TU Delft

Delft University of Technology

Efficient Qubit Calibration by Binary-Search Hamiltonian Tracking

Berritta, Fabrizio; Benestad, Jacob; Pahl, Lukas; Mathews, Melvin; Krzywda, Jan A.; Assouly, Réouven; Sung, Youngkyu; Kim, David K.; Chatterjee, Anasua; More Authors

DOI
10.1103/77qg-p68k

Publication date
2025

Document Version
Final published version

Published in
PRX Quantum

Citation (APA)
Berritta, F., Benestad, J., Pahl, L., Mathews, M., Krzywda, J. A., Assouly, R., Sung, Y., Kim, D. K., Chatterjee, A., & More Authors (2025). Efficient Qubit Calibration by Binary-Search Hamiltonian Tracking. *PRX Quantum*, *6*(3), Article 030335. https://doi.org/10.1103/77qg-p68k

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Efficient Qubit Calibration by Binary-Search Hamiltonian Tracking

Fabrizio Berritta,[1,2] Jacob Benestad,[3] Lukas Pahl,[1,4] Melvin Mathews,[1,5] Jan A. Krzywda,[6] Réouven Assouly,[1] Youngkyu Sung,[1,4] David K. Kim,[7] Bethany M. Niedzielski,[7] Kyle Serniak,[1,7] Mollie E. Schwartz,[7] Jonilyn L. Yoder,[7] Anasua Chatterjee,[2,8] Jeffrey A. Grover,[1] Jeroen Danon,[3] William D. Oliver,[1,4,9] and Ferdinand Kuemmeth[2,10,11,*]

[1] *Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

[2] *Center for Quantum Devices, Niels Bohr Institute, University of Copenhagen, 2100 Copenhagen, Denmark*

[3] *Center for Quantum Spintronics, Department of Physics, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway*

[4] *Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

[5] *Department of Information Technology and Electrical Engineering, ETH Zürich, 8093 Zürich, Switzerland*

[6] *⟨aQa^L⟩ Applied Quantum Algorithms—Lorentz Institute for Theoretical Physics and Leiden Institute of Advanced Computer Science, Universiteit Leiden, The Netherlands*

[7] *Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts 02421, USA*

[8] *QuTech and Kavli Institute of Nanoscience, Delft University of Technology, Delft, The Netherlands*

[9] *Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

[10] *Institute of Experimental and Applied Physics, University of Regensburg, 93040 Regensburg, Germany*

[11] *QDevil, Quantum Machines, 2750 Ballerup, Denmark*

(Received 10 January 2025; revised 26 May 2025; accepted 14 July 2025; published 26 August 2025)

We present and experimentally implement a real-time protocol for calibrating the frequency of a resonantly driven qubit, achieving exponential scaling in calibration precision with the number of measurements, up to the limit imposed by decoherence. The real-time processing capabilities of a classical controller dynamically generate adaptive probing sequences for qubit-frequency estimation. Each probing evolution time and drive frequency are calculated to divide the prior probability distribution into two branches, following a locally optimal strategy that mimics a conventional binary search. The scheme does not require repeated measurements at the same setting, as it accounts for state preparation and measurement errors. Its use of a parametrized probability distribution favors numerical accuracy and computational speed. We show the efficacy of the algorithm by stabilizing a flux-tunable transmon qubit, leading to improved coherence and gate fidelity. As benchmarked by gate-set tomography, the field-programmable gate array (FPGA) powered control electronics partially mitigates non-Markovian noise, which is detrimental to quantum error correction. The mitigation is achieved by dynamically updating and feeding forward the qubit frequency. Our protocol highlights the importance of feedback in improving the calibration and stability of qubits subject to drift and can be readily applied to other qubit platforms.

DOI: 10.1103/77qg-p68k

I. INTRODUCTION

Quantum processing units (QPUs) with tens of qubits are becoming increasingly common [1]. Efforts to reduce error rates have resulted in the first demonstrations of quantum error correction [2,3], though these advances come with costly calibration overhead as the number of qubits increases. Temporal instability in QPUs arises from numerous stochastic noise channels [4] and the lowest-performing outlier qubits often limit overall performance [5]. Tackling time-dependent fluctuations and outlier qubits in large QPUs requires scalable and efficient calibration methods to ensure stable fault-tolerant operation while minimizing calibration down time. Here, we introduce a protocol that features exponential scaling of the

*Contact author: ferdinand.kuemmeth@ur.de

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

calibration precision versus the number of measurements until it is limited by decoherence. This scaling is the result of a locally optimal strategy that maximizes the expected precision for the next measurement.

Online Hamiltonian learning [6] offers a promising strategy for addressing drifts in stochastic qubit parameters through real-time estimation [6–13], facilitated by modern field-programmable gate array (FPGA) hardware advancements. Although various estimation methods [14–17] have been proposed to enhance calibration efficiency, no experimental implementation has yet achieved real-time adaptive estimation in a resonantly driven qubit with a parametrized probability distribution. This work fills the gap by the experimental demonstration of a real-time Bayesian-estimation approach applied to a superconducting qubit.

To demonstrate the adaptive and efficient calibration protocol, we employ a flux-tunable transmon qubit [18,19]. Transmons are largely insensitive to charge noise by design. The tunable qubit frequency is controlled by a magnetic field and it is susceptible to random flux variations, known as flux noise. Typically, flux-tunable transmons are thus operated at bias points (sweet spots) that are first-order insensitive to small flux changes. However, operating them away from these bias points may be necessary to reduce frequency crowding in large qubit arrays [20,21] or minimize coupling to two-level fluctuators. To maintain high-fidelity operation despite the increased sensitivity to flux noise, one may use real-time frequency estimation and stabilization [22].

The so-called frequentist approach is the most common method for estimating fluctuations in Hamiltonian parameters of superconducting qubits. Frequentist estimation methods infer an observable from the observed frequencies of measurement outcomes in repeated experiments and they do not yield an exponential scaling with the number of measurements. For instance, in Ref. [22], the qubit-frequency fluctuation is computed in real time by repeated Ramsey-interferometry cycles. In the probing Ramsey cycle, the Bloch vector is initialized on the equator of the Bloch sphere, evolves under its Hamiltonian, and is projected back onto the $\hat{z}$ axis after a fixed (*nonadaptive*) evolution time. The measurement outcome is averaged over many probing cycles. The fraction of times the system is measured in, e.g., the excited state is linearly mapped to the qubit-frequency fluctuation, as further detailed in the next section. The drawbacks of such frequentist methods are that (i) they are not optimally efficient, (ii) there is a trade-off between frequency sensitivity and probing range, and (iii) the qubit must already be approximately calibrated for the stabilization to work. In contrast, a Bayesian-estimation approach assigns probabilities to hypotheses and naturally incorporates real-time optimization techniques [6,8,9,12,13]. At any given time, the current knowledge of the parameters can be used to *adaptively* select the optimal experimental settings for the next probing cycle. While such techniques have been widely adopted in nitrogen-vacancy centers and semiconductor spin qubits over the past decade, they have not been similarly applied to superconducting qubits.

