined in the text

Mitigation strategy	State fidelity
Bare	0.818
RC + NOX	0.951
Bare + Purified	0.912
RC + Purified	0.998
RC + NOX + Purified	0.995

We note that the reported state fidelity for each experiment employs RCAL.

Qutrit random circuit sampling. RCS with qubits has recently garnered significant interest due to its role in demonstrating quantum advantage[2,4], and it has been speculated that RCS may also have pragmatic use cases[85,86]. In this work, we extend the task of RCS to qutrits and study the two and three-qutrit experiments at a number of depths. Notably, leveraging qutrits for RCS, in principle, could allow one to probe the regime of quantum supremacy with significantly fewer qutrits (~35) than qubits (~55).

In both the two and three-qutrit RCS experiments, Haar-random single qutrit gates are interleaved with $CZ^\dagger$ gates (hard cycles). We study 20 separate RCS instances for each circuit depth. In Fig. 4, the bare RCS results are compared to the mitigated results measured using RC (20 randomizations) and NOX. To quantify circuit performance, we calculate the variation distance from the ideal trit string distributions, defined as

$$\mathrm{VD} := \frac{1}{2}\sum_{\bar{s}} |p_{\mathrm{id}}(\bar{s}) - p_{\mathrm{exp}}(\bar{s})| \tag{11}$$

where $\bar{s} \in (0,1,2)^{\otimes n}$ and p_{id} and p_{exp} are the ideal and experimentally measured trit string distributions, respectively. The best results for the

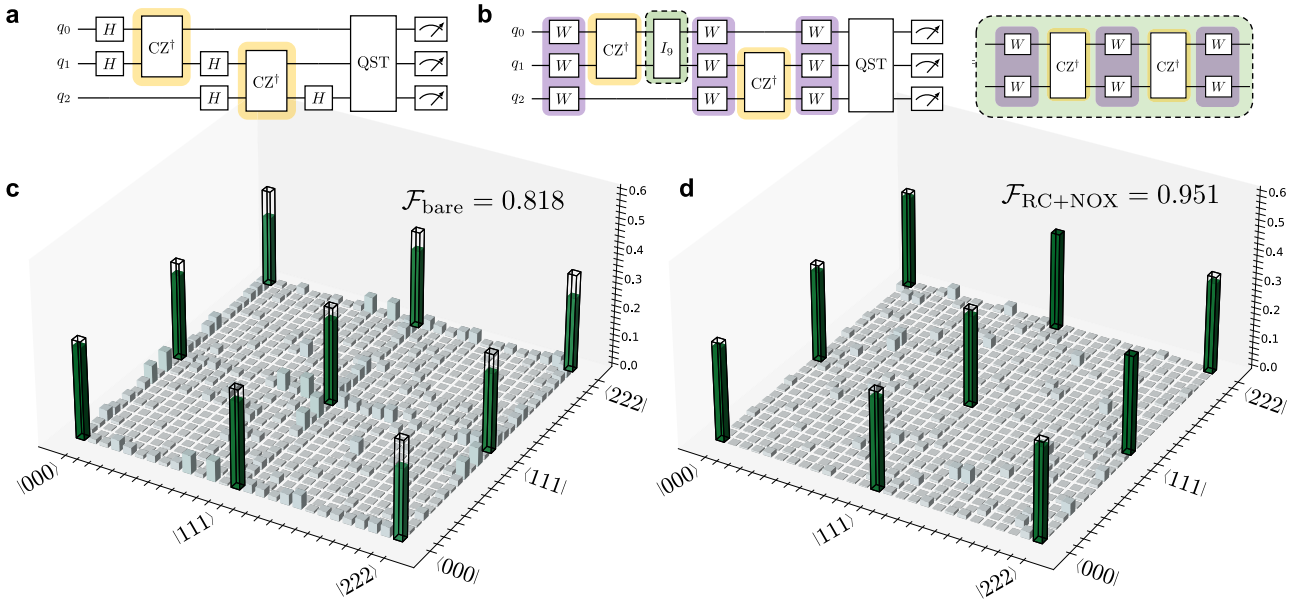

Fig. 3 | Experimentally reconstructed qutrit GHZ density matrices. a The circuit diagram for experimentally reconstrucing the density matrix of a 3-qutrit GHZ state. The state can be efficiently generated with two-qutrit $CZ^\dagger$ gates (yellow) and single qutrit Hadamard gates (H). **b** An example circuit diagram with RC + NOX, including Weyl twirling (purple), where the identity insertion (green) is placed after the $CZ^\dagger$ (yellow) between q_0 and q_1. An example of a twirled decomposition of the identity (green) is provided to the right. **c** The experimentally reconstructed density matrix (plotting $|\rho|$) of the 3-qutrit GHZ state with only readout assignment errors corrected. The ideal density matrix is shown with a black outline. **d** The density matrix generated by the noiseless expectation values produced by RC+NOX.

