

Mitigating errors in state preparation and measurement with noncomputational states

Conrad J. Haupt^{1,2}, Almudena Carrera Vazquez¹, Laurin E. Fischer^{1,3}, Stefan Woerner¹,
and Daniel J. Egger^{1,*}

¹IBM Quantum, IBM Research Europe – Zurich, 8803 Rüschlikon, Switzerland

²Laboratory of Theoretical Physics of Nanosystems, École Polytechnique Fédérale de Lausanne,
1015 Lausanne, Switzerland

³Theory and Simulation of Materials, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland



(Received 30 June 2025; revised 3 December 2025; accepted 8 December 2025; published 25 February 2026)

Error mitigation has enabled quantum computing applications with more than a hundred qubits and deep circuits. Many error mitigation methods are noise-aware, relying on a faithful characterization of the noise channels of the hardware. However, fundamental limitations lead to unlearnable degrees of freedom of the underlying noise models when considering qubits. Here, we show how to leverage noncomputational states as an additional resource to learn state-preparation errors in superconducting qubits. This allows one to fully constrain the noise models. We can thus independently and accurately mitigate state-preparation errors, gate errors, and measurement errors. Our proposed method is also applicable to dynamic circuits with midcircuit measurements. This work opens the door to improved error mitigation for measurements, both at the end of the circuit and midcircuit.

DOI: [10.1103/jv9h-spzw](https://doi.org/10.1103/jv9h-spzw)

I. INTRODUCTION

Progress in quantum computing hardware and error mitigation has enabled utility-scale experiments [1–3]. At this scale, a brute-force classical simulation of the underlying quantum circuits is no longer possible. Moving forward, error mitigation may enable a quantum advantage before the onset of full fault tolerance [4]. In the fault-tolerant regime, error mitigation will remain relevant by enhancing performance and reducing residual logical errors beyond the capabilities of error correction alone [5,6]. Importantly, quantum computing is not only made of sequences of unitary gates. Midcircuit measurements (MCMs) offer a powerful computational extension. They form the bedrock of dynamic circuits and enable entanglement distribution [7], improved algorithmic execution [8–10], and circuit cutting [11–15].

Many error mitigation methods rely on accurate noise learning experiments, which also help characterize quantum processors at scale [1,16,17]. Gate noise is commonly characterized by repeating a gate layer to amplify its noise,

e.g., in cycle benchmarking [18,19]. A common assumption is to express the gate noise models as Pauli channels; this is, in practice, justified by applying Pauli twirling or randomized compiling [20–22]. Pauli noise learning is made scalable by “sparsifying” the noise generators [16]. This is typically done by dropping the noise generators that do not match the physical qubit couplings. In this framework, state-preparation and final measurement (SPAM) errors can be characterized with a model-free twirled readout circuit in a technique popularized as “twirled readout error extinction” (TREX) [23]. The SPAM errors on expectation value estimators are then mitigated in postprocessing. By contrast, learning the noise models of MCMs requires extending cycle benchmarking [24,25]. Here, readout-induced leakage can be characterized [26] and probabilistic error cancellation has been generalized to dynamic circuits [27–29]. Crucially, when separating SPAM errors from gate errors, there are fundamentally unlearnable degrees of freedom. These degrees of freedom in the noise model are rigorously understood as the cut space of a cycle graph [30]. The same happens when separating state from measurement errors. These unlearnable degrees of freedom can be circumvented by imposing symmetry assumptions on the noise models [16], which may be violated in real experiments. In some cases this comes with the cost of additional ancilla qubits and assumed ideal CNOT gates, implemented with error mitigation [31]. For successful noise-model-based

*Contact author: deg@zurich.ibm.com

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

error mitigation, it can be crucial to go beyond these naive symmetry assumptions [2].

Recent work has thus been focused on producing self-consistent noise characterization protocols that learn gate, state-preparation, and measurement noise. This approach results in a holistic noise model up to unlearnable *gauge* degrees of freedom, such that any measured observables in circuits composed of the given state-preparation, gate, and measurement layers are insensitive to the gauge [32]. While this approach is sufficient for noise-learning-based error mitigation of unitary circuits [33], it does not reveal the ground truth for all parameters of the noise model, such as state-preparation errors. This can limit certain uses of the noise model, such as combining noise models of different layers to form models for gate structures that have not been learned. Furthermore, its applicability to dynamic circuits with measurements and classical feed-forward still needs to be investigated. Other methods for separating state-preparation from measurement noise come with trade-offs, such as additional qubits [34,35], two-qubit gates to perform noise learning [31,36], or repeated measurements during mitigation [37].

Here, we show how noncomputational states help learn state-preparation errors and establish a ground truth for a noise model of midcircuit measurements. By using noncomputational states, we can overcome the existing no-go theorems for noise learning [30,32]. We demonstrate this in superconducting qubit hardware [38], and expect that the method should generalize to other hardware platforms where noncomputational states are addressable.

In Sec. II, we start by demonstrating the importance of characterizing state-preparation errors separately from measurement errors. In Sec. III, we propose a protocol to fix a ground truth in the noise model of MCMs. Finally, in Sec. IV we present numerical results of the proposed protocol. We discuss and conclude our work in Sec. V.

II. MOTIVATION

We now present an experiment, conducted on an IBM Quantum Eagle device, showing that jointly mitigating state-preparation errors and measurement errors fails to error mitigate global observables of up to eight qubits. A Greenberger-Horne-Zeilinger (GHZ) state-preparation circuit with a ladder structure of CNOT gates, shown in Fig. 1(a), turns the single-qubit Z_0 operator into a global $X^{\otimes n}$ operator. Hence, this $+1$ stabilizer is subject to measurement errors from all qubits, while only being affected by state-preparation errors from the first qubit. This makes the observable $O = X^{\otimes n}$ maximally sensitive to the separation of SPAM errors. We would thus like to study the circuit of Fig. 1(a). However, to account for gate noise we would need to learn the noise of the $n - 1$ distinct layers of CNOT gates [16,17]. To reduce this overhead, we convert the circuit in Fig. 1(a) to a similar circuit built from only two distinct layers of CNOT gates, see Fig. 1(b), which features the same behavior for the $X^{\otimes n}$ observable.

We mitigate the circuit in Fig. 1(b) with TREX [23] to illustrate how inaccurate noise models can result in unphysical mitigated results. We start by measuring the

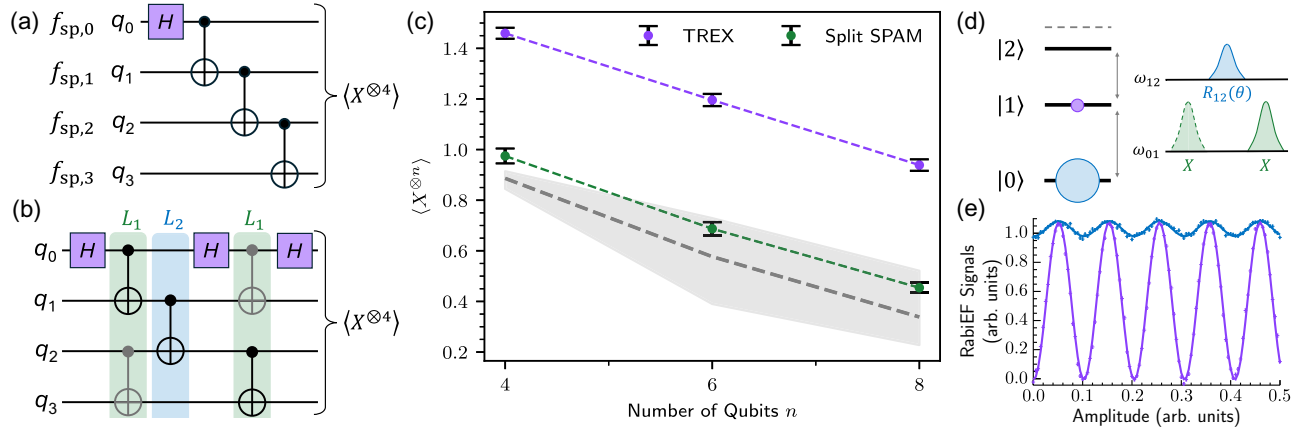


FIG. 1. Splitting state-preparation and final measurement (SPAM) errors with noncomputational states. (a) Greenberger-Horne-Zeilinger (GHZ) state-preparation circuit. (b) Circuit inspired from (a) but with two unique CNOT layers, L_1 and L_2 , to facilitate gate noise learning. (c) SPAM error mitigation on the circuit in (b) with twirled readout error extinction (TREX) (purple data) and split mitigation of SPAM errors (green data). Error bars, discussed in Appendix B, represent $\pm 2\sigma$. The shaded area represents possible values of $\langle X^{\otimes n} \rangle$ based on the CNOT noise model, see Appendix C. The dashed gray line represents a symmetric splitting of the conjugate pairs of fidelities in the learned CNOT noise. (d) First three levels of a superconducting transmon and the RabiEF pulse sequence. Circles indicate a thermal state, with populations corresponding to their sizes. (e) RabiEF experiment on IBM Quantum hardware. Markers are hardware measurements and curves are fits of $a \sin^2(b\theta) + c$ to the data. Here, θ is proportional to the pulse amplitude on the x -axis. Blue and purple are for the no- π (π) and π (π) RabiEF signals, respectively.

raw, unmitigated, expectation value $\langle O \rangle_{\text{raw}}$. All two-qubit gates are Pauli twirled so that the effective CNOT noise is a Pauli noise channel. We repeat this for $16n^2$ twirling configurations and collect 128 shots per configuration.

Next, we correct SPAM errors in $\langle O \rangle_{\text{raw}}$ with TREX by normalizing it by a correction factor $\langle Z^* \rangle$. This factor is found by measuring $Z^{\otimes n}$ on the all-zero initial state, with measurements twirled with X gates and an assumed ideal prepared state. Under ideal state preparation, the TREX correction factor is $\langle Z^* \rangle = \prod_{i=0}^{n-1} f_{m,i}$, where $f_{m,i}$ denotes the final measurement fidelity and subscript i denotes the corresponding qubit. However, in the presence of state-preparation errors, the measured TREX mitigator becomes $\langle Z^* \rangle = \prod_{i=0}^{n-1} f_{\text{sp},i} f_{m,i}$, where $f_{\text{sp},i}$ are the state-preparation fidelities. Normalizing $\langle O \rangle_{\text{raw}}$ by $\langle Z^* \rangle$ can thus cause TREX to overcompensate when the set of qubits whose state-preparation errors affect the observable (in our case, only qubit 0) differs from the support of the observable (in our case, all qubits), see, for example, Carrera Vazquez *et al.* [13] and Chen *et al.* [33].

Here, we measure $\langle Z^* \rangle$ after a passive reset into the thermal state with 5000 shots for each of the 16 randomizations. We observe that TREX does indeed overcorrect the expectation value $\langle X^{\otimes n} \rangle$, which we attribute to the factor $\prod_{i \neq 0} f_{\text{sp},i} \neq 1$; see the purple curve in Fig. 1(c). For our circuits with four and six qubits, this even results in unphysical expectation values. This calls for a SPAM error mitigation that accounts for state-preparation errors independently of measurement errors.

III. NOISE LEARNING WITH EXCITED STATES

Jointly mitigating SPAM errors can lead to nonphysical expectation values. We now demonstrate how to leverage noncomputational states to separate state-preparation from midcircuit measurement errors. The resulting learned noise model is applicable to advanced error mitigation methods, such as probabilistic error cancellation (PEC) [16,27]. Our work is based around the idea that well-controlled initial states, such as thermal states, can serve as a resource to this end.

A. Thermal states and passive reset

Learning the state-preparation error precisely necessitates stable state preparation in the computational basis. Therefore, we employ a simple passive reset by waiting more than $5 - 10 \times T_1$ when initializing qubits. This results in a thermal state with a population following a Boltzmann distribution [39]. Therefore, the probability of finding the transmon in the excited state is $p_{\text{sp}} = \exp(-\hbar\omega_{01}/k_B T_{\text{eff}})$. Here, ω_{01} is the energy difference between the first excited state and the ground state, see Fig. 1(d). The effective temperatures of superconducting qubits T_{eff} are typically above the 15 mK temperature of dilution refrigerators. Furthermore, the population of

the second excited state $|2\rangle$ is often negligible [40], see also Appendix A 1. After a passive reset, the qubit is thus in the thermal state $\rho_{\text{th}} = (1 - p_{\text{sp}}) |0\rangle\langle 0| + p_{\text{sp}} |1\rangle\langle 1|$, with typical p_{sp} values in the 1%–10% range. Crucially, p_{sp} can be precisely measured by driving an oscillation in the $|1\rangle \leftrightarrow |2\rangle$ subspace with the $R_{12}(\theta)$ gate and comparing it with a reference oscillation [40,41], see Fig. 1(e). The probability p_{sp} is estimated from the amplitudes of the no-pi (π) signal and the reference, or pi (π), signal, see Appendix A. This procedure, also known as a “RabiEF” experiment [42], provides a direct measurement of the state-preparation error p_{sp} of a passive reset. Through comparison with a reference oscillation, which is equally affected by measurement errors, the RabiEF experiment is mostly insensitive to measurement noise, see Appendix A 2. We can thus place ourselves in a situation where the state-preparation error is well known and measurable with sufficient accuracy by passively resetting the qubit.

B. Split mitigation of state and measurement errors

We can leverage our knowledge of state-preparation errors to prevent the TREX overcorrections shown in Sec. II. First, we measure the fidelities $f_{\text{sp},i} = 1 - 2p_{\text{sp},i}$ of each qubit with RabiEF following a 10 ms passive reset. Next, we multiply the correction factor $\langle Z^* \rangle$ by the product $\prod_{i=1}^{n-1} f_{\text{sp},i}$ to obtain the mitigated expectation value:

$$\langle O \rangle_{\text{mit}} = \frac{\langle O \rangle_{\text{raw}} \prod_{i=1}^{n-1} f_{\text{sp},i}}{\langle Z^* \rangle}. \quad (1)$$

Crucially, the circuits to compute $\langle Z^* \rangle$ must use the same passive reset as RabiEF, so that the state-preparation error is consistent across all experiments. This results in physical expectation values, see the green curve in Fig. 1(c). Furthermore, these values are consistent with the noise levels of the CNOT gates. We learn their noise model and compute their effect on the expectation value of $X^{\otimes n}$ via a Clifford simulation; details are in Appendix C. Since we can only learn the CNOT noise up to the unlearnable degrees of freedom, we compute upper and lower bounds for the ideal observable with only CNOT noise, represented as the shaded area in Fig. 1(c), assuming physical noise channels. This area represents the possible range for $\langle O \rangle_{\text{raw}}$ after removing SPAM errors; therefore, $\langle O \rangle_{\text{mit}}$ falling within this area indicates success. The two experiments with six and eight qubits fall well within this area. The experiment with four qubits does not. Out-of-model errors, such as non-Markovian noise and noise generators with a weight greater than two, may cause this [43].

