

# Error mitigation with stabilized noise in superconducting quantum processors

Received: 1 August 2024

Accepted: 31 July 2025

Published online: 26 September 2025



Youngseok Kim<sup>1</sup>✉, Luke C. G. Govia<sup>2</sup>, Andrew Dane<sup>1</sup>, Ewout van den Berg<sup>1</sup>, David M. Zajac<sup>1</sup>, Bradley Mitchell<sup>2</sup>, Yinyu Liu<sup>1</sup>, Karthik Balakrishnan<sup>1</sup>, George Keefe<sup>1</sup>, Adam Stabile<sup>1</sup>, Emily Pritchett<sup>1</sup>, Jiri Stehlik<sup>1</sup> & Abhinav Kandala<sup>1</sup>✉

Pre-fault tolerant quantum computers have already demonstrated the ability to estimate observable values accurately, at a scale beyond brute-force classical computation. This has been enabled by error mitigation techniques that often rely on a representative model of the device noise. However, learning and maintaining these models is complicated by fluctuations in the noise over unpredictable time scales, for instance, arising from resonant interactions between superconducting qubits and defect two-level systems (TLS). Such interactions affect the stability and uniformity of device performance as a whole, but also affect the noise model accuracy, leading to incorrect observable estimation. Here, we experimentally demonstrate that tuning of the qubit-TLS interactions helps reduce noise instabilities and consequently enables more reliable error-mitigation performance. These experiments provide a controlled platform for studying the performance of error mitigation in the presence of quasi-static noise. We anticipate that the capabilities introduced here will be crucial for the exploration of quantum applications on solid-state processors at non-trivial scales.

A common primitive in many near-term quantum algorithms is the accurate estimation of observable expectation values for a short-depth quantum circuit<sup>1,2</sup>. Since, in practice, these circuits are run on noisy quantum processors, the observable estimates tend to be biased away from the true values. In the absence of fault tolerance, quantum error mitigation methods provide a viable way to access more accurate observable estimates<sup>3–9</sup>. These methods typically rely on combining the results of several noisy quantum circuits in ways that cancel the effect of noise on observable estimates, with no qubit overhead. One of these methods, namely zero-noise extrapolation (ZNE), was recently demonstrated to work on circuit sizes that are beyond brute-force classical simulation<sup>7</sup>.

Several error-mitigation methods rely on access to a representative model of the device noise<sup>3,7,8</sup>. Learning and maintaining an accurate model, however, is complicated by fluctuations in physical device noise, which can occur over unpredictable time scales. In

superconducting quantum processors, one prominent source of such fluctuations is the interaction between qubits and defect two-level systems (TLS)<sup>10–16</sup>. It has been shown that the diffusion of TLS transition frequencies over time is a major contributor to fluctuations in qubit relaxation times<sup>13,15,17</sup>, and these dynamics have been extensively investigated through a number of experimental controls that modulate the qubit-TLS interaction, including: external electric fields<sup>18,19</sup>, structural deformation of the physical processor<sup>11,20</sup>, flux tuning<sup>13</sup>, and microwave drives<sup>15,17,21</sup>. Overall, fluctuations in qubit relaxation times have crucial consequences for the stability<sup>22–24</sup>, uniformity, and throughput of superconducting quantum computation. As an example, previous studies have shown that qubit-TLS interaction can degrade the performance of error-mitigation<sup>7</sup>, or even result in unphysical observable estimates<sup>6</sup>.

In this work, we present experiments that focus on noise instabilities in superconducting qubit hardware and their impact on

<sup>1</sup>IBM Quantum, IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, USA. <sup>2</sup>IBM Quantum, IBM Almaden Research Center, San Jose, CA 95120, USA.

✉ e-mail: [kim.10017@osu.edu](mailto:kim.10017@osu.edu); [akandala@us.ibm.com](mailto:akandala@us.ibm.com)

the performance of error mitigation. We employ a device with six fixed-frequency transmon qubits connected in a one-dimensional chain (Q1-Q2-...-Q6) with tunable couplers<sup>25–29</sup>. Each transmon qubit is equipped with an electrode that is placed above. The electrode is controlled by a separate control line to modulate the qubit–TLS interaction. Applying a bias on the electrodes, parametrized by  $k_{\text{TLS}}$  in arbitrary units, modifies the local electric field at defect sites and modulates the TLS resonance frequency<sup>19</sup>. Therefore, representative characteristics such as  $T_1$  can be modulated by  $k_{\text{TLS}}$  due to subsequent changes in qubit–TLS interaction. We demonstrate that, on average, controlling the qubit–TLS interaction can help improve and even stabilize worst-case  $T_1$  instances. We then study the impact these results have on noise characteristics and gate layer performance. Finally, we leverage improved stability of qubit noise in the context of observable estimation using quantum error mitigation.