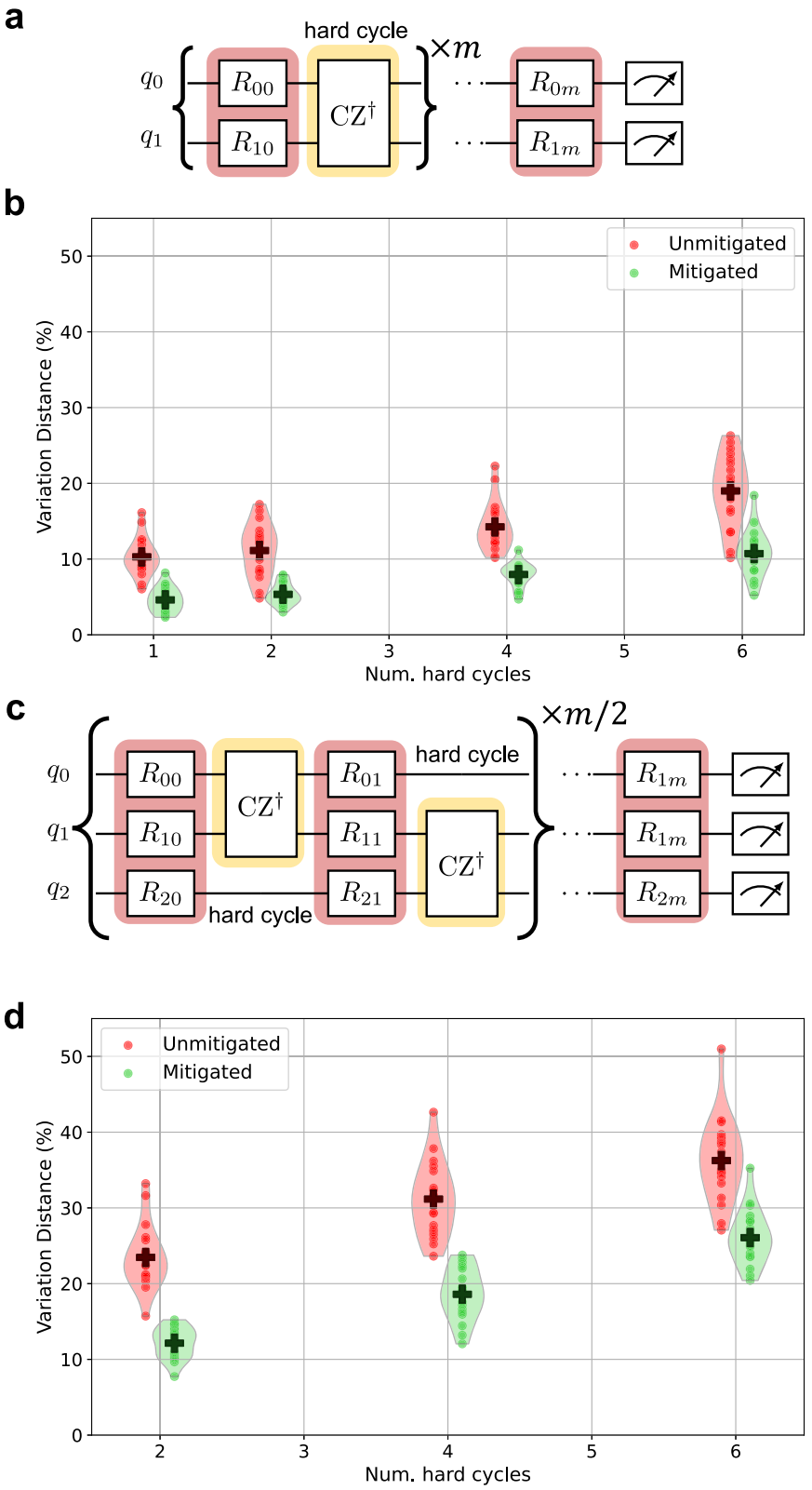

Fig. 4 | Random circuit sampling with qutrits. **a** The circuits for two-qutrit random circuit sampling. $CZ^\dagger$ (yellow) hard cycles are interleaved with Haar random SU(3) gates, R_{ij} (red). **b** Violin plots showing the distribution of the variation distances for the 20 RCS instances with (green) and without (red) RC + NOX at depths $m \in \{1, 2, 3, 6\}$, with mean values marked by crosses. **c** Circuits for three-qutrit RCS. **d** Results for three-qutrit RCS at $m \in \{2, 4, 6\}$.

mitigated case were found with 3 insertions of the identity (see Supplementary Note 3). Notably, when employing RC+NOX in both the two and three-qutrit RCS experiments, the variation distances at depth 6 were comparable to the unmitigated case at depth 2, and at all depths, we found at least a 30% fractional improvement in our results with RC + NOX.