In this work, we employ an adaptive Bayesian-estimation scheme to calibrate a transmon qubit, which circumvents the limitations of the frequentist approach mentioned above. Our protocol involves programming a commercial controller with an integrated FPGA to generate dynamic microwave pulses, enabling real-time qubit-frequency calibration through a binary-search algorithm inspired by Ref. [23]. We use the real-time capabilities of the controller (on the qubit coherence timescale) to select optimal parameters within an adaptive estimation cycle. Specifically, the controller dynamically adjusts the drive frequency and the evolution time of each probing cycle based on previous measurement results. We demonstrate the efficacy of the algorithm through improved effective qubit coherence T_2^* and fidelity.

While the majority of theoretical research concentrates on Markovian noise [24], non-Markovian noise sources such as $1/f$ flux noise, which exhibit memory effects, are common in solid-state qubit platforms [13]. They introduce significant overhead for error mitigation [25,26]; their dynamical decoupling [27,28] is not universally effective and may not align with specific experimental goals. In order to realize fault-tolerant quantum computing with error-corrected solid-state qubits, non-Markovian noise likely needs to be reduced or eliminated [29]. We show how our real-time frequency tracking protocol reduces such noise as validated by gate-set tomography (GST) [30], whereas Ref. [22] relied solely on randomized benchmarking (RB) [31]. While RB blends the non-Markovian noise with other error types by introducing random gate sequences, GST assumes Markovian noise and can therefore detect non-Markovianity in the system through model violations. Our scheme enables intermittent and efficient calibration of qubit frequencies, making it ideal for stable quantum circuit execution in the presence of drift.

II. METHOD

A. Device

We use a 2×2 superconducting qubit array operated at the mixing-chamber stage (below 30 mK) of a dilution refrigerator. The array design is based on Ref. [32] and we calibrate one of its flux-tunable transmon qubits. A commercial controller (Quantum Machines OPX+ and Octave) applies high-frequency waveforms to the lines for qubit control and single-shot readout. A Yokogawa GS200 provides the dc flux bias through a bias tee at room temperature (for details on the experimental setup, see the Supplemental Material [33]). The transmon comprises a dc superconducting quantum interference device (SQUID)

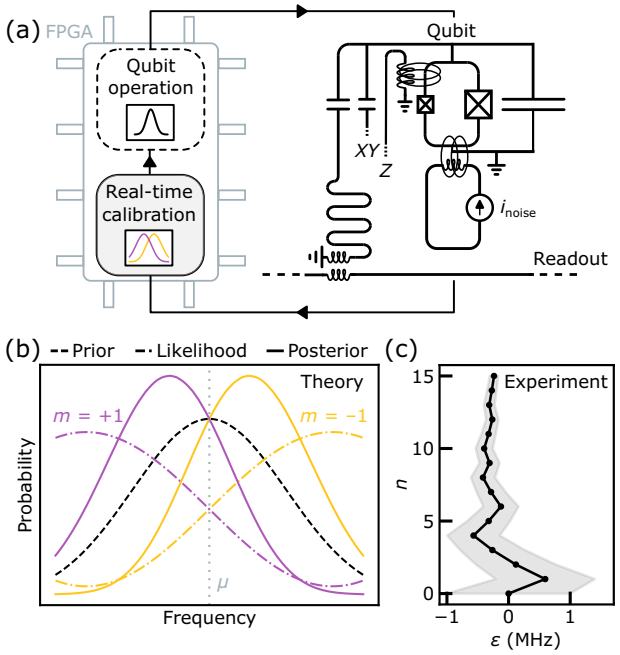

FIG. 1. Qubit implementation and frequency binary search. (a) The experiment schedule, alternating between periods of quantum information processing (dashed box) and periods for efficiently calibrating the qubit frequency (gray box). The transmon ("Qubit") is controlled by microwave pulses (XY) that are based on previous qubit measurements ("Readout"). (b) The evolution of the probability distribution $\mathcal{P}(\varepsilon)$ during the frequency binary-search algorithm. For each Ramsey probing cycle, the probing time and frequency detuning are chosen such that the two possible likelihood functions (dashed-dotted lines) multiplied with the prior distribution (dashed) yield posterior distributions (solid) that are shifted left or right, while the standard deviation is reduced (see the text). (c) The controller updates the probability distribution after each measurement. The dots show the resulting expectation value $\langle\varepsilon\rangle$ and the shaded area marks the 68% credible interval.

with asymmetric junctions, and on-chip Z and XY control lines [see Fig. 1(a)]. The qubit is dispersively coupled to a coplanar waveguide resonator for state readout [18,19].

The real-time calibration that we develop here can be embedded in an operation loop as sketched in Fig. 1(a). The calibration protocol is used to estimate the qubit frequency, after which the resulting estimate is used to adjust the frequency of the drive accordingly, leading to a decreased sensitivity to flux noise during qubit operation.

Our goal is to compensate for qubit-frequency fluctuations particularly when it is operated away from the sweet spot. Thus we tune the transmon where the effect of flux noise is the strongest, at $\Phi_{\text{ext}} = \Phi_0/4$, where $\Phi_0 = h/(2e)$ is the superconducting magnetic flux quantum and Φ_{ext} is the external magnetic flux applied via a mutual coupling to the Z line. Our qubit nominally has a transition frequency $f_{\text{q}} \approx 3.78\,\text{GHz}$ and in the rotating frame its Hamiltonian is

$$\frac{\mathcal{H}(t)}{h} = -\frac{\Delta f - \varepsilon(t)}{2}\sigma_z, \tag{1}$$

where σ_z is the Pauli z matrix, $\Delta f = f_{\text{d}} - f_{\text{q}}$ is the frequency detuning between f_{q} and the chosen rotating-frame frequency f_{d}, and $\varepsilon(t)$ represents the time-dependent shift of the qubit frequency due to flux noise.

B. Frequency binary search by Bayesian estimation

The frequency binary-search (FBS) algorithm involves a Ramsey probing cycle where the qubit is first prepared in a superposition state $|\psi\rangle = (|0\rangle - i|1\rangle)/\sqrt{2}$ using a $X_{\pi/2}$ rotation around the $\hat{x}$ axis of the Bloch sphere. The state preparation is followed by a period of free evolution for a duration τ, during which the qubit state acquires a phase $\phi = 2\pi\int_0^\tau \delta_{\text{q}}(t')\,\text{d}t'$, where $\delta_{\text{q}}(t) = \Delta f - \varepsilon(t)$. A second $X_{\pi/2}$ pulse is then applied and the state of the qubit is measured using dispersive readout. By the quasistatic noise approximation, we assume $\varepsilon(t)$ to be constant on the scale of a few probing cycles, resulting in $\phi = 2\pi[\Delta f - \varepsilon(t)]\tau$. In the following, we drop the time dependence of $\varepsilon(t)$ for ease of notation.