C. Noise model of the measurement

We now develop a framework to completely specify the noise model of (midcircuit) measurements, assuming

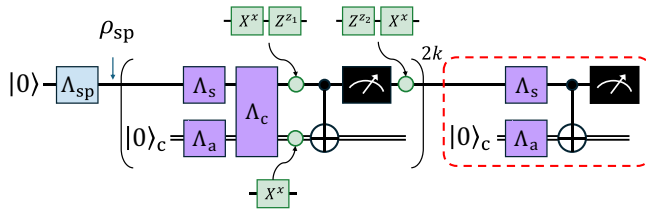


FIG. 2. Measurement noise model. The qubit and the classical bit are indicated by the single and double lines, respectively. Cycle benchmarking learns the noise model, shown in purple, by repeating the measurement $2k$ times. The twirling gates are indicated in green; the binary variables x, z_1, z_2 are drawn independently from a uniform distribution on $\{0, 1\}$ [44]. The dashed red rectangle indicates the final measurement, which is insensitive to correlated errors Λ_c .

projective Z basis measurements and ideal single-qubit and single-qutrit gates. As a result of appropriate twirling, state-preparation and measurement errors are bitflips, i.e., Pauli strings containing only I s or X s. We enforce this by applying random Z gates to the initial state with probability 50% and twirling measurements following Beale and Wallman [44]. This diagonalizes measurements and the initial state density matrix. It also permutes measurement outcomes while ensuring that the postmeasurement state is unchanged, see Fig. 3 of Beale and Wallman [44]. Note that we employ the convention that noise occurs before measurements without loss of generality, see Fig. 2.

After twirling, the Pauli transfer matrices (PTMs), introduced in Appendix D, of the measurement noise channels will only contain nonzero diagonal elements in the entries corresponding to Pauli strings containing only I s and Z s. Therefore, whenever we now write the PTM of such a noise channel, we omit rows and columns containing X s or Y s. Furthermore, we model the classical bit (cbit) as if it were a qubit existing in a corresponding Hilbert space, with the convention

$$\text{qubit} \otimes \text{cbit}. \quad (2)$$

For example, a bitflip error on the qubit is denoted XI and a bitflip on the classical bit is IX . The entries of any PTM will always appear in lexicographic order, e.g., II, IZ, ZI, ZZ . To construct the PTM, we represent a measurement as a CNOT gate targeting the classical bit controlled by the qubit, and a subsequent measurement on the qubit; see Fig. 2. The model of the noisy measurement is built using the principle of deferred measurement [45].

Our measurement noise model includes three possible error channels.

(1) A *state error*, modeled as a bitflip on the qubit only, with probability p_s and noise channel $\Lambda_s(\rho) = (1 -$

$p_s)\rho + p_s XI \rho XI$. The PTM of Λ_s is

$$\Gamma_s = \begin{pmatrix} 1 & 0 \\ 0 & 1 - 2p_s \end{pmatrix} \otimes I. \quad (3)$$

This may correspond to a qubit decay at the end of the readout through a T_1 event. With twirling, this becomes an X error instead of an amplitude damping channel.

(2) An *assignment error*, modeled as a bitflip on the classical bit only, with probability p_a and noise channel $\Lambda_a(\rho) = (1 - p_a)\rho + p_a IX \rho IX$. The PTM of Λ_a is

$$\Gamma_a = I \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 - 2p_a \end{pmatrix}. \quad (4)$$

This may correspond to the readout misclassifying the state without a quantum error occurring on the qubit.

(3) Finally, a *correlated error* is a bitflip on both the classical and quantum bits with probability p_c and noise channel $\Lambda_c(\rho) = (1 - p_c)\rho + p_c XX \rho XX$. The PTM of Λ_c is

$$\Gamma_c = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - 2p_c & 0 & 0 \\ 0 & 0 & 1 - 2p_c & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (5)$$

For example, this corresponds to a postmeasurement X error on the qubit, which is equivalent to the correlated error Λ_c by propagating the X back through the measurement.

The PTM Γ_m of the noise model is obtained by multiplying the PTMs of the three error channels and the PTM of the ideal CNOT between the classical and quantum bit such that

$$\Gamma_m = \Gamma_{cx} \Gamma_s \Gamma_a \Gamma_c = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & f_a f_s \\ 0 & 0 & f_c f_s & 0 \\ 0 & f_c f_a & 0 & 0 \end{pmatrix}. \quad (6)$$

For readability, we write the error fidelities $f_x := (1 - 2p_x)$ instead of their probabilities p_x . To learn the elements of Γ_m , we repeat the measurement $2k$ times in a cycle benchmarking experiment [24,25], resulting in

$$\Gamma_m^{2k} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & f_a^{2k} (f_c f_s)^k & 0 & 0 \\ 0 & 0 & (f_c f_s)^{2k} & 0 \\ 0 & 0 & 0 & f_a^{2k} (f_c f_s)^k \end{pmatrix}. \quad (7)$$

In practice, state-preparation errors may cause the qubit to start the circuit in the $|1\rangle$ state with probability p_{sp} , such

that the initial state is $\rho_{\text{sp}} \neq |0\rangle\langle 0|$. The PTM that prepares ρ_{sp} from $|0\rangle$ is

$$\Gamma_{\text{sp}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 - 2p_{\text{sp}} \end{pmatrix} \otimes I. \quad (8)$$

At the end of the execution, we measure the qubit and record its outcome on the same wire, which now functions as a classical bit, see Fig. 2. The PTM of the final measurement, derived in Appendix E, is

$$\Gamma_{\text{fm}} = \begin{pmatrix} 1 & 0 \\ 0 & f_a f_s \end{pmatrix} \otimes I. \quad (9)$$

Therefore, the PTM of the full measurement cycle benchmarking (MCB) circuit, shown in Fig. 2, is $\Gamma_{2k} = \Gamma_{\text{fm}} \Gamma_{\text{m}}^{2k} \Gamma_{\text{sp}}$, which equates to

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & f_a^{2k} (f_c f_s)^k & 0 & 0 \\ 0 & 0 & (f_c f_s)^{2k} f_a f_s f_{\text{sp}} & 0 \\ 0 & 0 & 0 & f_a^{2k} (f_c f_s)^k f_a f_s f_{\text{sp}} \end{pmatrix}. \quad (10)$$

We learn the entries of the PTM by measuring the corresponding noisy expectation values $\langle IZ \rangle$, $\langle ZI \rangle$, $\langle ZZ \rangle$ on the MCB circuit for different values of k . We then fit decaying exponential curves, of the form $A f^{2k}$, to the data. Thus, we can learn the offset $A = f_a f_s f_{\text{sp}}$ and the products $f = f_a \sqrt{f_c f_s}$ and $f = f_c f_s$. Consequently, we can learn (1) the product $f_s f_c$ of the state and correlated fidelities, (2) the probability p_a of a bitflip on the classical bit, and (3) the product $f_s f_{\text{sp}}$ of the state and state-preparation fidelities. This shows that we can fully specify the noise model if we can measure either of f_s, f_c , or f_{sp} individually. We therefore propose to estimate f_{sp} with a RabiEF experiment, thereby splitting the fidelity products and, therefore, also the SPAM error fidelities.

Here, we focus on correlated errors between qubits and their classical bits. Our model can be extended to capture correlations between qubits by increasing the support of the cycle benchmarking experiment, resulting in more Pauli-Lindblad generators [16,27,44]. As an illustrative example, Appendix G shows how our protocol learns the noise model for two correlated qubits undergoing measurement.

D. Mitigation workflow

In Sec. III B, we split the error mitigation of SPAM errors. This requires a slow passive reset when learning the TREX correction factor $\langle Z^* \rangle$ and when running the circuit in Fig. 1(b) to keep f_{sp} consistent across all executed circuits. Furthermore, this correction is possible because the circuit only has Clifford gates. We can thus compute

which qubits affect the raw observable resulting in Eq. (1). By contrast, an ideal error mitigation protocol (1) is fast and (2) mitigates both state-preparation and midcircuit measurement errors for arbitrary circuits.

We could avoid slow resets by executing a RabiEF experiment on the initial state resulting from a fast active reset, consisting of a measurement and a feedforward X gate conditioned on measuring a $|1\rangle$. For a RabiEF experiment to function properly, the reset must always prepare the same initial state of the form $(1 - p_{\text{sp}}) |0\rangle\langle 0| + p_{\text{sp}} |1\rangle\langle 1|$, i.e., with a negligible $|2\rangle$ -state population. If such a reliable fast reset is available, we propose a noise learning and mitigation protocol that (1) executes RabiEF with fast resets to learn f_{sp} ; (2) runs MCB to learn f_a, f_s , and f_c ; (3) runs gate noise learning; and (4) mitigates state-preparation, gate, and measurement errors in the intended circuit, with such schemes as probabilistic error cancellation (PEC) [16] or probabilistic error amplification (PEA) [1].

More generally, if such a fast reset—with *good* $|2\rangle$ -state reset—is not available, we can still use MCB from Sec. III C and RabiEF with both slow and fast resets to learn all the parameters of our noise model. First, we leverage thermal states to learn the slow reset fidelity $f_{\text{sp}}^{\text{slow}}$ with a RabiEF experiment. Second, we learn the measurement noise fidelities with a MCB experiment and slow resets. Next, we perform an additional MCB experiment with fast resets to obtain $f_{\text{sp}}^{\text{fast}}$, where we assume that the measurement noise is constant throughout all experiments and independent of the state-preparation method. Finally, we learn the remaining gate noise, with other noise learning protocols, and execute the intended quantum circuits with fast resets and an error mitigation scheme, such as PEC. Figure 3 summarizes the proposed noise learning protocol.

IV. SIMULATIONS

We now show, numerically, how to leverage a RabiEF experiment with a fast active qutrit reset to error mitigate measurements while accounting for errors in the reset. As RabiEF explicitly populates the $|2\rangle$ state, it is necessary to remove as much of the $|2\rangle$ -state population as possible in the state-preparation step. We therefore use qutrit resets. With simulations of MCB, we show correctly learned noise fidelities and how the full MCM noise learning protocol detailed in Sec. III C and Fig. 3 bypasses insufficient $|2\rangle$ -state resets. We also simulate the experiment described in Sec. II, without CNOT noise, and a three-qubit noisy quantum teleportation circuit with PEC mitigation.

A. RabiEF with fast resets

An active qubit reset is a qubit measurement followed by a feedforward X gate conditioned on measuring the $|1\rangle$ state that can be used for qubit initialization. Here, the qubit measurement misclassifies the $|2\rangle$ state as $|1\rangle$, owing

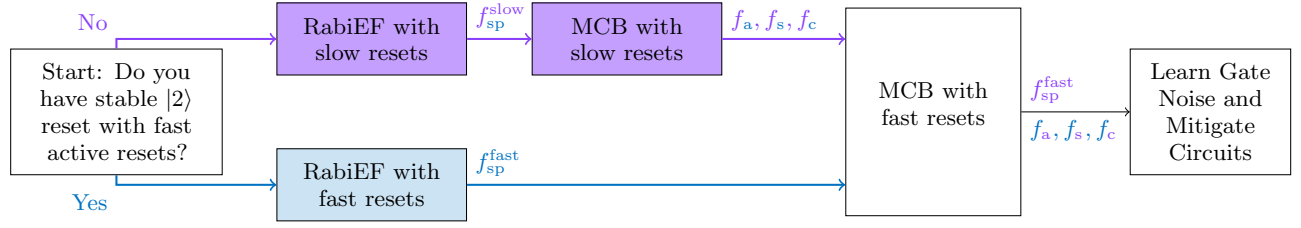


FIG. 3. Measurement noise learning protocol. Active resets that can and cannot remove the $|2\rangle$ -state population can both be supported. The learned parameters for a given step are shown on the outgoing arrows, e.g., f_a, f_s , and f_c are learned by the measurement cycle benchmarking (MCB) with slow resets. The preferred case is the lower blue path, as it has fewer experiments, but it requires active resets that can remove the $|2\rangle$ -state population. In the event that they cannot, the purple path should be taken, as this circumvents poor $|2\rangle$ -state reset and its impact on the RabiEF experiment. Regardless of the path taken, the final MCB experiment completes the set of learned state-preparation and final measurement (SPAM) fidelities f_a, f_s, f_c , and f_{sp}^{fast} .

to the overlap of their complex readout signals [26,46]. Such active resets are ineffective at removing the $|2\rangle$ -state population as they only implement binary classification and the feedforward X gate does not interact with $|2\rangle$. In the method described in Sec. II, this is overcome with a long passive reset before executing each circuit, which we call a “slow reset,” see Fig. 4(a). A “fast reset” forgoes this long delay, relying only on active resets in the repetition

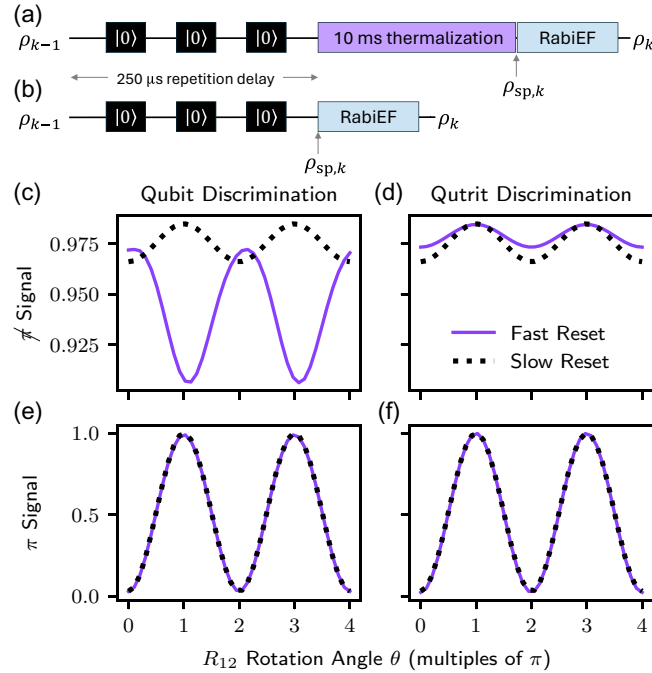


FIG. 4. RabiEF with different resets. (a) Slow reset: a 10 ms delay after the repetition delay thermalizes the transmon before the RabiEF circuits. (b) Fast reset followed by RabiEF circuits. (c–f) Classically simulated signals with (c, e) qubit and (d, f) qutrit discrimination in the active resets in the repetition delay. Fast and slow resets are shown as solid purple and dotted black lines, respectively.

delay, see Fig. 4(b). Qubit measurements with fast resets cause distortions in the RabiEF signals, owing to their poor $|2\rangle$ -state reset, which contaminates the estimates of p_{sp} in Eq. (8). Therefore, we consider an imperfect “qutrit reset,” which discriminates between $|0\rangle$, $|1\rangle$, and $|2\rangle$ [46,47] and applies the feedforward I , X , and XX_{12} gates, depending on the respective outcomes.