Discussion

We introduced generalized versions of two powerful methods for tailoring and mitigating noise in contemporary qudit systems: randomized compiling and noiseless output extrapolation. We tested the efficacy of these methods at generating noiseless expectation values on a system of three coupled transmon qutrits. Specifically, we explored the experimentally

reconstructed density matrix of a 3-qutrit GHZ state and the problem of multi-qutrit random circuit sampling. We found that despite the more complex noise environment, added noise sensitivity, and more difficult control requirements, our protocols proved to be a powerful tool for significantly improving our results for all of the aforementioned experiments. Specifically, we effectively tailored coherent errors to stochastic errors on both gate and spectator qutrits via the first demonstration of randomized compiling for qudit dimension $d \geq 3$, and extrapolated beyond the noise in our system via the first demonstration of noiseless output extrapolation in $d \geq 3$.

As higher-dimensional quantum devices begin to mature and compete with qubit systems, the ability to perform longer-depth algorithms without significant errors will be critical to convincing the community of their feasibility and scalability. To this end, our work opens the door to explore many of the advantages leveraged by qudit devices in both quantum algorithms and gate based quantum simulation in the near term on contemporary devices.

Methods

Calculating the proportion of coherent errors in qudit PTMs

In Fig. 2d, we show the numerical results of twirling away (using qudit RC) off-diagonal elements (or coherent errors) in qudit PTMs. Here, we briefly comment on how one calculates the proportion of coherent errors present in a qudit PTM, following ref. 87.

Consider a quantum error channel with PTM denoted as $\mathcal{E}$. The process fidelity is defined as the normalized trace of the $(d^2 \times d^2)$ PTM:

$$F(\mathcal{E}) := \mathrm{tr}\,\mathcal{E}/d^2 \tag{12}$$

Let's define the decoherent process fidelity of $\mathcal{E}$ as

$$F_{\text{decoh.}}(\mathcal{E}) := \sqrt{\mathrm{tr}\,\mathcal{E}^\dagger\mathcal{E}/d^2} = \|\mathcal{E}\|_F/d. \tag{13}$$

In ref. 87, it is shown that as long as the error channel is reasonably close to the identity, that is if $F(\mathcal{E}) > 1/2$ and $F_{\text{decoh.}}(\mathcal{E}) > 1/\sqrt{2}$, the error channel has a well-defined coherent-decoherent polar decomposition:

$$\mathcal{E} = \mathcal{U}\mathcal{D}, \tag{14}$$

where $\mathcal{U}$ is a coherent (i.e., unitary) error process and $\mathcal{D}$ is a purely decoherent process. The precise definition of a decoherent channel is elaborated and justified in ref. 87.

In realistic physical scenarios, the process fidelity can be expressed as a simple product:

$$F(\mathcal{E}) = F(\mathcal{U})F(\mathcal{D}) + h.o., \tag{15}$$

where the higher order term (*h. o.*) is of second order in the infidelity, $O\big((1-F(\mathcal{E}))^2\big)$. For reference, Eq. (15) does not hold in pathological cases where a significant part of the error process originates from specially crafted high-body interactions (e.g., Hamiltonian/Lindbladian terms that are made of tensor products of a large number of subsystems). A discussion of these pathological cases is included in ref. 87.

Notice that by combining eq. (13) and eq. (14), we get:

$$\begin{aligned} F_{\text{decoh.}}(\mathcal{E}) &= F_{\text{decoh.}}(\mathcal{U}\mathcal{D}) \\ &= \|\mathcal{U}\mathcal{D}\|_F/d \\ &= \|\mathcal{D}\|_F/d \\ &= F_{\text{decoh.}}(\mathcal{D}) \\ &= F(\mathcal{D}) + O\big((1-F(\mathcal{D}))^2\big) \end{aligned} \tag{16}$$

where the last line comes from the fact that non-pathological decoherent errors obey ([87])

$$\mathrm{tr}\,\mathcal{D}^\dagger\mathcal{D}/d^2 = \big(\mathrm{tr}\mathcal{D}/d^2\big)^2 + O\big((1-F(\mathcal{D}))^2\big). \tag{17}$$

The above is essentially a corollary from the fact that decoherent errors build up according to a multiplicative decay.