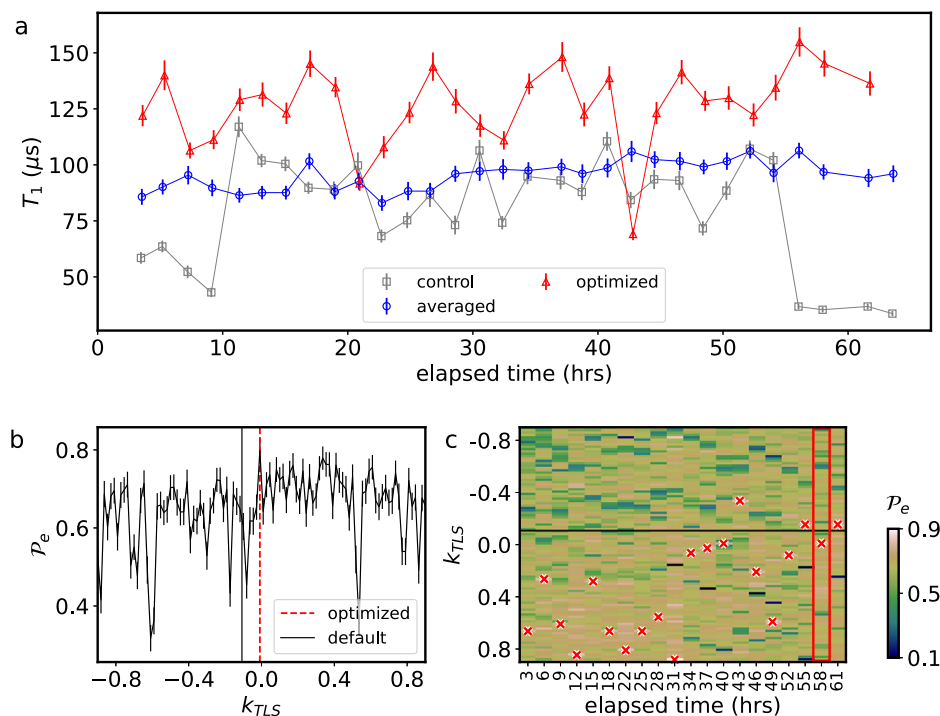
## Results

### Stabilizing

**$T_1$**  In this section, we focus on temporal fluctuations of the qubit–TLS interaction and different strategies to minimize its impact on  $T_1$ . Over a 60 h period, the qubit  $T_1$  values on the device are seen to fluctuate on average by over 300% (see Supplementary SII). Figure 1a depicts an example of these fluctuations for one of these qubits. We now consider modulation of the qubit–TLS interaction by means of a control parameter  $k_{\text{TLS}}$ . To study its effect, we use the excited state population of the qubit,  $\mathcal{P}_e$ , measured after a fixed delay time of 40  $\mu\text{s}$  as a quick proxy for  $T_1$ <sup>15</sup>. The resulting  $\mathcal{P}_e$  values obtained for a range of  $k_{\text{TLS}}$  parameters are shown in Fig. 1b. The peaks and dips in the plot are a reflection of the qubit–TLS interaction landscape as modulated by  $k_{\text{TLS}}$

at the given time instance. In the meantime, we observe that no further calibration is needed for the qubit itself, as we mainly modulate TLS characteristics unless we had a pronounced qubit–TLS interaction during initial qubit calibration. Repeating the same experiment at different times, Fig. 1c illustrates the temporal fluctuation of the qubit–TLS interaction.

The value of the  $k_{\text{TLS}}$  parameter clearly has a strong effect on  $\mathcal{P}_e$ , which raises the natural question of how best to select it. One way to improve qubit coherence is to actively monitor the temporal snapshot of the TLS landscape and choose  $k_{\text{TLS}}$  that produces the best  $\mathcal{P}_e$ . The results show a clear benefit from this optimization, improving the overall  $T_1$  in Fig. 1a. This strategy, which we refer to as the optimized noise strategy, requires active monitoring of the TLS environment. Between monitoring events, the qubit remains exposed to random fluctuations in the qubit–TLS interaction. Alternatively, we mitigate the impact of these fluctuations by averaging over randomly sampled TLS environments per shot. This is achieved by applying slowly varying sinusoidal (or triangular) amplitude modulation on  $k_{\text{TLS}}$ . The applied modulation frequency (1 Hz) is much lower than the shot repetition rate (1 kHz). Therefore, the modulation is effectively quasi-static within each shot, but samples a different quasi-static TLS environment for each shot over the duration of the entire experiment. We refer to this approach as the averaged noise strategy. Figure 1a illustrates that the averaged  $T_1$  value is more stable than that of the control and optimized experiments. Since this strategy only requires passive sampling of the TLS environment from shot to shot, it does not require constant monitoring. A natural question to consider is whether the  $T_1$  decay is well approximated by a single-exponential in this case—this is seen to hold for this data set (see Supplementary Fig. S2). However, this may



**Fig. 1 | Stabilizing  $T_1$ .** **a** Experimentally measured  $T_1$  for a given  $k_{\text{TLS}}$ . Black rectangles show  $T_1$  fluctuation without any optimization procedure. Red triangles illustrate  $T_1$  after optimization is carried out, as illustrated in **(b)**. Blue circles show  $T_1$  measured by averaging over the TLS landscapes. Averaging is achieved by modulating  $k_{\text{TLS}}$  with 1 Hz sine wave with amplitude 0.5 on top of the default value. The corresponding  $T_1$  decay curves are shown in the Supplementary material. **b** At each  $k_{\text{TLS}}$ , we prepare an excited state for a given qubit and then probe the probability of measuring the excited state,  $\mathcal{P}_e$ , after 40  $\mu\text{s}$ . Ideally, the qubit would stay in the excited state; however,  $T_1$  decay results in  $\mathcal{P}_e < 1$ <sup>15</sup>. The decay of  $\mathcal{P}_e$  therefore serves as a proxy for  $T_1$  decay. By changing the magnitude  $k_{\text{TLS}}$  of the TLS

modulation, we can alter the qubit–TLS interaction strength. This causes  $\mathcal{P}_e$  to vary as a function of  $k_{\text{TLS}}$ . When a strong qubit–TLS interaction exists, a pronounced decay of  $\mathcal{P}_e$  is observed<sup>15</sup>, as seen by the dips in the curve. The vertical lines indicate the choice of  $k_{\text{TLS}}$  that maximizes  $\mathcal{P}_e$  (dashed), compared to an arbitrarily chosen default  $k_{\text{TLS}}$  value (solid). **c** We monitor the TLS landscape over time to illustrate the fluctuation of qubit–TLS interaction dynamics. The black horizontal line indicates  $k_{\text{TLS}}$  of the control experiment, the red crosses indicate the values used for the optimized experiment. The red rectangular box indicates the data illustrated in **(b)**. The experiments are carried out with 1 kHz repetition rate. Measured values and error bars are obtained from 300 single-shot measurements.

not always be the case, as is often seen in scenarios with strong couplings to TLS, and the implications of this for noise learning and mitigation are discussed later and detailed in Supplementary SIII.

### Stabilizing noise in gate layers

Having demonstrated different modulation strategies to stabilize  $T_1$ , we now extend our study to the characterization of noise associated with layers of concurrent entangling two-qubit gates. An accurate characterization of the noise enables us to remove its effect on observable estimation by means of probabilistic error cancellation (PEC)<sup>3</sup>. Previous experiments have shown that noise on our devices can be tailored in such a way that a sparse Pauli–Lindblad (SPL) model<sup>8</sup> accurately captures the noise and facilitates the estimation of accurate observable values using PEC, and even enables ZNE experiments exceeding 100 qubits<sup>7</sup>. This motivates us to study the impact on the optimized and averaged noise strategies have on the learned SPL model parameters and their stability.

