Modeling Low- and High-Frequency Noise in Transmon Qubits with Resource-Efficient Measurement

Vinay Tripathi,[1,2,*] Huo Chen,[2,3] Eli Levenson-Falk,[1,2] and Daniel A. Lidar[1,2,3,4]

[1] *Department of Physics & Astronomy, University of Southern California, Los Angeles, California 90089, USA*
[2] *Center for Quantum Information Science & Technology, University of Southern California, Los Angeles, California 90089, USA*
[3] *Department of Electrical & Computer Engineering, University of Southern California, Los Angeles, California 90089, USA*
[4] *Department of Chemistry, University of Southern California, Los Angeles, California 90089, USA*

(Received 16 March 2023; revised 26 October 2023; accepted 16 January 2024; published 7 February 2024)

Transmon qubits experience open-system effects that manifest as noise at a broad range of frequencies. We present a model of these effects using the Redfield master equation with a hybrid bath consisting of low- and high-frequency components. We use two-level fluctuators to simulate $1/f$-like noise behavior, which is a dominant source of decoherence for superconducting qubits. By measuring quantum state fidelity under free evolution with and without dynamical decoupling (DD), we can fit the low- and high-frequency noise parameters in our model. We train and test our model using experiments on quantum devices available through IBM quantum experience. Our model accurately predicts the fidelity decay of random initial states, including the effect of DD pulse sequences. We compare our model with two simpler models and confirm the importance of including both high frequency and $1/f$ noise in order to accurately predict transmon behavior.

DOI: 10.1103/PRXQuantum.5.010320

I. INTRODUCTION

Quantum computing based on superconducting qubits has made significant progress. Starting with the first implementations of superconducting qubits [1,2], the field has developed several flavors of qubits, broadly classified as charge, flux, and phase qubits [3]. However, the real workhorse behind many of the recent critical developments [4–13] in gate-based quantum computing is the transmon qubit [14]. Transmons are designed by adding a large shunting capacitor to charge qubits, the result being that they are almost insensitive to charge noise. Transmon-based cloud quantum computers (QCs) are now widely available to the broad research community for proof-of-principle quantum computing experiments [15–25].

Quantum computers in their current form have high error rates. This includes coherent errors (originating from imperfect gates), state preparation and measurement errors (SPAM), and incoherent errors (environment-induced noise) [26,27]. The latter, which results in dephasing and relaxation errors, is a pernicious problem in quantum information processing. Characterizing and modeling these open quantum system effects is crucial for advancing the field and improving the prospects of fault-tolerant quantum computation [28–30]. Various procedures for modeling decoherence and control noise affecting idealized qubits have been discussed [31–33]. Still, modeling noise effects from first principles, i.e., starting at the circuit level of transmons and including $1/f$ noise, is relatively unexplored [34,35].

In this work, we develop a framework to model environment-induced noise effects on a transmon qubit using the master-equation formalism. We use a hybrid quantum bath with an Ohmic-like noise spectrum to model dephasing and relaxation processes in multilevel transmons. We also include classical fluctuators and use a hybrid Redfield model [36] to account for both low- ($1/f$) and high-frequency noise. We develop a simple noise learning procedure relying on dynamical decoupling (DD) [37–41] to obtain the noise parameters (see Ref. [42] for early experimental work in this area). Our procedure relies only on measurements of quantum state fidelity with and without a single type of DD sequence, and so is quite resource-efficient compared to protocols requiring full quantum state tomography or DD-based spectral analysis.

*vinaytri@usc.edu

We test our noise model via fidelity preservation experiments on IBMQE processors [43] for random initial states and find that the model can correctly capture these experiments. The model is, moreover, capable of reproducing the effects of time-dependent dynamical decoupling pulses on the main qubit. Finally, we compare the predictions based on our model with two simpler models using ideal two-level qubits, excluding the fluctuators and assuming ideal, zero-width DD pulses. In contrast to our complete model, these simpler models fail to capture noise simultaneously in both the low- and high-frequency regimes. As a result, whether with or without DD, they underperform in capturing fidelity preservation experiments. Our model is tailored to transmon qubits; however, our approach is extremely general and can be adapted to any qubit experiencing the ubiquitous combinations of high frequency and $1/f$ noise.

This paper is organized as follows. In Sec. II, we develop our numerical method focusing on simulating multilevel transmon qubit and single-qubit gates, which form the DD sequences. Next, we discuss our open quantum system in Sec. III and describe our noise learning method in Sec. IV. We then test our learned model on Quito using DD experiments with random initial single-qubit states in Sec. V. We extend our method to Lima, which relies on a different calibration procedure, in Sec. VI. We finally extend our method to two qubits in Sec. VII and conclude in Sec. VIII. The Appendix provides additional details and calculations in support of the main text.

II. NUMERICAL MODEL OF TRANSMONS

In this section, we focus on the circuit-level description of the transmon qubit that we use in our model. We start with the transmon Hamiltonian and find an effective Hamiltonian to simulate single-qubit time-dependent microwave gates. We include the derivative removal of adiabatic gates (DRAG) [44] technique in our numerical model. By including DRAG—which is used in the IBMQE devices—and considering the residual errors it cannot suppress, we more accurately model the transmon behavior.

A. Transmon Hamiltonian

The Hamiltonian of a fixed-frequency transmon qubit is [14]:

$$H_{\text{trans}} = 4E_C\left(\hat{n} - n_g\right)^2 - E_J\cos\hat{\varphi}. \tag{1}$$

We work in units where $\hbar = 1$. $E_C = e^2/(2C)$ is the charging energy (C is the capacitance, and e is the electron charge), $E_J = I_C/(2e)$ is the potential energy of the Josephson junction (I_C is the critical current of the junction) and n_g represents the charge offset number, which can result in charge noise. In the operating regime of a transmon qubit, i.e., $E_J/E_C \gg 1$, the lowest few energy levels of the transmon are almost immune to charge noise, in which case n_g can be safely ignored. The two operators $\hat{n}$ and $\hat{\varphi}$ are a canonically conjugate pair analogous to momentum and position. They satisfy the commutation relation $\left[\hat{n}, \hat{\varphi}\right] = i$; $\hat{n}$ is the number operator for the Cooper pairs transferred between the superconducting islands of the Josephson junction, and $\hat{\varphi}$ is the gauge invariant phase difference across the Josephson junction, i.e., between the islands. Note that this commutation relation is exact only if we restrict $\hat{\varphi}$ to a single 2π range; this is a good approximation for the lower-energy states of transmons, which are almost completely confined to a single well.

B. Time-dependent drives

To numerically simulate the time-dependent drive pulses or gates, we start with Eq. (1) and write it in the charge basis (the eigenbasis of $\hat{n}$) such that the number of Cooper pairs takes values from $-n_{\max}$ to $n_{\max}$. Equation (1) thus reduces to

$$H_{\text{trans}} = 4E_C\sum_{-n_{\max}}^{n_{\max}} n^2|n\rangle\langle n| - \frac{E_J}{2}\sum_{-n_{\max}}^{n_{\max}}\left(|n\rangle\langle n+1| + |n+1\rangle\langle n|\right), \tag{2}$$

where we have taken $n_g = 0$. This is justified since we are in the transmon regime: we have $E_J/E_C \approx 47$, which, by virtue of the charge dispersion $\epsilon_{n_g} \propto e^{-\sqrt{8E_J/E_C}}$, means that the eigenenergies of the transmon Hamiltonian have essentially no dependence on n_g [14, Fig. 2(d)]. We truncate to $n_{\max}$ (later we set $n_{\max} = 50$) and diagonalize the resulting Hamiltonian:

$$H_{\text{trans}}^{\text{eigen}} = SH_{\text{trans}}S^\dagger = \sum_{k\geq 0}\omega_k|k\rangle\langle k|, \tag{3}$$

where ω_k for $k = 0, 1, \ldots$ represents the energy of the k'th level in the transmon eigenbasis, and S is the unitary similarity transformation. The eigenfrequencies are $\omega_{ij} \equiv \omega_i - \omega_j$. The *bare qubit frequency* is $\omega_q \equiv \omega_{10}$ and the *anharmonicity* is $\eta_q \equiv \omega_{10} - \omega_{21}$. Since ω_q and η_q are the two quantities accessible via experiments, we use these values to obtain the fitting parameters E_C and E_J in Eq. (1), which is the starting point of our transmon model [45].

Next, we add the coupling to the microwave drive, which couples to the transmon charge operator. The total system Hamiltonian can then be written as

$$H_{\text{sys}} = H_{\text{trans}}^{\text{eigen}} + \varepsilon(t)\cos\left(\omega_{\text{d}}t + \phi_{\text{d}}\right)\hat{n}, \tag{4}$$

where $\varepsilon(t)$ is the pulse envelope, ω_{d} is the drive frequency, and ϕ_{d} is the phase of the drive. We can simplify Eq. (4) by

first writing the charge operator in the transmon eigenbasis of Eq. (3), i.e., $\hat{n} = \sum_{k,k'} \langle k|\hat{n}|k'\rangle |k\rangle\langle k'|$ and considering the charge-coupling matrix elements. Only nearest-level couplings $\langle k|\hat{n}|k\pm 1\rangle$ are found to be non-negligible, allowing us to ignore all higher-order terms:

$$\hat{n} \approx \sum_{k\geq 0} \langle k|\hat{n}|k+1\rangle |k\rangle\langle k+1| + \text{h.c.} \tag{5}$$

Transforming into a frame rotating with the drive and employing the rotating-wave approximation (RWA), we obtain, for $\phi_d = 0$, the effective Hamiltonian

$$\tilde{H}_{\text{sys}} = \sum_{k\geq 0} (\omega_k - k\omega_d) |k\rangle\langle k| + \frac{\varepsilon(t)}{2} \sum_{k\geq 0} g_{k,k+1}(|k\rangle\langle k+1| + |k+1\rangle\langle k|) , \tag{6}$$

where $g_{k,k+1} \equiv \langle k|\hat{n}|k+1\rangle$. By tuning ϕ_d, we can implement a rotation about any axis in the (x,y) plane of the qubit subspace (after an additional projection). In particular, taking $\phi_d = 0$ or $\pi/2$ corresponds to a rotation about the x or y axis, respectively. Appendix A provides a derivation of Eq. (6) from Eq. (4).

The pulse envelope $\varepsilon(t)$ plays a vital role in the final implementation of the gate. Since we are interested mainly in applying π pulses, we choose

$$\int_0^{t_g} \varepsilon(t)dt = \pi , \tag{7}$$

where t_g is the pulse or gate duration. For our numerical simulations, we choose Gaussian-shaped pulses with envelopes given by

$$\varepsilon(t) = \varepsilon\left[G(t,t_g,\sigma) - G(0,t_g,\sigma)\right]\left(\Theta(t) - \Theta\left(t-t_g\right)\right) , \tag{8}$$

where

$$G(t,t_g,\sigma) = \exp\left(-\frac{\left(t-t_g/2\right)^2}{2\sigma^2}\right). \tag{9}$$

Here ε is the maximum drive amplitude during the pulse, $\Theta(t)$ is the step function, and σ is the standard deviation of the Gaussian pulse.

An essential aspect of gate design is that population should not leak to higher levels of the transmon, i.e., the drive pulses should be bandwidth limited (adiabatic). An accurate measure of these off-resonant excitations is the Fourier transform of the pulse envelope at the detuning frequencies [46,47]. For example, consider a Gaussian pulse with standard deviation σ. Its Fourier transform has a standard deviation proportional to $1/\sigma$, which means that the drive pulse applied at the qubit frequency ω_q has a frequency spread close to $1/\sigma$ about ω_q. If $1/\sigma$ is of the order of the anharmonicity of the transmon, the pulse spectral width will overlap with some of the higher-level transitions. Figure 1 shows the Gaussian pulse envelope and its Fourier transform, and illustrates how choosing a shorter gate time results in a larger frequency spread and vice versa. The use of DRAG pulses mitigates this leakage, as discussed further below.

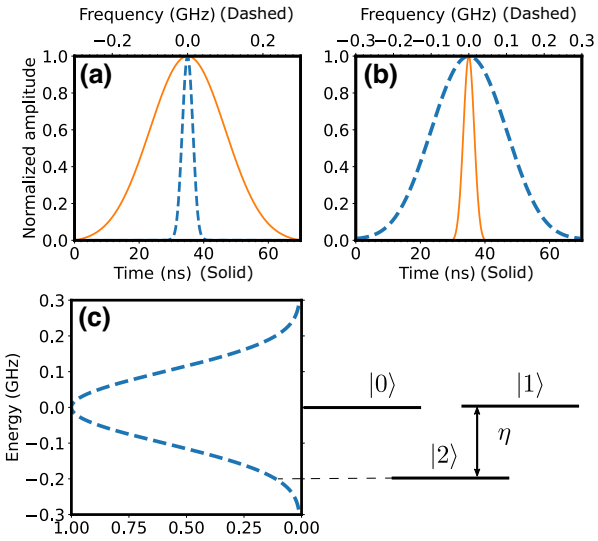

FIG. 1. Gaussian pulse envelope (solid, orange) [see Eq. (8)] and its Fourier transform (dashed, blue) with amplitude ε chosen to keep both in the range [0, 1]. (a),(b) show the pulse with gate time $t_g = 70$ and 10 ns, respectively, and $\sigma = t_g/6$. The bottom horizontal axis represents time in ns, and the top horizontal axis represents frequency in GHz. Shorter gate times result in a larger frequency spread of the spectrum, with associated larger leakage, as illustrated in (c), which shows the frequency spectrum corresponding to a $t_g = 10$ ns gate (left), compared to the energy levels (right) $|0\rangle$, $|1\rangle$, and $|2\rangle$ of the transmon. The energy levels are shown in the rotating frame such that $E_{|0\rangle} = E_{|1\rangle}$ and $E_{|2\rangle} = -\eta_q = -200$ MHz. As indicated by the dashed horizontal line, the spectrum overlaps with level $|2\rangle$, resulting in leakage into this level from the $\{|0\rangle, |1\rangle\}$ qubit subspace. The sampling frequency used to compute the Fourier transform is 10 GS/s, which is state of the art in experiments; the pulses that control the IBM processors used in this work have a sampling frequency of approximately 5 GS/s.