We assume that the probability of measuring an outcome $m \in \{-1, 1\}$ corresponding to the states $|0\rangle$ and $|1\rangle$ is given by the likelihood function

$$P(m|\varepsilon, \Delta f, \tau) = \frac{1}{2} + \frac{m}{2}\{\alpha + \beta e^{-\tau/T}\cos[2\pi(\Delta f - \varepsilon)\tau]\}, \tag{2}$$

where the parameters $\alpha = -0.02$ and $\beta = 0.6$ are coefficients that, together, capture state preparation and readout errors, while T describes a coherence time that limits the duration of the evolution time τ (for further details on how they impact the algorithm performance, see the Supplemental Material [33]). We set $T = 10\,\mu\text{s}$ based on the Hahn-echo coherence time $T_{2\text{E}} \approx 12\,\mu\text{s}$. We choose a value slightly lower than the estimated $T_{2\text{E}}$ to be on the safe side in case the intrinsic decoherence rate fluctuates on longer timescales. For reference, $T_2 \approx 42\,\mu\text{s}$ at $\Phi_{\text{ext}} = 0$. The qubit fluctuation ε is what we want the controller to estimate.

We emphasize that in this experiment, the controller *adaptively* computes in real time both the Ramsey evolution time τ and the drive frequency f_{d} (and thus Δf) for each probing cycle. This differs from previously employed approaches in which (i) the evolution time τ is fixed [22] and (ii) only either the evolution time τ *or* the effective rotating-frame frequency is chosen adaptively [8,12,34,35].

Adaptive control of the phase theoretically allows for one bit of information to be gained per experiment in an error-free qubit. The estimation begins with the longest

sensing time, which corresponds to the smallest frequency sensitivity [36], and proceeds with shorter sensing times that yield larger frequency resolutions. To improve robustness against experimental errors, the strategy can be modified by repeating measurements with reduced information gain per measurement [37,38], which has been experimentally demonstrated using nitrogen-vacancy centers [34, 39,40]. However, this phase-estimation approach presents some limitations: (i) an extensive parameter search is required to determine the optimal number of repetitions for each sensing time [34]; (ii) a Bayesian approach would yield a multimodal probability distribution that requires many more than two parameters to be stored in memory; and (iii) the estimate becomes only useful after the very last cycle.

We also highlight that using a fixed evolution time τ in nonadaptive probing cycles [22] limits the frequency range to $(-1/2\tau, +1/2\tau]$, which results in a trade-off between range and frequency sensitivity. This trade-off does not apply to the binary-search algorithm introduced here.

In the quasistatic approximation, we want to find the optimal sequence of τ and f_{d} values to estimate ε in as few measurements as possible. Using Bayes' rule, we write

$$\mathcal{P}_{n+1}(\varepsilon) \propto \mathcal{P}_n(\varepsilon) P(m_{n+1}|\varepsilon, \Delta f_{n+1}, \tau_{n+1}), \tag{3}$$

where $\mathcal{P}_n(\varepsilon)$ denotes the probability distribution for ε after the nth probing cycle, which thus depends on all previously used Δf_i and τ_i and all measurement outcomes m_i (i.e., $i \leq n$).

In the following, we assume $\mathcal{P}_n(\varepsilon)$ to be a Gaussian, which well approximates the actual distribution (for more details, see the Supplemental Material [33]). The advantage of using the Gaussian approximation is that it allows the distribution to be conveniently described using just two parameters: the mean μ and the standard deviation σ.

For each individual probing cycle, to minimize the expected posterior variance, in a greedy approach the optimal experiment $(\tau, \Delta f)$ is the one that divides the prior distribution into equal (or approximately equal) left and right portions [see Fig. 1(b)], so that a half period of Eq. (2) is comparable to the width of the prior distribution, performing something similar to a binary-search partitioning [41]. The approach of using Gaussians and partitioning the posterior distributions in a similar fashion has been introduced by Ref. [23]. However, in that case, the lack of phase control in the likelihood prevents the algorithm from determining the sign of ε, making it ineffective for priors where $\mu/\sigma \to 0$ [42]. While fitting to a bimodal Gaussian has been suggested as a solution [43], this approach still does not resolve the ambiguity of the sign. In this work, we address both aspects by dynamically updating f_{d} (and thus Δf) in real time on the controller. This allows the controller to consistently use the optimal likelihood function and perform the most efficient estimation using only a Gaussian distribution.

Two steps are needed to implement the binary search in the controller. The first step is to determine optimal experiment parameters τ and Δf based on the prior distribution. Then, one must approximate the resulting posterior $\mathcal{P}_{n+1}(\varepsilon)$ to a Gaussian. This can be done using equations implemented directly on the controller in real time.

The prior distribution [see the dashed black line in Fig. 1(b)] in the Gaussian approximation is given by

$$\mathcal{P}_n(\varepsilon) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left\{-\left(\frac{\varepsilon - \mu_n}{\sqrt{2}\sigma_n}\right)^2\right\}, \tag{4}$$

where μ_n is our guess for what ε is with uncertainty σ_n. In order to separate the prior at the center by the likelihood functions, the detuned frequency must satisfy

$$\Delta f_{n+1} = \frac{1 + 2l}{4\tau_{n+1}} + \mu_n, \tag{5}$$

where l can be any integer number; we set it to be zero. While this is formally the optimal solution for $\alpha = 0$, it is a reasonable approximation for the optimum at other realistic values of α. Given this constraint, the optimal evolution time τ that minimizes the posterior variance is given by

$$\tau_{n+1} = \frac{\sqrt{16\pi^2\sigma_n^2 + 1/T^2} - 1/T}{8\pi^2\sigma_n^2}. \tag{6}$$

After measuring $m_{n+1} = \pm 1$, the posterior distribution is obtained by using Eq. (3), after which it is fitted to a Gaussian using the method of moments, i.e., by calculating its mean and variance [23],

$$\mu_{n+1} = \mu_n - \frac{2\pi m_{n+1}\beta\sigma_n^2\tau_{n+1}e^{-\tau_{n+1}/T - 2\pi^2\sigma_n^2\tau_{n+1}^2}}{1 + m_{n+1}\alpha}, \tag{7a}$$

$$\sigma_{n+1}^2 = \sigma_n^2 - \frac{4\pi^2\beta^2\sigma_n^4\tau_{n+1}^2 e^{-2\tau_{n+1}/T - 4\pi^2\sigma_n^2\tau_{n+1}^2}}{(1 + m_{n+1}\alpha)^2}, \tag{7b}$$

and using these values to construct the new prior to be a Gaussian distribution with mean and variance μ_{n+1} and σ_{n+1}^2 [see Fig. 1(b)]. This scheme is repeated N times to obtain a sufficiently narrow distribution with exponential scaling, with the number of measurements given by Eq. (7b) [44].

In Fig. 1(c), we illustrate the resulting evolution of $\mathcal{P}_n(\varepsilon)$ as a function of the measurement number n for one representative estimation sequence, with an initial prior with $\sigma_0 = 1\,\mathrm{MHz}$ and $\mu_0 = 0$, and $N = 15$. The black dots show the expectation value $\langle\varepsilon\rangle$ and the shaded area indicates its 68% credible interval, narrowing here down to 67 kHz after 15 measurements in less than 100 μs. With

similar parameters, this efficient frequency estimation could replace currently used periodic calibration routines by narrowing the qubit frequency within a user-defined error budget. As the FBS relies only on single-qubit operations, it can be extended to multiqubit systems without introducing additional overhead.