We simulate RabiEF experiments on qutrits with ideal qubit gates embedded in $SU(3)$, e.g., $X = |1\rangle\langle 0| + |0\rangle\langle 1| + |2\rangle\langle 2|$. This is justified in practice as experimental single-qubit and -qutrit gates typically have much higher fidelities than two-qubit gates and measurements [48,49]. Thermal relaxation and measurement errors are included as separate operations, attached to appropriate circuit instructions. To investigate how state preparation removes the $|2\rangle$ -state population we perform density matrix simulations of RabiEF with $N = 1000$ shots and 40 rotation angles $\theta \in [0, 4\pi]$ for $R_{12}(\theta)$. This is done for fast qubit reset, slow qubit reset, fast qutrit reset, and slow qutrit reset. The projected postmeasurement state ρ_{k-1} of the previous RabiEF circuit and shot is the input to the repetition delay of the next circuit, as was done in Ref. [50], see Figs. 4(a) and 4(b). This allows us to observe how postmeasurement states and imperfect qutrit reset contaminate subsequent shots. The measurement outcomes are saved, alongside all input states $\rho_{sp,k}$, see Fig. 4. See Appendix F 1 for more details on the simulation setup.

Our simulations include three active resets interspaced in a 250 μs repetition delay, since multiple reset instructions increase the fidelity of initializing the transmon [51]. The long delay for passive resets is set to 10 ms. We simulate thermal relaxation of the qutrit during all idle times and measurements. Measurement errors are modeled as thermal relaxation, measurement-induced leakage, and readout signal misclassifications. Misclassifications are simulated with a readout assignment error matrix R , storing probabilities q_{jk} of misclassifying $|k\rangle$ as $|j\rangle$. For a state with ideal measurement probabilities (α, β, γ) , for $(|0\rangle, |1\rangle, |2\rangle)$, the noisy final measurement probabilities

TABLE I. State-preparation errors from simulated RabiEF. The estimated probabilities $\hat{p}_{\text{sp}}$ for fast and slow resets are given, with the corresponding *true* probabilities p_{sp} . The standard errors for $\hat{p}_{\text{sp}}$ are those reported by the curve fitting software. True values are averages over the ensemble of initial states $\rho_{\text{sp},k}$ from all shots k in the qutrit simulations, e.g., $p_{\text{sp}} = \text{avg}_k \text{Tr} [|1\rangle\langle 1| \rho_{\text{sp},k}]$. True values $p_{\text{sp}}^{(2)}$ for the $|2\rangle$ state populations are obtained with $p_{\text{sp}}^{(2)} = \text{avg}_k \text{Tr} [|2\rangle\langle 2| \rho_{\text{sp},k}]$. Problematic cases are given in bold. This includes fits that give unphysical results, probabilities p_{sp} and $\hat{p}_{\text{sp}}$ that fluctuate widely, and non-negligible $p_{\text{sp}}^{(2)}$ values.

		Fast reset		Slow reset	
		Qubit reset	Qutrit reset	Qubit reset	Qutrit reset
$\hat{p}_{\text{sp}}$	Fit	-7.375×10^{-2}	1.138×10^{-2}	1.899×10^{-2}	1.899×10^{-2}
	Standard error	2.366×10^{-3}	2.066×10^{-6}	4.354×10^{-7}	3.982×10^{-10}
p_{sp}	Mean	2.843×10^{-2}	1.206×10^{-2}	1.946×10^{-2}	1.946×10^{-2}
	Standard deviation	1.840×10^{-2}	1.387×10^{-5}	3.914×10^{-6}	3.228×10^{-9}
$p_{\text{sp}}^{(2)}$	Mean	2.429×10^{-2}	6.518×10^{-4}	4.912×10^{-4}	4.910×10^{-4}
	Standard deviation	2.662×10^{-2}	2.393×10^{-5}	1.924×10^{-7}	1.587×10^{-10}

are $R(\alpha, \beta, \gamma)^T$, where

$$R = \begin{pmatrix} q_{00} & q_{01} & q_{02} \\ q_{10} & q_{11} & q_{12} \\ q_{20} & q_{21} & q_{22} \end{pmatrix} = \begin{pmatrix} 0.991 & 0.009 & 0 \\ 0.009 & 0.931 & 0.06 \\ 0 & 0.06 & 0.94 \end{pmatrix}. \quad (11)$$

The normalization constraint $q_{0j} + q_{1j} + q_{2j} = 1$ applies to all $j = 0, 1, 2$. For example, the probability of measuring a $|1\rangle$ for a given state is $q_{10}\alpha + q_{11}\beta + q_{12}\gamma$. The values in R are chosen to match hardware experiments on superconducting qubits [46,52,53]. As a result, our R satisfies $q_{02} = q_{20} = 0$, $q_{jk} = q_{kj}$, and $q_{01} < q_{12}$. Furthermore, the simulated measurement errors correspond to measurement noise channel fidelities $f_a = 0.99096$ and $f_s = 0.99096$; the associated derivation is in Appendix F 2. The readout assignment error matrix for qubit discrimination and resets is obtained from R with transformations $q_{10} \rightarrow q_{10} + q_{20}$, $q_{11} \rightarrow q_{11} + q_{21}$. This is equivalent to qutrit discrimination followed by a $|2\rangle$ - $|1\rangle$ misclassification with 100% probability.

These simulation results are plotted in Figs. 4(c)–4(f), with the true and estimated state-preparation probabilities given in Table I. Both RabiEF signals *should* be of the form $a \sin^2(b\theta) + c$ with non-negative amplitudes a being physical, and all estimates $\hat{p}_{\text{sp}}$ should be valid probabilities, i.e., $0 \leq \hat{p}_{\text{sp}} \leq 1$. However, the non-negligible $|2\rangle$ -state population with *fast qubit* reset results in a negative amplitude of the no-pi signal (π). This gives an incorrect negative probability estimate, see Fig. 4(c) and $\hat{p}_{\text{sp}}$ in Table I. Indeed, initial states of the form

$$\rho_{\text{sp}} = (1 - p_{\text{sp}} - p_{\text{sp}}^{(2)}) |0\rangle\langle 0| + p_{\text{sp}} |1\rangle\langle 1| + p_{\text{sp}}^{(2)} |2\rangle\langle 2| \quad (12)$$

bias our estimate $\hat{p}_{\text{sp}}$ as

$$\hat{p}_{\text{sp}} = \frac{a_{\pi}}{a_{\pi} + a_{\pi}} = \frac{p_{\text{sp}} - p_{\text{sp}}^{(2)}}{1 - 3p_{\text{sp}}^{(2)}}. \quad (13)$$

If $p_{\text{sp}}^{(2)}$ is sufficiently large, the no-pi signal has a negative amplitude and our estimate becomes negative and unphysical [54]. In contrast, our simulations of fast *qutrit* resets and slow resets are physical. Furthermore, if we look at the standard deviation of the true value p_{sp} , shown in Table I, we see that the prepared state is more stable with fast qutrit resets and slow resets than with fast resets based on qubit discrimination. Crucially, the relative accuracy of the RabiEF with fast qutrit reset is -5.6% on a true p_{sp} of 0.01206% .

These simulations demonstrate how unstable state preparation, where $|2\rangle$ is not reset correctly, is insufficient for RabiEF. Without an accurate estimate of p_{sp} , separating SPAM errors in our model to mitigate dynamic circuits is not possible. Our full MCM noise learning protocol bypasses *bad* fast resets to obtain the estimate $\hat{p}_{\text{sp}}$ using slow resets. Though slow resets can be used for mitigation, they result in expensive circuit execution, owing to their slow rates. Fast *qutrit* resets—even if imperfect—allow for the state-preparation characterization required to split SPAM errors quickly.

B. Cycle benchmarking for readout

We now show that including an estimate of the state-preparation error in a MCB experiment results in accurate estimates of the measurement noise fidelities f_a , f_s , and f_c . We run two simulations with slow and fast resets to illustrate both noise learning paths in the full MCM noise learning protocol shown in Fig. 3.

The true measurement noise fidelities are set to $f_a = 0.99096$, $f_s = 0.99096$, and $f_c = 0.9950$ to match the values given in Sec. IV A; see Appendix F 2. The first simulation learns f_a , f_s , and f_c with slow state-preparation error probability $p_{\text{sp}}^{\text{slow}} = 0.01946$. Then the fast state-preparation error probability $p_{\text{sp}}^{\text{fast}} = 0.01206$ is learned with a second MCB experiment, using the now learned measurement fidelities f_a , f_s , and f_c . The first simulation splits the products of fidelities in Eq. (10) with the slow

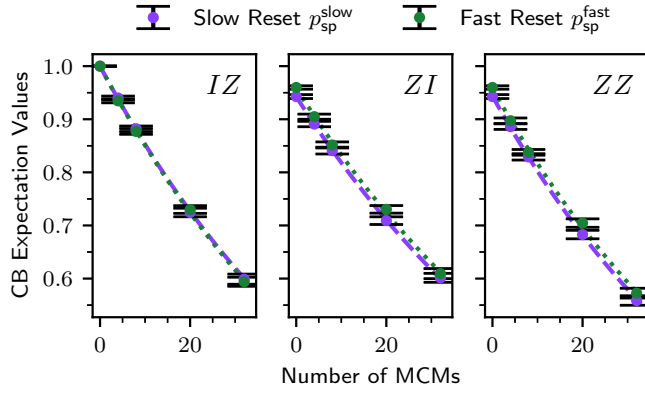


FIG. 5. Measurement cycle benchmarking on a qubit. Fits to these data determine the noise channel fidelities. Error bars show $\pm 2\sigma$ standard errors for each expectation value. Dotted and dashed lines are obtained from nonlinear least-square fits of the equation Af^{2k} to the data, where k is a function of the depth. CB, cycle benchmarking; MCM, midcircuit measurement.

reset estimate from RabiEF given in Table I, i.e., $\hat{p}_{\text{sp}}^{\text{slow}} = 0.01899$.

We simulated MCB with 256 twirled-circuit randomizations for each depth, with $k = 0, 1, 2, 5, 8$, and 128 shots per randomization. The simulations were run with QISKIT AER [55], the details of which are given in Appendix F 3. The resulting six decay curves are shown in Fig. 5: three per reset type, corresponding to the observables IZ , ZI , and ZZ , identifying the PTM entries in Eq. (10). We estimated the measurement noise fidelities by fitting decaying exponentials to the slow-reset data. The same fitting procedure was carried out for the fast reset simulation data, except that the fast reset fidelity $f_{\text{sp}}^{\text{fast}}$ was estimated using the learned measurement noise fidelities from the first simulation.

Not only does our learning procedure effectively learn the noise fidelities to less than 0.01% relative error but we can also estimate the fast reset fidelity $f_{\text{sp}}^{\text{fast}}$ with the second MCB experiment; see Table II. The state-preparation method can thus be different between full MCM noise learning and error mitigation, as long as an

TABLE II. Comparison of noise fidelities from simulated MCB. Estimated fidelities $\hat{f}$ are given, with their corresponding true fidelities f and the relative error $|f - \hat{f}|/f$.

	True f	Estimated $\hat{f}$	Relative error
MCB with slow resets			
f_a	0.990 96	0.991 01	5.637×10^{-5}
f_s	0.990 96	0.990 22	7.460×10^{-4}
f_c	0.995 00	0.995 70	7.055×10^{-4}
Second MCB experiment with fast resets			
$f_{\text{sp}}^{\text{fast}}$	0.975 88	0.976 76	9.061×10^{-4}

additional MCB experiment is run to estimate the fast state-preparation error $p_{\text{sp}}^{\text{fast}}$. This demonstrates that both paths in our full MCM noise learning protocol result in an estimate of the full noise model in Eq. (10) with split SPAM errors. An extension of our scheme to high-rate and high-fidelity state preparation is discussed in Appendix H.

C. Example mitigation of final readout errors

We now show classical simulations of our method for the stabilizer circuit illustrated in Fig. 1(a). The simulations use the same slow reset fidelity as described in Secs. IV A and IV B, and implement both TREX and the mitigation from Eq. (1). These simulations, run using QISKIT AER, do not simulate leakage, as the $|2\rangle$ -state is assumed to be negligible for all experiments except RabiEF.

The TREX mitigator $\langle Z^* \rangle$ is learned with 16 twirling randomizations and 5000 shots per randomization, for $n = 4, 6, 8, 10$ qubits, i.e., the same as in Sec. II. The measurement and state-preparation noise fidelities were taken from Sec. IV B. The stabilizer circuit was simulated for $n = 4, 6, 8, 10$ qubits, with 128 shots per twirling randomization and $2^{10} \times n^2$ total shots. With the hardware experiments, we compare our mitigated results with Clifford simulations with CNOT noise. By contrast, here we assume ideal gates to compare the mitigated expectation values with their ideal value of +1, the dashed green line in Fig. 6. The expected result with TREX is f_{sp}^{1-n} , shown by the dashed purple line in Fig. 6.

As seen in the hardware experiments, TREX overshoots the ideal expectation values by a factor f_{sp}^{1-n} ; see the purple dots in Fig. 6. This is expected since $\langle Z^* \rangle$ picks up the

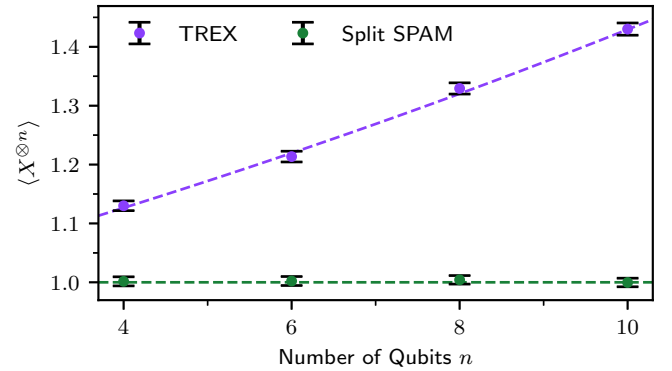


FIG. 6. Mitigation of $\langle X^{\otimes n} \rangle$ expectation values. Values are mitigated with twirled readout error extinction (TREX) and the modified mitigator from the hardware experiments, which splits state-preparation and final measurement (SPAM) errors (“split SPAM”). Error bars show $\pm 2\sigma$ standard errors for each mitigated expectation value. Dashed lines show expected expectation values, given knowledge of the simulated Greenberger-Horne-Zeilinger state-preparation circuit. The green dashed line shows the noiseless expectation value of +1.

same factor from imperfect state-preparation, i.e., the purple dashed line. If, however, the values are mitigated with Eq. (1), the expectation values do not overshoot, instead lying around the ideal value of $\langle O \rangle = +1$; see the green dots in Fig. 6. This shows a correct mitigation of SPAM errors, further validating the error mitigation described in Sec. III.

D. Example mitigation of dynamic circuits

Splitting SPAM errors is most beneficial with mid-circuit measurements, where measurement noise can be mitigated separately from state-preparation and gate noise. We demonstrate that our MCM noise learning protocol helps mitigate errors in dynamic circuits. In particular, we simulate a noisy three-qubit teleportation circuit and mitigate SPAM errors with PEC [16,27] built on our noise learning protocol. The circuit is constructed in the CNOT picture, i.e., including the classical bits, as in Fig. 2, and simulated with QISKIT AER [55]. The teleportation circuit, shown in Fig. 7, includes explicit noise channels for state-preparation Λ_{sp} , measurement assignment Λ_a , state Λ_s , and correlated Λ_c errors, as defined in Sec. III. The underlying fidelities f_x are taken from Sec. IV B, with a state-preparation fidelity $f_{\text{sp}} = 0.961\,08$ equivalent to the slow resets.