By substituting eq. (16) in Eq. (15), we get, up to second order in the infidelity,

$$F(\mathcal{E}) = F(\mathcal{U})F_{\text{decoh.}}(\mathcal{E}), \tag{18}$$

or in terms of infidelity (up to second order in the infidelity)

$$\underbrace{1-F(\mathcal{E})}_{\text{Tot. infid.}} = \underbrace{1-F(\mathcal{U})}_{\text{Coh. infid.}} + \underbrace{1-F_{\text{decoh.}}(\mathcal{E})}_{\text{Decoh. infid.}}. \tag{19}$$

As such, the contribution of coherence to the process of infidelity is given by (up to the second order in the infidelity):

$$1 - F(\mathcal{U}) = F_{\text{decoh.}}(\mathcal{E}) - F(\mathcal{E}). \tag{20}$$

The relative coherent contribution to the infidelity is therefore obtained via

$$\frac{F_{\text{decoh.}}(\mathcal{E}) - F(\mathcal{E})}{1-F(\mathcal{E})} = \frac{\mathrm{tr}\,\mathcal{E}/d^2 - d\,\|\,\mathcal{E}\|_F/d}{1-\mathrm{tr}\,\mathcal{E}/d^2}. \tag{21}$$

The above is the formula used in Fig. 2d to quantify the effect of twirling on the error channel.

Error amplification via unitary folding (UF)

Noise amplification through unitary folding (UF) is performed by repeating a desired noisy cycle $\mathcal{W}\mathcal{H}$[54]. When the error channel $\mathcal{W}$ commutes with the cycle $\mathcal{H}$, we get $(\mathcal{W}\mathcal{H})^{\alpha+1} = \mathcal{W}^{\alpha+1}\mathcal{H}$. In the more general situation where $\mathcal{W}\mathcal{H} \neq \mathcal{H}\mathcal{W}$, we instead get a remainder term

$$(\mathcal{W}\mathcal{H})^{\alpha+1} = \mathcal{W}^{\alpha+1}\mathcal{H} + \mathcal{R}\mathcal{H}. \tag{22}$$

In this section, we briefly argue why one may expect the remainder $\mathcal{R}\mathcal{H}$ to be of little impact in some instances.

Given a cyclicity c representing the smallest positive integer such that $\mathcal{H}^c = I$, we can express the error channel $\mathcal{W}$ as

$$\mathcal{W} = e^{\sum_{k=0}^{c-1} L_k}, \tag{23}$$

where the dissipator terms $\{L_0, L_1, \cdots, L_{c-1}\}$ obey:

$$L_k\mathcal{H} = \omega^k\mathcal{H}L_k, \tag{24}$$

where $\omega = e^{2i\pi/c}$. In our experiments, the cycle $\mathcal{H}$ is a tensor product of $CZ^\dagger$ gates and has a cyclicity of $c = 3$. This yields $\mathcal{W} = e^{L_0+L_1+L_2}$ with $L_k\mathcal{H} = \omega^k\mathcal{H}L_k$ where $\omega = e^{2i\pi/3}$. In this case, UF yields:

$$\begin{aligned} (\mathcal{W}\mathcal{H})^{\alpha+1} &= \left(e^{L_0+L_1+L_2}\mathcal{H}\right)^{\alpha+1} \\ &= \left(\prod_{k=0}^{k=\alpha} e^{\mathcal{H}^k(L_0+L_1+L_2)\mathcal{H}^{-k}}\right)\mathcal{H} \\ &= \left(\prod_{k=0}^{k=\alpha} e^{L_0+\omega^k L_1+\omega^{2k}L_2}\right)\mathcal{H} \\ &\approx e^{(\alpha+1)L_0+L_1+L_2}\mathcal{H}, \end{aligned} \tag{25}$$

where in the last line we truncated the second-order terms from the BCH expansion[88] for simplicity. The purpose behind the derivation of Eq. (25) is to show that, up to a second-order approximation, only the dissipator term