In the following, we optimize the number of single-shot measurements N that we use based on the specific experiment we are performing. We test the performance of the FBS by performing Ramsey repetitions to verify that it finds the correct qubit frequency and we use RB and GST to assess the impact of the FBS on the qubit fidelity and drift. We emphasize again that the goal of these experiments is not to increase qubit coherence time or fidelity as much as possible, but simply to confirm that the information gained from the FBS is most likely correct.

III. RESULTS

A. Qubit coherence time T_2^*

We thus validate the FBS by rapid estimation of the shift of the qubit frequency ε and by demonstrating an extended coherence of the flux-tunable transmon qubit.

The fluctuating parameter ε is estimated from the probing sequence shown in the top part of Fig. 2(a). For each probe cycle, the Bloch vector is positioned on the equator using an $X_{\pi/2}$ pulse. After precessing for a time τ_i, a virtual $Z(\Delta f_i)$ gate [45] is performed, introducing a phase offset $\phi = 2\pi\,\Delta f_i \tau_i$. Here, virtual means that the phase offset is added to the subsequent $X_{\pi/2}$ pulse, which projects the Bloch vector back onto the $\hat{z}$ axis. The virtual Z gate controls the phase $\Delta f\,\tau$ of the likelihood function [Eq. (2)] based on Eq. (5). This is followed by a measurement, after which the qubit state s_i, ground ($s_i = 0$) or excited ($s_i = 1$), is assigned by thresholding the demodulated dispersive readout signal on the controller.

The Bayesian probability distribution of ε is updated after comparing the measurement outcome s_i to the previous one s_{i-1}, as we do not initialize the qubit to the ground state at the beginning of each cycle to reduce the probing cycle period (thus yielding higher feedback bandwidth). It follows that in Eq. (2), $m_i = 2|s_i - s_{i-1}| - 1$. For instance, if in the previous measurement the qubit was in the ground state $s_{i-1} = 0$ and now it is in the excited state ($s_i = 1$), then $m_i = +1$. The duration of the probe cycle is short compared to the measured $T_1 \approx 80\,\mu$s at quarter flux: the average Ramsey evolution time is approximately $4.36\,\mu$s, the time used to read out the qubit is $1.44\,\mu$s, and the time used to subsequently cool down the resonator again (to deplete it from any residual photons after the readout pulse) is approximately $2\,\mu$s.

The controller is programmed to start from a prior with $\mu_0 = 0$ and $\sigma_0 = 30\,$kHz, based on previously measured qubit-frequency fluctuations. After each measurement, the controller updates the probability distribution, resulting in a narrowing of the width of the distribution [cf. Fig. 1(c)]. In the middle panel of Fig. 2(a), we plot the experimentally found final posterior probability distributions $\mathcal{P}_N(\varepsilon)$ from all estimations done in the full experiment, where we consistently use $N = 8$ measurements per estimation sequence. For plotting clarity, we down-sample to approximately 14 ms.

To confirm that these narrowed distribution functions indeed represent more accurate and correct knowledge about ε, we interleave the estimations with a series of Ramsey repetitions, half of them making use of the outcome of the estimations and half of them not, as shown schematically in the upper part of Fig. 2. After each estimation sequence of ε, the controller updates the qubit-frequency parameter f_q (in software) to compensate for the estimated shift $\langle\varepsilon\rangle$, so that the total expected detuning becomes $\Delta f_j - \langle\varepsilon\rangle = 1\,$MHz. It then performs one Ramsey cycle with evolution time τ_j, after which it again estimates ε. The intentional detuning of 1 MHz applied via the virtual Z gate improves the visibility of the Ramsey fringes, making them easier to fit. This cycle is repeated $M = 50$ times, while τ_j is increased linearly from 0 to $7\,\mu$s. Each row in the middle panel of Fig. 2(b) shows the result of one such set of 50 Ramsey cycles, where we plot all single-shot measurement outcomes p_j as white or black pixels. After each set of $M = 50$ Ramsey cycles with feedback, the controller resets f_q (in software) to the offline-calibrated value, tuning to $\Delta f_k = 1\,$MHz. It then performs the same series of 50 Ramsey cycles as before, linearly stepping the evolution time τ_k from 0 to $7\,\mu$s. The rows in the middle panel of Fig. 2(c) show the single-shot measurement outcomes of this part of the protocol. After completing the Ramsey cycles without feedback, the whole protocol is repeated, as indicated in the top of Fig. 2, starting with a new estimation of ε. In the new estimation sequence, the controller starts from a prior with μ_0 equal to the estimated ε of the previous sequence.

To highlight the effect of the feedback, we plot the averages of all Ramsey repetitions in the lower panels of Figs. 2(b) and 2(c). In both cases, we fit the decaying signal using Gaussian envelopes (solid line), yielding $T_2^* = 3.73(0.11)\,\mu$s without feedback and $T_2^* = 5.57(0.09)\,\mu$s with feedback, corresponding to an approximately 49% improvement when including feedback. This increased coherence of the qubit demonstrates that the narrowed distribution function obtained from the estimation procedure indeed reflects more accurate knowledge about the fluctuating parameter ε. Notably, our approach requires fewer single-shot measurement outcomes ($N = 8$) compared to $N = 20$ of Ref. [22]. The T_2^* coherence time achieved with feedback is ultimately limited by the Hahn-echo value due to the lower bandwidth of the feedback loop (approximately 16 kHz for $N = 8$), compared to the much higher bandwidth of the dynamically decoupled Ramsey ($1/\tau \gtrapprox 100\,$kHz). This explains why the Hahn