In the two-state (qubit) subspace, Eq. (6) reduces to

$$H_X(t) = \frac{\varepsilon(t)}{2}(|0\rangle\langle 1| + |1\rangle\langle 0|) = \frac{\varepsilon(t)}{2}\sigma^x, \tag{10}$$

where $g_{0,1} = g$ has been absorbed into $\varepsilon(t)$ and σ^x is the Pauli X matrix. When $H_X(t)$ is evolved for a time t_g such that Eq. (7) is satisfied, the resulting unitary is an ideal X_π gate. To include the effect of higher levels, we first use the full gate Hamiltonian from Eq. (6) and then project the result to the qubit subspace.

The gate fidelity averaged over all input states in the qubit Hilbert space can be written as the average over the six polar states (i.e., the six eigenstates of σ^x, σ^y, and σ^z) [44,48]:

$$F_g = \frac{1}{6} \sum_{j=\pm x,\pm y,\pm z} \mathrm{Tr}\left[U_{\mathrm{ideal}} \rho_j^{1\mathrm{q}} U_{\mathrm{ideal}}^\dagger \Pi[\rho(t_g)]\right], \quad (11)$$

where $\rho_j^{1\mathrm{q}}$ is the single qubit density matrix, and U_{ideal} represents the ideal unitary corresponding to the gate we wish to study. $\Pi[\rho(t_g)]$ is the projection of the full density matrix onto the single qubit subspace. F_g compares the density matrix ρ after application of the gate (i.e., at $t = t_g$) with the expected density matrix in the qubit subspace.

To reduce phase errors caused by the presence of additional levels, a commonly used trick to implement single-qubit gates such as X_π is to break the gate into two halves where each half performs a $\pi/2$ rotation, accompanied by some virtual Z rotations [49,50]. Numerically, we observe that with four levels included in the transmon Hamiltonian and a total gate duration $t_g = 70$ ns, with $\sigma = t_g/6$, the average single-qubit gate error $1 - F_g$ can be suppressed by around 20% if we use two such pulses instead of a single long pulse. The exact quantitative improvement depended on other model parameters. We observed this error reduction in a closed system setting with no environmental coupling, and so any fidelity improvement may be counteracted by open-system effects. For a detailed numerical study on the error of time-dependent gates with transmon qubits in the open-system settings, see Ref. [51].

We also include DRAG [44,52] in our single-qubit gates. DRAG is routinely used to improve the performance of single-qubit gates by suppressing leakage and phase errors. The former refers to a population exchange between the qubit subspace and the noncomputational levels at the end of a pulse, while the latter is a type of coherent error that results from a nonzero amplitude of the noncomputational levels *during* the pulse: the admixture of such levels leads to a phase shift of the computational levels, resulting in a net phase error at the end of the pulse. For more details see Appendix B.

III. OPEN QUANTUM SYSTEM SIMULATION

This section describes the noise model and discusses the hybrid Redfield equation used for the open quantum system simulations. For all simulations we truncate to the lowest four levels of the transmon qubit.

A. Interaction Hamiltonian

The single-qubit system bath interaction Hamiltonian in the lab frame can be written as

$$H_{\mathrm{SB}} = \sum_{i=x,y,z} g_i A_i \otimes B_i, \quad (12)$$

where A_i and B_i represent the dimensionless system and bath coupling operators, respectively, and the coupling strengths g_i have dimensions of energy. There are several contributions to decoherence and noise for a multilevel transmon circuit. With fixed-frequency architectures, charge noise and fluctuations in the critical current contribute most to decoherence. In the flux-tunable variants of transmon qubits, flux noise becomes an additional noise channel [14]. These considerations determine which coupling operators are needed to describe the noise model for a given architecture. We can define similar coupling operators for fluxonium qubits [53]. In the IBMQE processors used here, the transmons are of the fixed-frequency type. We choose the appropriate noise operators below.

We consider the following system-bath interaction Hamiltonian:

$$H_{\mathrm{SB}} = g_x A_x \otimes B_x + A_z \otimes \left(g_z B_z + \sum_k b_k \chi_k(t)\, I_{\mathrm{B}}\right), \quad (13)$$

where the coupling operators A_x and A_z correspond to the charge-coupling operator and to the Josephson energy operator and are defined as

$$A_x = \mathrm{c}_1 \hat{n}, \quad (14a)$$

$$A_z = \mathrm{c}_2 \cos\hat{\varphi}, \quad (14b)$$

where c_1 and c_2 are fixed constants that depend on the charge energy E_C and the Josephson energy E_J of the transmon qubit, respectively. We expect, based on the discussion in Sec. II B—and observe in our simulations—that A_x and A_z act like σ^x and σ^z when projected into the qubit subspace. We find numerically that Eq. (13) is an adequate model accounting for the nearly equal decay of the $|+\rangle$ and $|i\rangle$ states, which is why we do not include a separate σ^y coupling term. Note, however, that a (dependent) σ^y component appears when we transform Eq. (13) from the lab frame into a frame rotating with the drive.

Previous studies have found that noise in the superconducting circuit can be separated into high- and low-frequency components [54]. To account for this observation, we combine two noise models. We choose the bath operators B_x and B_z in Eq. (13) to be bosonic bath operators, which generally represent the high-frequency component of the noise. However, this is not always the case, as we argue in Sec. IV.

To account for the low-frequency noise component, which is a dominant noise source for superconducting qubits [55], we include a sum over classical fluctuators in Eq. (13), via the term proportional to the bath identity operator I_{B}. This semiclassical contribution, when parameterized properly, can simulate the behavior of $1/f$ noise. We model the fluctuators as having equal coupling

strengths, i.e., we set $b_k = b$ (with dimensions of energy) for $k = 1, \ldots, 10$. Each fluctuator can be characterized by a stochastic process $\chi_k(t)$ that switches between ± 1 with a frequency γ_k, which is log uniformly distributed between $\gamma_{\min}$ and $\gamma_{\max}$ [56]. We find ten fluctuators suffice to reproduce the $1/f$ spectrum ubiquitous in these devices.

B. Hybrid Redfield model

To simulate the reduced system dynamics of the interaction Hamiltonian in Eq. (13), we use a hybrid form of the Redfield (or TCL2) master equation [36,58]. We first define the standard bath correlation function

$$C_{ij}(t) = \mathrm{Tr}\{U_{\mathrm{B}}(t) B_i U_{\mathrm{B}}^{\dagger}(t) B_j \rho_{\mathrm{B}}\}, \tag{15}$$

where $U_{\mathrm{B}}(t) = e^{-iH_{\mathrm{B}}t}$ is the unitary evolution operator generated by the bath Hamiltonian H_{B}, and the reference state ρ_{B} is the Gibbs state of H_{B}:

$$\rho_{\mathrm{B}} = e^{-\beta H_{\mathrm{B}}}/\mathrm{Tr}\left(e^{-\beta H_{\mathrm{B}}}\right), \tag{16}$$

where $\beta = 1/T$ is the inverse temperature. Assuming the bath operators B_x and B_z are uncorrelated, i.e., $C_{xz}(t) = C_{zx}(t) = 0$, we construct the following hybrid Redfield equation:

$$\frac{\partial \rho_{\mathrm{S}}}{\partial t} = -i[H_{\mathrm{sys}} + A_z \sum_{k=1}^{10} b_k \chi_k(t), \rho_{\mathrm{S}}] + \mathcal{L}_{\mathrm{R}}(\rho_{\mathrm{S}}), \tag{17}$$

where $\mathcal{L}_{\mathrm{R}}$ is the Redfield Liouville superoperator

$$\mathcal{L}_{\mathrm{R}}(\rho_{\mathrm{S}}) = -\sum_{i=x,z} [A_i, \Lambda_i(t)\rho_{\mathrm{S}}(t)] + \mathrm{h.c.}, \tag{18}$$

and

$$\Lambda_i(t) = \int_0^t C_i(t-\tau) U_{\mathrm{sys}}(t,\tau) A_i U_{\mathrm{sys}}^{\dagger}(t,\tau) \mathrm{d}\tau, \tag{19}$$

where $C_j(\tau) \equiv C_{jj}(\tau)$ [from Eq. (15)] and $U_{\mathrm{sys}}(t)$ is the unitary evolution operator generated by the system Hamiltonian H_{sys}. The reduced system dynamics are obtained by averaging the solution of Eq. (17) over all the realizations of $\chi_k(t)$ for $k = 1, \ldots, 10$.

The bath component correlation functions $C_j(\tau)$ are the Fourier transforms of the bath component noise spectra

$$\gamma_j(\omega) = \int_{-\infty}^{\infty} C_j(\tau) e^{i\omega\tau} \mathrm{d}\tau \,. \tag{20}$$

We choose the bath to be Ohmic, which means that the component noise spectra have the form

$$\gamma_j(\omega) = 2\pi \eta g_j^2 \frac{\omega \mathrm{e}^{-|\omega|/\omega_j^c}}{1 - \mathrm{e}^{-\beta\omega}}, \tag{21}$$

where ω_j^c is the cut-off frequency for bath operator B_j, and η is a positive constant with dimensions of time squared arising in the specification of the Ohmic spectral function.

Lastly, the hybrid Redfield Eq. (17) can be transformed into a frame rotating with the drive frequency ω_{d} (see Appendix A for details) by replacing every operator with the interaction-picture one [specifically, the A_i operator in Eq. (19) needs to be replaced by $A_i(\tau)$]. We simulate the Redfield master equation in this rotating frame in the methodology and results we discuss next.

IV. METHODOLOGY AND FITTING RESULTS

This section discusses our methodology for modeling a transmon qubit's open quantum system behavior in a multiqubit processor. We refer to the qubit of interest as the main qubit and all the others as spectator qubits. The goal is to extract the bath parameters in our open quantum system model and then use this model to predict the outcomes of experiments on the main qubit, including dynamical decoupling sequences. We treat qubit 1 (Q1) of the Quito processor as our main qubit. We are interested only in the main qubit's behavior here; hence, we measure only the main qubit. Appendix D describes the procedure to extract and analyze the experimental data.

A. Free and DD evolution experiments

Our procedure involves two types of experiments, as shown in Fig. 2. The first type, which we call a *free-evolution* experiment, consists of initializing all the qubits in a given state by applying a particular unitary operation $U3(\theta, \phi, \lambda)$ [59] (denoted as U in Fig. 2) to each of the qubits and is given by

$$U3(\theta,\phi,\lambda) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -e^{i\lambda}\sin\left(\frac{\theta}{2}\right) \\ e^{i\phi}\sin\left(\frac{\theta}{2}\right) & e^{i(\phi+\lambda)}\cos\left(\frac{\theta}{2}\right) \end{pmatrix}. \tag{22}$$

We then apply a sequence of identity gates on the main qubit, which we vary in number. Simultaneously we also apply the $XY4$ DD sequence to all the other (spectator) qubits (i.e., $Xf_\tau Yf_\tau Xf_\tau Yf_\tau$, where f_τ denotes free evolution in the absence of pulses for a duration of τ [41]) for the same total duration as that of the identity gates on the main qubit. As shown in Ref. [57], DD sequences applied to spectator qubits suppress unwanted ZZ interactions, i.e., ZZ crosstalk between the main qubit and the spectator qubits. Without crosstalk suppression, we observe oscillations in the probability decay as a function of time [20,31]; see Appendix E. With the crosstalk suppression scheme, i.e., DD applied to the spectator qubits, these oscillations disappear, and the main qubit is now primarily affected only by environment-induced noise. Finally, we apply the inverse of $U3(\theta, \phi, \lambda)$ and measure in the

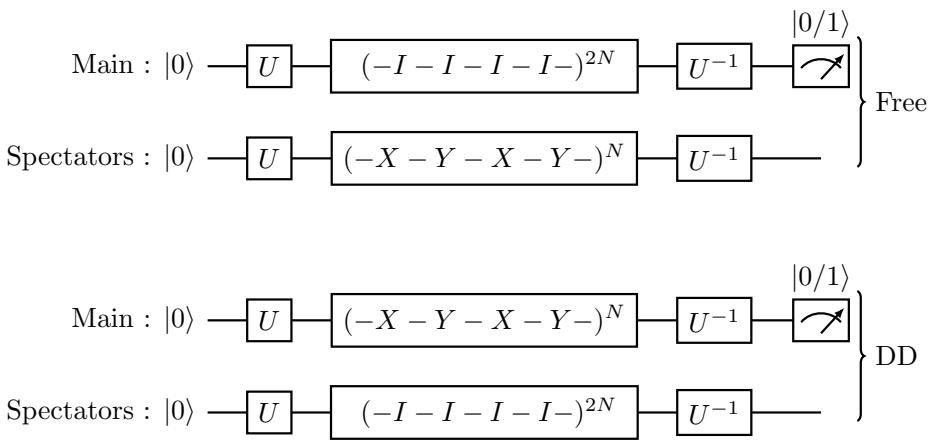

FIG. 2. The circuit schematics for the free-evolution and DD-evolution types of experiments. For the free-evolution case, we apply N cycles of the $XY4$ DD sequence on all the spectator qubits and $2N$ cycles of the I_4 sequence (here I_4 means four identity gates) on the main qubit, which suppresses crosstalk errors [57]. Note that an X or Y gate is twice as long as an identity gate on the IBM cloud quantum devices, hence the extra factor of 2. For the DD-evolution case, we apply the DD sequence only to the main qubit and apply identity gates to all the spectator qubits. This suppresses both crosstalk and environment-induced noise. We measure only the main qubit.