L_0 that commutes with the cycle $\mathcal{H}$ gets properly amplified by UF. One reason that makes UF an effective error amplification strategy despite this limitation is that L_0 is often a dominating source of error compared to the other dissipators. Indeed, following the derivation of[78], given a noisy process $\mathcal{WH}$ obtained from integrating a Lindbladian,

$$\mathcal{WH} = \int_0^T -iH(t) + L(t)\, dt, \tag{26}$$

where $H(t)$ and $L(t)$ correspond to the super-operator versions of the ideal and noisy dynamics, respectively, we can approximate $\mathcal{W}$ as

$$\mathcal{W} \approx \mathcal{H} \exp\left(\int_0^T dt\, \mathcal{U}^\dagger(t) L(t) \mathcal{U}(t)\right) \mathcal{H}^{-1}, \tag{27}$$

where $\mathcal{U}(t)$ is the ideal evolution at time t. This approximation corresponds to the truncation of the Magnus expansion to first order. The integral in Eq. (27) is close to a twirl operation, and has the effect of partially suppressing the matrix components that do not commute with $\mathcal{W}$. To see this more clearly, let's assume that $H(t)$ and $L(t)$ are time-independent and let's perform column-vectorization on the integral in Eq. (27):

$$\mathrm{vec}\left(\int_0^T dt\, \mathcal{U}^\dagger(t) L\, \mathcal{U}(t)\right) = \underbrace{\left(\int_0^T dt\, \mathcal{U}(t)^T \otimes \mathcal{U}^\dagger(t)\right)}_{\text{Partial twirl operator.}} \mathrm{vec}(L) \tag{28}$$

where we used the identity $\mathrm{vec}(ABC) = C^T \otimes A\mathrm{vec}(B)$. We denote the operator acting on $\mathrm{vec}(L)$ as a "partial twirl"; a full twirl would involve the self-inverting evolution from 0 to τ where $\mathcal{U}(0) = \mathcal{U}(\tau)$. It is straightforward to show that the full twirl projects $\mathrm{vec}(L)$ unto the subspace of operators that commute with $\mathcal{U}(t) = \exp(iHt)$ for any time t. The partial twirl operator instead dampens the components of L that do not commute with the unitary evolution evolution.

Data availability

All data are available from the corresponding author upon reasonable request.

Code availability

All front-end code and jupyter notebooks used for analysis are available upon request. The mitigation routines and circuit compilation were performed using True-Q, a proprietary software package.

Received: 18 December 2023; Accepted: 23 September 2024;
Published online: 14 October 2024