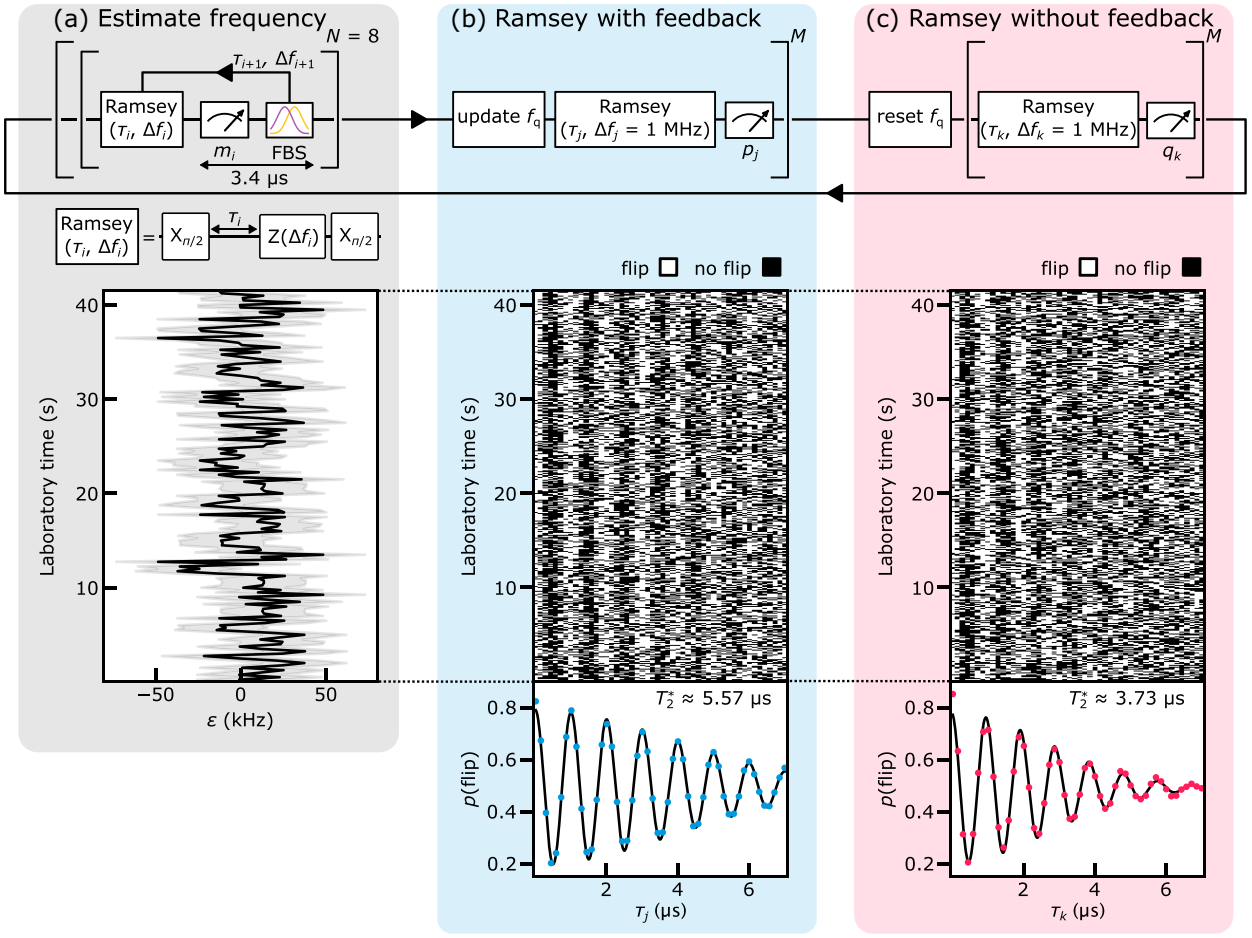

FIG. 2. Frequency binary-search validation by suppressed dephasing of a qubit in a feedback-controlled rotating frame. In the upper diagram, one loop (solid arrows) represents one repetition of the protocol. (a) For each estimate of the frequency shift ε, the controller uses $N = 8$ Ramsey probing cycles with adaptive probe times τ_i and detuned frequencies Δf_i, computed in real time. The bottom panel shows $\langle\varepsilon\rangle$ (solid line) and the 68% credible interval (shaded area) of the final probability distribution $\mathcal{P}_8(\varepsilon)$ of all estimates performed during the approximately 41.5 s of the experiment. (b) After each estimation, we task the controller to perform a Ramsey cycle with evolution time τ_j, while setting $\Delta f_j = 1\,\mathrm{MHz}$ by adjusting f_d using the latest estimate $\langle\varepsilon\rangle$. These interleaved FBS estimations and Ramsey cycles are repeated $M = 50$ times with evenly spaced τ_j. Single-shot measurement outcomes p_j are plotted in the middle panel as white or black pixels and the fraction of flipped outcomes in each column is shown as a blue dot in the lower panel. (c) Subsequently, the controller resets all frequencies to the offline-calibrated values, i.e., assumes $\langle\varepsilon\rangle = 0$, and again performs $M = 50$ evenly spaced Ramsey cycles of set $\Delta f_k = 1\,\mathrm{MHz}$. The single-shot measurement outcomes q_k are plotted as white or black pixels in the middle panel and the fraction of flipped outcomes in each column is shown as a red dot in the lower panel. Comparing (b) to (c), the coherence time improves by approximately 49% with feedback.

echo is less sensitive to quasistatic fluctuations than the feedback-stabilized Ramsey experiment.

We perform another experiment for 6 h, from which we extract the noise power spectral density and show that the FBS also keeps track of the qubit frequency over longer timescales (see the Supplemental Material [33]). Overall, the results presented in this section show how the FBS can efficiently find and stabilize the qubit frequency.

B. Single-qubit gate fidelity

In this section, we further validate the FBS calibration by showing improvement of the single-qubit gate fidelity by RB [31].

The pulse sequence is shown in Fig. 3(a): the controller resets the qubit-frequency parameter f_q (in software) and it performs the FBS with $N = 15$ probing sequences, starting with $\mu_0 = 0$ and $\sigma_0 = 200\,\mathrm{kHz}$ [46]. At the end of the estimation, f_q is updated (in software) and an RB sequence of depth L_s is performed. An interleaved RB measurement without feedback follows by resetting the qubit frequency to the offline-calibrated value. The maximum circuit depth (L_max) is 2300 and the repetition is averaged 10 000 times. We implement derivative reduction by adiabatic gate (DRAG) for our 20-ns-long pulses to suppress leakage errors [18,19] in the RB experiment. As with the previous experiment, the qubit is not initialized in the ground state at the beginning of each cycle; the

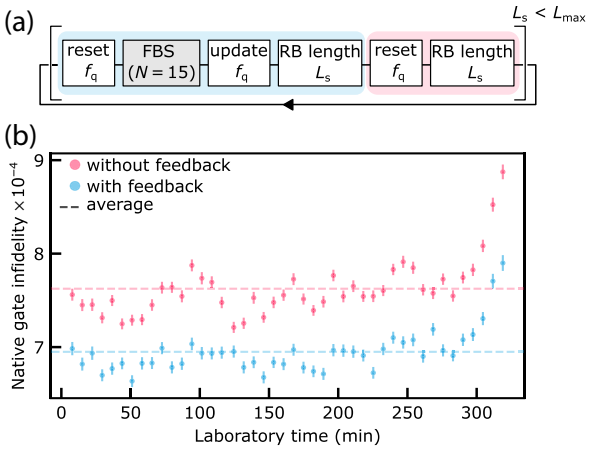

FIG. 3. Frequency binary-search validation by randomized benchmarking (RB). (a) An RB repetition of length L_s combined with the FBS feedback (with $N = 15$ probing cycles). (b) RB is repeated 44 times with the feedback (blue dots) and without (red dots). The native gate errors are extracted from the fit to the data, which consists of 10 000 realizations of 30 random Clifford sequences with and without the feedback. The error bars show the 68% confidence interval for the fitted gate error. The dashed lines are the averaged gate infidelities.

controller keeps track of whether or not the state is different compared to the previous measurement. Every 1000 averages, the threshold that classifies the demodulated dispersive readout signal is updated online in the controller by taking the average of 10 000 single-shot measurements after performing an $X_{\pi/2}$ pulse to the qubit.