Z basis. The result is how we compute the initial states' decay probability.

Everything remains the same in the second type of experiment, which we call *DD evolution*, except that we now apply the $XY4$ sequence to the main qubit and identity gates on the spectator qubits. As discussed in Ref. [57], when we apply the $XY4$ sequence only to the main qubit, we suppress the ZZ crosstalk between the latter and all the spectator qubits and also decouple unwanted interactions between the main qubit and the environment. We note that, in contrast to experiments using DD to perform noise spectroscopy, here we use only a single type of DD sequence and do not vary any of its parameters.

B. Fitting procedure

We perform the free-evolution and DD-evolution experiments for the six Pauli states as initial states, i.e., we choose $U3(\theta,\phi,\lambda)$ to prepare $|0\rangle$, $|1\rangle$, $|+\rangle$, $|-\rangle$, $|i\rangle$, and $|-i\rangle$, and use the hybrid Redfield equation described in Sec. III B to simulate the dynamics of these experiments. We sweep over different values of the bath parameters and obtain the simulated probability decay as a function of time. To identify the simulation parameters that optimally match the experimental results, we define a cost function C for a given initial state $|\psi\rangle$ as the l_2-norm distance between the experimental probabilities $P^{\text{expt}}_{|\psi\rangle,s}(t_i)$ and the simulation probabilities $P^{\text{sim}}_{|\psi\rangle,s}(t_i)$ for every instant t_i:

$$C_{|\psi\rangle,s} = \frac{1}{N}\sqrt{\sum_{i=0}^{N-1}\left(P^{\text{sim}}_{|\psi\rangle,s}(t_i) - P^{\text{expt}}_{|\psi\rangle,s}(t_i)\right)^2}, \tag{23}$$

where $s \in \{\text{free}, \text{DD}\}$ and N is the total number of instants. Note that we compensate for state preparation and measurement (SPAM) errors by shifting the experimental results such that in all cases, $P^{\text{expt}}_{|\psi\rangle,s}(0) = 1$.

We limit the number of free parameters requiring fitting to six: the coupling strengths g_x, g_z, and $b_k \equiv b$ [Eq. (13)], and the cut-off frequencies $\gamma_{\max}$ (for $1/f$ noise), ω_x^c, and ω_z^c [Eq. (21)]. We set the bath temperature $T = 20$ mK (approximately the fridge temperature), $\gamma_{\min} = 10^{-4}$ GHz, and $\eta = 10^{-4}$ GHz^{-2}; these values are the same as in our previous work [57], which showed strong agreement between open-system simulations and experiments using other IBMQE devices, and remain fixed throughout our fitting procedure. This procedure consists of three steps, which we detail next.

1. Step I: free evolution for $|1\rangle$

We first focus on the free-evolution experiment for the initial state $|1\rangle$, the first excited state in the transmon eigenbasis. Since the free evolution for this state is only affected by charge noise, i.e., noise along the x axis, the only contribution to the decay of $|1\rangle$ should come from the $g_x A_x \otimes B_x$ term in Eq. (13). Thus, we consider only this term in our numerical simulations for this initial state. For a given set of values of the coupling strength g_x and bath cut-off frequency ω_x^c, we compute the cost function $C_{|1\rangle,\text{free}}$ using Eq. (23), and obtain the contour plot shown in Fig. 3(a).

In our simulations, we vary $\omega_x^c/(2\pi)$ from 0.5 to 3 GHz and g_x from 0 to 10^{-2} GHz, each with 20 equidistant points so that the contour plot has a total of 400 data points. We take the position of the global minimum of the cost function on this grid as the optimal set of bath parameter values. To reduce the resulting discretization uncertainty,

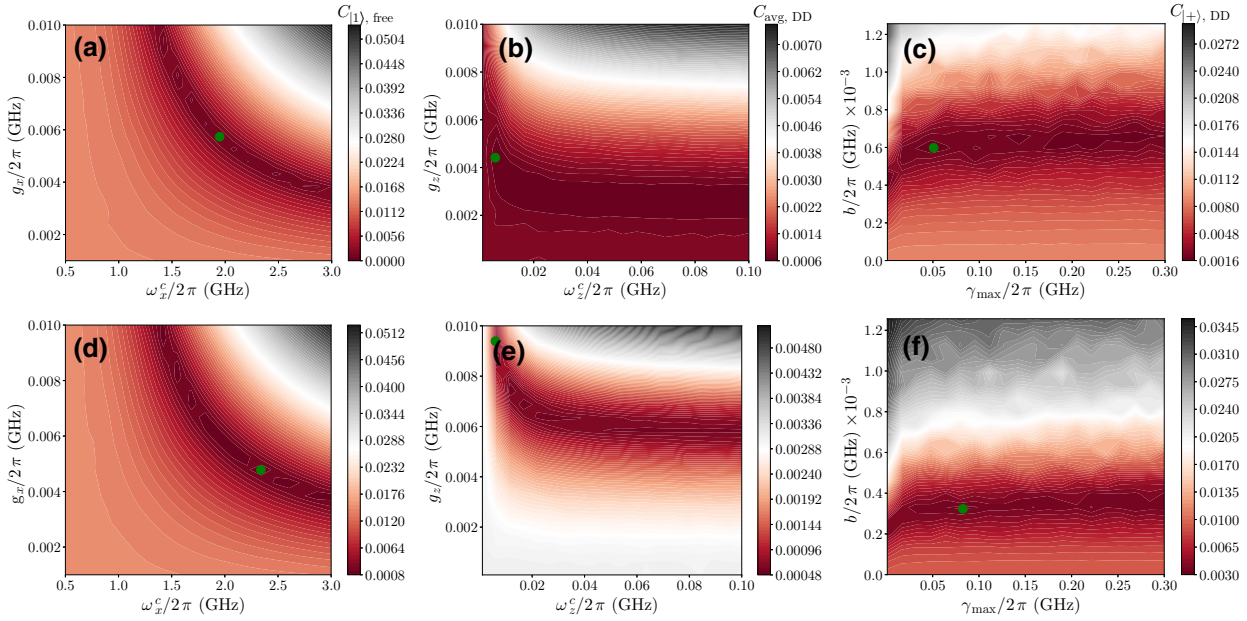

FIG. 3. Top: results for the Quito processor. Bottom: results for the Lima processor. Left: the cost function defined as the l_2-norm distance between the experimental and simulation results [Eq. (23)], averaged over $N = 70$ time instants, as a function of the bath parameters ω_x^c and g_x for free evolution of the $|1\rangle$ initial state. Middle: the average of the cost function over the six Pauli states for DD evolution as a function of ω_z^c and g_z. Right: the cost function for free evolution of the $|+\rangle$ initial state, as a function of $\gamma_{\max}$ and b. The green circles indicate the positions of the global minima in all the panels.

we interpolate the contour plot and use the Nelder-Mead optimization method to locate the minima. We numerically find the global minimum at $\omega_x^c/(2\pi) = 1.948$ GHz and $g_x/(2\pi) = 0.573 \times 10^{-2}$ GHz, denoted by the green circle in Fig. 3(a). With this, we have two out of the six bath parameters, and we use these learned parameters in the subsequent steps.

2. Step II: DD evolution for all six Pauli states

The second step involves the DD-evolution experiment for all six Pauli states. This requires including the term $g_z A_z \otimes B_z$ in Eq. (13), along with the first term whose bath parameters we already obtained. We do not include the semiclassical term in Eq. (13) consisting of fluctuators since it is expected to be strongly suppressed when DD is applied to the main qubit. We simulate time-dependent gates with DRAG corrections to model the DD pulses, as discussed in Sec. II. We are again left with just two bath parameters to optimize: ω_z^c and g_z. Figure 3(b) shows the average of the cost function [Eq. (23)] over the six Pauli states. The global minimum is found at $\omega_z^c/(2\pi) = 0.569 \times 10^{-2}$ GHz and $g_z/(2\pi) = 0.441 \times 10^{-2}$ GHz.

3. Step III: free evolution for $|+\rangle$

The final step requires optimizing the two remaining free parameters associated with the fluctuators: $\gamma_{\max}$ and b. Here we focus on the free-evolution experiment for initial state $|+\rangle$. We now employ the full system-bath Hamiltonian in Eq. (13) with the optimal parameters found in steps I and II. Figure 3(c) shows the contour plot for the cost function [Eq. (23)], where, as in step I, we again use 20 different values of $\gamma_{\max}$ and b each. The global minimum is found at $\gamma_{\max}/(2\pi) = 0.051$ GHz and $b/(2\pi) = 0.598 \times 10^{-3}$ GHz.

C. Methodology wrap up

Let us briefly summarize our methodology and add a few technical details. As explained above, we extract the bath parameters by performing free-evolution experiments for two initial states ($|1\rangle$ and $|+\rangle$) and DD-evolution experiments for up to six initial states (the Pauli states). Our optimization procedure is iterative and is thus not guaranteed to yield the globally optimal values of all the bath parameters, but this is by design: we choose initial states that allow us to isolate the bath parameters one pair at a time, which renders the optimization problem tractable.

This methodology is quite general and can be used to characterize all the transmon qubits on a given transmon processor or, much more broadly, to characterize single qubits on any quantum information processing platform capable of supporting individual qubit gates and measurements, provided a sufficiently accurate and descriptive model of the qubits and the system-bath interaction is available. Our procedure inherently suppresses the effects of crosstalk due to the neighboring qubits via DD applied either to the spectator qubits (free-evolution experiments) or the main qubit (DD-evolution experiment), which

reduces the number of free parameters of the noise model by eliminating the need to model crosstalk.

To obtain the contour plots shown in Fig. 3, we solve the Redfield master equation (Sec. III B) for each point (i.e., each set of model parameters), requiring a total of 400 simulation runs for each optimization. In the final step, including classical fluctuators to obtain Fig. 3(c), we use the trajectory version of the Redfield model introduced in Sec. III B to simulate a total of 600 trajectories at each point. This is large enough to yield negligible error bars ($< 2 \times 10^{-2}$). The experimental results are obtained using the standard bootstrap method (see Appendix D). In defining the cost function [Eq. (23)], we use the mean value of the experimental fidelity obtained after bootstrapping (see Appendix E) and ignore the associated tiny error bars ($\leq 6 \times 10^{-3}$). These error bars are much smaller than the error induced by the discrete nature of our 40×40 parameter grid, and so we can safely ignore them. We confirmed that varying the probabilities to the extremes of the error bars does not affect the values of the bath parameters we have extracted to the least significant digit we report. Table I summarizes the extracted values and the parameters we have fixed.

The accuracy of our results depends on the number of points in the contour plots in Fig. 3 (we used a 20×20 grid for each panel). Even though we interpolate the otherwise discrete contour plots and find the minima over the resulting smooth surface, the limited number of points affects the precision of the learned bath parameters. Increasing this precision requires more sophisticated optimization techniques to speed up the process of obtaining the bath parameters. This becomes especially acute when extending the model to learning a multiqubit system-bath Hamiltonian with correlated noise, as in this case, the number of bath parameters increases significantly. Here, we aim to demonstrate the model and methodology and illustrate both via the example of a single transmon qubit, and so we perform a simple brute-force search of the parameter space. Note that our methods for extracting the bath parameters also work with density matrices (from state tomography) instead of just probabilities. In that case, the l_2-norm distance in the cost function of Eq. (23) can be replaced by the trace-norm distance between the density matrices obtained from the simulation and the experiment. However, quantum state tomography imposes a much higher cost in terms of the number of required experiments and is thus less practical to scale up with a larger number of qubits. Our protocol requires only fidelity measurements and so is more resource efficient.

V. MODEL PREDICTION RESULTS

A. Full model

We now test our model for different initial states of the main qubit of the Quito processor. Since we always apply the DD sequence to the spectator qubits during the free-evolution experiments, the initial states of the latter do not matter due to the suppressed ZZ coupling. This section considers a total of 16 initial states, consisting of the six Pauli states and ten Haar-random states. We model the experimental results using the bath parameters we extracted in the previous section (Table I). This serves as a stringent test of the model: we now use the previously fitted model to predict the outcome of experiments not included in steps I–III of Sec. IV B, i.e., the results with different initial states. The data for all the experiments (both fitting and testing) was obtained in one batch. We used only the data for the six Pauli initial states needed for steps I–III to perform the fitting. We used the data for all 16 initial states in the testing phase.