The ideal teleported state

$$\rho_{\text{in}}^{\text{ideal}} = |\psi\rangle\langle\psi| = R_X(\theta) |0\rangle\langle 0| R_X^\dagger(\theta) \quad (14)$$

is prepared by an R_X gate with rotation angle $\theta \in [0, 2\pi]$. However, SPAM noise introduce errors into the final teleported state on qubit q_2 .

To apply PEC, we generate 128 PEC circuit realizations, replacing the inverse channels $\tilde{\Lambda}^{-1}$ with Paulis from their quasiprobability distribution representations. Each circuit is then simulated, resulting in 128 noisy teleported states

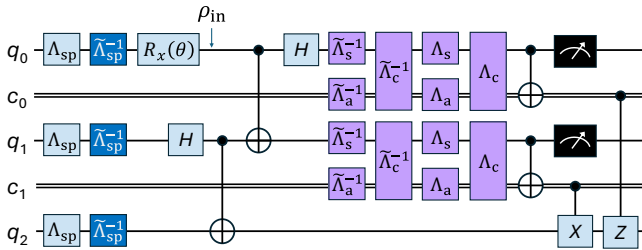


FIG. 7. Teleportation circuit in the CNOT picture. The circuit includes state-preparation noise Λ_{sp} and measurement noises Λ_s , Λ_a , and Λ_c . Inverse channels $\tilde{\Lambda}^{-1}$ are implemented by sampling their quasiprobability distributions for probabilistic error cancellation (PEC). The output state ρ_{out} is the state of the last qubit q_2 , i.e., with the other qubits and classical bits traced out. The state to be teleported is ρ_{in} , which is affected by the state-preparation noise Λ_{sp} of qubit 0 and its inverse channel $\tilde{\Lambda}_{\text{sp}}^{-1}$.

$\rho_{\text{out},j}$ on qubit q_2 , with $j \in [0, 127]$. As PEC mitigates expectation values and not states, we obtain mitigated states via state tomography. From each noisy teleported state $\rho_{\text{out},j}$, we obtain three ideal Pauli expectation values,

$$\langle P \rangle_j = \text{Tr}[P \rho_{\text{out},j}], \quad (15)$$

for Paulis $P = X, Y, Z$. We emulate shot noise by sampling 100 times per circuit realization, from the corresponding Bernoulli distributions, resulting in expectation values $E_{P,j}$ affected by both SPAM and shot noise. The PEC mitigated expectation values are then

$$\langle P \rangle_{\text{mit}} = \gamma \sum_j (-1)^{m_j} E_{P,j}, \quad (16)$$

where γ is the noise factor for the inverse noise channels and m_j is an integer from PEC, controlling how the quasiprobability inverse channels are engineered [16,27]. Physical states are obtained from state tomography by minimizing a mean-squared error, as

$$\rho_{\text{mit}} = \arg \min_{\rho_t} \sum_{P=X,Y,Z} (\text{Tr}[P \rho_t] - \langle P \rangle_{\text{mit}})^2. \quad (17)$$

Here, the state ρ_t is parameterized as

$$\rho_t = \rho(t_1, t_2, t_3, t_4) = \frac{T^T T}{\text{Tr}[T^T T]}, \quad (18)$$

with real parameters t_1, t_2, t_3 , and t_4 , and triangular matrix

$$T = \begin{pmatrix} t_1 & 0 \\ t_3 + it_4 & t_2 \end{pmatrix}. \quad (19)$$

This ensures a physical mitigated state ρ_{mit} for a given ensemble of PEC circuit realizations. We then compute the state fidelity

$$\mathcal{F}(\rho_{\text{in}}^{\text{ideal}}, \rho_{\text{mit}}) = \text{Tr} \left[\sqrt{\sqrt{\rho_{\text{in}}^{\text{ideal}}} \rho_{\text{mit}} \sqrt{\rho_{\text{in}}^{\text{ideal}}}} \right]^2 \quad (20)$$

as a metric of success. This is done for 300 realizations of the simulation and 15 rotation angles in the input state, see Fig. 8. The unmitigated case does not need PEC and is thus simulated as a single circuit, directly resulting in a physical state. Therefore, the state fidelity is computed on the output state ρ_{out} directly, instead of using state tomography.

The PEC mitigated state fidelities are significantly higher than the unmitigated values, showing that we are mitigating SPAM errors, see Fig. 8. The majority of the 300 simulation realizations result in state fidelities above 98%, whereas the best-case unmitigated fidelity is less than 96%. We also confirmed that PEC effectively mitigates

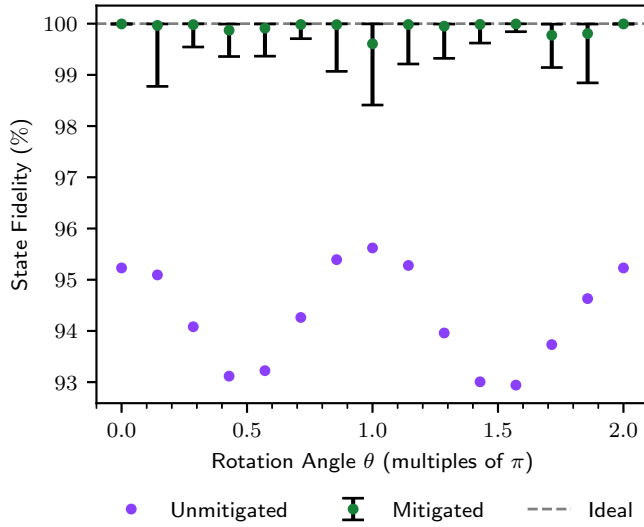


FIG. 8. State fidelities for unmitigated and probabilistic error correction (PEC) mitigated teleportation circuits. PEC is used to mitigate Pauli expectation values, which are then used with state tomography to estimate the state infidelity. The median mitigated state infidelities are shown in green, over 300 random realizations of the simulation. Error bars show the interquartile range over the 300 realizations. The unmitigated fidelities, in purple, are computed directly on the unmitigated state ρ_{out} without PEC samples or state tomography.

both state-preparation and measurement errors separately by rerunning the simulations with only one of the error sources, still observing improvements to the state fidelities with PEC (data not shown). Our MCM noise learning procedure thus effectively splits SPAM errors, facilitating PEC mitigation of dynamic circuits. Splitting SPAM errors was made possible by learning the state-preparation fidelity p_{sp} using noncomputational states with a RabiEF experiment.

V. DISCUSSION AND CONCLUSION

Our experimental data show how state-preparation errors prevent the accurate mitigation of SPAM errors for certain expectation values, corroborating recent results [33]. A passive reset of the qubits into a thermal state, whose population we measure with noncomputational states, allows us to resolve this by mitigating state-preparation errors independently of measurement errors. This prevents the unphysical behavior seen in some TREX-mitigated expectation values.

We introduce a noise model of (mid)circuit measurements, which fully accounts for state-preparation errors and errors in the final measurement. This extends recent noise learning protocols, such as that of Zhang *et al.* [25], in that the state-preparation error is learned independently of measurement noise. Our work also complements a recently introduced noise learning framework for

learning state-preparation, gate, and measurement noise in a self-consistent way up to unlearnable gauge degrees of freedom [32]. While this protocol also resolves the aforementioned issues of TREX [33], the unconstrained degrees of freedom render it unsuitable for dynamic circuits.

With the added resource of noncomputational states, we learn the ground truth of the state-preparation error. Crucially, our noise learning then allows us to mitigate errors that occur in dynamic circuits that explicitly require MCMs and feedforward gates. We successfully demonstrate this with simulations of a teleportation circuit, mitigating errors with PEC. The noise model resulting from our protocol is applicable to other noise-aware error mitigation methods that utilize twirled-noise channels, such as PEA [1].

As an outlook, we propose to combine RabiEF with other gate and measurement noise learning techniques and, more generally, the formalism from Chen *et al.* [32,33]. This will fully constrain the self-consistent noise models by anchoring the gauge, thus significantly broadening their scope.

We propose to learn the state-preparation error with a RabiEF experiment that requires either a slow passive reset or a fast active qutrit reset. This fits well with recent developments, since error correction requires a strong suppression of leakage that can be achieved through appropriate qutrit resets [56,57]. Importantly, our work requires the transmon to stay within its first three levels. Therefore, we do not capture effects where the transmon may escape from this subspace [58], or even confined states [59]. Here, we believe that operating transmons in such regimes should be avoided.

Our work leverages a noncomputational state of transmon qubits. However, we expect that the work should generalize to other hardware platforms that have addressable noncomputational states. These may be found in, for instance, solid state electron spin qutrits [60,61] and trapped ions [62]. An analytical treatment of the effect of all errors in the measurements and how they propagate to the mitigated observables would be a useful extension of the numerical analysis in Appendix A 2. In addition, future work may also explore generalizations of the qubit readout model to a full qutrit model, including measurement-induced leakage, and the subsequent effect on noise learnability [26,48]. In our work, we show noise models for up to two correlated qubits undergoing measurement. Generalizing our protocol and noise model to include correlations between more qubits and with idle qubits should allow for more accurate error mitigation on systems where strong correlations occur [16,27,43,44].

In summary, noncomputational states allow us to overcome no-go theorems that hold when the learning framework is restricted to the computational subspace. As a result, we can fully specify degrees of freedom of noise models that are fundamentally unlearnable when restricted

to qubit circuits. We illustrate this by separately mitigating state and preparation errors. Crucially, this is also possible without a slow passive reset of the quantum computer when a noisy qutrit reset is available.

Note added. Recently, we became aware of related work by S. Chen *et al.* [63], in which the authors also leverage noncomputational states of the transmon.

ACKNOWLEDGMENTS

We acknowledge M. Mergenthaler, A. Fuhrer, L.C.G. Govia, M. Takita, and A. Seif for useful discussions. This work was supported as a part of NCCR SPIN, a National Centre of Competence in Research, funded by the Swiss National Science Foundation (Grant No. 225153). D.J.E. acknowledges funding within the HPQC project by the Austrian Research Promotion Agency (FFG, Project No. 897481) supported by the European Union – NextGenerationEU. L.E.F. acknowledges funding from the European Union’s Horizon 2020 research and innovation program under Marie Skłodowska-Curie, Grant Agreement No. 955479 (MOQS—Molecular Quantum Simulations).

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors on reasonable request.

APPENDIX A: MEASURING THERMAL STATE POPULATION

In the RabiEF experiment, the population α is measured in the excited state of the qubit for a state $\rho = (1 - \alpha)|0\rangle\langle 0| + \alpha|1\rangle\langle 1|$ [41]. RabiEF drives oscillations in the $\{|1\rangle, |2\rangle\}$ subspace with a $R_{12}(\theta)$ gate that rotates around the X -axis of the corresponding Bloch sphere. Measuring α requires two independent experiments, called π and no- π , the latter indicated by π . In the π -experiment we apply the gate sequence $X-R_{12}(\theta)-X$, while in the no- π -experiment, the first X gate is omitted. The difference between both experiments is thus the π -pulse on the thermal state at the beginning of the gate sequence. Following this gate sequence, which we assume is ideal, we apply a noisy measurement with discrimination errors. This results in two oscillating signals, $s_\pi(\theta)$ and $s_{\pi'}(\theta)$, which we fit to two independent functions of the form $a \sin^2(\theta + c) + b$, see Fig. 1(e). Here, a , b , and c are the fitting parameters. Finally, the excited state population α is estimated by

$$\hat{\alpha} = \frac{a_{\pi'}}{a_{\pi'} + a_{\pi}}, \quad (\text{A1})$$

where $a_{\pi'}$ and a_{π} are the amplitudes of the function that fits $s_{\pi'}(\theta)$ and $s_{\pi}(\theta)$, respectively. In practice, we implement $R_{12}(\theta)$ with a Rabi pulse $R_{12}(A)$ of amplitude $A \propto \theta$.

The exact relationship between A and θ is irrelevant, as it only controls the frequency of the RabiEF signals, not their amplitudes.

Superconducting qubit measurements return complex values in the in-phase–quadrature (IQ) plane [38], to which we apply a principal component analysis to forgo classification and instead use the projected signals $s'_\pi(A) \propto s_\pi(\theta)$ and $s'_{\pi'}(A) \propto s_{\pi'}(\theta)$. These signals have arbitrary units and can be larger than 1, owing to the location of the reference points in the IQ plane, see Fig. 1(e). Crucially, gate and measurement errors tend to cancel out in Eq. (A1), as they affect a_{π} and $a_{\pi'}$ in the same way.

1. Thermal state populations

Here, we discuss the experimental thermal state data presented in Fig. 1, which are taken on qubits 114 to 121 on the IBM Quantum device `ibm_pinguino3`. The thermal populations of these qubits are measured using RabiEF after a 10 ms passive reset. The resulting $|1\rangle$ state population α ranges from 3.06% to 8.75% (see Table III) and creates a significant state-preparation error. From the measured population α in the $|1\rangle$ state, we compute an effective qubit temperature T_{eff} , following the relation $\alpha = \exp[-\hbar\omega_{01}/(k_b T_{\text{eff}})]$. The resulting effective temperatures, shown in Table III, are larger than the ≈ 15 mK of dilution refrigerators. We then use this relation to estimate the population in the second excited state by replacing ω_{01} with $2\omega_{01} + \Delta$, where Δ is the transmon anharmonicity. The resulting populations β are at least an order of magnitude lower than their corresponding populations α in the first excited state. This justifies the assumption that the $|2\rangle$ -state population is negligible when slow resets are used.

2. RabiEF and assignment errors

We provide numerical evidence that the RabiEF measurement of α is accurate despite measurement assignment errors. We construct a three-level model with states $|0\rangle$, $|1\rangle$, and $|2\rangle$. The initial state is the thermal state $\rho_{\text{th}} = (1 - \alpha)|0\rangle\langle 0| + \alpha|1\rangle\langle 1|$, where we assume that the thermal population in $|2\rangle$ is negligible. The gates in the RabiEF circuits are assumed to be ideal. Crucially, we add finite sampling effects and discrimination errors in the readout process, with a discrimination in a one-dimensional space with overlapping Gaussians.

Here, the state $|0\rangle$ is mapped to a Gaussian distribution $\mathcal{N}(\mu = -1, \sigma_m)$ with mean -1 and standard deviation σ_m . The $|1\rangle$ and $|2\rangle$ states are mapped to a single Gaussian distribution $\mathcal{N}(\mu = 1, \sigma_m)$ with mean $+1$ and standard deviation σ_m . To draw a shot from a state ρ , we first choose a random number μ from $\{-1, 1\}$, where the probability of -1 is $\text{Tr}[\rho|0\rangle\langle 0|]$ and the probability of $+1$ is $\text{Tr}[\rho(|1\rangle\langle 1| + |2\rangle\langle 2|)]$. Next, we mimic the readout process by sampling from $\mathcal{N}(\mu, \sigma_m)$ and assigning a count

TABLE III. Properties of qubits used for the data presented in Fig. 1. Qubit frequency, ω_{01} ; transmon anharmonicity, Δ ; population in $|1\rangle$ state measured using RabiEF, α ; inferred population in second excited state, β .