The SPL noise model proposed in ref. 8 provides a scalable framework for learning the noise associated with a layer of gates. The sparsity of the noise model is achieved by imposing assumptions on the noise in real hardware. We tailor and learn the noise employing protocols described in ref. 8 and references therein. First, by applying Pauli twirling<sup>30–34</sup>, we ensure that the noise can be described by a Pauli channel. We then model the noise as  $\mathcal{E}(\rho) = \exp[\mathcal{L}](\rho)$ , where  $\mathcal{L}$  represents a Lindbladian with Pauli jump terms  $P_k$  weighted by non-negative model coefficients  $\lambda_k$ . Second, we can obtain a sparse noise model by making the reasonable assumption that noise originates locally on individual or connected pairs of qubits. This allows us to restrict the set of generators  $\mathcal{K}$  to one- and two-local Pauli terms in accordance with the qubit topology. The model parameters  $\lambda_k$  are characterized by measuring the channel fidelities of Pauli operators using a procedure described in the Methods section. The fact that the individual  $P_k$  terms in  $\mathcal{L}$  commute make the noise model ideal for probabilistic error cancellation<sup>3,8</sup>, since the channels generated by each of these terms can be inverted independently. The inverse channels, however, are non-physical, and the observable is therefore reconstructed through post-processing. In the absence of model inaccuracy, we obtain an unbiased estimate for the expectation value of any Pauli operator. The variance of the estimator, however, is amplified by a multiplicative factor  $\gamma = \exp(\sum_{k \in \mathcal{K}} 2\lambda_k)$ , which can be compensated by increasing the number of samples<sup>8</sup>. We therefore refer to  $\gamma$  as the sampling overhead. In the following, we experimentally monitor individual model parameters,  $\lambda_k$ , and connect the overall noise strength to runtime overhead for error mitigation by using the sampling overhead,  $\gamma$ .

We now experimentally characterize the optimized and averaged noise channels over time with a goal of assessing whether the TLS modulation strategies succeed in stabilizing the noise, and therefore the model parameters. For this, we learn model parameters  $\lambda_k$  associated with two different gate layers for a one-dimensional chain of six qubits, covering all local CZ gate pairs. Figure 2a shows the parameter values obtained for the optimized noise channel for a single learning experiment. To quantify fluctuations in the noise, we repeat the noise characterization over time and track the learned model parameters over ~50 h of monitoring for the control, optimized, and averaged noise channels in Fig. 2b–d.

The control experiment depicts large fluctuations in the model parameters around 13 h of elapsed time, which shows reasonable correlation with a strong Q2–TLS interaction occurring around the same time (Fig. 2f). Meanwhile, the optimized experiment selects  $k_{\text{TLS}}$  that aims to avoid impact of nearby qubit–TLS impact while maximizing  $\mathcal{P}_e$ , indicated by red cross symbols in Fig. 2f. This optimization is performed right before learning experiment and help us to avoid configurations that have particularly strong qubit–TLS interaction. Aside from the relatively smaller aberrations associated with short-term fluctuations that are fixed by the next optimization round, the

model parameters are seen to be largely stable over the duration of the experiment, Fig. 2c. The smaller aberrations are seen to be further stabilized in the averaged case, Fig. 2d.

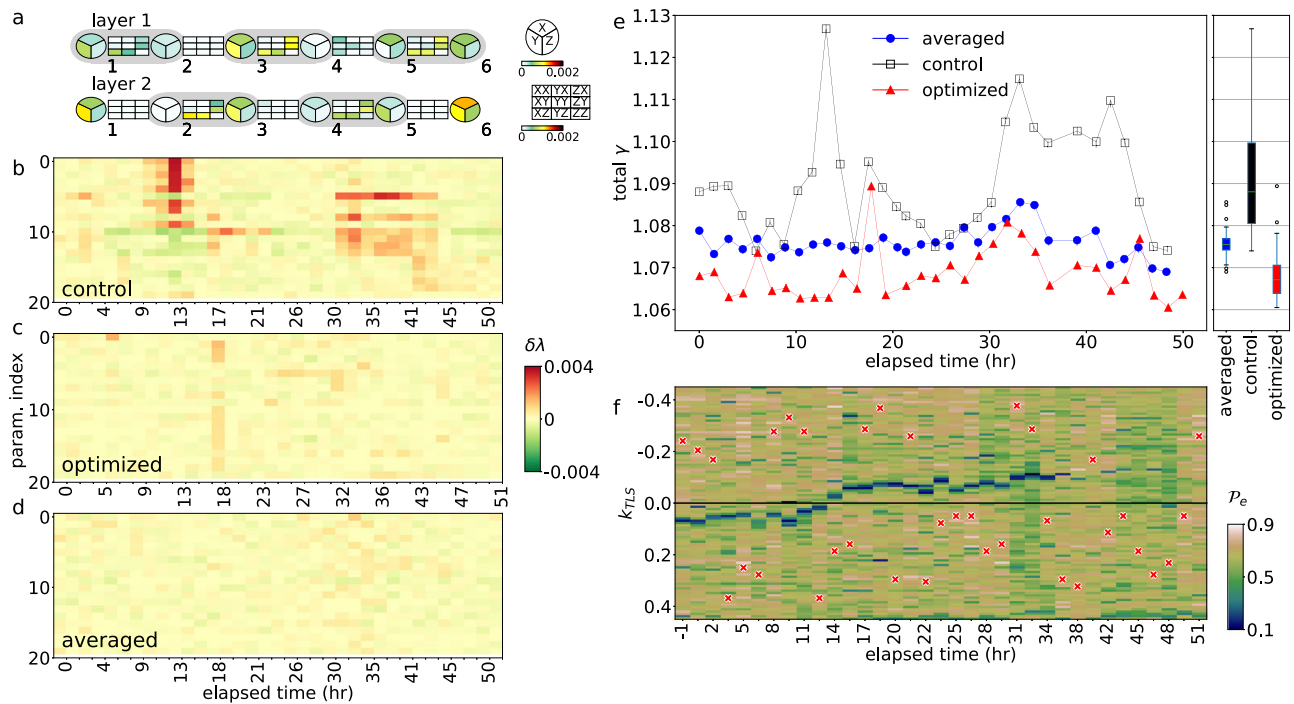
The time dependence of the model parameters can be further extended to track the stability of the total sampling overhead  $\gamma$ , which is given by the product of  $\gamma_1$  and  $\gamma_2$ , the sampling overhead values for each of the two layers, respectively. Figure 2e shows that the optimized strategy attains the lowest overall sampling overhead, while the averaged noise channel experiment exhibits better stability compared to the other two experiments. In addition, the stability we obtain from this scheme does not require frequent monitoring, and we expect more stable operation during periods when the experiment is running and monitoring jobs are not.

Additional analysis in Fig. 2b–f shows that the observed stability  $\gamma$  reflects the stability of single- and two-qubit model parameters. This is an important observation since temporal fluctuations or noise control strategies may induce discrepancies between learned parameters at one time and actual parameters at a later time, which would greatly reduce the ability to perform meaningful error mitigation. For instance, one might worry that, after each  $T_1$  optimization stage, the new  $T_1$  and, along with it, the noise channel may be completely different. Figure 2c experimentally shows that this is largely not the case. Aside from the relatively smaller aberrations associated with short-term fluctuations that are fixed by the next optimization round, most of the model coefficients remain relatively consistent for most of the time. Figure 2d similarly visualizes that the averaged noise models exhibit consistent model parameter values throughout the duration of the experiment.