The controller performs the RB experiment for 6 h, yielding a native gate infidelity of $(7.6 \pm 0.3) \times 10^{-4}$ without feedback and $(6.9 \pm 0.2) \times 10^{-4}$ with feedback [dashed lines in Fig. 3(b)]. The single-qubit gate infidelities are higher than the decoherence limit [47] approximated by $t_{\text{gate}}(1/T_1 + 1/T_\phi)/3 \approx 6 \times 10^{-4}$, given $t_{\text{gate}} \approx 20\,\text{ns}$, $T_1 \approx 80\,\mu\text{s}$, and the exponential part of the pure dephasing time $T_\phi \approx 13\,\mu\text{s}$. The feedback protocol always performs better than without feedback and with less spread around the mean value as a result of the stabilization. Some of the drifts of the infidelity remain correlated, which we tentatively attribute to other factors (e.g., changes in T_1). The improved fidelities of single-qubit gates by feedback by interleaved RB are presented in the Supplemental Material [33]. We attribute the smaller relative improvement in the single-qubit gate qubit fidelity, compared to coherence in the previous section, to the larger final value of $\sigma_{N=15} \approx 90\,\text{kHz}$, as opposed to $\sigma_{N=8} \approx 24\,\text{kHz}$, which resulted from different initial prior distributions.

C. Reduction of non-Markovian noise

So far, we have demonstrated an efficient qubit-calibration protocol by instructing a controller to generate adaptive probing sequences in real time. The calibration has been validated by improved coherence and fidelity. Next, we investigate if our estimation protocol also reduces the non-Markovian noise in the system, this time estimated by GST [30]. While GST has been employed to validate the mitigation of drifts in semiconductor spin qubits [13] by real-time estimation, we present its application for validating drift mitigation in a superconducting qubit by real-time feedback.

GST is a calibration-free method for benchmarking and characterizing operations in a quantum processor and we implement it by using the pyGSTi [48] software package. GST relies on running circuits designed to amplify certain types of errors. Each circuit, as shown in the right-hand panel of Fig. 4(a), consists of the gate to be characterized, $\{G_\text{I}, G_\text{x}, G_\text{y}, G_\text{x}G_\text{y}, G_\text{x}G_\text{x}G_\text{y}\}$ (the germs), sandwiched between two fiducial circuits from the set $\{I, X_{\pi/2}, Y_{\pi/2}, X_{\pi/2}X_{\pi/2}, X_{\pi/2}X_{\pi/2}X_{\pi/2}, Y_{\pi/2}Y_{\pi/2}Y_{\pi/2}\}$. These are applied to the qubit after initialization, to generate the state ρ_i, and before the measurement in the computational basis, to measure the operator E_j. The variable L_s is the maximum depth used to construct the base circuit from the germ [30]. By running the circuit multiple times, we extract the measurement probabilities $P_{ji} = \text{Tr}E_j\,G[\rho_i]$, which are then fitted to the error model, providing a maximum-likelihood estimator of the transfer matrix for all of the gates. The quality of the fit can be used to detect violations of the Markovian model of the noise [30,49].

We incorporate the feedback loop inside the GST protocol using the scheme shown in the left-hand panel of Fig. 4(a), implemented on the controller. We find suitable parameters for testing the FBS to be $N = 6$ per estimation sequence with feedback, with initial $\mu_0 = 0$ and $\sigma_0 = 30\,\text{kHz}$. Each repetition is executed 100 times for each of the 616 sequences constructed for GST. In the new estimation sequence, the controller starts from a prior with μ_0 equal to the estimated ε of the previous repetition.

The measured discrepancy between the Markovian model and the data ($2\Delta \log L_s$ as defined in Ref. [30]) is shown in Fig. 4(b), using the color of the squares: The gray scale is used to reflect the total discrepancy and a red color signals a model violation, indicating non-Markovian noise at a 95% confidence interval. The rows and columns of the inner matrices represent the fiducial operations used for state preparation and measurement, respectively. For the data with feedback, represented in the right-hand panel of Fig. 4(b), a clear decrease in the number of red squares in comparison to the feedback-free result of the left-hand panel is visible.

To further quantify the reduction of non-Markovian noise, the controller repeats the GST experiment 30 times, spanning a laboratory time of 40 min. In Fig. 4(c), we plot, as a function of the maximum length, the total violation of the model. While there is no clear improvement for maximum lengths of 1 and 2, it is evident for longer

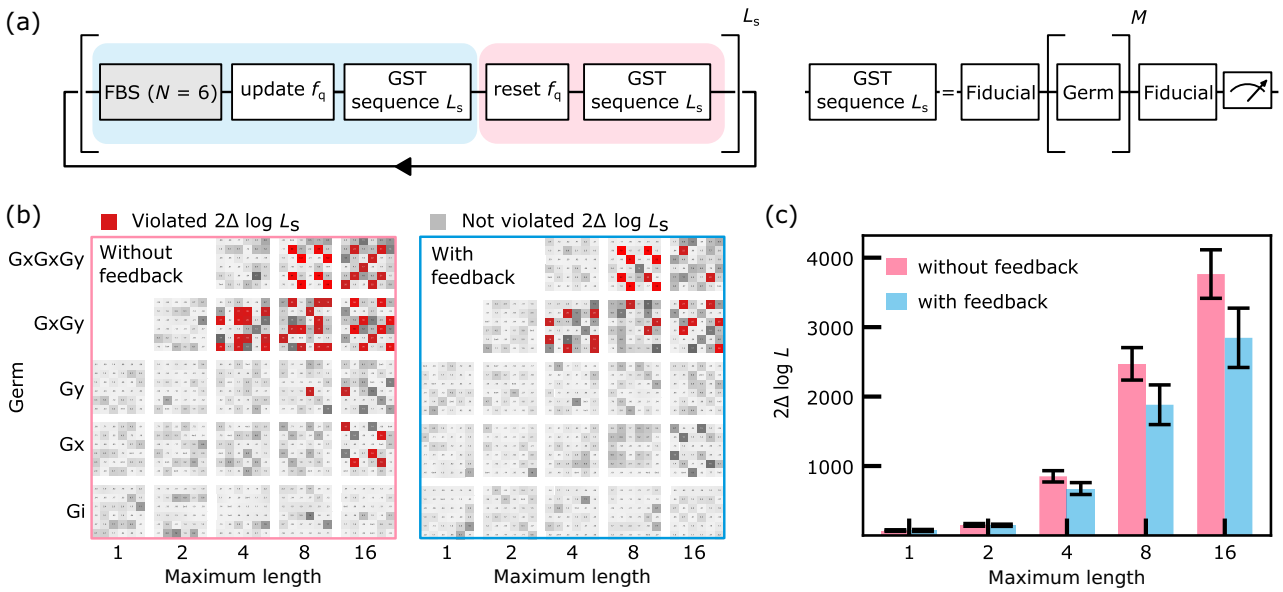

FIG. 4. Gate-set tomography (GST) and model violation. (a) A GST sequence combined with the FBS feedback (with $N = 6$ probing cycles). (b) A model violation plot without (left) and with (right) FBS feedback. The red marks reveal the detection of model violation at a confidence level of more than 95% and the gray boxes indicate statistical fluctuations. (c) The FBS feedback decreases the total amount of the log-likelihood ratios $2\Delta \log L$ at maximum lengths of 4, 8, and 16, extracted from 30 GST runs. The error bars show the 68% credible interval.

sequences. As shown in the Supplementary Material [33], feedback allows to significantly decrease the amplitude of non-Markovian noise, identified as frequency fluctuations below the inverse of the feedback timescale. It complements the previous analysis of GST data, which is consistent with the expectation that longer circuits are more sensitive to non-Markovian noise [30]. We attribute the residual discrepancy between the model and the data, especially where the G_x germ is used, to larger $|\varepsilon|$ fluctuations not captured by the chosen initial prior distribution $\mathcal{P}_0(\varepsilon)$. Further improvements would require a more quantitative analysis of the trade-off between the initially chosen uncertainty, the estimation bandwidth, and the white-noise floor of the estimation method (see the Supplemental Material [33]).