Qubit	114	115	116	117	118	119	120	121
$\omega_{01}/(2\pi)$ (GHz)	5.14	5.22	5.08	5.22	5.28	5.16	5.03	5.15
$\Delta/(2\pi)$ (MHz)	−303	−303	−304	−302	−303	−304	−305	−302
α (%)	4.99	8.75	4.01	6.00	4.17	3.06	5.42	2.79
T_{eff} (mK)	82	103	76	89	80	71	83	69
β (%)	0.30	0.88	0.19	0.42	0.21	0.11	0.35	0.10
T_1 (μ s)	281	259	207	222	199	293	175	261

of 1 if the result is greater than 0. Therefore, the overlap between the two Gaussian distributions $\mathcal{N}(\mu = -1, \sigma_m)$ and $\mathcal{N}(\mu = 1, \sigma_m)$ —controlled by σ_m —creates measurement assignment errors, see inset in Fig. 9(b). The relation between σ_m and the readout fidelity $F = 1 - [P(1|0) + P(0|1)]/2$ is shown in Fig. 9(b).

Even in the presence of strong assignment errors, e.g., $\sigma_m > 0.4$, for which $F < 99\%$, RabiEF accurately measures α . The estimates $\hat{\alpha}$ center around the ideal value set to 1.5%, see Fig. 9(a). Furthermore, the magnitude of the error bars and, correspondingly, the standard deviation of the distribution of independent measures, increases with the assignment error, see Fig. 9(b). Crucially, we can reduce these errors by increasing the number of shots, see Fig. 10. Here, the distribution of measurements is again centered around the ideal value. Importantly, increasing the number of shots reduces the variance of the distribution.

These results are expected, given that assignment errors, and potentially other imperfections, affect both a_π and a_ϕ . Therefore, as α is estimated with a ratio where both numerator and denominator depend linearly on the fitted amplitudes, we can expect the measurement to be first-order insensitive to assignment errors.

APPENDIX B: ERROR PROPAGATION

Here, we give details on the computation of the error bars shown in Fig. 1. The uncertainty is calculated via standard error propagation. For example, the propagated standard deviation on $\langle O \rangle_{\text{mit}}$ is

$$\sigma_{\text{mit}}^2 = \left(\frac{\partial \langle O \rangle_{\text{mit}}}{\partial \langle O \rangle_{\text{raw}}} \sigma_{\text{raw}} \right)^2 + \left(\frac{\partial \langle O \rangle_{\text{mit}}}{\partial \langle Z^* \rangle} \sigma_{\langle Z^* \rangle} \right)^2 + \sum_{i=1}^{n-1} \left(\frac{\partial \langle O \rangle_{\text{mit}}}{\partial f_{\text{sp},i}} \sigma_{\text{sp},i} \right)^2. \quad (\text{B1})$$

Here, $\sigma_{\langle Z^* \rangle}$ denotes the standard deviation on the TREX denominator. The standard deviations for observables O , i.e., σ_{raw} and $\sigma_{\langle Z^* \rangle}$, are computed from the samples based on $\langle O^2 \rangle - \langle O \rangle^2$. The standard deviation $\sigma_{\text{sp},i}$ from the RabiEF experiment on qubit i is based on the uncertainties stemming from the underlying curve fits.

APPENDIX C: CNOT NOISE

The shaded area in Fig. 1(c) represents the range of possible values for the observable when only CNOT noise

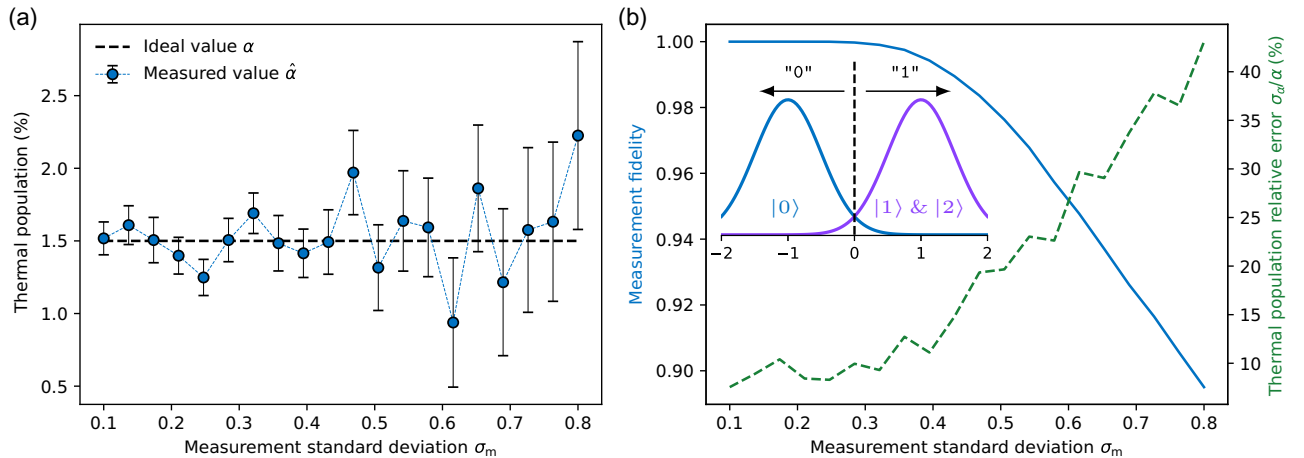


FIG. 9. Sensitivity of RabiEF to assignment errors. (a) Simulations of estimates of $|1\rangle$ state population α with 300 shots per data point. The error bars are obtained by propagating fit errors on a_π and a_ϕ through Eq. (A1). The dashed horizontal line shows the ideal value, $\alpha = 1.5\%$. (b) The solid line shows how the readout fidelity F varies with σ_m . The dashed green line shows how the error bars from panel (a), normalized to α , grow with the strength of assignment errors σ_m for a fixed 300 shots per $\hat{\alpha}$. The inset shows how the three states are assigned to labels “0” and “1.”

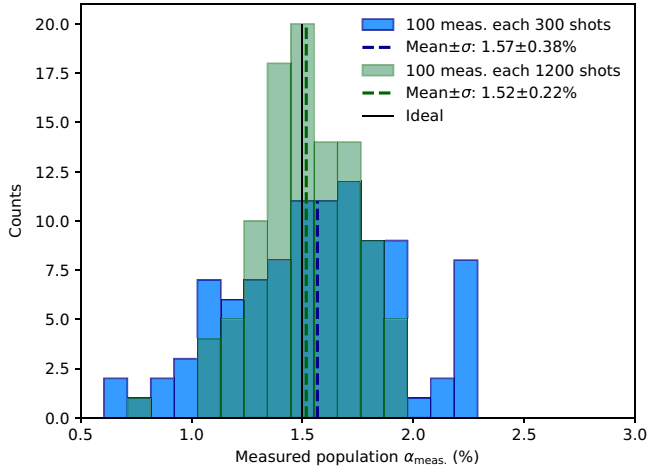


FIG. 10. Increasing the number of shots improves RabiEF precision. Distribution of 100 individual measurements of α with 300 and 1200 shots. Here, $\sigma_m = 0.6$, which corresponds to a readout error of 4.8%. By comparison, the mean readout error on `ibm_fez` is 1.7%.

is present, i.e., the raw observable after removing SPAM errors and assuming no out-of-model errors. To obtain this shaded area, first, we twirl the two-qubit gates and learn a sparse Pauli-Lindblad noise model for the two layers L_1 and L_2 in Fig. 1(b), following van den Berg *et al.* [16]. Next, we use the resulting noise model to compute the noisy observable via a classical Clifford simulation, which contains only the noise for the CNOT gates. Crucially, the error rates of conjugate Pauli pairs, i.e., P and $UPU^\dagger$ (where P is a Pauli and U is the unitary of the CNOT layer) can only be inferred as a product using this noise learning protocol. However, the simulation requires the specification of individual rates, which introduces degrees of freedom. Initially, we split the product symmetrically, as supported by recent theoretical work up to leading order [64]. Finally, we solve two optimization problems that adjust the splits of conjugate Pauli pairs to either maximize or minimize the measured observable under the constraint of all Pauli fidelities remaining physical, i.e., ≤ 1 . This results in the shaded area in Fig. 1(c). As these optimization problems are nonconvex, we do not know whether the solutions that we find are the global optima. This may also explain why the data on four qubits in Fig. 1(c) lie outside of the gray area. We have opted for this procedure out of simplicity as it suffices to highlight the importance of separating state-preparation and measurement noise. The resulting spread in possible observable values is large. With additional noise learning circuits, the noise model of the CNOT gate layers could be further constrained by implementing interleaved cycle benchmarking [30], multi-layer cycle benchmarking [65], or a self-consistent gate set learning scheme [32].

APPENDIX D: PAULI TRANSFER MATRICES

For completeness, we now provide a definition of Pauli transfer matrices (PTMs) [66]. They provide a useful representation of quantum channels. The elements of the PTM of an n -qubit quantum channel Λ are

$$(\Gamma_\Lambda)_{ij} = \frac{1}{2^n} \text{Tr}[P_i \Lambda(P_j)], \quad (\text{D1})$$

where $[P_0, \dots, P_{4^n-1}]$ denotes the n -qubit Pauli basis arranged in lexicographic order. The PTM of a composite map $H \circ G$ is the matrix product of the PTMs of the individual maps $\Gamma_H \Gamma_G$. Therefore, the convention that noise Λ_G occurs before a gate G translates to PTMs as $\Gamma_G \Gamma_{\Lambda_G}$.

Next, the PTM of a measurement is a matrix whose only nonzero entries are 1s in the top-left and bottom-right corners, with all other entries equal to 0. As a result, it annihilates any prior X or Y components of an operation. Consequently, our PTMs are defined as before, but with the basis restricted to the set $\{I, Z\}^{\otimes n}$ arranged in lexicographic order. As an example, consider the one-qubit channel $\Lambda(\rho) = (1-p)\rho + pX\rho X$. Its PTM in the full basis $\{I, X, Y, Z\}$ is

$$\Gamma_\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1-2p & 0 \\ 0 & 0 & 0 & 1-2p \end{pmatrix}. \quad (\text{D2})$$

Since we consider only the $\{I, Z\}$ part of the basis, we express this PTM as

$$\Gamma_\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & 1-2p \end{pmatrix}. \quad (\text{D3})$$

Finally, an important property of PTMs combined with twirling is that twirling transforms any noise channel into a Pauli noise channel, resulting in a diagonal PTM.

APPENDIX E: PTM OF FINAL MEASUREMENTS

In a final measurement—i.e., measuring the control qubit in the CNOT picture—the outcome is written back onto the same wire, which then acts as a classical bit. To derive the corresponding PTM, we temporarily introduce an ancillary classical line and decompose the measurement procedure into the following sequence, also shown in Fig. 11. (1) Apply an ideal reset to the ancillary classical line to initialize it in the 0 state. (2) Apply the noisy measurement in the CNOT picture by transferring the control qubit outcome onto the ancillary line. (3) Reset the qubit wire, which from this point on acts as a classical line. (4) Swap the value from the ancillary classical line back onto the (now classical) qubit wire. (5) Remove the ancillary classical line. We compute the PTMs of each

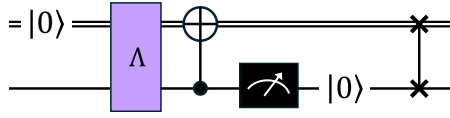


FIG. 11. In the CNOT picture, the final measurement corresponds to measuring the control qubit. The outcome is written back onto the qubit (control) wire, which thereafter functions as a classical line.

component in this circuit and multiply them to obtain the PTM for the final noisy measurement. The PTM of a reset operator $\mathcal{R}(\rho) = E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger$, where $E_0 = |0\rangle\langle 0|$ and $E_1 = |0\rangle\langle 1|$, is

$$\Gamma_{|0\rangle} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}. \quad (\text{E1})$$

The PTM of the swap is

$$\Gamma_{\text{swap}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (\text{E2})$$

The PTM of a final measurement can then be calculated as $\Gamma_{\text{fm}} = \Gamma_{\text{swap}} (\Gamma_{|0\rangle} \otimes I) \Gamma_{\text{CX}} \Gamma_{\Lambda} (I \otimes \Gamma_{|0\rangle})$, where $\Gamma_{\text{CX}} \Gamma_{\Lambda}$ is defined in Eq. (6), and $\Gamma_{|0\rangle}$ and Γ_{swap} are defined in Eqs. (E2) and (E1), respectively. The result is

$$\Gamma_{\text{fm}} = \begin{pmatrix} 1 & 0 \\ 0 & f_a f_s \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}. \quad (\text{E3})$$

Here, f_a and f_s are, as before, the fidelity of the assignment and state errors, respectively. Thus, after discarding the ancillary classical line, the PTM of a final noisy measurement is $\text{diag}(1, f_a f_s)$. The absence of f_c in this PTM is understood as follows. The correlated error simultaneously flips the qubit and the classical bit. For example, a qubit, initially in state $|0\rangle$, is measured as 1, and the qubit ends up in $|1\rangle$. However, in final measurements, we are only concerned with the correctness of the reported outcome. Thus, such an event is effectively indistinguishable from a simple bitflip error and can be interpreted as either a readout assignment error or a state-preparation error.

APPENDIX F: SETUP OF SIMULATIONS

Our simulations fall into two categories, qubit and qutrit simulations. As the RabiEF experiment uses gates that interact with the $|2\rangle$ state, it is the only experiment simulated with qutrits. Appendix F1 describes our qutrit simulator and its application to the RabiEF experiment. Appendix F2 translates the readout assignment matrix R to the measurement noise fidelities. Appendix F3 discusses the setup of qubit simulations for the results shown in

Secs. IV B and IV C. Appendix F4 gives details on the simulation of the teleportation circuit, including PEC.