We consider the same two kinds of experiments: free evolution and DD evolution. Figure 4 (top row) shows our model's prediction accuracy for the ten random and six Pauli states. The top left panel [Fig. 4(a)] corresponds to the free-evolution case, where we present the relative error in the prediction of our model as a function of time compared with the experimental results. The relative error is defined as $|\text{mean}_{\text{expt}} - \text{mean}_{\text{sim}}|/\text{mean}_{\text{expt}}$, where $\text{mean}_{\text{expt}}$ is the bootstrapped average over 8192 experimental repetition and mean_{sim} is the average over 600 trajectory simulations of the hybrid Redfield model for any given time instant. The box plot contains the spread in the relative error over all 16 states, showing that the relative error of the model is always below 8% over the total time considered here. The median and the mean over the 16 states are confined well below 3% for every instant.

The performance of our model for the DD-evolution experiments is shown in the top right panel [Fig. 4(b)]. Here the relative error is always below 2%. The median and the mean are below 1%. The closer agreement of the model with the DD-evolution experiments is expected, given that in contrast to the free-evolution experiments, DD suppresses the low-frequency noise affecting the main qubit, and the limitations of our fluctuator model of this noise is a likely source of modeling error.

The first and fourth columns of Fig. 5 show the relative error of our model over the 16 states and all 70 instants of the free and DD-evolution experiments, respectively. The results of the latter are better, as expected from Fig. 4. In both cases, however, we observe that the model has a relative prediction error of just a few percent.

B. Simplified models

As discussed in Sec. II, our numerical simulations use the circuit model Hamiltonian of a transmon qubit truncated to the four lowest transmon eigenstates. The gates are applied with time-dependent pulses of nonzero duration. To test the robustness of our detailed model and learning procedure, we compare it with two simpler models,

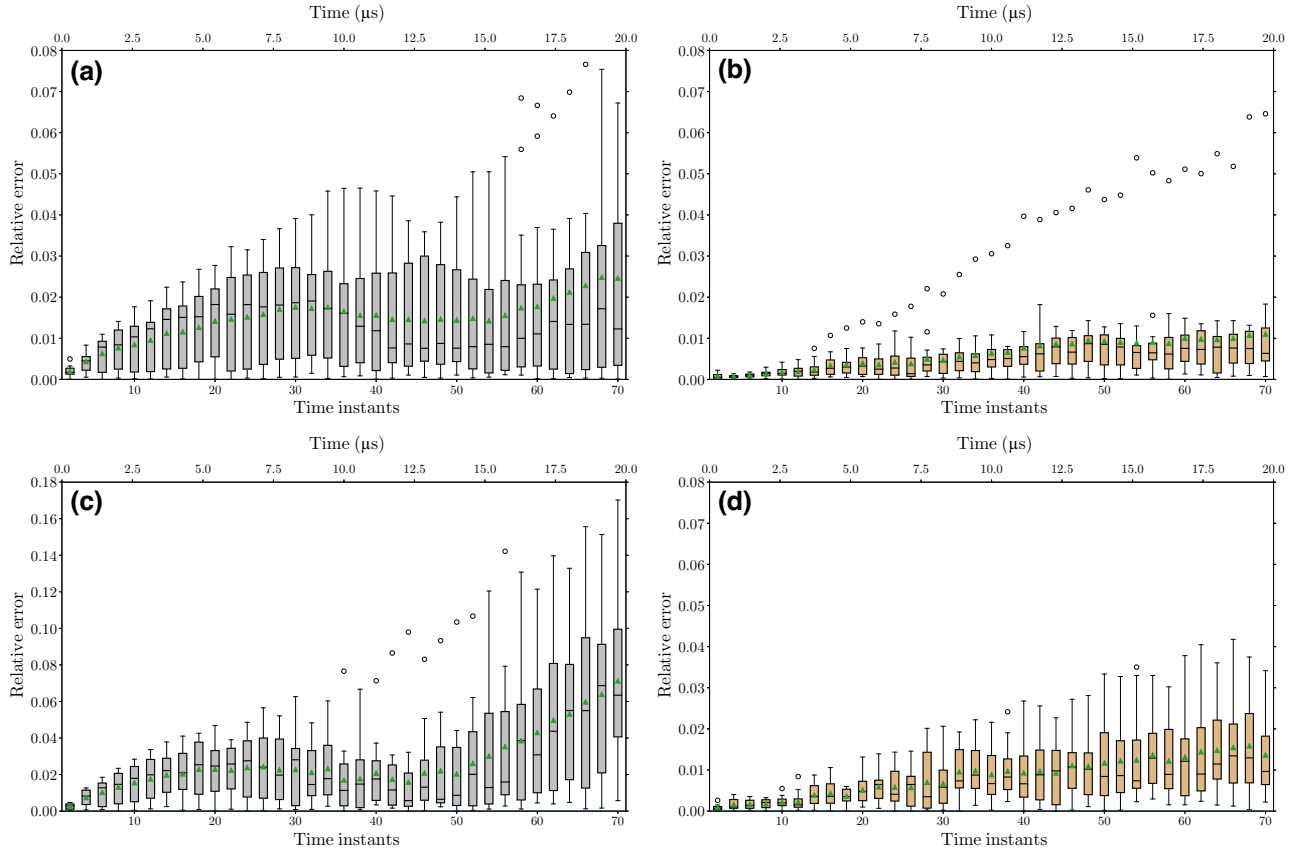

FIG. 4. Results for the Quito (top row) and Lima (bottom row) processors. Left: box plots showing the relative error of our model for the free-evolution experiments as a function of time for 16 different initial states containing six Pauli states and ten Haar-random states. Right: the same as on the left, but for the DD-evolution experiments. We measured a total of 70 time instants, up to a total evolution time of 19.6 μs, but display only every other instant to avoid overcrowding. Green triangles indicate the mean over the 16 initial states, black horizontal lines are the median, gray boxes represent the [25, 75] percentiles, the whiskers (black lines extending outside the boxes) represent the [0, 25] and [75, 100] percentiles, and circles are outliers.

SM1 and SM2, derived from our detailed model. The simpler models use a more straightforward qubit description where we truncate the transmon Hamiltonian to only two levels. The time-dependent gates are replaced with instantaneous (zero-duration), ideal gates. Moreover, we focus only on the Ohmic bath terms in Eq. (13), thus simplifying the noise model by removing the classical fluctuators. To test these simpler models' predictive power, we follow the same procedure as in Sec. IV, but using only steps I and II. The difference between SM1 and SM2 lies in step II, where SM1 uses the DD-evolution experiments for the six Pauli states, whereas SM2 uses the free-evolution experiments for the same states. In both cases, we extract the model parameters and then use the resulting learned models to predict the outcomes of both the free-evolution and the DD-evolution experiment.

Figure 5 shows the comparison between our detailed model and the simpler models SM1 and SM2. We observe that SM1 has the largest relative error for the free-evolution experiments, whereas SM2 has the largest relative error for the DD-evolution case. Our full model has the smallest relative error among the three models considered here for *both* the free and DD-evolution experiments. However, the performance of SM1 and SM2 is essentially indistinguishable from the full model results in the DD- and free-evolution cases, respectively. This is not unexpected, given that SM1 (SM2) is trained on the DD (free) evolution experiments and predicts these well. In other words, SM1 (SM2) captures the high- (low-) frequency noise well, as expected since for SM1, the use of DD suppresses most of the low-frequency noise, while for SM2, the use of free evolution means that the low-frequency noise remains a dominant source of decoherence. The added value of the detailed model and the use of step III is that this provides enough information to capture both the low- and high-frequency components of the noise, which yields a more complete model with better predictive power. We do note that taking the qubit approximation and treating DD pulses as instantaneous does not seem to appreciably worsen the performance of the simple models

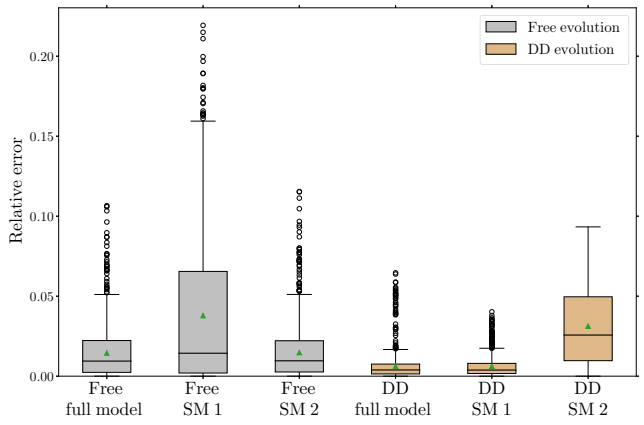

FIG. 5. Relative error results for the Quito processor. We display a comparison of the relative errors between the full model (which uses a three-step learning procedure and consists of four energy levels per transmon and realistic pulses) with the simplified models SM1 and SM2 (which are based on a two-step learning procedure and use just two energy levels and instantaneous pulses) for free-evolution and DD-evolution experiments. SM1 (SM2) is trained on the DD- (free) evolution experiments. Each box contains a total of 16 initial states and all 70 time instants varying from 0 to 19.6 μs.

in their regime of accuracy, as SM1 (SM2) are roughly as accurate as the full model in the DD- (free) evolution case. This suggests that an intermediate model, taking the qubit and instantaneous-pulse approximations but retaining the fluctuators, may be accurate and computationally efficient.

It is reasonable to ask whether our detailed simulation models are necessary, in the sense that perhaps a simple Markovian model might suffice. We address this in Appendix C, where we discuss results from using the Lindblad master equation with the backend-reported T_1 and T_2 values for the Quito processor. We simulate the free and DD evolutions for all 16 initial states and compute the relative error with respect to the experimental results. The Lindblad results (see Fig. 7) have significantly larger relative error than our detailed simulation model [Figs. 4(a), 4(b)] for the Quito processor, highlighting the need for the latter methodology.

TABLE I. System-bath parameter values extracted using the fitting procedure of Sec. IV B, and corresponding to the minima indicated by the green circles in Fig. 3 for Quito (top row) and Lima (bottom row).

Params/(2π)	Quito	Lima
g_x [MHz]	5.734	4.782
g_z [MHz]	4.413	9.393
ω_x^c [GHz]	1.948	2.340
ω_z^c [MHz]	5.690	5.979
b [MHz]	0.598	0.323
$\gamma_{\max}$ [GHz]	0.051	0.083

VI. CALIBRATION-INDEPENDENT LEARNING

For multiqubit superconducting processors, calibrating single-qubit drive frequencies is crucial for gate operations. In the presence of ZZ coupling, the state of the spectator qubits modifies the eigenfrequency of the main qubit. This results in different choices of calibration frequencies depending on the spectator qubits' state [57]. So far, we have focused on one particular device (Quito), which is calibrated while keeping the spectator qubits in the $|+\rangle$ state (see Appendix E and Ref. [57]). The all $|+\rangle$ or all $|0\rangle$ are usually the two preferred choices for the spectators' state while calibrating a given qubit in a multiqubit processor. When we perform state protection experiments on the main qubit initialized in the $|+\rangle$ state while keeping all the spectator qubits in $|0\rangle$, there exists a frequency mismatch, which results in ZZ-crosstalk oscillations (see Appendix E); to remove these oscillations, we applied DD to the spectator qubits before starting our noise-learning procedure. When device calibration is performed while keeping the spectators in the $|0\rangle$ state, a similar state-protection experiment does not result in any oscillations, as evidenced by our Lima results (see Appendix E).

Therefore, we extend our noise-learning method to Lima in this section. Following our procedure from Sec. IV, we again perform free-evolution (no DD is applied to any qubit) and DD-evolution ($XY4$ is applied just to the main qubit) experiments. The only difference from the Quito case is that, for the reasons explained above, the free-evolution experiment does not require the application of DD to the spectator qubits to suppress crosstalk oscillations. Figure 3 (bottom row) shows the contour plots for each of the three steps involved in our learning methodology as described in Sec. IV B. We find the global minima at the parameter values given in Table I. Comparing the Quito and Lima parameters in Table I, we observe that the coupling strength g_z of the Ohmic bath along the z axis is roughly double for Lima, whereas the strength of the fluctuators is roughly double for Quito. This indicates that Quito is more prone to low-frequency ($1/f$) noise.