Regarding gate performance, the GST protocol yields an $X_{\pi/2}$-pulse infidelity of $(2.6 \pm 0.3) \times 10^{-2}$ and for the $Y_{\pi/2}$ pulse $(1.6 \pm 0.3) \times 10^{-2}$, both with and without feedback. These values are 2 orders of magnitude larger than those reported by RB, which further suggests the dominant role of coherent noise, to which RB is less sensitive [50]. In summary, the decrease in Markovian-model violation for longer GST sequences indicates that the FBS feedback protocol mitigates a significant portion of non-Markovian noise.

IV. CONCLUSIONS AND OUTLOOK

Our work presents the experimental demonstration of a greedy adaptive Bayesian-estimation protocol for the frequency of a resonantly driven qubit, improving its performance using only $N = 6$ single-shot measurements leveraging the exponential scaling of the algorithm with the number of measurements. This small number is to be compared with the tens [22] or hundreds of single-shot measurements of previous works [6,13,51] in resonantly driven qubits. Our approach reduces qubit-frequency drift by real-time Bayesian estimation and feedback, implemented via a low-latency FPGA-based qubit control system.

The binary-search estimation algorithm allows control pulses to compensate for qubit-frequency fluctuations caused by flux noise, improving coherence and fidelity without sacrificing frequency sensitivity or range. We validate the protocol using GST, which further corroborates our claim that our adaptive feedback loop reduces the effects of non-Markovian noise. This may facilitate quantum error-correction methods, which generally assume Markovian noise.

Our protocol assumes that a Gaussian distribution can approximate the frequency fluctuations and that these remain constant over the timescale of a few measurements. In this work, the controller updates the drive frequency in real time but the scheme can be adapted to adjust the qubit frequency by modifying the flux bias instead. The estimation bandwidth is limited by the relatively slow measurement time of a few microseconds. From this perspective, our work represents a worst-case scenario, underlying the efficacy of our experimental technique.

Traditional real-time approaches for continuous parameter estimations have relied on simple linear or

trigonometric functions or precomputed lookup tables [6,8,9,12,13,51]. In contrast, our FPGA implementation performs real-time Bayesian inference and processes widely used Gaussian distributions, resulting in a numerical complexity higher than previous works. At the same time, the parametrization of Gaussian distributions allows the FPGA to perform fewer and faster computations compared to traditional approaches using histograms and particle filtering.

We anticipate further improvements possibly by the implementation of a physics-informed prior [12] that accounts for $1/f$ noise. Also, the online calibration of single-qubit gates is a possible direction with available hardware, for single-qubit corrections in two-qubit gates by actual (not virtual) Z gates.

Adaptive Bayesian techniques could be implemented for frequency tracking in the presence of two-level fluctuators [52,53]. Specific to our flux-tunable transmon qubit, while the feedback bandwidth is limited by readout and resonator cooldown time, higher bandwidth can be achieved by adding a Purcell filter [18,19] which protects the qubit from relaxing into its environment. Our protocol may favor using symmetric junctions to increase the frequency range of the qubit, without worrying about increased sensitivity to flux noise as shown in Ref. [22].

Although qubit calibration by Bayesian inference can be relatively easily integrated with existing classical control hardware, it requires prior knowledge of the parametrization of the system. We expect that quantum model learning agents [6], which are more challenging to implement efficiently in real time, will gain wider adoption. We anticipate further research that merges theoretical and hardware advances, ultimately eliminating the need for a case-to-case modeling of experimental parameters.

Beyond superconducting qubits, our scheme offers new insights into real-time calibration of any qubits manipulated by resonant pulses. The protocol is not restricted to flux noise but is compatible with any source of low-frequency noise. We envision future applications in mitigating quasistatic electrical and nuclear noise in other solid-state qubit platforms.

This work advances quantum control by implementing an adaptive Bayesian technique to calibrate the qubit frequencies in real time. Our algorithm is a locally optimal solution by minimizing the expected estimator variance under a Gaussian distribution approximation, making it appealing for real-time calibration in large QPUs.

ACKNOWLEDGMENTS

We gratefully acknowledge Patrick Harrington and Max Hays for fruitful discussions. This work received funding from the European Union Horizon 2020 research and innovation program under Grant Agreements No. 101017733 (QuantERA II) and No. 951852 (QLSI), from EUREKA Eurostars 3 (ECHIDNA), from the European Research Council (ERC) (Grant No. 856526), from the Novo Nordisk Foundation under Challenge Programme NNF20OC0060019 (SolidQ), from the Inge Lehmann Programme of the Independent Research Fund Denmark, from the U.S. Army Research Office (ARO) under Award No. W911NF-24-2-0043, from the Norwegian University of Science and Technology (NTNU) QuSpin Mobility Grant, from the INTFELLES-Project No. 333990, which is funded by the Research Council of Norway (RCN), from the Dutch National Growth Fund (NGF) as part of the Quantum Delta NL program, and from the Danish Agency for Higher Education and Science (DAHES) (Grant No. 2076-00014B). It was also funded in part by the ARO Multidisciplinary University Research Initiative (MURI) W911NF-18-1-0218, in part by the ARO under Award No. W911NF-23-1-0045, and in part under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Air Force or the U.S. Government.

F.B., J.B., and J.A.K. conceptualized the experiment. F.B. led the measurements and data analysis, and wrote the manuscript, with input from all authors. F.B., L.P., M.M., R.A., A.C., J.A.G., W.D.O., and F.K. performed the experiment, with theoretical contributions from J.B., J.A.K., and J.D. The device was designed by Y.S. and fabricated by D.K.K. and B.M.N., under the supervision of K.S., M.E.S., and J.L.Y. The project was supervised by J.A.G., J.D., W.D.O., and F.K.

DATA AVAILABILITY

The data that support the findings of this paper are available from the corresponding author upon request.