References

1. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
2. Arute, F. et al. Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505–510 (2019).
3. Zhong, H.-S. et al. Quantum computational advantage using photons. *Science* **370**, 1460–1463 (2020).
4. Wu, Y. et al. Strong quantum computational advantage using a superconducting quantum processor. *Phys. Rev. Lett.* **127**, 180501 (2021).
5. Madsen, L. S. et al. Quantum computational advantage with a programmable photonic processor. *Nature* **606**, 75–81 (2022).
6. Wang, Y., Hu, Z., Sanders, B. C. & Kais, S. Qudits and high-dimensional quantum computing. *Front. Phys*. **8** (2020).
7. Bocharov, A., Roetteler, M. & Svore, K. M. Factoring with qutrits: Shor's algorithm on ternary and metaplectic quantum architectures. *Phys. Rev. A* **96**, 012306 (2017).
8. Nikolaeva, A. S., Kiktenko, E. O. & Fedorov, A. K. Efficient realization of quantum algorithms with qudits. *EPJ Quantum Technol.* **11**, 43 (2024).
9. Bullock, S. S., O'Leary, D. P. & Brennen, G. K. Asymptotically optimal quantum circuits for *d*-level systems. *Phys. Rev. Lett.* **94**, 230502 (2005).
10. Lanyon, B. P. et al. Simplifying quantum logic using higher-dimensional hilbert spaces. *Nat. Phys.* **5**, 134–140 (2009).
11. Gedik, Z. et al. Computational speed-up with a single qudit. *Sci. Rep.* **5**, 14671 (2015).
12. Blok, M. S. et al. Quantum information scrambling on a superconducting qutrit processor. *Phys. Rev. X* **11**, 021010 (2021).
13. Morvan, A. et al. Qutrit randomized benchmarking. *Phys. Rev. Lett.* **126**, 210504 (2021).
14. Goss, N. et al. High-fidelity qutrit entangling gates for superconducting circuits. *Nat. Commun.* **13**, 7481 (2022).
15. Luo, K. et al. Experimental realization of two qutrits gate with tunable coupling in superconducting circuits. *Phys. Rev. Lett.* **130**, 030603 (2023).
16. Roy, T., Li, Z., Kapit, E. & Schuster, D. I. Realization of two-qutrit quantum algorithms on a programmable superconducting processor. *Phys. Rev. Applied* **19**, 064024 (2023).
17. Cao, S. et al. Emulating two qubits with a four-level transmon qudit for variational quantum algorithms. *Quantum Sci. Technol.* **9**, 035003 (2024).
18. Litteken, A. et al. Dancing the quantum waltz: Compiling three-qubit gates on four level architectures. In *Proceedings of the 50th Annual International Symposium on Computer Architecture.* https://doi.org/10.1145/3579371.3589106 (ACM, 2023).
19. Ringbauer, M. et al. A universal qudit quantum processor with trapped ions. *Nat. Phys.* **18**, 1053–1057 (2022).
20. Hrmo, P. et al. Native qudit entanglement in a trapped ion quantum processor. *Nat. Commun.* **14**, 2242 (2023).
21. Lanyon, B. P. et al. Manipulating biphotonic qutrits. *Phys. Rev. Lett.* **100**, 060504 (2008).
22. Chi, Y. et al. A programmable qudit-based quantum processor. *Nat. Commun.* **13**, 1166 (2022).
23. Wang, Y., Snizhko, K., Romito, A., Gefen, Y. & Murch, K. Dissipative preparation and stabilization of many-body quantum states in a superconducting qutrit array. *Phys. Rev. A.* **108**, 013712 (2023).
24. Senko, C. et al. Realization of a quantum integer-spin chain with controllable interactions. *Phys. Rev. X* **5**, 021026 (2015).
25. Gustafson, E. Noise improvements in quantum simulations of sqed using qutrits. Preprint at *arXiv* https://arxiv.org/abs/2201.04546 (2022).
26. Bechmann-Pasquinucci, H. & Peres, A. Quantum cryptography with 3-state systems. *Phys. Rev. Lett.* **85**, 3313–3316 (2000).
27. Bruß, D. & Macchiavello, C. Optimal eavesdropping in cryptography with three-dimensional quantum states. *Phys. Rev. Lett.* **88**, 127901 (2002).
28. Vaziri, A., Weihs, G. & Zeilinger, A. Experimental two-photon, three-dimensional entanglement for quantum communication. *Phys. Rev. Lett.* **89**, 240401 (2002).
29. Gokhale, P. et al. Asymptotic improvements to quantum circuits via qutrits. In *Proceedings of the 46th International Symposium on Computer Architecture*, ISCA '19, 554–566. https://doi.org/10.1145/3307650.3322253 (Association for Computing Machinery, New York, NY, USA, 2019).
30. Fedorov, A., Steffen, L., Baur, M., da Silva, M. P. & Wallraff, A. Implementation of a toffoli gate with superconducting circuits. *Nature* **481**, 170–172 (2012).
31. Chu, J. et al. Scalable algorithm simplification using quantum and logic. *Nat. Phys.* **19**, 126–131 (2023).
32. Nguyen, L. B. et al. Programmable heisenberg interactions between floquet qubits. *Nat. Phys.* **20**, 240–246 (2024).
33. Galda, A., Cubeddu, M., Kanazawa, N., Narang, P. & Earnest-Noble, N. Implementing a ternary decomposition of the toffoli gate on fixed-