1. Qutrit simulations for RabiEF

To simulate the RabiEF experiments while taking into account the $|2\rangle$ -state population, we implemented a qutrit simulator with QISKIT [67]. All qubit instructions are embedded as ideal gates in $\text{SU}(3)$, and thus do not interact with the $|2\rangle$ state. Only the $R_{12}(\theta)$ and X_{12} gates, noise channels, and measurements interact with the $|2\rangle$ -state. The circuits are converted into a series of operators, i.e., SuperOp class instances with qutrit dimensions. For example, a single qutrit error channel is a 9×9 superoperator matrix. A state ρ is evolved by a circuit c by repeatedly calling the `DensityMatrix.evolve(op, qargs)` QISKIT method, where the `op` argument is the operator for each gate and `qargs` is a list of qutrit indices identifying the support of the operator. A RabiEF experiment consists of three circuits: the two π and π , circuits c_{π} and c_{π} , respectively, and a repetition delay circuit c_{rep} consisting of three active resets and an optional delay for slow resets. The delay is not simulated when fast resets are used. The simulator evolves an initial state $\rho = (|0\rangle\langle 0|)^{\otimes n}$ by simulating the circuits in the following loop for N samples many times: (1) apply c_{rep} , (2) apply c_{π} , (3) apply c_{rep} , and (4) apply c_{π} . One loop corresponds to a single sample per π and π circuit. Both RabiEF circuits c_{π} and c_{π} produce measurement outcomes, which are stored in two length- N arrays. The repetition delay also produces measurement outcomes, but they are only used to control the feedforward instructions in the active resets and are not stored long-term. The initial states for each circuit, $\rho_{\text{in},k}$, over all k samples are saved to compute the *true* state-preparation probability p_{sp} , shown in Table I, as

$$p_{\text{sp}} := \text{avg}_{\substack{s=\pi, \pi \\ k=0, \dots, N-1}} \text{Tr} \left[|1\rangle\langle 1| \rho_{\text{in},k}^{(s)} \right]. \quad (\text{F1})$$

We simulate thermal relaxation during slow resets, measurements, and idle times in the repetition delay. Noise channels for measurements are applied before the ideal measurement. We numerically compute the thermal relaxation channel with the Lindblad master equation and six Lindblad operators, defining relaxation, heating, and dephasing. Each Lindblad operator L_i has a corresponding rate γ_i , controlling the strength of the operator. The operators and rates are

$$\begin{aligned} L_1 &= |0\rangle\langle 1|, & L_2 &= |1\rangle\langle 0|, \\ L_3 &= |1\rangle\langle 2|, & L_4 &= |2\rangle\langle 1|, \\ L_5 &= |0\rangle\langle 0| - |1\rangle\langle 1|, & L_6 &= |1\rangle\langle 1| - |2\rangle\langle 2|, \end{aligned} \quad (\text{F2})$$

$$\begin{aligned}
\gamma_1 &= \frac{1}{T_1}, \quad \gamma_2 = \gamma_1 \exp(-\hbar\omega_{01}/k_B T_{\text{eff}}), \\
\gamma_3 &= \frac{1}{T_1^{12}}, \quad \gamma_4 = \gamma_3 \exp(-\hbar(\omega_{01} + \Delta)/k_B T_{\text{eff}}), \quad (\text{F3}) \\
\gamma_5 &= \frac{1}{T_2}, \quad \gamma_6 = \frac{1}{T_2^{12}}.
\end{aligned}$$

The solution to the Lindblad master equation $\Lambda_{\text{th}}(\rho, t)$ for a given evolution time t is computed with QISKIT DYNAMICS [68] and included as a SuperOp instruction in the circuits. The rates γ_1 , γ_3 , γ_5 , and γ_6 are chosen from typical superconducting qubit decoherence times [69], with their definitions given in Eq. (F3) and their values given in Table IV. The heating rates γ_2 and γ_4 are calculated using the energy differences between $|0\rangle$, $|1\rangle$, and $|2\rangle$, and an assumed effective qutrit temperature $T_{\text{eff}} = 60$ mK, following the Boltzmann distribution. We further assume that the qutrits decay sequentially, i.e., $|2\rangle \rightarrow |0\rangle$ transitions are naturally suppressed by the quantum hardware [46]. We include measurement-induced leakage with an additional error channel prior to the ideal measurements, after the measurement-associated thermal relaxation. This is modeled as a projection to the $|2\rangle\langle 2|$ state with probability $p_{\text{leak}} = 0.2\%$. The error channel on a qubit undergoing measurement, with duration $t_{\text{meas}} = 1244$ ns, is thus

$$\Lambda_{\text{meas}}(\rho) = (1 - p_{\text{leak}})\Lambda_{\text{th}}(\rho, t_{\text{meas}}) + p_{\text{leak}}|2\rangle\langle 2|. \quad (\text{F4})$$

Following this noise channel, our simulator applies an ideal projective measurement and then the readout assignment matrix R , see Appendix F 2 and Eq. (11). The simulator then records the noisy measurement outcome as the integer $o = 0, 1, 2$ and uses the postmeasurement state as the input to the next circuit. Qutrit discrimination is only used for measurements in the repetition delay, as RabiEF requires binary classification.

TABLE IV. Parameters used in the simulations.

Symbol	Value	Description
T_{eff}	60 mK	Qubit effective temperature
T_1	200 μs	Relaxation time in qubit subspace
T_2	100 μs	Dephasing time in qubit subspace
T_1^{12}	100 μs	Relaxation time in qutrit subspace
T_2^{12}	50 μs	Dephasing time in qutrit subspace
ω_{01}	4.9 GHz	Qubit 0–1 frequency
Δ	−0.3 GHz	Qubit anharmonicity
$t_{\text{rep. delay}}$	250 μs	Repetition delay time
t_{meas}	1244 ns	Measurement time
p_{leak}	0.2%	Probability of leak during measurement

2. Assignment error fidelities for qutrits

Errors in final measurements can be specified as a “read-out assignment error” matrix where entry j, k is the probability q_{jk} of misclassifying state $|k\rangle$ as $|j\rangle$ [70]. The 3×3 qutrit readout assignment matrix

$$R = \begin{pmatrix} q_{00} & q_{01} & q_{02} \\ q_{10} & q_{11} & q_{12} \\ q_{20} & q_{21} & q_{22} \end{pmatrix} \quad (\text{F5})$$

used in the RabiEF simulation is based on existing hardware experiments [46,52,53]. We thus neglect the overlap between the $|0\rangle$ and $|2\rangle$ readout signals, i.e., $q_{02} = q_{20} = 0$, and assume that the qutrit misclassification probability for $|1\rangle \rightarrow |2\rangle$ is approximately an order of magnitude larger than that for $|0\rangle \rightarrow |1\rangle$. Therefore, we set $q_{12} = 6\%$ and $q_{01} = 0.9\%$, resulting in the readout assignment matrix shown in Eq. (11).

When qubit reset is used, we require a readout assignment matrix that always misclassifies $|2\rangle$ as $|1\rangle$, see Eq. (11). The qubit readout assignment errors p_{jk} are related to the qutrit ones, following

$$p_{00} = q_{00}, \quad (\text{F6a})$$

$$p_{11} = q_{11} + q_{21}, \quad (\text{F6b})$$

$$p_{01} = q_{01}, \quad (\text{F6c})$$

$$p_{10} = q_{10} + q_{20}, \quad (\text{F6d})$$

with $q_{20} = 0$ as discussed previously. From the values in Eq. (11) we obtain qubit assignment matrix probabilities $p_{00} = p_{11} = 99.1\%$ and $p_{01} = p_{10} = 0.9\%$.

To ensure consistency between the RabiEF simulations in Secs. IV A and IV B, we now connect the qubit readout assignment matrix p_{jk} to the measurement noise fidelities f_a , f_s , and f_c of the model in Sec. III C. We apply the PTM of the readout to the different qubit initial states. In the qubit $\otimes$ cbit notation, the initial state $|00\rangle = (II + IZ + ZI + ZZ)/4$, for example, corresponds to the vector $\Gamma_{00} = (1, 1, 1, 1)^T$ in the basis $\{II, IZ, ZI, ZZ\}$. The PTM of the readout Γ_m , given in Eq. (6), transforms this vector according to

$$\begin{aligned}
\Gamma_m \Gamma_{00} &= \frac{1}{4} (II + f_a f_s IZ + f_c f_s ZI + f_c f_a ZZ) \\
&= \frac{1}{4} (1 + f_a f_s + f_c f_s + f_c f_a) |00\rangle\langle 00| \\
&\quad + \frac{1}{4} (1 + f_a f_s - f_c f_s - f_c f_a) |10\rangle\langle 10| \\
&\quad + \frac{1}{4} (1 - f_a f_s + f_c f_s - f_c f_a) |01\rangle\langle 01| \\
&\quad + \frac{1}{4} (1 - f_a f_s - f_c f_s + f_c f_a) |11\rangle\langle 11|.
\end{aligned}$$

Since we consider a final measurement, we only care about the state of the classical bit and thus trace-out the qubit. Therefore, the measurement acts on the qubit $|0\rangle$ state as

$$|0\rangle\langle 0| \rightarrow \frac{1+f_a f_s}{2} |0\rangle\langle 0| + \frac{1-f_a f_s}{2} |1\rangle\langle 1|.$$

This implies $p_{00} = (1+f_a f_s)/2$ and $p_{10} = (1-f_a f_s)/2$. Similarly, for the excited qubit state, we consider $|10\rangle$ in the qubit $\otimes$ cbit notation and apply the same reasoning. In summary, for a noisy measurement, the corresponding assignment error probabilities p_{jk} are

$$\begin{aligned} p_{00} &= \frac{1+f_a f_s}{2}, & p_{10} &= \frac{1-f_a f_s}{2}, \\ p_{10} &= \frac{1-f_a f_s}{2}, & p_{11} &= \frac{1+f_a f_s}{2}. \end{aligned} \quad (\text{F7})$$

This result also shows that the correlated error f_c is irrelevant when considering final measurements, see also Appendix E. The values $p_{00} = p_{11} = 99.1\%$ and $p_{01} = p_{10} = 0.9\%$ thus require $f_a f_s = 0.982$. In our simulations, we assume that assignment and state errors are equally likely, such that $f_a f_s$ splits equally into $f_a = f_s = 0.99096$. As the readout assignment matrix R does not capture correlated errors, this analysis cannot determine the value of f_c . However, as the correlated error is of a higher weight than the state and assignment errors, we assume that it has a higher fidelity and set $f_c = 0.995$.

3. Qubit simulation setup for noise learning and mitigation

This section covers the setup for simulations in Secs. IV B to IV D. We used QISKIT AER [55] to simulate MCB and the mitigation circuits. We inserted additional dummy gates to engineer the noise model in Sec. III C. As QISKIT AER applies noise before measurements, implementation of state errors requires only an X error on the measurement qubit with probability $p_s = (1-f_s)/2$. QISKIT AER has native support for readout assignment matrices, which we leveraged to implement assignment errors on the classical bit. The assignment error PTM in Eq. (4) is simulated with the readout assignment matrix R_a :

$$R_a = \begin{pmatrix} 1-p_a & p_a \\ p_a & 1-p_a \end{pmatrix}. \quad (\text{F8})$$

QISKIT AER does not support correlated errors between qubits and classical bits. Therefore, we implemented Λ_c in Fig. 2 with a postmeasurement dummy delay instruction, to which we attached an X error. This postmeasurement error is equivalent to a correlated error premeasurement, which can be seen by back propagating the X Pauli through the measurement in the CNOT picture. Therefore, an X gate postmeasurement with probability $p_c = (1-f_c)/2$ is

equivalent to a correlated error premeasurement with the same probability. We set the duration of the delay to 0 to ensure that no additional error channels were attached to it.

To simulate a given prepared initial state, we prepended reset instructions to the MCB and mitigation circuits, and attached an X error to them with probability p_{sp} . This forced QISKIT AER to prepare the state $\rho_{sp} = (1-p_{sp})|0\rangle\langle 0| + p_{sp}|1\rangle\langle 1|$. The measurement outcomes returned by QISKIT AER were postprocessed to obtain the expectation values in Figs. 5 and 6.

4. Error mitigation of the teleportation circuit

The teleportation circuit in Sec. IV D is simulated with density matrices using QISKIT AER [55]. State-preparation and measurement noise are mitigated with probabilistic error cancellation [16,27]. The simulation is carried out as follows. (1) We generate $N = 128$ teleportation circuit realizations, where the inverse noise channels $\tilde{\Lambda}^{-1}$ in Fig. 7 are replaced with Paulis sampled from their quasiprobability distributions. (2) The circuits are classically simulated, obtaining density matrices $\rho_{out,j}$ on the target qubit q_2 , for $j \in [0, N-1]$. (3) We obtain a mitigated physical state ρ_{mit} with state tomography. Since PEC is applied to expectation values and not states, this is done as follows. (3.a) We compute the expectation values $\langle P \rangle_{raw,j} = \text{Tr}[P \rho_{out,j}]$, for $P = X, Y, Z$, without shot noise. However, since PEC is very sensitive to shot noise, we emulate it by sampling 100 times from a Bernoulli distribution with probability $(1 + \langle P \rangle_{raw,j})/2$ for outcome “0,” approximating $\langle P \rangle_{raw,j}$ as $E_{P,j}$ with both shot noise and SPAM errors. (3.b) The mitigated expectation values $\langle P \rangle_{mit}$ are computed using PEC, i.e., Eq. (16), over all N circuit realizations and $P = X, Y, Z$ Paulis. (3.c) We minimize the mean-squared error (MSE) in Eq. (17) over all expectation values to obtain a physical state ρ_{mit} based on the mitigated expectation values. (4) We compute the state fidelity between the mitigated state ρ_{mit} and the ideal one with Eq. (20). Steps (1)–(4) are repeated for 15 equidistant rotation angles $\theta \in [0, 2\pi]$ to study different states to teleport. The mitigated fidelities in Sec. IV D display the interquartile range and the median obtained from 300 realizations of this simulation. To avoid simulating 300×128 circuits, we simulate 1000 PEC circuit realizations and compute the raw expectation values $\langle P \rangle_{raw,j}$ on all of them. Next, we bootstrap these results into 300 datasets by sampling $N = 128$ circuit realizations from the 1000 circuit executions.

The 15 unmitigated fidelities, shown as purple dots in Fig. 7, are obtained using the same procedure with $N = 1$ circuit and no PEC sampling. Furthermore, since there is only one circuit, and thus only one density matrix $\rho_{out,0}$, MSE minimization is not necessary.

APPENDIX G: EXTENSION OF THE NOISE MODEL TO CORRELATED QUBITS UNDERGOING MEASUREMENT

Our proposed noise model in Sec. III C assumes that noise is local to the qubit and associated classical bit. Accounting for correlations between neighboring qubits requires increasing the support of the Pauli-Lindblad noise model and measurement cycle benchmarking experiments. The same noise learning procedure as before can be employed, but now with more MCB decay curves and model fidelities to learn. This is similar to how gate noise learning with cycle benchmarking is extended to include higher-weight noise generators [16,27]. The chosen support determines which correlations are assumed to be negligible.

As an example, we derive the twirled-noise PTM for two neighboring qubits undergoing measurement and describe the noise learning procedure, following our protocol in Fig. 3. We assume that correlations between qubits arise from measurements and that thermal states are uncorrelated. Given that fast resets use measurements, this model accounts for correlations in prepared states using fast resets.