### Stability of quantum error mitigation

In the previous section, it was shown that noise models for the optimized and averaged noise channels remain consistent over time. We now study their impact on the error mitigation of a benchmark circuit. The circuit we use is a mirror circuit on a chain of six qubits with alternating layers of CZ gates that cover all neighboring qubit pairs, as illustrated in Fig. 3. Since the circuit is mirrored, it effectively implements an identity operator; the ideal expectation value of all Pauli-Z observables is equal to 1. Each experiment consists of a noise learning and a mitigation stage. During the first stage, we learn the noise associated with the two layers of CZ gates for each of the control, optimized, and averaged noise channel settings. We then run error mitigation using the learned noise models in their respective settings. For each stage, we interleave the circuits for all three settings to ensure they all run within the same time window.

Figure 4a–c shows the mitigated and unmitigated values for the weight-6 observable  $\langle ZZZZZZ \rangle$  for the control, optimized, and averaged noise channels, respectively, with independent runs over a ~50 h period. In all three settings, the unmitigated values are  $0.341 \pm 0.052$ ,  $0.446 \pm 0.036$ , and  $0.371 \pm 0.027$  for control, optimized, and averaged, respectively—a large deviation from the ideal value of 1. While fluctuations in the mitigated observable values are seen in all three settings, they are most pronounced in the control setting, where they correlate with periods of strong TLS interaction, as shown earlier in Fig. 2f. In principle, one could run separate background circuits to monitor TLS interactions and trigger disposal of data acquired during periods of large fluctuation to improve performance. However, the limited shot budget for such monitoring circuits greatly reduces the ability to accurately detect all but the largest fluctuation events, as discussed in Supplementary SIV and SV. Another strategy to reduce fluctuations in the observable values is to average the results of multiple independent learning-mitigation cycles, as seen from the cumulative average trace of Fig. 4a. However, such an approach is increasingly expensive for larger circuits, as discussed in Supplementary SIV. Meanwhile, Fig. 4b, c demonstrates that both the optimized and averaged noise strategies help stabilize the error mitigation results and enable smaller



**Fig. 2 | Stabilizing noise in gate layers.** **a** Graphical representation of the coefficients of the sparse Pauli-Lindblad noise models of two layers. The layers consist of CZ gates covering different qubit pairs (shaded box): {(1, 2), (3, 4), (5, 6)} for layer 1 and {(2, 3), (4, 5)} for layer 2. The model parameters apply to Pauli terms on each of the six qubits, as well as weight-two Pauli terms on connected qubits. The model parameters  $\{\lambda_k\}_{k \in \mathcal{K}}$  are determined by applying the learning protocol separately to each of the two layers. The inset on the right depicts the position of the Pauli coefficients and the color bars for the one- and two-qubit terms. **b–d** Provides a more detailed picture of model parameter instability by computing the model coefficient fluctuation  $\delta\lambda_k(t) = \lambda_k(t) - \text{median}[\lambda_k(t)]$ , where  $\text{median}[\cdot]$  computes a median value of the time-varying model coefficient. The plot shows  $\delta\lambda_k(t)$  at a specific time  $t$  (x-axis) for various Pauli jump terms  $P_k$  (y-axis). We show the first 20 parameters sorted by maximum fluctuation. **e** Shows the total sampling overhead,  $\gamma$ , monitored over time for three different scenarios: (i) a baseline control experiment with  $k_{\text{TLS}} = \kappa$  held at a constant neutral point  $\kappa$  for all qubits; (ii) an averaged experiment where  $k_{\text{TLS}}$  is set to  $\kappa$  plus a slowly varying triangular wave with 1 Hz

frequency and amplitude of  $\pm 0.2$ ; and (iii) an optimized experiment carried out by periodically update  $k_{\text{TLS}}$  to a value that maximizes  $\mathcal{P}_e$ . Optimization is performed just prior to each learning experiment for the optimized noise channel. The right inset illustrates the median value and the first and third quartiles of each experiment. **f** For the optimized experiment, the qubit-TLS interaction landscape of Q2 is probed using TLS control parameter  $k_{\text{TLS}}$ . The plot displays the resulting  $\mathcal{P}_e$  over time, along with the optimized control parameters, indicated by the red crosses. Strong qubit-TLS interactions appear as dark green boxes, and can be seen to drift close to the neutral point  $\kappa$  (horizontal black line) at an elapsed time of ~13 h. This coincides with the elevated noise level  $\gamma$  and noticeably larger fluctuations over time in plot (**e**). The data used for this plot is the same as that used for plot (**b–d**). Error bars in **e** are determined using 100 bootstrapped instances of the experimental data. Likewise, error bars for  $\lambda_k$  are obtained from 100 bootstrapped instances, and the maximum fluctuation of each row in (**b–d**) is all larger than the error bar. Further details on experiment conditions are described in the “Methods” section.

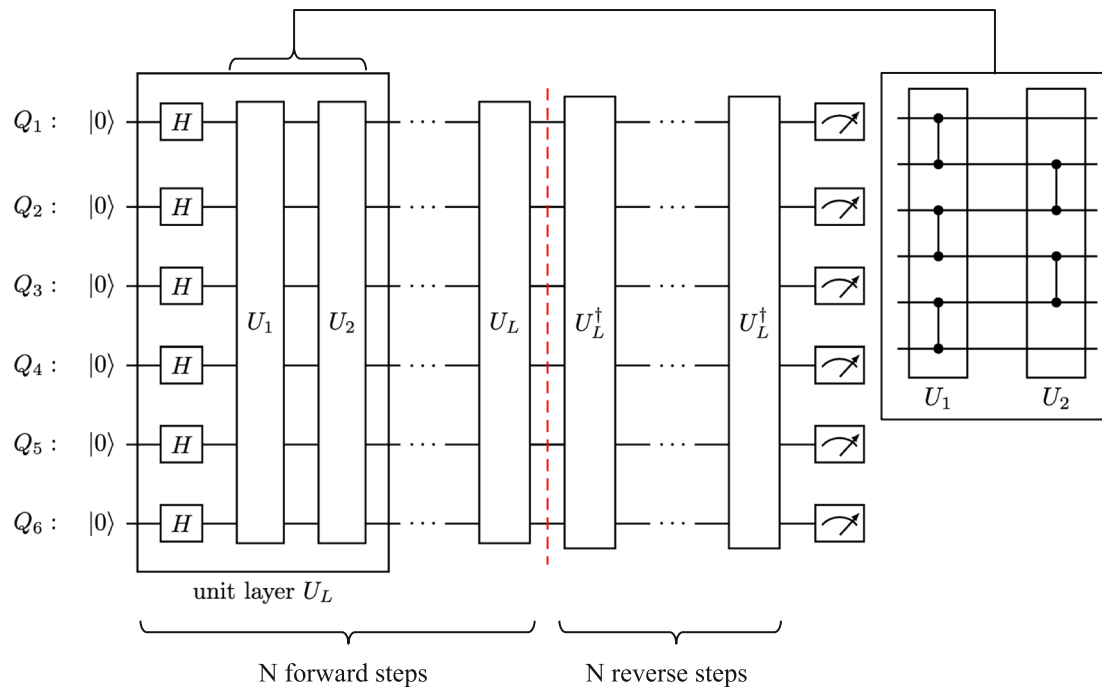
fluctuations than observed in the control experiment. The improvement is predicted to be even more prominent at larger depths, see Supplementary Fig. S7.