frequencytransmon qutrits. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2109.00558 (2021).
34. Hill, A. D., Hodson, M. J., Didier, N. & Reagor, M. J. Realization of arbitrary doubly-controlled quantum phase gates. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2108.01652 (2021).
35. Elder, S. S. et al. High-fidelity measurement of qubits encoded in multilevel superconducting circuits. *Phys. Rev. X* **10**, 011001 (2020).
36. Jurcevic, P. et al. Demonstration of quantum volume 64 on a superconducting quantum computing system. *Quantum Sci. Technol.* **6**, 025020 (2021).
37. Chen, L. et al. Transmon qubit readout fidelity at the threshold for quantum error correction without a quantum-limited amplifier. *npj Quantum Inf.* **9**, 26 (2023).
38. Duclos-Cianci, G. & Poulin, D. Kitaev's $\mathbb{Z}_d$-code threshold estimates. Phys. Rev. A 87, 062338 (2013)..
39. Anwar, H., Brown, B. J., Campbell, E. T. & Browne, D. E. Fast decoders for qudit topological codes. *N. J. Phys.* **16**, 063038 (2014).
40. Andrist, R. S., Wootton, J. R. & Katzgraber, H. G. Error thresholds for abelian quantum double models: Increasing the bit-flip stability of topological quantum memory. *Phys. Rev. A* **91**, 042331 (2015).
41. Ma, Y., Hanks, M. & Kim, M. S. Non-Pauli errors can be efficiently sampled in qudit surface codes. *Phys. Rev. Lett.* **131**, 200602 (2023).
42. Muralidharan, S., Zou, C.-L., Li, L., Wen, J. & Jiang, L. Overcoming erasure errors with multilevel systems. *N. J. Phys.* **19**, 013026 (2017).
43. Kubica, A., Haim, A., Vaknin, Y., Brandão, F. & Retzker, A. Erasure qubits: overcoming the T_1 limit in superconducting circuits. *Phys. Rev. X* **13**, 041022 (2023).
44. Campbell, E. T., Anwar, H. & Browne, D. E. Magic-state distillation in all prime dimensions using quantum reed-muller codes. *Phys. Rev. X* **2**, 041021 (2012).
45. Campbell, E. T. Enhanced fault-tolerant quantum computing in *d*-level systems. *Phys. Rev. Lett.* **113**, 230501 (2014).
46. Majumdar, R., Basu, S., Ghosh, S. & Sur-Kolay, S. Quantum error-correcting code for ternary logic. *Phys. Rev. A* **97**, 052302 (2018).
47. Kapit, E. Hardware-efficient and fully autonomous quantum error correction in superconducting circuits. *Phys. Rev. Lett.* **116**, 150501 (2016).
48. Li, Z. et al. Autonomous error correction of a single logical qubit using two transmons. *Nat Commun.* **15**, 1681 (2024).
49. Li, Y. & Benjamin, S. C. Efficient variational quantum simulator incorporating active error minimization. *Phys. Rev. X* **7**, 021050 (2017).
50. Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.* **119**, 180509 (2017).
51. Endo, S., Benjamin, S. C. & Li, Y. Practical quantum error mitigation for near-future applications. *Phys. Rev. X* **8**, 031027 (2018).
52. Kandala, A. et al. Error mitigation extends the computational reach of a noisy quantum processor. *Nature* **567**, 491 (2019).
53. Strikis, A., Qin, D., Chen, Y., Benjamin, S. & Li, Y. Learning-based quantum error mitigation. *PRX Quantum* **2**, 040330 (2021).
54. Giurgica-Tiron, T., Hindy, Y., LaRose, R., Mari, A. & Zeng, W. Digital zero noise extrapolation for quantum error mitigation. *IEEE International Conference on Quantum Computing and Engineering (QCE), Denver, CO, USA* (2020).
55. LaRose, R., Mari, A., Karalekas, P., Shammah, N. & Zeng, W. Mitiq: a software package for error mitigation on noisy quantum computers. *Quantum* **6**, 774 (2022).
56. Endo, S., Cai, Z., Benjamin, S. & Yuan, X. Hybrid quantum-classical algorithms and quantum error mitigation. *J. Phys. Soc. Jpn*. **90**, 032001 (2021).
57. Kim, Y. et al. Scalable error mitigation for noisy quantum circuits produces competitive expectation values. *Nat. Phys.* **19**, 752–759 (2023).
58. Koczor, B. Exponential error suppression for near-term quantum devices. *Phys. Rev. X* **11**, 031057 (2021).
59. He, A., Nachman, B., de Jong, W. & Bauer, C. Zero-noise extrapolation for quantum-gate error mitigation with identity insertions. *Phys. Rev. A* **102**, 012426 (2020).
60. Huggins, W. et al. Virtual distillation for quantum error mitigation. *Phys. Rev. X* **11**, 041036 (2021).
61. Pascuzzi, V. et al. Computationally efficient zero noise extrapolation for quantum gate error mitigation. *Phys. Rev. A* **105**, 042406 (2022).
62. Song, C. et al. Quantum computation with universal error mitigation on a superconducting quantum processor. *Sci. Adv.* https://www.science.org/doi/10.1126/sciadv.aaw5686 (2019).
63. Zhang, S. et al. Error-mitigated quantum gates exceeding physical fidelities in a trapped-ion system. *Nat. Commun.* https://www.nature.com/articles/s41467-020-14376-z (2020).
64. Czarnik, P., Arrasmith, A., Coles, P. & Cincio, L. Error mitigation with clifford quantum-circuit data. *Quantum* **5**, 592 (2021).
65. Van Den Berg, E., Minev, Z., Kandala, A. & Temme, K. Probabilistic error cancellation with sparse pauli-lindblad models on noisy quantum processors. *Nat. Phys.* **19**, 1116–1121 (2023).
66. Ferracin, S. et al. Efficiently improving the performance of noisy quantum computers. *Quantum* **8**, 1410 (2024).
67. Jain, A., Iyer, P., Bartlett, S. D. & Emerson, J. Improved quantum error correction with randomized compiling (2023). 2303.06846.
68. Hashim, A. et al. Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor. *Phys. Rev. X* **11**, 041039 (2021).
69. Sung, Y. et al. Multi-level quantum noise spectroscopy. *Nat. Commun.* **12**, 967 (2021).
70. Wallman, J. J. & Emerson, J. Noise tailoring for scalable quantum computation via randomized compiling. *Phys. Rev. A* **94**, 052325 (2016).
71. Graydon, M., Skanes-Norman, J. & Wallman, J. Designing stochastic channels. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2201.07156 (2022).
72. Appleby, D. M. Symmetric informationally complete-positive operator valued measures and the extended clifford group. *J. Math. Phys.* https://doi.org/10.1063/1.1896384 (2005).
73. Ferracin, S., Kapourniotis, T. & Datta, A. Accrediting outputs of noisy intermediate-scale quantum computing devices. *N. J. Phys.* **21**, 113038 (2019).
74. Ferracin, S., Merkel, S., McKay, D. & Datta, A. Experimental accreditation of outputs of noisy quantum computers. *Phys. Rev. A* **104**, 042603 (2021).
75. Hashim, A. et al. Benchmarking quantum logic operations relative to thresholds for fault tolerance. *npj Quantum Inf.* **9**, 109 (2023).
76. Hoeffding, W. Probability inequalities for sums of bounded random variables. *J. Am. Stat. Assoc.* **58**, 13–30 (1963).
77. Carignan-Dugas, A. et al. The error reconstruction and compiled calibration of quantum computing cycles. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2303.17714 (2022).
78. Henao, I., Santos, J. P. & Uzdin, R. Adaptive quantum error mitigation using pulse-based inverse evolutions. *npj Quantum Inf.* https://doi.org/10.1038/s41534-023-00785-7 (2023).
79. Beale, S. et al. *True-Q*. https://doi.org/10.5281/zenodo.3945250.
80. Hensen, B. et al. Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* **526**, 682–686 (2015).
81. Cervera-Lierta, A., Krenn, M., Aspuru-Guzik, A. & Galda, A. Experimental high-dimensional greenberger-horne-zeilinger entanglement with superconducting transmon qutrits. *Phys. Rev. Appl.* **17**, 024062 (2022).
82. Bianchetti, R. et al. Control and tomography of a three level superconducting artificial atom. *Phys. Rev. Lett.* **105**, 223601 (2010).
83. Ville, J.-L. et al. Leveraging randomized compiling for the quantum imaginary-time-evolution algorithm. *Phys. Rev. Res.* **4**, 033140 (2022).
84. Truflandier, L. A., Dianzinga, R. M. & Bowler, D. R. Communication: generalized canonical purification for density matrix minimization.