We model correlations between two subsystems $j = 0, 1$, each composed of a qubit-classical bit pair (q_j, c_j) . In addition to the errors possible in our original noise model in Sec. III C, we include correlated errors between the qubits and correlated errors between the classical bits. The resulting error model is shown in Fig. 12 for the measurement cycle benchmarking circuit. Employing the same twirling as before [16,27,44], all errors are bitflips, i.e., Pauli strings composed of I s and X s. We order the subsystems so that qubits q_0 and q_1 are before classical bits c_0 and c_1 , i.e., $q_0 q_1 c_0 c_1$. Therefore, Pauli string $IXIX$ corresponds to the original noise model's correlated error between qubit q_1 and classical bit c_1 ; and Pauli string $XXII$ corresponds to a weight-two error between qubits q_0 and q_1 . Errors local to subsystems $j = 0, 1$ are denoted with superscripts (0) and (1), respectively. New errors between subsystems

are denoted with superscript (2), and are described by the following channels.

- (1) Weight-two correlated state errors between qubits:

$$\Lambda_s^{(2)} = (1 - p_s^{(2)})\rho + p_s^{(2)}XXII\rho XXII, \quad (G1)$$

where $f_s^{(2)} = 1 - 2p_s^{(2)}$. These errors can be caused by crosstalk between qubits, potentially induced by measurements [29].

- (2) Weight-four correlated errors between both qubits and both classical bits:

$$\Lambda_c^{(2)} = (1 - p_c^{(2)})\rho + p_c^{(2)}XXXX\rho XXXX, \quad (G2)$$

where $f_c^{(2)} = 1 - 2p_c^{(2)}$. This arises from a weight-two postmeasurement error on both qubits $\tilde{\Lambda}_c^{(2)}$, commuted backward through the measurements:

$$\tilde{\Lambda}_c^{(2)} = (1 - p_c^{(2)})\rho + p_c^{(2)}XXII\rho XXII. \quad (G3)$$

This is similar to how the correlated error Λ_c in Fig. 2 and Eq. (5) is derived from a postmeasurement qubit-error. Similarly to $\Lambda_s^{(2)}$, these errors can be caused by crosstalk between qubits, only after measurement.

- (3) Weight-two correlated assignment errors between classical bits:

$$\Lambda_a^{(2)} = (1 - p_a^{(2)})\rho + p_a^{(2)}IIXX\rho IIXX, \quad (G4)$$

where $f_a^{(2)} = 1 - 2p_a^{(2)}$. These errors can occur because of frequency collisions in readout resonators.

- (4) Weight-two correlated state-preparation errors between qubits:

$$\Lambda_{sp}^{(2)} = (1 - p_{sp}^{(2)})\rho + p_{sp}^{(2)}XXII\rho XXII, \quad (G5)$$

where $f_{sp}^{(2)} = 1 - 2p_{sp}^{(2)}$. As active resets utilize a number of measurements, correlated errors can manifest as a result of measurement-induced correlations.

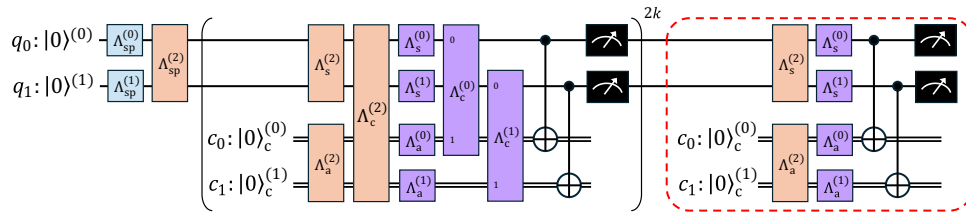


FIG. 12. Measurement noise model for two correlated qubits. Qubits and classical bits are indicated by single and double lines, respectively. Noise channels from the original noise model in the main text are shown in purple and blue. New noise channels are in light orange. The support of some channels differs from the wires they cross. For these, indices are used to denote the qubits and classical bits on which they are applied. For example, $\Lambda_c^{(0)}$ is supported by q_0 and c_0 but not q_1 . Superscripts (0) and (1) denote errors that arise from the original noise model on qubit-classical bit pairs (q_0, c_0) and (q_1, c_1) , respectively. New errors that model correlations between the qubit-classical bit pairs are denoted with superscript (2). The dashed red rectangle indicates the final measurement, which is insensitive to postmeasurement qubit errors, equivalent to $\Lambda_c^{(j)}$ for $j = 0, 1, 2$.

The full PTM $\Gamma_{2k}^{(2q)}$ for this noise model in a MCB circuit is diagonal and has 12 fidelities to learn: four per qubit ($f_{\text{sp}}^{(j)}$, $f_s^{(j)}$, $f_a^{(j)}$, and $f_c^{(j)}$ for $j = 0, 1$), from our original

noise model, and four from the new correlated errors ($f_{\text{sp}}^{(2)}$, $f_s^{(2)}$, $f_a^{(2)}$, and $f_c^{(2)}$). The entries y_P of the diagonal, for $P \in \{I, Z\}^4$, are given as follows:

$$y_{IIII} = 1, \quad (\text{G6a})$$

$$y_{IIIZ} = (f_a^{(1)} f_a^{(2)})^{2k} (f_c^{(1)} f_c^{(2)} f_s^{(1)} f_s^{(2)})^k, \quad (\text{G6b})$$

$$y_{IIZI} = (f_a^{(0)} f_a^{(2)})^{2k} (f_c^{(0)} f_c^{(2)} f_s^{(0)} f_s^{(2)})^k, \quad (\text{G6c})$$

$$y_{IIZZ} = (f_a^{(0)} f_a^{(1)})^{2k} (f_c^{(0)} f_c^{(1)} f_s^{(0)} f_s^{(1)})^k, \quad (\text{G6d})$$

$$y_{IZII} = f_a^{(1)} f_a^{(2)} f_{\text{sp}}^{(1)} f_{\text{sp}}^{(2)} (f_s^{(1)} f_s^{(2)})^{2k+1} (f_c^{(1)} f_c^{(2)})^{2k}, \quad (\text{G6e})$$

$$y_{IZIZ} = f_{\text{sp}}^{(1)} f_{\text{sp}}^{(2)} (f_a^{(1)} f_a^{(2)})^{2k+1} (f_c^{(1)} f_c^{(2)})^k (f_s^{(1)} f_s^{(2)})^{k+1}, \quad (\text{G6f})$$

$$y_{IZZI} = f_a^{(1)} f_{\text{sp}}^{(1)} f_{\text{sp}}^{(2)} (f_a^{(2)} f_s^{(1)})^{2k+1} (f_a^{(0)} f_c^{(1)})^{2k} (f_c^{(0)} f_c^{(2)} f_s^{(0)})^k (f_s^{(2)})^{k+1}, \quad (\text{G6g})$$

$$y_{IZZI} = f_a^{(2)} f_{\text{sp}}^{(1)} f_{\text{sp}}^{(2)} (f_a^{(1)} f_s^{(2)})^{2k+1} (f_a^{(0)} f_c^{(2)})^{2k} (f_c^{(0)} f_c^{(1)} f_s^{(0)})^k (f_s^{(1)})^{k+1}, \quad (\text{G6h})$$

$$y_{ZIII} = f_a^{(0)} f_a^{(2)} f_{\text{sp}}^{(0)} f_{\text{sp}}^{(2)} (f_s^{(0)} f_s^{(2)})^{2k+1} (f_c^{(0)} f_c^{(2)})^{2k}, \quad (\text{G6i})$$

$$y_{ZIIZ} = f_a^{(0)} f_{\text{sp}}^{(0)} f_{\text{sp}}^{(2)} (f_a^{(2)} f_s^{(0)})^{2k+1} (f_a^{(1)} f_c^{(0)})^{2k} (f_c^{(1)} f_c^{(2)} f_s^{(1)})^k (f_s^{(2)})^{k+1}, \quad (\text{G6j})$$

$$y_{ZIZI} = f_{\text{sp}}^{(0)} f_{\text{sp}}^{(2)} (f_a^{(0)} f_a^{(2)})^{2k+1} (f_c^{(0)} f_c^{(2)})^k (f_s^{(0)} f_s^{(2)})^{k+1}, \quad (\text{G6k})$$

$$y_{ZIZZ} = f_a^{(2)} f_{\text{sp}}^{(0)} f_{\text{sp}}^{(2)} (f_a^{(0)} f_s^{(2)})^{2k+1} (f_a^{(1)} f_c^{(2)})^{2k} (f_c^{(0)} f_c^{(1)} f_s^{(1)})^k (f_s^{(0)})^{k+1}, \quad (\text{G6l})$$

$$y_{ZZII} = f_a^{(0)} f_a^{(1)} f_{\text{sp}}^{(0)} f_{\text{sp}}^{(1)} (f_s^{(0)} f_s^{(1)})^{2k+1} (f_c^{(0)} f_c^{(1)})^{2k}, \quad (\text{G6m})$$

$$y_{ZZIZ} = f_a^{(0)} f_{\text{sp}}^{(0)} f_{\text{sp}}^{(1)} (f_a^{(1)} f_s^{(0)})^{2k+1} (f_a^{(2)} f_c^{(0)})^{2k} (f_c^{(1)} f_c^{(2)} f_s^{(2)})^k (f_s^{(1)})^{k+1}, \quad (\text{G6n})$$

$$y_{ZZZI} = f_a^{(1)} f_{\text{sp}}^{(0)} f_{\text{sp}}^{(1)} (f_a^{(0)} f_s^{(1)})^{2k+1} (f_a^{(2)} f_c^{(1)})^{2k} (f_c^{(0)} f_c^{(2)} f_s^{(2)})^k (f_s^{(0)})^{k+1}, \quad (\text{G6o})$$

$$y_{ZZZZ} = f_{\text{sp}}^{(0)} f_{\text{sp}}^{(1)} (f_a^{(0)} f_a^{(1)})^{2k+1} (f_c^{(0)} f_c^{(1)})^k (f_s^{(0)} f_s^{(1)})^{k+1}. \quad (\text{G6p})$$

To identify any unlearnable degrees of freedom in $\Gamma_{2k}^{(2q)}$, we compute the observability matrix with respect to the logarithm of y_P and all fidelities [71]. The observability matrix is rank 9, which, with 12 unknown fidelities, indicates that a single MCB experiment is insufficient to learn the full noise model. Learning $f_{\text{sp}}^{(0)}$ and $f_{\text{sp}}^{(1)}$ with RabiEF on thermal states, assuming $f_{\text{sp}}^{(2)} = 1$, reduces the number of unknown fidelities to 9. This indicates that one RabiEF experiment and one MCB experiment are sufficient to learn $\Gamma_{2k}^{(2q)}$ with slow resets. To illustrate the strength of our full noise learning protocol, we now show how all 12 fidelities can be learned when fast resets introduce non-negligible correlations during state preparation.

(1) First, learn $f_{\text{sp}}^{(0),\text{slow}}$ and $f_{\text{sp}}^{(1),\text{slow}}$ with separate RabiEF experiments, using slow resets. As we assume

no correlations during state preparation with slow resets, we set $f_{\text{sp}}^{(2),\text{slow}} = 1$. This reduces the number of unknown fidelities to 9. (2) Run a MCB experiment with slow resets to learn the remaining 9 unknown fidelities $f_s^{(j)}$, $f_a^{(j)}$, and $f_c^{(j)}$ for $j = 0, 1, 2$. (3) The last step, as per our protocol, is a MCB experiment with fast resets. Assuming that all fidelities remain the same, other than those for state preparation, we need only learn $f_{\text{sp}}^{(0),\text{fast}}$, $f_{\text{sp}}^{(1),\text{fast}}$, and $f_{\text{sp}}^{(2),\text{fast}}$. The PTM $\Gamma_{2k}^{(2q)}$ for fast resets has a rank 3 observability matrix, indicating that the last MCB experiment allows us to complete the full noise model on both subsystems. This shows that our full noise learning protocol with RabiEF can also fix the gauge in the noise model of two correlated qubits undergoing measurement.

APPENDIX H: EXTENSION FOR HIGH-FIDELITY RESETS AND HIGH-RATE LEARNING

As coherence times and state-preparation fidelities improve, it is expected that p_{sp} should decrease. This poses a problem for RabiEF, as the error in the estimate of p_{sp} is additive, owing to sampling error on the π and π RabiEF signals. To reduce the effect of this error on the learned measurement noise fidelities, we can amplify p_{sp} by inserting random local X gates on the initial state with probability p_{amp} . The effective state-preparation error probability $\tilde{p}_{\text{sp}}$ is then

$$\tilde{p}_{\text{sp}} = p_{\text{amp}} + p_{\text{sp}}(1 - 2p_{\text{amp}}) = p_{\text{amp}}(1 - 2p_{\text{sp}}) + p_{\text{sp}}, \quad (\text{H1})$$

which is greater than p_{sp} for $p_{\text{sp}} < 1/2$. As a result, one can utilize our MCM noise learning protocol, specifically the purple path in Fig. 3, to (1) learn $\tilde{p}_{\text{sp}}$ with a RabiEF experiment, (2) learn the measurement fidelities with the amplified state-preparation error, and finally (3) learn the high-fidelity reset probability p_{sp} with a second MCB experiment. For the RabiEF experiment, the random X gate is equivalent to relabeling π and π shots with probability p_{amp} , which can be done in postprocessing. For the measurement cycle benchmarking and mitigation experiments, the circuits are already twirled; therefore, the random X gates can be included in the normal twirling samples by applying a bias to the X component in the initial state twirling.