Figure 4 (bottom row) shows the Lima prediction results for the 16 different initial states described above, using the learned noise parameters. The bottom left panel [Fig. 4(c)] shows the relative error in the prediction of our model of the free-evolution experiments as a function of time compared to the experimental results for all 16 states. The bottom-right panel [Fig. 4(d)] shows the same for the DD-evolution experiments. The relative error is always below 17% and 4% for free evolution and DD evolution, respectively. Similar to the Quito case, the relative error is significantly lower for the DD-evolution experiments. For longer evolution times ($\gtrsim 14$ μs), the agreement worsens for the free-evolution experiments. As for the Quito results, the closer agreement of the model with the DD-evolution experiments is likely because DD suppresses

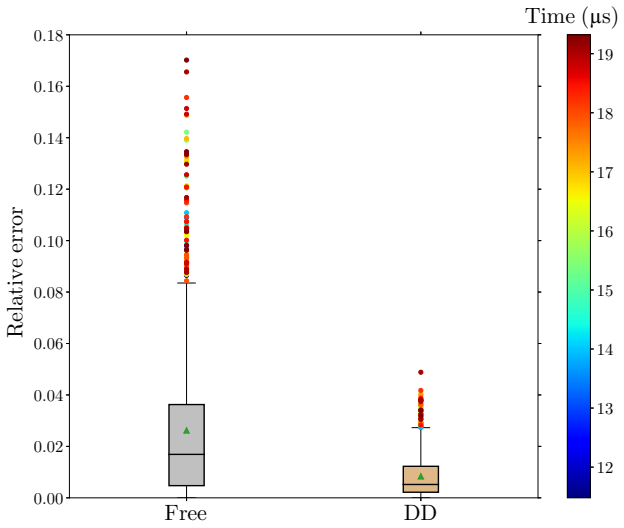

FIG. 6. Integrated relative error results for the Lima processor. We compare the relative errors between the free-evolution and DD-evolution experiments. Each box contains 16 initial states and all 70 time instants varying from 0 to 19.6 μs. The color bar indicates the time evolved for the outliers. All the outliers correspond to times longer than 11 μs.

the low-frequency noise affecting the main qubit, which dominates the free-evolution case's simulation error.

Figure 6 shows the time-integrated version of the Lima results of Figs. 4(c), 4(d), where we have combined all 16 states and 70 time instants into one box each for the free-evolution and DD-evolution experiments. Except for a few outliers, almost all the data points for the free-evolution case have relative errors below 9%. The outliers all correspond to evolution times longer than 11 μs, as indicated by the color bar. Similarly, for DD evolution, all data points have relative errors below 3%, except for a few outliers. There are two main reasons for the larger errors at longer evolution times. First, the Redfield equation is based on the weak coupling approximation, and its accuracy degrades as we increase the simulation time (for rigorous error bounds, see Ref. [60]). Second, as time increases, the effect of distant fluctuators is increasingly felt, thus reducing the accuracy of our fluctuator model, which relies on a fixed number of fluctuators. Adding more fluctuators and exploring a distribution of fluctuator strengths, as opposed to our assumption of a fixed fluctuator strength b, is expected to improve the predictive power of our model.

Finally, as another test of our learning procedure, we computed T_1 from the learned models, using the values we report in Table II. We did this by simulating the fidelity decay of the $|1\rangle$ state for 19.6 μs and fitting $\exp(-t/T_1)$ to estimate T_1. We find $T_1 = 92.5$ μs for Quito and 86.5 μs for Lima, compared to the reported $T_1 = 98.6$ μs and 105 μs, respectively. Our result gives the correct order of magnitude and is particularly reasonable for Quito. The discrepancy may be in part due to the relatively short simulation time of 19.6 μs (larger times become prohibitively expensive). In addition, the discrepancy for Quito is smaller because we use three rounds of experiments within the same calibration to model the noise, which removes short-time fluctuations via bootstrap averaging, whereas for Lima, we used only one repetition. As T_1 often drifts significantly over hour-long time scales, we would not expect the Lima prediction to line up exactly with the reported T_1.

VII. EXTENSION TO TWO QUBITS

Building upon the preceding analysis and validation of our model for characterizing the interaction of individual qubits with a bath, we now extend our methodology to the two-qubit case. For simpler models relying on Lindblad equations based on measured T_1 and T_2 data, noise modeling for two-qubit systems has been studied in Ref. [61]. Learning higher-order couplings, i.e., terms involving three or more qubits coupled to the bath, is unnecessary unless a circuit directly involves such higher-order interactions [62].

For a system comprising two qubits, we can write the system-bath interaction as $H_{\mathrm{SB}}^{\mathrm{tot}} = H_{\mathrm{SB}}^{1} + H_{\mathrm{SB}}^{2} + H_{\mathrm{SB}}^{2\mathrm{q}}$, where H_{SB}^{i} represents the single-qubit system-bath interaction for the ith qubit as described by Eq. (13). The two-qubit system-bath interaction term can be expressed as

$$H_{\mathrm{SB}}^{2\mathrm{q}} = \sum_{\alpha,\beta\in\{x,z\}} g_{\alpha\beta} A_\alpha \otimes A_\beta \otimes B_{\alpha\beta} \tag{24a}$$

$$+ A_z \otimes A_z \otimes \sum_k b_{k,12}\chi_k(t) I_{\mathrm{B}} \,. \tag{24b}$$

Recall that A_x and A_z correspond to σ^x and σ^z, respectively, when reduced to two-level systems. As in the single-qubit case, our model assumes a combination of an Ohmic bath [Eq. (24a)] and a fluctuator bath [Eq. (24b)].

As with the single-qubit noise learning method, we can conduct both free-evolution and DD-evolution experiments, this time with selected two-qubit initial states. The free-evolution experiments still necessitate the application of DD sequences to spectator qubits, in case one or both of the primary qubits are coupled to spectators. This is done to suppress ZZ crosstalk [57]. For the DD-evolution experiments on the primary qubits, we utilize the ZZ-crosstalk robust variant of the $XY4$ sequences, represented as $XIf_\tau IXf_\tau YIf_\tau IYf_\tau XIf_\tau IXf_\tau YIf_\tau IYf_\tau$ [63], where f_τ denotes free evolution in the absence of pulses for a duration of τ. This sequence aims to suppress the ZZ crosstalk between the main qubits as well as between the main and spectator qubits. Additionally, it suppresses the low-frequency noise resulting from the fluctuator term in the two-qubit system-bath interaction Hamiltonian, as described by Eq. (24b). It is worth noting that the ZZ crosstalk occurring between the two transmons can be measured directly. Consequently,

this should be incorporated into the closed-system Hamiltonian.

There are four distinct system-bath couplings in Eq. (24) that necessitate learning. Specifically, these encompass the following:

(1) Assuming that $g_{xz} = g_{zx}$ and $B_{xz} = B_{zx}$, there are three coupling strengths associated with the Ohmic bath, given by g_{zx}, g_{xx}, and g_{zz}. These have corresponding cut-off frequencies ω^c_{zx}, ω^c_{xx}, and ω^c_{zz}.
(2) A coupling strength related to the fluctuators, given by $b_{k,12}$. For simplificity, we assume that $b_{k,2q} = b_{2q}\ \forall k$, with a high-frequency cut-off $\gamma^{\max}_{zz}$.

To learn this set of four coupling strengths and the associated set of four cut-off frequencies, we propose a sequence of four steps. Each step is tailored to determine a subset of these unknown parameters. The methodology involves leveraging specific initial states, combined with either free or DD evolution, ensuring that each term is distinctly accentuated.

A. Step I: free evolution to determine g_{zx} and ω^c_{zx}

We commence with the free-evolution experiment for the initial state $1/\sqrt{2}(|00\rangle + |11\rangle)$. This state is invariant under $A_x \otimes A_x$ and $A_z \otimes A_z$. Therefore, the only noise contribution is due to $A_z \otimes A_x + A_x \otimes A_z$, with coupling strength g_{zx} under an Ohmic bath with a cut-off frequency ω^c_{zx}. By comparing the corresponding Redfield simulations to the experiment, and using the cost function from Eq. (23), we can deduce the two unknown parameters g_{zx} and ω^c_{zx}. Note that for evolution times of the order of $T_1 \sim 100\ \mu$s, the $A_x \otimes A_x$ error term generates a decay of the form $\sigma^- \otimes \sigma^-$ in the Markovian limit, under which $1/\sqrt{2}(|00\rangle + |11\rangle)$ is not invariant. However, as long we limit the free evolution to times much shorter than T_1 (e.g., 20 μs as in our experiments), we can safely ignore this effect.

B. Step II: free evolution to determine g_{xx} and ω^c_{xx}

We now initialize the system in the $|11\rangle$ state and include the $A_x \otimes A_x$ in the Redfield dynamics simulation of free evolution. The $|11\rangle$ state remains invariant under $A_z \otimes A_z$. Having learned the $A_z \otimes A_x + A_x \otimes A_z$ term in the previous step, we can now isolate and extract the coupling strength g_{xx} and the corresponding cut-off frequency ω^c_{xx} by comparing the simulation results to the experimental data for the same initial state.

C. Step III: DD evolution to determine g_{zz} and ω^c_{zz}

In this step, we wish to isolate the Ohmic $A_z \otimes A_z$ term and learn g_{zz} and ω^c_{zz}. Toward this end, we employ crosstalk-robust DD, as explained above, which suppresses the low-frequency $A_z \otimes A_z$ term, Eq. (24b). The Ohmic contributions, in contrast, are not strongly suppressed, so we include all of them in our simulations, but not the fluctuator term. As in the single-qubit case, we use an overcomplete set of initial states (the nine computational states $|\psi_i\rangle\,|\psi_i\rangle$, where $|\psi_i\rangle \in \{|0\rangle, |1\rangle, |+\rangle\}$) to determine g_{zz} and ω^c_{zz}, ensuring that state-dependent effects are averaged out. However, it may not be necessary to use all nine states to achieve a reliable estimate of g_{zz} and ω^c_{zz}, thus further saving in terms of the required number of experiments and simulations.

D. Step IV: free evolution to determine b_{2q} and $\gamma^{\max}_{zz}$

As the final step, we isolate the fluctuator term and learn b_{2q} and $\gamma^{\max}_{zz}$. We do this by simulating free evolution under the complete H^{2q}_{SB} of Eq. (24), after initializing in the $|+\rangle\,|+\rangle$ state. Given that we have already learned all the terms other than the fluctuator term, comparison between the simulations and the experimental data suffices to extract b_{2q} and $\gamma^{\max}_{zz}$ using the cost function approach.

VIII. SUMMARY AND CONCLUSIONS

This work presents a detailed noise model for transmon qubits consisting of both low- ($1/f$ -like) and high-frequency noise components based on a hybrid Redfield master equation. We designed an iterative three-step procedure to extract the unknown system-bath and bath parameters from a few simple "free-evolution" and "DD-evolution" experiments, as illustrated in Fig. 2, using the six Pauli matrix eigenstates. In both cases, we used DD to suppress diagonal (ZZ) qubit crosstalk [57] so that the remaining dominant noise effect on the main qubit (the qubit of interest) is decoherence. Our model treats the transmon qubit as a four-level system based on the circuit-model description of transmons (Sec. II) and treats the DD pulses as realistic time-dependent gates subject to quantum control (DRAG).

Once the unknown system-bath and bath parameters are extracted, we compare the model predictions with new experiments and a larger set of initial states and demonstrate that the model predicts the experimental results of free evolution and DD evolution with a relative error below 8% and 2% for Quito, and below 9% and 3% for Lima, respectively. This is based on a test with the six Pauli matrix eigenstates and ten random states for a total duration of up to 19.6 μs. The relative errors are higher for larger times (see Fig. 4), as expected because the Redfield model is based on the weak-coupling approximation, and its accuracy degrades as the simulation time is increased [60].

To test the robustness of our model, we performed a comparison with two simpler, two-level models with instantaneous pulses; while these models capture either the low- or the high-frequency noise, the full model captures both types of noise. Furthermore, our method is applicable

independently of the particular device-calibration procedure followed, as witnessed by the agreement we find for both Quito and Lima—devices with different drive-frequency calibrations.

The low relative error we find in the case of DD-evolution experiments ($< 2\%$ and $< 3\%$ for Quito and Lima, respectively) suggests that our full noise model helps model gate dynamics under the influence of decoherence. The model can also be used to study several qubits in parallel as long as there is no direct crosstalk between these qubits.

Finally, we proposed a simple extension of our methodology to learn the two-qubit couplings to the bath by performing simple to implement free-evolution and DD-evolution experiments. We leave the experimental implementation of this extension as a future direction.

We hope this work will benefit experimental groups working with superconducting qubits by helping them understand and learn experimental noise using a first-principles approach, which requires only a set of straightforward experiments. Extending our noise model beyond weak coupling and using it to analyze and improve entangling gates are additional promising future directions.

ACKNOWLEDGMENTS

We thank Mostafa Khezri, Ka-Wa Yip, and Humberto Munoz Bauza for several helpful discussions. This material is based upon work supported by the National Science Foundation, the Quantum Leap Big Idea under Grant No. OMA-1936388, the ARO MURI grant W911NF-22-S-0007, the Defense Advanced Research Projects Agency under Agreement HR00112230006, and the Office of Naval Research under Grant No. N00014-21-1-2688. We acknowledge the use of IBM Quantum services for this work. The views expressed are those of the authors and do not reflect the official policy or position of IBM or the IBM Quantum team.

APPENDIX A: DERIVATION OF TIME-DEPENDENT DRIVES

We start with Eq. (4) from the main text, and to obtain Eq. (5), we focus on the charge-coupling matrix elements for transmon qubits. The selection rules of the transmon qubit due to its cosine potential dictate that $g_{k,k\pm2} = 0\ \forall k$ [14]. The next order coupling terms, $g_{k,k\pm3}$, are proportional to the ratio of the anharmonicity and the qubit frequency: η_q/ω_q [64]. Therefore, $g_{k,k\pm1}$ (the coupling between nearest levels) is dominant, and all other couplings can be ignored, giving Eq. (5).