A leading source of fluctuations in the mitigated results in all three settings could be attributed to drifts in the device noise between the noise-learning step and the subsequent mitigation step. To quantify such deviations, without relearning the noise model, we proceed as follows. First, note that Clifford circuits subject to Pauli noise can be simulated efficiently. This allows us to predict the expected noisy observable value  $\langle \tilde{O} \rangle_{\text{pred}} = f_{\text{pred}}(\langle O \rangle)$ , where  $\langle O \rangle$  is an ideal observable value. Second, by running unmitigated benchmark circuits interleaved with the mitigation circuits, we can measure both the noisy observable value  $\langle \tilde{O} \rangle = f_{\text{exp}}(\langle O \rangle)$ , and the error-mitigated estimate  $\langle O \rangle_{\text{mit}}$ . In the absence of noise-model fluctuations, we have  $f_{\text{exp}} = f_{\text{pred}}$ , which allows us to recover the ideal observable  $\langle O \rangle = \langle \tilde{O} \rangle / f_{\text{pred}}$ . In the presence of the noise fluctuations, however, this no longer holds. The noise fluctuation may lead to under- or over-estimation on the target observable, and sometime even results in unphysical value. We quantify this known source of deviation on the mitigated observable as  $\delta_{\text{pred}} = \langle \tilde{O} \rangle / f_{\text{pred}} - \langle O \rangle = \langle \tilde{O} \rangle / f_{\text{pred}} - 1$ , where  $\langle O \rangle = 1$  for our benchmarking mirror circuit. In Fig. 4d, we plot an expected deviation due to the noise fluctuation ( $\delta_{\text{pred}}$ ) on the x-axis and an observed deviation ( $\delta_{\text{mit}} = \langle O \rangle_{\text{mit}} - 1$ ) on the y-axis. The plot shows a clear correlation

between  $\delta_{\text{pred}}$  and  $\delta_{\text{mit}}$ . This quantifies that time fluctuation plays a major role in the observed spread of the error-mitigated observable. A similar analysis applies to the optimized and averaged noise channels in Fig. 4e, f, albeit with a distribution that is packed more closely around the origin. The tighter histograms in the inset of Fig. 4e, f highlight that both optimized and averaged noise channels effectively stabilize the temporal fluctuation of the error mitigation performance. In addition to the Z parity analyzed here, we extend the comparison to observables of all weights in Supplementary Fig. S6.

We note that additional sources of bias in the error-mitigated observables may remain, even for the average and optimized experiments. One source of bias that is important to consider, particularly for the averaged noise case, is the effect of quasi-static noise on learning and mitigation, as detailed in Supplementary SIII. For instance, quasi-static noise can lead to noise learning circuit fidelities that do not follow a clean single-exponential decay with increasing depth, introducing bias in the mitigation that relies on an assumption of exponential decay. These effects are also relevant in the absence of any modulation, due to the natural temporal fluctuations in the TLS landscape and data collection over long periods of time. This essentially implies that the noise channel is, in practice, always quasi-static to some degree. As such, one remaining question of broad interest is how the quasi-static nature of noise, resulting from modulation at a shot-to-





**Fig. 3 | Benchmark circuit for error mitigation.** We use a 6-qubit mirror circuit to benchmark the error-mitigation performance. The circuit features two unique layers,  $U_1$  and  $U_2$ , of two-qubit entangling  $R_{ZZ}(\pi/2) = CZ$  gates, as shown in the inset on the right. We next define a compound block  $U_L$  of single-qubit Hadamard gates

on each qubit, followed by the two layers. The benchmark circuit is then constructed by repeating the  $U_L$  circuit  $N = 10$  times, followed by an equal number of reverse operations  $U_L^\dagger$ .

shot basis or intrinsic fluctuations, affects observable estimates and error mitigation. We explore this partially in Supplementary SIII, but a general investigation of these other sources and further closing the gap between ideal and mitigated observables remains a question for future work.

## Discussion

We experimentally demonstrate that noise in superconducting quantum processors can be stabilized by modulating the qubit–TLS interaction to improve the performance of error mitigation. Amongst the considered modes of operation, the optimized strategy gives the best performance, but remains exposed to random temporal fluctuations of qubit–TLS interaction between parameter re-optimizations, which could be particularly frequent at large qubit counts. By contrast, the averaged strategy smooths out small fluctuations and produces a more stable device noise model, albeit at the cost of a slightly increased sampling overhead for error mitigation and potential bias in observable estimates. While these modulation strategies are complementary to the continued development of cleaner devices and novel designs to reduce the density and impact of defect two-level systems, we expect that the methods discussed in this work will be crucial for the reliability of error-mitigated quantum computation with solid-state processors, particularly at scales beyond exact verification.

For interested readers, we point to ref. 35 for additional details of our implementation and control of the qubit–TLS modulation. We also note subsequent work<sup>36</sup> on multi-qubit noise stabilization, using voltage control of TLS, that appeared during peer review of our paper.

## Methods

### Optimized noise channel

One way to improve qubit coherence is to actively monitor the temporal snapshot of the TLS landscape and choose  $k_{\text{TLS}}$  that produces the best qubit property of interest. Assuming that we sample a TLS environment  $i$ , with  $i \in \{1, 2, \dots, N_t\}$  for  $N_t$  different TLS environments with corresponding  $k_{\text{TLS}}$  values, the ideal unitary  $\mathcal{U}$  is then exposed to a

given TLS environment. Along with other noise sources, the given qubit–TLS interaction incurs a noise channel  $\mathcal{E}_i$ , resulting in a noisy operator  $\tilde{\mathcal{U}}_i = \mathcal{U} \circ \mathcal{E}_i$ . The observable expectation value is then computed as

$$\langle \mathcal{O} \rangle_i = \text{Tr} [\mathcal{O} \tilde{\mathcal{U}}_i(\rho)] \quad (1)$$

for a given initial state  $\rho$ . One may sample an optimal TLS environment  $i$  that minimizes the bias of Eq. (1) from its ideal value. In practice, we choose a TLS control parameter that optimizes a representative metric, such as  $T_1$  or a single-qubit randomized benchmarking result. This requires active monitoring of the TLS environment via this representative metric. Throughout our study, we choose  $\mathcal{P}_e$  (a proxy of  $T_1$ ) as an optimization metric. We aim for the best target operation, assuming the best coherence can be achieved by minimizing the qubit–TLS interaction. However, the qubit is still exposed to random fluctuations in qubit–TLS interaction between monitoring events.