J. Chem. Phys. https://doi.org/10.1063/1.4943213 https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.4943213/13330452/091102_1_online.pdf (2016).
85. Aaronson, S. & Hung, S.-H. Certified randomness from quantum supremacy. *STOC 2023: Proceedings of the 55th Annual ACM Symposium on Theory of Computing,* 933-944 (2023).
86. Gokhale, P. et al. Supercheq: quantum advantage for distributed databases. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2212.03850 (2022).
87. Carignan-Dugas, A., Alexander, M. & Emerson, J. A polar decomposition for quantum channels (with applications to bounding error propagation in quantum circuits). *Quantum* **3**, 173 (2019).
88. Hall, B. C. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction* (Springer).

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 2210391. Additional support was provided by the Office of Advanced Scientific Computing Research, Testbeds for Science program, Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. A.H. acknowledges financial support from the Berkeley Initiative for Computational Transformation Fellows Program. N.G. acknowledges Long Nguyen, Karthik Siva, and Brian Marinelli for useful discussions. S.F. acknowledges Joel Wallman and Anthony Chytros for useful discussions.

Author contributions

N.G. conceptualized and performed the experiment. S.F. developed the error mitigation protocols. A.H. and A.C.D. aided with the design of the experiments and analysis. J.M.K. fabricated the superconducting chip. R.K.N., D.I.S., and I.S. supervised all experimental work. N.G. and S.F. wrote the paper with input from all authors.

Competing interests

S.F. and A.C.D. have a financial interest in Keysight Technologies and the use of True-Q software[79]. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41534-024-00892-z.

Correspondence and requests for materials should be addressed to Noah Goss.