We can write the approximated charge operator $\hat{p} \equiv \sum_{k\geq0} \langle k|\,\hat{n}\,|k+1\rangle\,|k\rangle\langle k+1| + \text{h.c.}$ of Eq. (5) in terms of effective creation and annihilation operators for the eigenstates of the transmon, i.e., $\hat{p} = i(a^\dagger - a)$, where

$$a \equiv i\sum_{k\geq0} g_{k,k+1}|k\rangle\langle k+1| = i\sum_{k\geq0} \tilde{g}_k\sqrt{k+1}|k\rangle\langle k+1|, \tag{A1}$$

with

$$g_{k,k+1} \equiv \langle k|\,\hat{n}\,|k+1\rangle = g\frac{\langle k|\,\hat{n}\,|k+1\rangle}{\langle 0|\,\hat{n}\,|1\rangle} \tag{A2a}$$

$$\equiv \tilde{g}_k\sqrt{k+1}, \tag{A2b}$$

and $g \equiv g_{0,1} = \langle 0|\,\hat{n}\,|1\rangle = \tilde{g}_0$. We note that $g_{k,k+1} = g^*_{k,k+1}$, which follows from the fact that $\hat{n}$ is the number operator for Cooper pairs. We note further that to first order in η_q/ω_q [64]

$$\tilde{g}_k \approx g\left(1 - \frac{k}{2}\frac{\eta_q}{\omega_q}\right). \tag{A3}$$

We defined $\tilde{g}_k(=\tilde{g}^*_k)$ in Eq. (A2b) to include all higher-order perturbative corrections and do not use the approximation Eq. (A3) in our numerical calculations. However, we note that the leading-order correction to $\tilde{g}_k$ is of the same order as $g_{k,k\pm3}$, i.e., proportional to η_q/ω_q.

Equation (4) becomes

$$H_{\text{sys}} = H^{\text{eigen}}_{\text{trans}} + H_{\text{drive}}, \tag{A4a}$$

$$H_{\text{drive}} \equiv i\varepsilon(t)\cos(\omega_{\text{d}}t + \phi_{\text{d}})\left(a^\dagger - a\right). \tag{A4b}$$

For simplicity, we consider the case when $\phi_{\text{d}} = 0$ in Eq. (A4b), and transform the Hamiltonian $H^{\text{eigen}}_{\text{sys}}$ into a frame rotating with $U_{\text{d}} = e^{i\omega_{\text{d}}\hat{N}t}$. Here $\hat{N} = \sum_{k\geq0} k|k\rangle\langle k|$ is the transmon number operator. In this frame, rotating with the drive frequency, the effective Hamiltonian is given by

$$\tilde{H}_{\text{sys}} = U_{\text{d}}H_{\text{sys}}U^\dagger_{\text{d}} + i\dot{U}_dU^\dagger_{\text{d}} \tag{A5a}$$

$$= H^{\text{eigen}}_{\text{trans}} + U_{\text{d}}H_{\text{drive}}U^\dagger_{\text{d}} + i\dot{U}_dU^\dagger_{\text{d}} \tag{A5b}$$

$$= \sum_{k\geq0}(\omega_k - k\omega_{\text{d}})\,|k\rangle\langle k| + U_{\text{d}}H_{\text{drive}}U^\dagger_{\text{d}}, \tag{A5c}$$

where in going from Eq. (A5a) to Eq. (A5b), we used the fact that $\hat{N}$ commutes with $H^{\text{eigen}}_{\text{trans}}$ [Eq. (3)].

Note that a and $a^\dagger$ should not be confused with the harmonic oscillator raising and lowering operators since

$$[a, a^\dagger] = \sum_{k\geq1}\left((k+1)\tilde{g}^2_k - k\tilde{g}^2_{k-1}\right)|k\rangle\langle k| + \tilde{g}_0|0\rangle\langle 0|$$

$$\neq I. \tag{A6}$$

For the harmonic oscillator case, $\tilde{g}_k = 1$ for all k and we obtain the usual commutation relation $\left[a, a^\dagger\right] = I$. Despite

this, we have, as for the harmonic oscillator:

$$[a,\hat{N}] = [i\sum_{k\geq 0}\tilde{g}_k\sqrt{k+1}|k\rangle\langle k+1|, \sum_{k'\geq 0}k'|k'\rangle\langle k'|]$$
$$= i\sum_{k\geq 0}\tilde{g}_k(k+1)\sqrt{k+1}|k\rangle\langle k+1| \quad \text{(A7a)}$$
$$- i\sum_{k\geq 0}\tilde{g}_k k\sqrt{k+1}|k\rangle\langle k+1| \quad \text{(A7b)}$$
$$= i\sum_{k\geq 0}\tilde{g}_k\sqrt{k+1}|k\rangle\langle k+1|, \quad \text{(A7c)}$$

so that

$$[a,\hat{N}] = a\,, \quad [a^\dagger,\hat{N}] = -a^\dagger. \quad \text{(A8)}$$

Let us write the last term in Eq. (A5c) as

$$U_{\rm d}H_{\rm drive}U_{\rm d}^\dagger = i\frac{\varepsilon(t)}{2}\left(e^{i\omega_{\rm d}t}+e^{-i\omega_{\rm d}t}\right)U_{\rm d}\left(a^\dagger - a\right)U_{\rm d}^\dagger. \quad \text{(A9)}$$

Using Eq. (A8), we then have

$$U_{\rm d}\left(a^\dagger - a\right)U_{\rm d}^\dagger = e^{i\omega_{\rm d}t}a^\dagger - e^{-i\omega_{\rm d}t}a. \quad \text{(A10)}$$

Making the rotating-wave approximation, we ignore the fast-oscillating terms with frequencies $\pm 2\omega_{\rm d}$, and Eq. (A9) reduces to

$$U_{\rm d}H_{\rm drive}U_{\rm d}^\dagger \approx i\frac{\varepsilon(t)}{2}\left(a^\dagger - a\right). \quad \text{(A11)}$$

Combining this with Eqs. (A1) and (A5), the effective Hamiltonian for a drive that results in a rotation about the x axis (in the qubit subspace) is given by

$$\tilde{H}_{\rm sys} = \sum_{k\geq 0}(\omega_k - k\omega_{\rm d})\,|k\rangle\langle k|$$
$$+ \frac{\varepsilon(t)}{2}\sum_{k\geq 0}\tilde{g}_k\sqrt{k+1}(|k\rangle\langle k+1| + |k+1\rangle\langle k|), \quad \text{(A12)}$$

which is Eq. (6) of the main text. By tuning $\phi_{\rm d}$, we can implement a rotation about any axis in the (x,y) plane in the qubit subspace.

APPENDIX B: DRAG

DRAG [44,52] is a useful technique to reduce both the leakage and the phase errors, which accumulate during the operation of single-qubit gates. The most standard DRAG technique is to add the derivative of the pulse envelope to the quadrature component such that the final form of the pulse envelope $\tilde{\varepsilon}(t)$ is given by

$$\tilde{\varepsilon}(t) = \varepsilon(t) + i\alpha\frac{\dot{\varepsilon}(t)}{\eta_q}\,, \quad \text{(B1)}$$

where η_q is the transmon anharmonicity and α is a constant that can have different values depending on which errors need to be suppressed, e.g., $\alpha = 1$ to suppress leakage and $\alpha = 1/2$ to suppress coherent phase errors [65]. Equation (B1) implies that if we need to apply a single X gate whose pulse envelope is given by $\varepsilon(t)$, then $\alpha\dot{\varepsilon}(t)/\eta_q$ needs to be applied along the y axis.

IBMQE devices use DRAG, but the exact pulse parameters are not available to users. Instead, the value of α in our simulations can be optimized numerically to match the experimentally reported gate fidelity. This is the approach we take here, with the goal being to model experiments on IBMQE devices with single-qubit gate errors of the order of 10^{-3}. This is the value reported using randomized benchmarking [66], which equals the average gate infidelity $(1 - F_g)$ when the gate set has gate-independent errors [67,68], an assumption we make here to justify the use of 10^{-3} as our target gate infidelity.

We perform closed-system simulations and find that without DRAG, the gate infidelity (due to leakage and phase errors) is in the range 10^{-2}–10^{-3}. With DRAG, we

TABLE II. Specifications of the Quito (top row) and Lima (bottom row) devices accessed on September 1, 2021, and January 1, 2023, respectively. The sx ($\sqrt{\sigma^x}$) gate forms the basis of all the single qubit gates, and any single-qubit gate of the form $U3(\theta,\phi,\lambda)$ is composed of two sx and three rz(λ) $= \exp(-i\frac{\lambda}{2}\sigma^z)$ gates (which are error free and take zero time, as they correspond to frame updates).

Quito	Q0	Q1	Q2	Q3	Q4
Qubit freq. (GHz)	5.3006	5.0806	5.3220	5.1637	5.0524
η_q (MHz)	331.5	319.2	332.3	335.1	319.3
T_1 (μs)	86.7	98.6	61.5	111.5	85.7
T_2 (μs)	132.5	149.0	78.9	22.7	136.7
sx gate error $[10^{-2}]$	0.0302	0.0243	0.1042	0.0629	0.0884
sx gate length (ns)	35.556	35.556	35.556	35.556	35.556
readout error $[10^{-2}]$	3.91	2.10	6.42	2.28	2.00
Lima	**Q0**	**Q1**	**Q2**	**Q3**	**Q4**
Qubit freq. (GHz)	5.0297	5.1277	5.2474	5.3026	5.0920
η_q (MHz)	335.7	318.3	333.6	331.2	334.5
T_1 (μs)	125.2	105.7	88.1	59.9	23.2
T_2 (μs)	194.3	136.2	123.3	16.8	21.1
sx gate error $[10^{-2}]$	0.0230	0.0189	0.0308	0.0251	0.0578
sx gate length (ns)	35.556	35.556	35.556	35.556	35.556
readout error $[10^{-2}]$	1.73	1.40	1.69	2.42	4.820

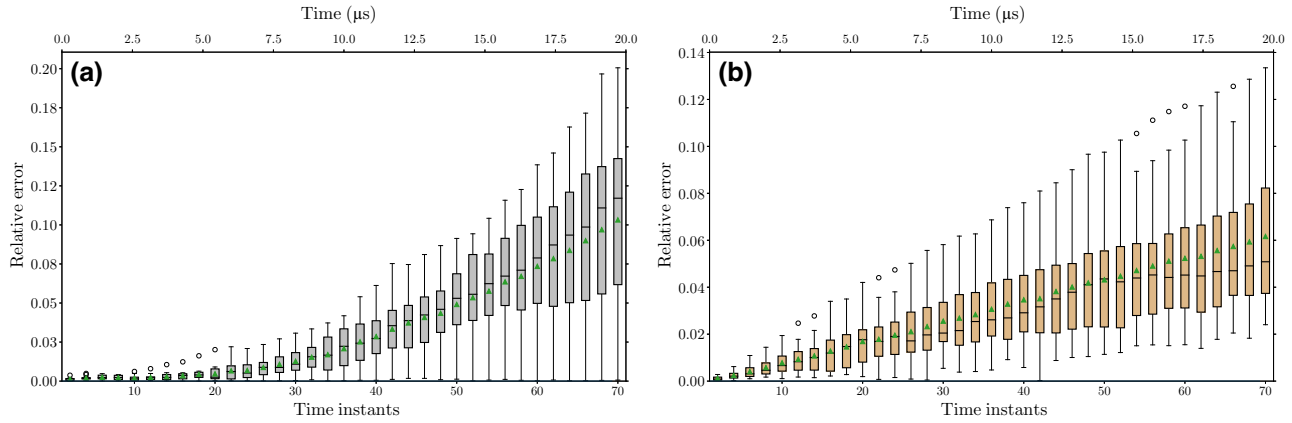

FIG. 7. Lindblad equation simulation results for the Quito processor. Left: box plots showing the relative error of the Lindblad simulations for the free-evolution experiments as a function of time for 16 different initial states containing six Pauli states and ten Haar-random states. Right: the same as on the left, but for the DD-evolution experiments. We measured a total of 70 time instants, up to a total evolution time of 19.6 μs, but display only every other instant to avoid overcrowding. Green triangles indicate the mean over the 16 initial states, black horizontal lines are the median, gray boxes represent the [25, 75] percentiles, the whiskers (black lines extending outside the boxes) represent the [0, 25] and [75, 100] percentiles, and circles are outliers.

find that varying α from 1/2 to 1 increases the gate infidelity from approximately 10^{-6} to 10^{-3}, respectively. We thus choose $\alpha = 1$ to match the reported fidelity. This suggests, as described in Ref. [65], that the remaining errors are mostly phase error. Indeed, we find that even without DRAG, single-qubit X and Y gates (both implemented using two $\pi/2$ pulses as explained in the main text) have leakage errors well below 10^{-5}. This is unsurprising, as these long gates are quite narrowband compared to the transmon anharmonicity. We note that this attributes all errors to coherent closed-system effects rather than decoherence. We expect incoherent errors to be of the order of $t_g/T_2 \approx 5 \times 10^{-4}$, suggesting that this choice is defensible (see Table II). Note that the above expression for incoherent errors comes from the short-time expansion of e^{-t_g/T_2}, which is how we would expect decoherence to manifest, and hence the error term can be approximated as $1 - e^{-t_g/T_2} = t_g/T_2 + O((t_g/T_2)^2)$. Furthermore, given that both gate fidelity and coherence times drift over hour-long timescales, we focus only on matching the correct order of magnitude for fidelity with our coherent error model.