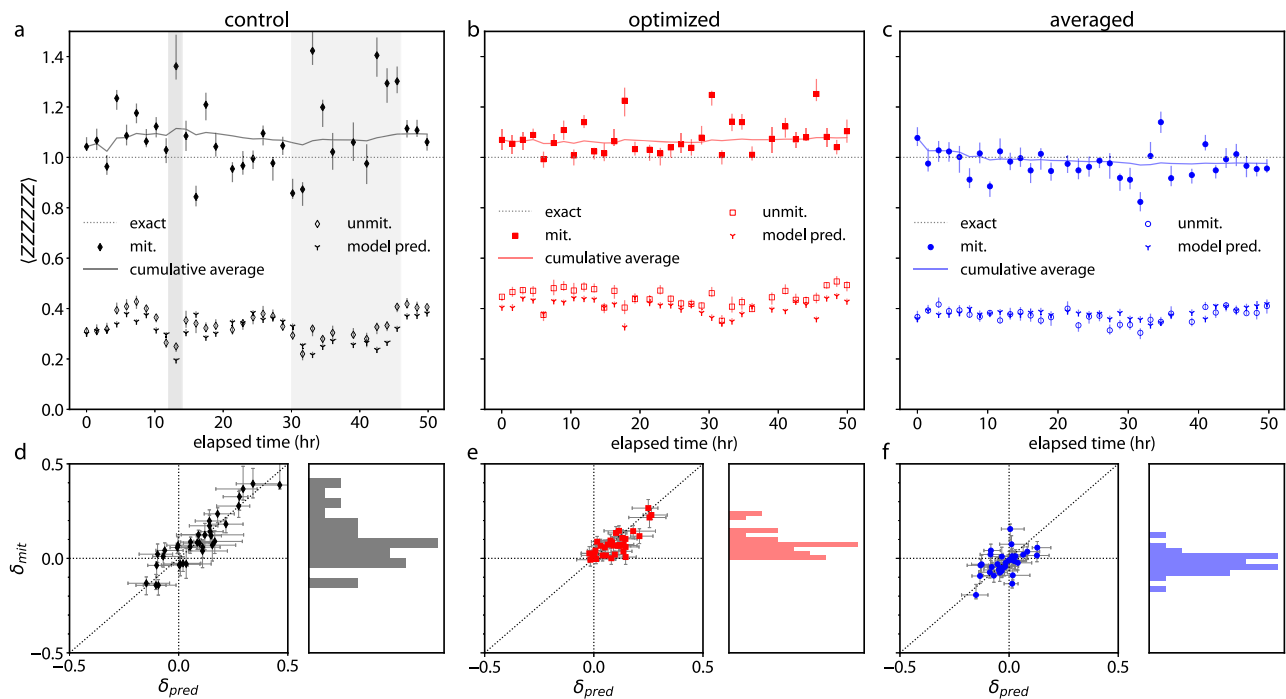
### Averaged noise channel

Alternatively, we may obtain an observable by averaging over randomly sampled TLS environments for each single-shot measurement. In the limit of infinite shots, and assuming the TLS environment is static for the duration of a shot, the observable expectation value obtained by averaging over sampled TLS environments is given by

$$\begin{aligned} \langle \bar{\mathcal{O}} \rangle &= \sum_i p_i \langle \mathcal{O} \rangle_i = \sum_i p_i \text{Tr} [\mathcal{O} \tilde{\mathcal{U}}_i(\rho)], \\ &= \text{Tr} \left[ \mathcal{O} \sum_i p_i \tilde{\mathcal{U}}_i(\rho) \right] = \text{Tr} [\mathcal{O} \bar{\mathcal{U}}(\rho)], \end{aligned} \quad (2)$$

where  $p_i$  is the probability of having noise realization  $i$  at a given shot, and we define an averaged noise channel

$$\bar{\mathcal{U}} = \sum_i p_i \tilde{\mathcal{U}}_i. \quad (3)$$



**Fig. 4 | Error mitigation results for different qubit–TLS modulation strategies.** **a–c** Weight-6 observable ( $\langle ZZZZZZ \rangle$ ) estimates as a function of time using the three different strategies with (filled markers) and without (open markers) error mitigation, along with the cumulative average of the mitigated observable values (solid line). The experiment is performed following the schedule described in Fig. 2 for the **a** control, **b** optimized, and **c** averaged modulation strategies. For reference, the ideal observable of 1 is indicated as a dashed line. The mitigated observables can be seen to fluctuate near the ideal value, and the shaded regions highlight time windows with high fluctuations as determined by the analysis in Fig. 2d. Each data point is obtained from 4096 random circuit instances with 32 shots per circuit. During

the experiment, we interleave 2048 random circuit instances for readout-error mitigation<sup>38</sup> and 512 random circuit instances for estimating the unmitigated observable. Readout-error mitigation is applied to both the unmitigated and mitigated observable estimates. Error bars for the unmitigated and mitigated results are obtained by bootstrapping the PEC result 25 times. **d–f** Scatter plots of the predicted ( $\delta_{pred}$ ) and observed ( $\delta_{mit}$ ) deviations of the observable from the ideal expectation value. The correlation between  $\delta_{pred}$  and  $\delta_{mit}$  confirms that the temporal fluctuation of the noise model plays a role in the mitigation error observed in the PEC protocol. The histograms along the y-axis show the respective distributions of  $\delta_{mit}$ .

The averaged noise channel may not be optimal in terms of minimizing the bias on the observable of interest. Nevertheless, the resulting variation in the observable value due to a local fluctuation of the TLS environment is reduced by averaging over the sampled TLS environments. The method aims to provide a stable observable estimate in the presence of natural fluctuations of the TLS environment on timescales shorter than the total duration of data collection for an experiment. This method only requires passive sampling of the TLS environment from shot to shot; thus, constant monitoring activities are not required, unlike the optimized noise channel scenario.