-
- [1] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. van den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, and A. Kandala, Evidence for the utility of quantum computing before fault tolerance, *Nature* **618**, 500 (2023).
 - [2] L. E. Fischer, M. Leahy, A. Eddins, N. Keenan, D. Ferracin, M. A. C. Rossi, Y. Kim, A. He, F. Pietracaprina, B. Sokolov, S. Dooley, Z. Zimborás, F. Tacchino, S. Maniscalco, J. Goold, G. García-Pérez, I. Tavernelli, A. Kandala, and S. N. Filippov, Dynamical simulations of many-body quantum chaos on a quantum computer, *Nature Phys.*, **1** (2026).
 - [3] B. Fuller, M. C. Tran, D. Lykov, C. Johnson, M. Rossmannek, K. X. Wei, A. He, Y. Kim, D. Vu, K. Sharma, Y. Alexeev, A. Kandala, and A. Mezzacapo, Improved quantum computation using operator backpropagation, *arXiv:2502.01897*.
 - [4] Z. Zimborás *et al.*, Myths around quantum computation before full fault tolerance: What no-go theorems rule out and what they don't, *arXiv:2501.05694*.
 - [5] C. Piveteau, D. Sutter, S. Bravyi, J. M. Gambetta, and K. Temme, Error mitigation for universal gates on encoded qubits, *Phys. Rev. Lett.* **127**, 200505 (2021).
 - [6] D. Aharonov, O. Alberton, I. Arad, Y. Atia, E. Bairey, Z. Brakerski, I. Cohen, O. Golan, I. Gurwich, O. Kenneth, E. Leviatan, N. H. Lindner, R. A. Melcer, A. Meyer, G. Schul, and M. Shutman, On the importance of error mitigation for quantum computation, *arXiv:2503.17243*.
 - [7] E. Bäumer, V. Tripathi, D. S. Wang, P. Rall, E. H. Chen, S. Majumder, A. Seif, and Z. K. Mineev, Efficient long-range entanglement using dynamic circuits, *PRX Quantum* **5**, 030339 (2024).
 - [8] A. D. Córcoles, M. Takita, K. Inoue, S. Lekuch, Z. K. Mineev, J. M. Chow, and J. M. Gambetta, Exploiting dynamic quantum circuits in a quantum algorithm with superconducting qubits, *Phys. Rev. Lett.* **127**, 100501 (2021).
 - [9] E. Bäumer, V. Tripathi, A. Seif, D. Lidar, and D. S. Wang, Quantum Fourier transform using dynamic circuits, *Phys. Rev. Lett.* **133**, 150602 (2024).
 - [10] E. Bäumer, D. Sutter, and S. Woerner, Approximate quantum Fourier transform in logarithmic depth on a line, *arXiv:2504.20832*.
 - [11] C. Piveteau and D. Sutter, Circuit knitting with classical communication, *IEEE Trans. Inf. Theory* **70**, 2734 (2024).
 - [12] L. Brenner, C. Piveteau, and D. Sutter, Optimal wire cutting with classical communication, *IEEE Trans. Inf. Theory* **71**, 7742 (2025).
 - [13] A. Carrera Vazquez, C. Tornow, D. Ristè, S. Woerner, M. Takita, and D. J. Egger, Combining quantum processors with real-time classical communication, *Nature* **636**, 75 (2024).
 - [14] K. Mitarai and K. Fujii, Constructing a virtual two-qubit gate by sampling single-qubit operations, *New J. Phys.* **23**, 023021 (2021).
 - [15] A. P. Singh, K. Mitarai, Y. Suzuki, K. Heya, Y. Tabuchi, K. Fujii, and Y. Nakamura, Experimental demonstration of a high-fidelity virtual two-qubit gate, *Phys. Rev. Res.* **6**, 013235 (2024).
 - [16] E. van den Berg, Z. K. Mineev, A. Kandala, and K. Temme, Probabilistic error cancellation with sparse Pauli–Lindblad models on noisy quantum processors, *Nat. Phys.* **19**, 1116 (2023).
 - [17] D. C. McKay, I. Hincks, E. J. Pritchett, M. Carroll, L. C. G. Govia, and S. T. Merkel, Benchmarking quantum processor performance at scale, *arXiv:2311.05933*.
 - [18] A. Erhard, J. J. Wallman, L. Postler, M. Meth, R. Stricker, E. A. Martinez, P. Schindler, T. Monz, J. Emerson, and R. Blatt, Characterizing large-scale quantum computers via cycle benchmarking, *Nat. Commun.* **10**, 5347 (2019).
 - [19] A. Carignan-Dugas, D. Dahlen, I. Hincks, E. Ospadov, S. J. Beale, S. Ferracin, J. Skanes-Norman, J. Emerson, and J. J. Wallman, The error reconstruction and compiled calibration of quantum computing cycles, *arXiv:2303.17714v1*.
 - [20] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Purification of noisy entanglement and faithful teleportation via noisy channels, *Phys. Rev. Lett.* **76**, 722 (1996).
 - [21] E. Knill, Fault-tolerant postselected quantum computation: Threshold analysis, *arXiv:quant-ph/0404104*.
 - [22] J. J. Wallman and J. Emerson, Noise tailoring for scalable quantum computation via randomized compiling, *Phys. Rev. A* **94**, 052325 (2016).
 - [23] E. Van Den Berg, Z. K. Mineev, and K. Temme, Model-free readout-error mitigation for quantum expectation values, *Phys. Rev. A* **105**, 032620 (2022).

- [24] J. Hines and T. Proctor, Pauli noise learning for mid-circuit measurements, *Phys. Rev. Lett.* **134**, 020602 (2025).
- [25] Z. Zhang, S. Chen, Y. Liu, and L. Jiang, Generalized cycle benchmarking algorithm for characterizing midcircuit measurements, *PRX Quantum* **6**, 010310 (2025).
- [26] S. Hazra, W. Dai, T. Connolly, P. D. Kurilovich, Z. Wang, L. Frunzio, and M. H. Devoret, Benchmarking the readout of a superconducting qubit for repeated measurements, *Phys. Rev. Lett.* **134**, 100601 (2025).
- [27] R. S. Gupta, E. van den Berg, M. Takita, D. Ristè, K. Temme, and A. Kandala, Probabilistic error cancellation for dynamic quantum circuits, *Phys. Rev. A* **109**, 062617 (2024).
- [28] J. M. Koh, D. E. Koh, and J. Thompson, Readout error mitigation for mid-circuit measurements and feedforward, *PRX Quantum* **7**, 10317 (2026).
- [29] A. Hashim, A. Carignan-Dugas, L. Chen, C. Jünger, N. Fruitwala, Y. Xu, G. Huang, J. J. Wallman, and I. Siddiqi, Quasiprobabilistic readout correction of midcircuit measurements for adaptive feedback via measurement randomized compiling, *PRX Quantum* **6**, 010307 (2025).
- [30] S. Chen, Y. Liu, M. Otten, A. Seif, B. Fefferman, and L. Jiang, The learnability of Pauli noise, *Nat. Commun.* **14**, 52 (2023).
- [31] H. Yu and T.-C. Wei, Efficient separate quantification of state preparation errors and measurement errors on quantum computers and their mitigation, *Quantum* **9**, 1724 (2025).
- [32] S. Chen, Z. Zhang, L. Jiang, and S. T. Flammia, Efficient self-consistent learning of gate set Pauli noise, *PRX Quantum* **7**, 10305 (2026).
- [33] E. H. Chen, S. Chen, L. E. Fischer, A. Eddins, L. C. G. Govia, B. Mitchell, A. He, Y. Kim, L. Jiang, and A. Seif, Disambiguating Pauli noise in quantum computers, *arXiv:2505.22629*.
- [34] W. J. Huggins, S. McArdle, T. E. O'Brien, J. Lee, N. C. Rubin, S. Boixo, K. B. Whaley, R. Babbush, and J. R. McClean, Virtual distillation for quantum error mitigation, *Phys. Rev. X* **11**, 41036 (2021).
- [35] B. Koczor, Exponential error suppression for near-term quantum devices, *Phys. Rev. X* **11**, 31057 (2021).
- [36] R. Laflamme, J. Lin, and T. Mor, Algorithmic cooling for resolving state preparation and measurement errors in quantum computing, *Phys. Rev. A* **106**, 12439 (2022).
- [37] J. P. Santos and R. Uzdin, Drift-resilient mid-circuit measurement and state preparation error mitigation for dynamic circuits, *arXiv:2506.11270*.
- [38] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, A quantum engineer's guide to superconducting qubits, *Appl. Phys. Rev.* **6**, 021318 (2019).
- [39] D. Ristè, C. C. Bultink, K. W. Lehnert, and L. DiCarlo, Feedback control of a solid-state qubit using high-fidelity projective measurement, *Phys. Rev. Lett.* **109**, 240502 (2012).
- [40] X. Y. Jin, A. Kamal, A. P. Sears, T. Gudmundsen, D. Hover, J. Miloshi, R. Slaterry, F. Yan, J. Yoder, T. P. Orlando, S. Gustavsson, and W. D. Oliver, Thermal and residual excited-state population in a 3D transmon qubit, *Phys. Rev. Lett.* **114**, 240501 (2015).
- [41] K. Geerlings, Z. Leghtas, I. M. Pop, S. Shankar, L. Frunzio, R. J. Schoelkopf, M. Mirrahimi, and M. H. Devoret, Demonstrating a driven reset protocol for a superconducting qubit, *Phys. Rev. Lett.* **110**, 120501 (2013).
- [42] The E and F in RabiEF come from an alternative transmon level naming, where the states $|0\rangle$, $|1\rangle$, and $|2\rangle$ are labeled $|g\rangle$, $|e\rangle$, and $|f\rangle$, respectively. Here, g and e stand for ground and excited, respectively.
- [43] L. Govia, S. Majumder, S. Barron, B. Mitchell, A. Seif, Y. Kim, C. Wood, E. Pritchett, S. Merkel, and D. McKay, Bounding the systematic error in quantum error mitigation due to model violation, *PRX Quantum* **6**, 010354 (2025).
- [44] S. J. Beale and J. J. Wallman, Randomized compiling for subsystem measurements, *arXiv:2304.06599*.
- [45] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, United Kingdom, 2010), 10th ed.
- [46] L. E. Fischer, D. Miller, F. Tacchino, P. K. Barkoutsos, D. J. Egger, and I. Tavernelli, Ancilla-free implementation of generalized measurements for qubits embedded in a qudit space, *Phys. Rev. Res.* **4**, 033027 (2022).
- [47] R. Bianchetti, S. Filipp, M. Baur, J. M. Fink, C. Lang, L. Steffen, M. Boissonneault, A. Blais, and A. Wallraff, Control and tomography of a three level superconducting artificial atom, *Phys. Rev. Lett.* **105**, 223601 (2010).
- [48] M. S. Blok, V. V. Ramasesh, T. Schuster, K. O'Brien, J. M. Kreikebaum, D. Dahlen, A. Morvan, B. Yoshida, N. Y. Yao, and I. Siddiqi, Quantum information scrambling on a superconducting qutrit processor, *Phys. Rev. X* **11**, 21010 (2021).
- [49] M. AbuGhanem, IBM Quantum computers: Evolution, performance, and future directions, *J. Supercomput.* **81**, 687 (2025).
- [50] C. J. Haupt and D. J. Egger, Leakage in restless quantum gate calibration, *Phys. Rev. A* **108**, 022614 (2023).
- [51] S. Brandhofer, I. Polian, and K. Krsulich, in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, Piscataway, NJ, 2023), Vol. 01, pp. 859–869.
- [52] L. Chen, H.-X. Li, Y. Lu, C. W. Warren, C. J. Krizan, S. Kosen, M. Rommel, S. Ahmed, A. Osman, J. Biznárová, A. Fadavi Roudsari, B. Lienhard, M. Caputo, K. Grigoras, L. Grönberg, J. Govenius, A. F. Kockum, P. Delsing, J. Bylander, and G. Tancredi, Transmon qubit readout fidelity at the threshold for quantum error correction without a quantum-limited amplifier, *npj Quantum Inf.* **9**, 1 (2023).
- [53] N. Kanazawa, H. Emori, and D. C. McKay, Qutrit state discrimination with mid-circuit measurements, *arXiv:2309.11303*.
- [54] This argument holds for an initial state with non-negligible $|2\rangle$ -state population that is the same for each shot. However, this is not the case in our simulations as the initial state changes from shot to shot. This manifests as large standard deviations in p_{sp} and $p_{sp}^{(2)}$ with fast qubit resets, see Table I. The large fluctuations in $|1\rangle$ - and $|2\rangle$ -state populations is the

- cause of distortions in the RabiEF signals, deviating from the form $a \sin^2(b\theta) + c$ entirely.
- [55] Qiskit Aer Contributors, Qiskit/qiskit-aer: QISKIT AER 0.16.0, 2025, <https://github.com/Qiskit/qiskit-aer>.
 - [56] F. Battistel, B. Varbanov, and B. Terhal, Hardware-efficient leakage-reduction scheme for quantum error correction with superconducting transmon qubits, *PRX Quantum* **2**, 030314 (2021).
 - [57] K. C. Miao *et al.*, Overcoming leakage in quantum error correction, *Nat. Phys.* **19**, 1780 (2023).
 - [58] M. Khezri, A. Opremcak, Z. Chen, K. C. Miao, M. McEwen, A. Bengtsson, T. White, O. Naaman, D. Sank, A. N. Korotkov, Y. Chen, and V. Smelyanskiy, Measurement-induced state transitions in a superconducting qubit: Within the rotating-wave approximation, *Phys. Rev. Appl.* **20**, 054008 (2023).
 - [59] R. Lescanne, L. Verney, Q. Ficheux, M. H. Devoret, B. Huard, M. Mirrahimi, and Z. Leghtas, Escape of a driven quantum Josephson circuit into unconfined states, *Phys. Rev. Appl.* **11**, 014030 (2019).
 - [60] Y. Fu, W. Liu, X. Ye, Y. Wang, C. Zhang, C.-K. Duan, X. Rong, and J. Du, Experimental investigation of quantum correlations in a two-qutrit spin system, *Phys. Rev. Lett.* **129**, 100501 (2022).
 - [61] Y. Guo, W. Ji, X. Kong, M. Wang, H. Sun, J. Zhou, Z. Chai, X. Rong, F. Shi, Y. Wang, and J. Du, Single-shot readout of a solid-state electron spin qutrit, *Phys. Rev. Lett.* **132**, 060601 (2024).
 - [62] A. B. Klimov, R. Guzmán, J. C. Retamal, and C. Saavedra, Qutrit quantum computer with trapped ions, *Phys. Rev. A* **67**, 062313 (2003).
 - [63] S. Chen, A. Hashim, N. Goss, A. Seif, I. Siddiqi, and L. Jiang, Enhancing quantum noise characterization via extra energy levels, [arXiv:2506.09131](https://arxiv.org/abs/2506.09131).
 - [64] M. Malekakhlagh, A. Seif, D. Puzzioli, L. C. G. Govia, and E. van den Berg, Efficient Lindblad synthesis for noise model construction, *npj Quantum Inf.* **11**, 191 (2025).
 - [65] A. Calzona, M. Papič, P. Figueroa-Romero, and A. Auer, Multi-layer cycle benchmarking for high-accuracy error characterization, [arXiv:2412.09332](https://arxiv.org/abs/2412.09332) [quant-ph].
 - [66] D. Greenbaum, Introduction to quantum gate set tomography, [arXiv:1509.02921](https://arxiv.org/abs/1509.02921) [quant-ph].
 - [67] A. Javadi-Abhari, M. Treinish, K. Krsulich, C. J. Wood, J. Lishman, J. Gacon, S. Martiel, P. D. Nation, L. S. Bishop, A. W. Cross, B. R. Johnson, and J. M. Gambetta, Quantum computing with QISKIT, [arXiv:2405.08810](https://arxiv.org/abs/2405.08810).
 - [68] D. Puzzioli, C. J. Wood, D. J. Egger, B. Rosand, and K. Ueda, QISKIT DYNAMICS: A Python package for simulating the time dynamics of quantum systems, *J. Open Source Softw.* **8**, 5853 (2023).
 - [69] A. Dane, K. Balakrishnan, B. Wacaser, L.-W. Hung, H. J. Mamin, D. Rugar, R. M. Shelby, C. Murray, K. Rodbell, and J. Sleight, Performance stabilization of high-coherence superconducting qubits, [arXiv:2503.12514](https://arxiv.org/abs/2503.12514).
 - [70] J. Heinsoo, C. K. Andersen, A. Remm, S. Krinner, T. Walter, Y. Salathé, S. Gasparinetti, J.-C. Besse, A. Potočnik, A. Wallraff, and C. Eichler, Rapid high-fidelity multiplexed readout of superconducting qubits, *Phys. Rev. Appl.* **10**, 34040 (2018).
 - [71] R. S. Burns, *Advanced Control Engineering* (Butterworth-Heinemann, Oxford, United Kingdom, 2001).