APPENDIX C: COMPARISON WITH THE LINDBLAD MASTER EQUATION

As a sanity check, we also conducted a simple test for the Quito processor using the backend-reported T_1 and T_2 values (see Table II). Using the Lindblad equation in the standard form

$$\dot{\rho} = -i[H, \rho] + \sum_i \gamma_i \left(L_i \rho L_i^\dagger - \frac{1}{2} \left\{ L_i^\dagger L_i, \rho \right\} \right), \quad \text{(C1)}$$

we consider two Lindblad operators corresponding to relaxation ($L_{\text{relax}} = \sigma^- = |0\rangle\langle 1|$) and dephasing ($L_{\text{dephase}} = \sigma^z/2$). We take the corresponding rates to be $2\pi/T_1$ and $2\pi/T_2$, respectively. Here T_1 is the relaxation time and T_2 is the dephasing time obtained via a Hahn-echo-based Ramsey experiment. We treat the transmons as two-level systems and assume the DD pulses to be instantaneous. We then simulate both the free- and DD-evolution dynamics starting with all 16 states considered in the main text in the single-qubit case. The results, given in terms of the relative error as a function of time, are presented in Fig. 7.

We observe that the relative error of both the free- and DD-evolution cases is substantially larger than for our detailed simulation model results [see Figs. 4(a), 4(b)]. Qualitatively, the relative error using the Lindblad equation is as high as 20% and 14% for the free and DD evolutions, respectively, compared to our detailed model

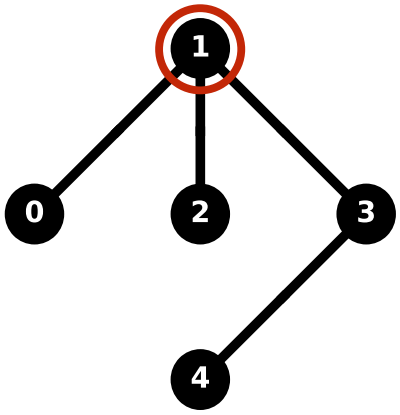

FIG. 8. Schematic layout of the Quito and Lima processors. The main qubit in our work is 1; we refer to the rest as spectators.

results, which always have corresponding relative errors below 8% and 2% for the Quito processor. This observation underscores the need for a methodology that goes beyond a simple Markovian model to capture noise in transmon qubits.

APPENDIX D: DATA COLLECTION AND ANALYSIS METHODOLOGY

We used the IBMQE processor ibmq_quito (Quito) and ibmq_lima (Lima), whose layout is shown schematically in Fig. 8. For both Quito and Lima, we use qubit 1 (Q1) as the main qubit. These are five-qubit processors consisting of superconducting transmon qubits. Various calibration details and hardware specifications relevant to the qubits and gates used in this work are provided in Table II.

For each initial state, we selected 70 equidistant time instants (19.6 μs/70), with each such instant corresponding to an integer number of cycles of the $XY4$ DD sequences. We generated a circuit according to the scheme given in Fig. 2 for each such initial state and each such instant. We sent all 70 such circuits in one job (the maximum allowed number of circuits per job is 75), and each job was repeated 8192 times. We ensured that all the jobs were sent consecutively within the same calibration cycle to avoid charge-noise-dependent fluctuations and variations in critical features over different calibration cycles.

We measured only the main qubit in the Z basis, each measurement yielding either 0 or 1. We computed the empirical fidelity $F^{(e)}$ as the number of favorable outcomes (0) to the total number of experiments (8192 per initial state and per measurement time instant). This is a proxy for the $F_{|\psi\rangle} = \mathrm{Tr}\left[U^{-1}\mathcal{E}\left(U|0\rangle\langle 0|U^{-1}\right)U\right]$, where U represents $U3(\theta,\phi,\lambda)$ and $\mathcal{E}$ represents the quantum map of the main qubit corresponding to any of the three types of

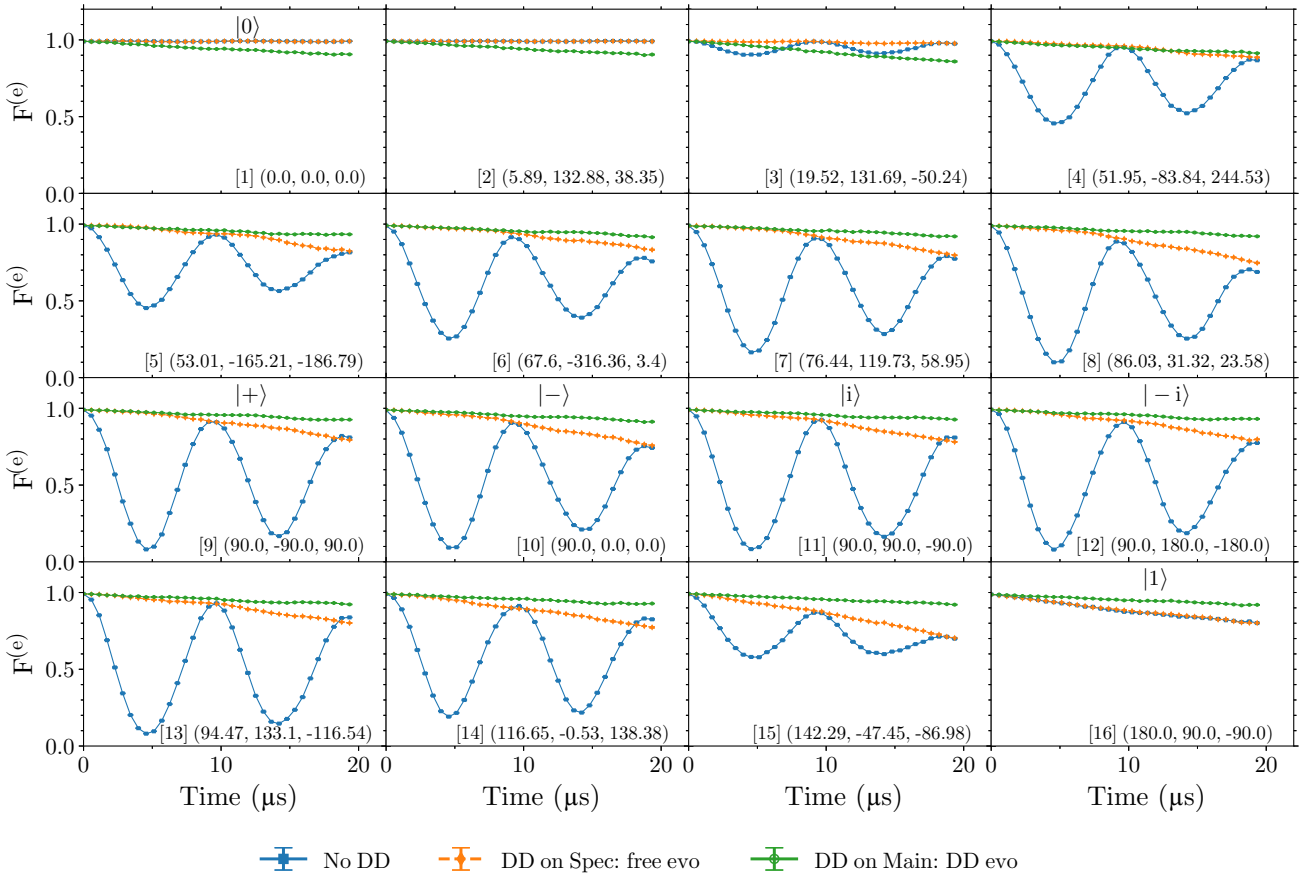

FIG. 9. Fidelity results for the Quito processor, for the 16 different initial states of the main qubit. The caption of each of the panels gives (θ,ϕ,λ) in degrees, parametrizing the initial state $|\psi\rangle = U3(\theta,\phi,\lambda)\,|0\rangle$ (panels are arranged in increasing order of θ, the polar angle with the z axis). These are the six Pauli states (panels 1,9-12,16) and the ten Haar-random states. Blue curves (squares): no DD is applied, resulting in coherent oscillations due to crosstalk. Orange curves (diamonds): DD ($XY4$) is applied just to the spectator qubits; the resulting suppression of crosstalk between the main qubit and the spectator qubits removes the oscillations. These are what we call the free-evolution experiments in the main text. Green curves (circles): DD ($XY4$) is applied just to the main qubit, suppressing both crosstalk and errors due to environment-induced noise. Results are averaged over three different runs of experiments. All of the data was acquired on September 1, 2021. Error bars are smaller than the markers.

experiments described in the main text and in Appendix E below.

Error bars were then generated using the standard bootstrapping procedure, where we resample (with replacement) counts out of the experimental counts' dictionary (i.e., the list of 0/1 measurement outcomes per state and instant) and create several new dictionaries. The final fidelity and error bars are obtained by calculating the mean and standard deviations over the fidelities of these newly resampled dictionaries. Using ten such resampled dictionaries of the counts sufficed to give small error bars. We report the final fidelity with 2σ error bars, corresponding to 95% confidence intervals.

APPENDIX E: EXPERIMENTAL FIDELITY RESULTS

For the Quito device, Fig. 9 shows the results of three different types of experiments for the 16 states consisting of six Pauli states and ten Haar-random states. In the first type of experiment, we apply a series of identity gates separated by barriers on all the qubits (main qubits and the spectator qubits). All the qubits always start in the $|0\rangle$ state. The second and third types are the experiments discussed in the main text: free evolution, where we apply an $XY4$ DD sequence to the spectator qubits and identity gates on the main qubit, and DD evolution, where we apply the $XY4$ DD sequences to the main qubit and identity gates to the spectator qubits (see Fig. 2).

For Lima, we show only two types of experiments in Fig. 10. The first is the free-evolution experiment, where we prepare some initial state of the main qubit, apply a series of identity gates, and measure in the computational basis. The second type is the usual DD experiment, where we apply a series of $XY4$ sequences to the main qubit and the identity operation to all the spectator qubits. In both types of experiments, the spectator qubits are always initialized in the ground state $|0\rangle$. In the Lima case, applying a DD sequence to the spectator qubits is unnecessary.

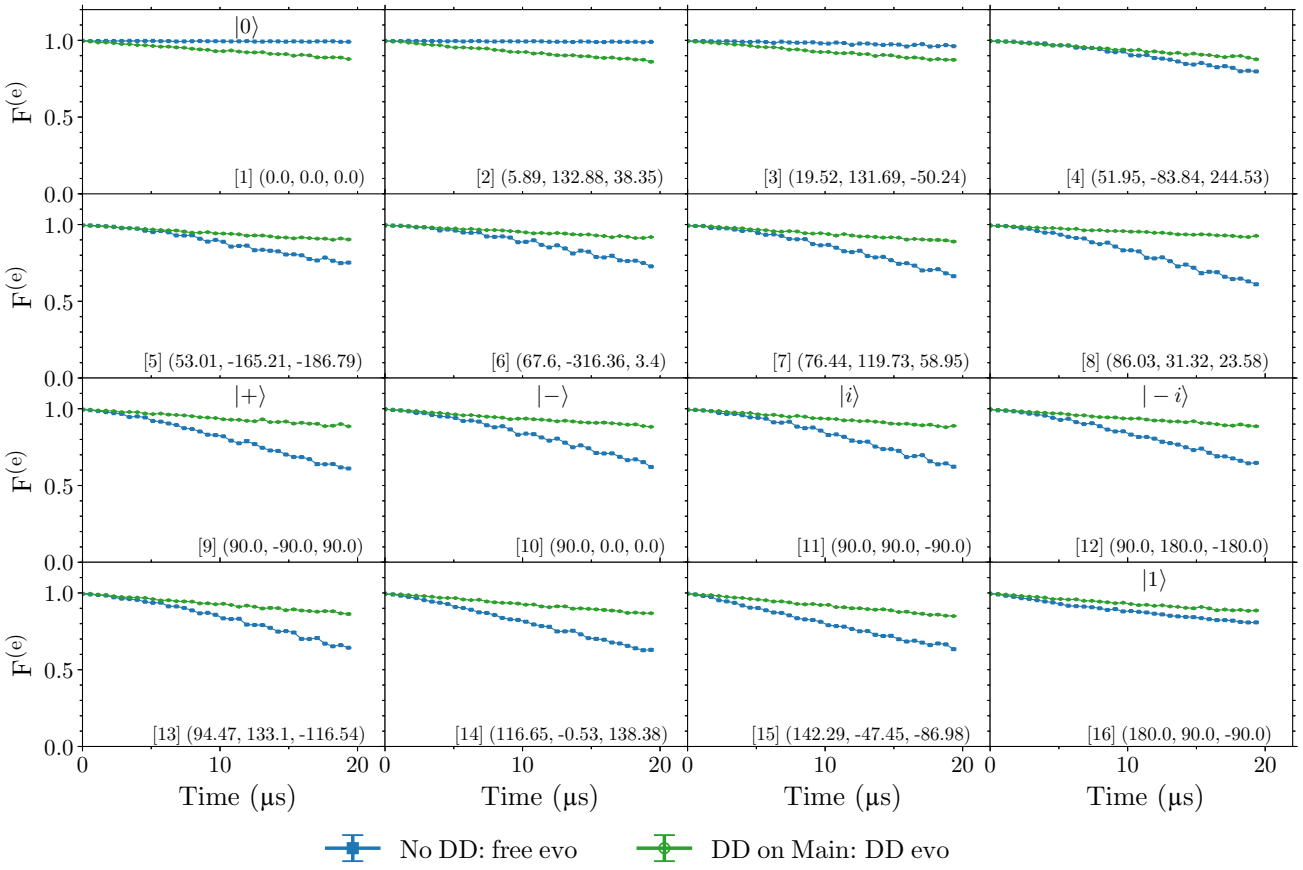

FIG. 10. Fidelity results for the Lima processor, for the 16 different initial states of the main qubit. The caption of each of the panels gives (θ, ϕ, λ) in degrees, parametrizing the initial state $|\psi\rangle = U3(\theta, \phi, \lambda)\,|0\rangle$ (panels are arranged in increasing order of θ, the polar angle with the z axis). These are the six Pauli states (panels 1,9-12,16) and the ten Haar-random states. Blue curves (squares): no DD is applied. In contrast to Fig. 9, there are no crosstalk oscillations. This is because all the spectator qubits are kept in the ground state $|0\rangle$, and calibration for Lima is done in the same spectators' state. These are what we call the free-evolution experiments for Lima in the main text. Green curves (circles): DD ($XY4$) is applied just to the main qubit, suppressing errors due to environment-induced noise. All of the data was acquired on January 1, 2023. Error bars are smaller than the markers.