### Characterizing the model parameters

By applying Pauli twirling, we can assume the noise channel to be a Pauli channel; that is  $\mathcal{E}_i(\rho) = \sum_k \alpha_{k,i} P_k(\rho) P_k^\dagger$  where the  $\alpha_{k,i} \geq 0$  terms sum up to one. Sampling TLS environment  $i$  by changing modulation parameter  $k_{TLS}$  changes the noise channel  $\mathcal{E}_i$ , and, consequently, the corresponding effective gate operation  $\mathcal{U}_i$ . In this section, we study how the different TLS modulation strategies affect the noise associated with two layers of concurrent CZ gates.

Learning a full  $n$ -qubit Pauli noise channel requires the characterization of  $4^n$  parameters  $\alpha_k$ , and is therefore feasible only for a small number of qubits. The sparse Pauli-Lindblad noise model proposed in<sup>8</sup> provides a scalable alternative obtained by imposing more structure on the noise. In particular, the noise is assumed to be of the form  $\mathcal{E}(\rho) = \exp[\mathcal{L}](\rho)$ , where  $\mathcal{L}$  represents a Lindbladian

$$\mathcal{L}(\rho) = \sum_{k \in \mathcal{K}} \lambda_k (P_k \rho P_k^\dagger - \rho),$$

with Pauli jump terms and non-negative coefficients  $\lambda_k$ . Sparsity is achieved by making the reasonable assumption that noise originates locally on individual or connected pairs of qubits, which is done by restricting the set of generators  $\mathcal{K}$  to all one- and two-local Pauli terms in accordance with the qubit topology. The set of model parameters  $\lambda_k$  are characterized by measuring the set of Pauli fidelities. The Pauli fidelity for a Pauli  $P_k$  is given by

$$f_k = \frac{1}{2^n} \text{Tr}(P_k \mathcal{E}(P_k)) = \exp\left(-2 \sum_{k' \in \mathcal{K}} \lambda_{k'} \langle k, k' \rangle_{sp}\right), \quad (4)$$

where the symplectic inner product  $\langle k, k' \rangle_{sp}$  is zero if Paulis  $P_k$  and  $P_{k'}$  commute, and one otherwise. As described in more detail in<sup>8</sup> and references therein, repeated application of the noisy layer of gates allows us to learn fidelity pairs  $f_k f_{k'}$ , where  $P_{k'} = \pm \mathcal{U}(P_k)$ . In general, it is not possible to measure the individual fidelities in a readout-error-free manner<sup>37</sup>, and a symmetry assumption that equates the fidelities appearing in each pair is therefore imposed. Given the fidelities, a nonnegative least-squares problem based on Eq. (4) can then be used to obtain the model parameters  $\lambda$ .

The sparse PL noise model can equivalently be expressed as a series of Pauli channels of the form  $\Lambda_k(\rho) = (w_k \rho + (1 - w_k) P_k \rho P_k^\dagger)$ , where  $w_k = \frac{1}{2}(1 + \exp(-2\lambda_k))$ . The fact that these Pauli channels commute makes the noise model ideal for probabilistic error cancellation<sup>3,8</sup>, since each channel can be inverted independently. Here, the inverse channel is non-physical, and the observable is therefore reconstructed through post-processing. Once we insert the appropriate canceling Paulis following the quasi-probability described

by  $w_k$ , the measured outcome is multiplied by a pre-factor given by

$$\gamma = \prod_{k \in K} (2w_k - 1)^{-1} = \exp\left(\sum_{k \in K} 2\lambda_k\right). \quad (5)$$

While we obtain an unbiased estimate, the variance is amplified by this pre-factor. Therefore, more sampling is required to compensate for the increased variance due to the scaling. For example, the sampling overhead is proportional to  $\gamma^2$  for PEC<sup>8</sup>, thus we refer  $\gamma$  to as a sampling overhead. The sampling overhead is a useful metric connecting the underlying noise strength to an experimental overhead for error mitigation. For instance, we track the sampling overhead to visualize the overall model coefficient fluctuation in Fig. 2e. The sampling overhead also provides a useful estimation of the relative runtime cost between different error rates. For instance, the worst (best) case sampling overhead for a given set of two-qubit layers in Fig. 2e is roughly  $\gamma_w = 1.13$  ( $\gamma_b = 1.06$ ). The ratio  $(\gamma_w^2/\gamma_b^2)^N$  could be translated as a relative extra sampling cost for PEC for the benchmark circuit with a repeated unit layer depth  $N$  between the worst and the best error rate scenario. As an example, the worst error rate scenario requires ~13 (~167) times more circuits to perform PEC than the best error rate scenario for  $N = 20$  ( $N = 40$ ).

The learning protocol consists of two steps. First, the learning procedure extracts a set of Pauli fidelities,  $\{f_b\}_{b \in B}$ , defined by a set of Pauli operators  $B$  much smaller than the set of all Pauli operators. Second, we use nonnegative least-squares minimization to extract the parameters  $\{\lambda_k\}_{k \in K}$  of our sparse PL model. The Pauli fidelities that are used to extract model parameters described in Fig. 2 are obtained by repeating the gate sequences an even number of times ( $n = 0, 4, 12, 24, 64$ ) for both layer 1 and layer 2, independently. At each depth, we generate 60 circuit instances with randomly sampled Pauli for Pauli twirling for gates and final measurements. All experiments were carried out with 1 kHz repetition rate and observable values for each circuit are estimated from 32 single shot measurements.

## Data availability

The data sets generated and analyzed during the current study will be available at a public repository, 10.6084/m9.figshare.29614709.

## References

- Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
- Bharti, K. et al. Noisy intermediate-scale quantum algorithms. *Rev. Mod. Phys.* **94**, 015004 (2022).
- Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.* **119**, 180509 (2017).
- Li, Y. & Benjamin, S. C. Efficient variational quantum simulator incorporating active error minimization. *Phys. Rev. X* **7**, 021050 (2017).
- Kandala, A. et al. Error mitigation extends the computational reach of a noisy quantum processor. *Nature* **567**, 491 (2019).
- Kim, Y. et al. Scalable error mitigation for noisy quantum circuits produces competitive expectation values. *Nat. Phys.* **19**, 752–759 (2023).
- Kim, Y. et al. Evidence for the utility of quantum computing before fault tolerance. *Nature* **618**, 500 (2023).
- van den Berg, E., Mineev, Z. K., Kandala, A. & Temme, K. Probabilistic error cancellation with sparse Pauli-Lindblad models on noisy quantum processors. *Nat. Phys.* **19**, 1116 (2023).
- Filippov, S., Leahy, M., Rossi, M. A. C. & García-Pérez, G., Scalable tensor-network error mitigation for near-term quantum computing. Preprint at *arXiv* <https://arxiv.org/abs/2307.11740> (2023).
- Burnett, J. et al. Evidence for interacting two-level systems from the 1/f noise of a superconducting resonator. *Nat. Commun.* **5**, 4119 (2014).
- Lisenfeld, J. et al. Observation of directly interacting coherent two-level systems in an amorphous material. *Nat. Commun.* **6**, 6182 (2015).
- de Graaf, S. E. et al. Direct identification of dilute surface spins on Al<sub>2</sub>O<sub>3</sub>: origin of flux noise in quantum circuits. *Phys. Rev. Lett.* **118**, 057703 (2017).
- Klimov, P. V. et al. Fluctuations of energy-relaxation times in superconducting qubits. *Phys. Rev. Lett.* **121**, 090502 (2018).
- De Graaf, S. et al. Two-level systems in superconducting quantum devices due to trapped quasiparticles. *Sci. Adv.* **6**, eabc5055 (2020).
- Carroll, M., Rosenblatt, S., Jurcevic, P., Lauer, I. & Kandala, A. Dynamics of superconducting qubit relaxation times. *npj Quantum Inf.* **8**, 132 (2022).
- Thorbeck, T., Eddins, A., Lauer, I., McClure, D. T. & Carroll, M. Two-level-system dynamics in a superconducting qubit due to background ionizing radiation. *PRX Quantum* **4**, 020356 (2023).
- Thorbeck, T., Xiao, Z., Kamal, A. & Govia, L. C. G. Readout-induced suppression and enhancement of superconducting qubit lifetimes. *Phys. Rev. Lett.* **132**, 090602 (2024).
- Sarabi, B., Ramanayaka, A. N., Burin, A. L., Wellstood, F. C. & Osborn, K. D. Projected dipole moments of individual two-level defects extracted using circuit quantum electrodynamics. *Phys. Rev. Lett.* **116**, 167002 (2016).
- Lisenfeld, J. et al. Electric field spectroscopy of material defects in transmon qubits. *npj Quantum Inf.* **5**, 105 (2019).
- Grabovskij, G. J., Peichl, T., Lisenfeld, J., Weiss, G. & Ustinov, A. V. Strain tuning of individual atomic tunneling systems detected by a superconducting qubit. *Science* **338**, 232 (2012).
- Abdurakhimov, L. V. et al. Driven-state relaxation of a coupled qubit-defect system in spin-locking measurements. *Phys. Rev. B* **102**, 100502 (2020).
- Etzezarreta Martinez, J., Fuentes, P., Crespo, P. & Garcia-Frias, J. Time-varying quantum channel models for superconducting qubits. *npj Quantum Inf.* **7**, 115 (2021).
- Dasgupta, S., Danageozian, A. & Humble, T. S. Adaptive mitigation of time-varying quantum noise. In *IEEE international conference on Quantum computing and Engineering (QCE)*. vol. 1, 99–110 (IEEE, 2023).
- Dasgupta, S. Stability of quantum computers. Preprint at *arXiv* <https://arxiv.org/abs/2404.19082> (2024).
- McKay, D. C. et al. Universal gate for fixed-frequency qubits via a tunable bus. *Phys. Rev. Appl.* **6**, 064007 (2016).
- Yan, F. et al. Tunable coupling scheme for implementing high-fidelity two-qubit gates. *Phys. Rev. Appl.* **10**, 054062 (2018).
- Mundada, P., Zhang, G., Hazard, T. & Houck, A. Suppression of qubit crosstalk in a tunable coupling superconducting circuit. *Phys. Rev. Appl.* **12**, 054023 (2019).
- Foxen, B. et al. Demonstrating a continuous set of two-qubit gates for near-term quantum algorithms. *Phys. Rev. Lett.* **125**, 120504 (2020).
- Stehlik, J. et al. Tunable coupling architecture for fixed-frequency transmon superconducting qubits. *Phys. Rev. Lett.* **127**, 080505 (2021).
- Bennett, C. H. et al. Purification of noisy entanglement and faithful teleportation via noisy channels. *Phys. Rev. Lett.* **76**, 722 (1996).
- Knill, E. Fault-tolerant postselected quantum computation: threshold analysis. <https://doi.org/10.48550/ARXIV.QUANT-PH/0404104> (2004).
- Kern, O., Alber, G. & Shepelyansky, D. L. Quantum error correction of coherent errors by randomization. *Eur. Phys. J. D.* **32**, 153 (2005).
- Geller, M. R. & Zhou, Z. Efficient error models for fault-tolerant architectures and the Pauli twirling approximation. *Phys. Rev. A* **88**, 012314 (2013).
- Wallman, J. J. & Emerson, J. Noise tailoring for scalable quantum computation via randomized compiling. *Phys. Rev. A* **94**, 052325 (2016).

35. Dane, A. et al. Performance stabilization of high-coherence superconducting qubits. Preprint at *arXiv* <https://arxiv.org/abs/2503.12514> (2025).
36. Chen, L. et al. Scalable and site-specific frequency tuning of two-level system defects in superconducting qubit arrays. Preprint at *arXiv* <https://arxiv.org/abs/2503.04702> (2025).
37. Chen, S. et al. The learnability of Pauli noise. *Nat. Commun.* **14**, 52 (2023).
38. van den Berg, E., Mineev, Z. K. & Temme, K. Model-free readout-error mitigation for quantum expectation values. *Phys. Rev. A* **105**, 032620 (2022).

## Acknowledgements

We acknowledge Dan Rugar, Robert M. Shelby, John Mamin, Jerry Tersoff, Sami Rosenblatt, Majo Lozano, Petar Jurcevic, Matthias Steffen and Oliver Dial for insightful discussions; Oliver Dial, Edward H. Chen, and Alireza Seif for feedback on the manuscript. For modeling the learning procedure under quasi-static noise, LCGG acknowledges support from the Army Research Office under Grant Number W911NF-21-1-0002. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes, notwithstanding any copyright notation herein.

## Author contributions

Y.K. performed error mitigation experiments. L.C.G.G. and E.P. performed numerical and theoretical analysis on quasi-static noise. A.D., D.M.Z., J.S. and K.B. established configurations to perform quantum experiments with quasi-static modulation on the TLS knob. B.M. and E.v.d.B. provided methods and code base to perform probabilistic error cancellation experiments. D.M.Z. and Y.L. configured the cryo-environment. D.M.Z., Y.L., and Y.K. calibrated the device. G.K., A.S., and J.S. contributed to the design and fabrication of the device. Y.K., L.C.G.G., A.D., and A.K. analyzed data and wrote the paper with input from all of the authors. A.K. managed the project.

## Competing interests

Elements of this work are included in a patent filed by the International Business Machines Corporation with the US Patent and Trademark Office. The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-62820-9>.

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

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

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

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025