[1] Y. Nakamura, Y. A. Pashkin, and J. S. Tsai, Coherent control of macroscopic quantum states in a single-cooper-pair box, Nature **398**, 786 EP (1999).

[2] J. E. Mooij, T. P. Orlando, L. Levitov, L. Tian, C. H. van der Wal, and S. Lloyd, Josephson persistent-current qubit, Science **285**, 1036 (1999).

[3] J. Clarke and F. K. Wilhelm, Superconducting quantum bits, Nature **453**, 1031 (2008).

[4] S. Hacohen-Gourgy, L. S. Martin, E. Flurin, V. V. Ramasesh, K. B. Whaley, and I. Siddiqi, Quantum dynamics of simultaneously measured non-commuting observables, Nature **538**, 491 (2016).

[5] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, Nature **549**, 242 EP (2017).

[6] Z. K. Minev, S. O. Mundhada, S. Shankar, P. Reinhold, R. Gutiérrez-Jáuregui, R. J. Schoelkopf, M. Mirrahimi, H. J. Carmichael, and M. H. Devoret, To catch and reverse a quantum jump mid-flight, Nature **570**, 200 (2019).

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

[8] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, Supervised learning with quantum-enhanced feature spaces, Nature **567**, 209 (2019).

[9] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Error mitigation extends the computational reach of a noisy quantum processor, Nature **567**, 491 (2019).

[10] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf, L. Frunzio, M. Mirrahimi, and M. H. Devoret, Quantum error correction of a qubit encoded in grid states of an oscillator, Nature **584**, 368 (2020).

[11] C. K. Andersen, A. Remm, S. Lazar, S. Krinner, N. Lacroix, G. J. Norris, M. Gabureac, C. Eichler, and A. Wallraff, Repeated quantum error detection in a surface code, Nat. Phys. **16**, 875 (2020).

[12] F. Arute *et al.*, Hartree-Fock on a superconducting qubit quantum computer, Science **369**, 1084 (2020).

[13] Y. Wu *et al.*, Strong quantum computational advantage using a superconducting quantum processor, Phys. Rev. Lett. **127**, 180501 (2021).

[14] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Charge-insensitive qubit design derived from the Cooper pair box, Phys. Rev. A **76**, 042319 (2007).

[15] S. J. Devitt, Performing quantum computing experiments in the cloud, Phys. Rev. A **94**, 032329 (2016).

[16] J. R. Wootton and D. Loss, Repetition code of 15 qubits, Phys. Rev. A **97**, 052313 (2018).

[17] C. Vuillot, Is error detection helpful on IBM 5Q chips?, Quantum Inf. Comput. **18**, 0949 (2018).

[18] J. Roffe, D. Headley, N. Chancellor, D. Horsman, and V. Kendon, Protecting quantum memories using coherent parity check codes, Quantum Sci. Technol. **3**, 035010 (2018).

[19] D. Willsch, M. Willsch, F. Jin, H. De Raedt, and K. Michielsen, Testing quantum fault tolerance on small systems, Phys. Rev. A **98**, 052348 (2018).

[20] B. Pokharel, N. Anand, B. Fortman, and D. A. Lidar, Demonstration of fidelity improvement using dynamical decoupling with superconducting qubits, Phys. Rev. Lett. **121**, 220502 (2018).

[21] R. Harper and S. T. Flammia, Fault-tolerant logical gates in the ibm quantum experience, Phys. Rev. Lett. **122**, 080504 (2019).

[22] D. Maslov, J.-S. Kim, S. Bravyi, T. J. Yoder, and S. Sheldon, Quantum advantage for computations with limited space, Nat. Phys. **17**, 894 (2021).

[23] K. Zhang, P. Rao, K. Yu, H. Lim, and V. Korepin, Implementation of efficient quantum search algorithms on NISQ computers, Quantum Inf. Proc. **20**, 233 (2021).

[24] B. Pokharel and D. A. Lidar, Demonstration of algorithmic quantum speedup, Phys. Rev. Lett. **130**, 210602 (2023).

[25] B. Pokharel and D. Lidar, Better-than-classical Grover search via quantum error detection and suppression, arXiv:2211.04543 [quant-ph] (2022).

[26] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, New York, 2010).

[27] D. Lidar and T. Brun, eds., *Quantum Error Correction* (Cambridge University Press, Cambridge, UK, 2013).

[28] P. Aliferis, D. Gottesman, and J. Preskill, Quantum accuracy threshold for concatenated distance-3 codes, Quantum Inf. Comput. **6**, 97 (2006).

[29] R. Chao and B. W. Reichardt, Quantum error correction with only two extra qubits, Phys. Rev. Lett. **121**, 050502 (2018).

[30] E. T. Campbell, B. M. Terhal, and C. Vuillot, Roads towards fault-tolerant universal quantum computation, Nature **549**, 172 EP (2017).

[31] A. M. Souza, Process tomography of robust dynamical decoupling with superconducting qubits, Quantum Inf. Process. **20**, 237 (2021).

[32] R. Harper, S. T. Flammia, and J. J. Wallman, Efficient learning of quantum noise, Nat. Phys. **16**, 1184 (2020).

[33] K. Georgopoulos, C. Emary, and P. Zuliani, Modeling and simulating the noisy behavior of near-term quantum computers, Phys. Rev. A **104**, 062432 (2021).

[34] T. McCourt, C. Neill, K. Lee, C. Quintana, Y. Chen, J. Kelly, J. Marshall, V. N. Smelyanskiy, M. I. Dykman, A. Korotkov, I. L. Chuang, and A. G. Petukhov, Learning noise via dynamical decoupling of entangled qubits, Phys. Rev. A **107**, 052610 (2023).

[35] B. G. Christensen, C. D. Wilen, A. Opremcak, J. Nelson, F. Schlenker, C. H. Zimonick, L. Faoro, L. B. Ioffe, Y. J. Rosen, J. L. DuBois, B. L. T. Plourde, and R. McDermott, Anomalous charge noise in superconducting qubits, Phys. Rev. B **100**, 140503 (2019).

[36] A. G. Redfield, in *Advances in Magnetic and Optical Resonance*, edited by J. S. Waugh (Academic Press, New York, 1965), Vol. 1, p. 1.

[37] L. Viola and S. Lloyd, Dynamical suppression of decoherence in two-state quantum systems, Phys. Rev. A **58**, 2733 (1998).

[38] L.-M. Duan and G.-C. Guo, Suppressing environmental noise in quantum computation through pulse control, Phys. Lett. A **261**, 139 (1999).

[39] D. Vitali and P. Tombesi, Using parity kicks for decoherence control, Phys. Rev. A **59**, 4178 (1999).

[40] P. Zanardi, Symmetrizing evolutions, Phys. Lett. A **258**, 77 (1999).

[41] L. Viola, E. Knill, and S. Lloyd, Dynamical decoupling of open quantum systems, Phys. Rev. Lett. **82**, 2417 (1999).

[42] J. Bylander, S. Gustavsson, F. Yan, F. Yoshihara, K. Harrabi, G. Fitch, D. Cory, Y. Nakamura, J. Tsai, and W. Oliver, Noise spectroscopy through dynamical decoupling with a superconducting flux qubit, Nat. Phys. **7**, 565 (2011).

[43] IBM Quantum (2021).

[44] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, Simple pulses for elimination of leakage in weakly nonlinear qubits, Phys. Rev. Lett. **103**, 110501 (2009).

[45] In more detail, we try different values of E_J and E_C in Eq. (1) by diagonalizing the corresponding H_{trans} and comparing the result with the experimental values of ω_q and η_q. Once we find the values of E_J and E_C yielding the closest match, we proceed to Eq. (3) to find the full spectrum.

[46] R. Freeman, *Spin Choreography* (Oxford University Press Oxford, Oxford, 1998).

[47] F. Motzoi and F. K. Wilhelm, Improving frequency selection of driven pulses using derivative-based transition suppression, Phys. Rev. A **88**, 062318 (2013).

[48] M. D. Bowdrey, D. K. L. Oi, A. J. Short, K. Banaszek, and J. A. Jones, Fidelity of single qubit maps, Phys. Lett. A **294**, 258 (2002).

[49] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, and J. M. Gambetta, Efficient z gates for quantum computing, Phys. Rev. A **96**, 022330 (2017).

[50] D. C. McKay, T. Alexander, L. Bello, M. J. Biercuk, L. Bishop, J. Chen, J. M. Chow, A. D. Córcoles, D. Egger, S. Filipp, J. Gomez, M. Hush, A. Javadi-Abhari, D. Moreda, P. Nation, B. Paulovicks, E. Winston, C. J. Wood, J. Wootton, and J. M. Gambetta, Qiskit backend specifications for openqasm and openpulse experiments, arXiv:1809.03452 [quant-ph] (2018).

[51] A. P. Babu, J. Tuorila, and T. Ala-Nissila, State leakage during fast decay and control of a superconducting transmon qubit, Npj Quantum Inf. **7**, 30 (2021).

[52] J. M. Gambetta, F. Motzoi, S. T. Merkel, and F. K. Wilhelm, Analytic control methods for high-fidelity unitary operations in a weakly nonlinear oscillator, Phys. Rev. A **83**, 012308 (2011).

[53] V. E. Manucharyan, J. Koch, L. I. Glazman, and M. H. Devoret, Fluxonium: Single Cooper-pair circuit free of charge offsets, Science **326**, 113 (2009).

[54] C. M. Quintana *et al.*, Observation of classical-quantum crossover of $1/f$ flux noise and its paramagnetic temperature dependence, Phys. Rev. Lett. **118**, 057702 (2017).

[55] E. Paladino, Y. M. Galperin, G. Falci, and B. L. Altshuler, $1/f$ noise: Implications for solid-state quantum information, Rev. Mod. Phys. **86**, 361 (2014).

[56] K. W. Yip, Ph.D. thesis, University of Southern California, 2021.

[57] V. Tripathi, H. Chen, M. Khezri, K.-W. Yip, E. M. Levenson-Falk, and D. A. Lidar, Suppression of crosstalk in superconducting qubits using dynamical decoupling, Phys. Rev. Appl. **18**, 024068 (2022).

[58] H. Chen and D. A. Lidar, Hamiltonian open quantum system toolkit, Commun. Phys. **5**, 112 (2022).

[59] qiskit.circuit.library.U3Gate (2021).

[60] E. Mozgunov and D. Lidar, Completely positive master equation for arbitrary driving and small level spacing, Quantum **4**, 227 (2020).

[61] D. Brand, I. Sinayskiy, and F. Petruccione, Markovian noise model fitting and parameter extraction of ibmq transmon qubits, arXiv:2202.04474 [quant-ph] (2023).

[62] T. Menke, W. P. Banner, T. R. Bergamaschi, A. Di Paolo, A. Vepsäläinen, S. J. Weber, R. Winik, A. Melville, B. M. Niedzielski, D. Rosenberg, K. Serniak, M. E. Schwartz, J. L. Yoder, A. Aspuru-Guzik, S. Gustavsson, J. A. Grover, C. F. Hirjibehedin, A. J. Kerman, and W. D. Oliver, Demonstration of tunable three-body interactions between superconducting qubits, Phys. Rev. Lett. **129**, 220501 (2022).

[63] Z. Zhou, R. Sitler, Y. Oda, K. Schultz, and G. Quiroz, Phys. Rev. Lett. **131**, 210802 (2023).

[64] M. Khezri, Ph.D. thesis, University of California, Riverside, 2018.

[65] Z. Chen *et al.*, Measuring and suppressing quantum state leakage in a superconducting qubit, Phys. Rev. Lett. **116**, 020501 (2016).

[66] J. Emerson, R. Alicki, and K. Życzkowski, Scalable noise estimation with random unitary operators, J. Opt. B: Quantum Semiclass. Opt. **7**, S347 (2005).

[67] E. Magesan, J. M. Gambetta, and J. Emerson, Scalable and robust randomized benchmarking of quantum processes, Phys. Rev. Lett. **106**, 180504 (2011).

[68] T. Proctor, K. Rudinger, K. Young, M. Sarovar, and R. Blume-Kohout, What randomized benchmarking actually measures, Phys. Rev. Lett. **119**, 130502 (2017).