[1] T. Ichikawa, *et al.*, Current numbers of qubits and their uses, Nat. Rev. Phys. **6**, 345 (2024).
[2] E. Campbell, A series of fast-paced advances in quantum error correction, Nat. Rev. Phys. **6**, 160 (2024).
[3] Google Quantum AI and Collaborators, Quantum error correction below the surface code threshold, Nature **638**, 920 (2025).
[4] H. Ball, W. D. Oliver, and M. J. Biercuk, The role of master clock stability in quantum information processing, npj Quantum Inf. **2**, 1 (2016).
[5] M. Mohseni, *et al.*, How to build a quantum supercomputer: Scaling challenges and opportunities, ArXiv:2411.10406.
[6] V. Gebhart, R. Santagati, A. A. Gentile, E. M. Gauger, D. Craig, N. Ares, L. Banchi, F. Marquardt, L. Pezzè, and C. Bonato, Learning quantum systems, Nat. Rev. Phys. **5**, 141 (2023).
[7] K. Reuer, J. Landgraf, T. Fösel, J. O'Sullivan, L. Beltrán, A. Akin, G. J. Norris, A. Remm, M. Kerschbaum, J.-C. Besse, *et al.*, Realizing a deep reinforcement learning agent

for real-time quantum feedback, Nat. Commun. **14**, 7138 (2023).

[8] M. J. Arshad, C. Bekker, B. Haylock, K. Skrzypczak, D. White, B. Griffiths, J. Gore, G. W. Morley, P. Salter, J. Smith, I. Zohar, A. Finkler, Y. Altmann, E. M. Gauger, and C. Bonato, Real-time adaptive estimation of decoherence timescales for a single qubit, Phys. Rev. Appl. **21**, 024026 (2024).

[9] F. Berritta, T. Rasmussen, J. A. Krzywda, J. van der Heijden, F. Fedele, S. Fallahi, G. C. Gardner, M. J. Manfra, E. van Nieuwenburg, J. Danon, A. Chatterjee, and F. Kuemmeth, Real-time two-axis control of a spin qubit, Nat. Commun. **15**, 1676 (2024).

[10] N. Dumoulin Stuyck, A. E. Seedhouse, S. Serrano, T. Tanttu, W. Gilbert, J. Y. Huang, F. Hudson, K. M. Itoh, A. Laucht, W. H. Lim, C. H. Yang, A. Saraiva, and A. S. Dzurak, Silicon spin qubit noise characterization using real-time feedback protocols and wavelet analysis, Appl. Phys. Lett. **124**, 114003 (2024).

[11] N. R. Vora, Y. Xu, A. Hashim, N. Fruitwala, H. N. Nguyen, H. Liao, J. Balewski, A. Rajagopala, K. Nowrouzi, Q. Ji, *et al.*, *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, (IEEE, Montreal, QC, Canada, 2024), pp. 414–415.

[12] F. Berritta, J. A. Krzywda, J. Benestad, J. van der Heijden, F. Fedele, S. Fallahi, G. C. Gardner, M. J. Manfra, E. van Nieuwenburg, J. Danon, A. Chatterjee, and F. Kuemmeth, Physics-informed tracking of qubit fluctuations, Phys. Rev. Appl. **22**, 014033 (2024).

[13] J. Park, H. Jang, H. Sohn, J. Yun, Y. Song, B. Kang, L. E. Stehouwer, D. D. Esposti, G. Scappucci, and D. Kim, Passive and active suppression of transduced noise in silicon spin qubits, Nat. Commun. **16**, 78 (2025).

[14] S. Kimmel, G. H. Low, and T. J. Yoder, Robust calibration of a universal single-qubit gate set via robust phase estimation, Phys. Rev. A **92**, 062315 (2015).

[15] R. D. McMichael and S. M. Blakley, Simplified algorithms for adaptive experiment design in parameter estimation, Phys. Rev. Appl. **18**, 054001 (2022).

[16] T. Hurant, K. Sun, Z. Jia, J. Kim, and K. R. Brown, Few-shot, robust calibration of single qubit gates using Bayesian robust phase estimation, ArXiv:2407.18339.

[17] B. de Neeve, A. V. Lebedev, V. Negnevitsky, and J. P. Home, Time-adaptive phase estimation, Phys. Rev. Res. **7**, 023070 (2025).

[18] 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).

[19] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, Circuit quantum electrodynamics, Rev. Mod. Phys. **93**, 025005 (2021).

[20] J. B. Hertzberg, E. J. Zhang, S. Rosenblatt, E. Magesan, J. A. Smolin, J.-B. Yau, V. P. Adiga, M. Sandberg, M. Brink, J. M. Chow, *et al.*, Laser-annealing Josephson junctions for yielding scaled-up superconducting quantum processors, npj Quantum Inf. **7**, 129 (2021).

[21] C. Berke, E. Varvelis, S. Trebst, A. Altland, and D. P. DiVincenzo, Transmon platform for quantum computing challenged by chaotic fluctuations, Nat. Commun. **13**, 2495 (2022).

[22] A. Vepsäläinen, R. Winik, A. H. Karamlou, J. Braumüller, A. D. Paolo, Y. Sung, B. Kannan, M. Kjaergaard, D. K. Kim, A. J. Melville, B. M. Niedzielski, J. L. Yoder, S. Gustavsson, and W. D. Oliver, Improving qubit coherence using closed-loop feedback, Nat. Commun. **13**, 1932 (2022).

[23] C. Ferrie, C. E. Granade, and D. G. Cory, How to best sample a periodic probability distribution, or on the accuracy of Hamiltonian finding strategies, Quantum Inf. Process. **12**, 611 (2013).

[24] The commonly used term *non-Markovian noise* refers to the nonunitary evolution of a quantum system that cannot be modeled by the Markovian master equation. Such non-Markovian decoherence arises from the finite memory of the environment, which introduces correlations between consecutive measurement outcomes.

[25] H. Hakoshima, Y. Matsuzaki, and S. Endo, Relationship between costs for quantum error mitigation and non-Markovian measures, Phys. Rev. A **103**, 012611 (2021).

[26] J. F. Kam, S. Gicev, K. Modi, A. Southwell, and M. Usman, Detrimental non-Markovian errors for surface code memory, Quantum Sci. Technol. **10**, 035060 (2025).

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

[28] P. Szańkowski, G. Ramon, J. Krzywda, D. Kwiatkowski, and Ł. Cywiński, Environmental noise spectroscopy with qubits subjected to dynamical decoupling, J. Phys.: Condens. Matter **29**, 333001 (2017).

[29] D. Pataki, Á. Márton, J. K. Asbóth, and A. Pályi, Coherent errors in stabilizer codes caused by quasistatic phase damping, Phys. Rev. A **110**, 012417 (2024).

[30] E. Nielsen, J. K. Gamble, K. Rudinger, T. Scholten, K. Young, and R. Blume-Kohout, Gate set tomography, Quantum **5**, 557 (2021).

[31] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Randomized benchmarking of quantum gates, Phys. Rev. A **77**, 012307 (2008).

[32] Y. Sung, L. Ding, J. Braumüller, A. Vepsäläinen, B. Kannan, M. Kjaergaard, A. Greene, G. O. Samach, C. McNally, D. Kim, A. Melville, B. M. Niedzielski, M. E. Schwartz, J. L. Yoder, T. P. Orlando, S. Gustavsson, and W. D. Oliver, Realization of high-fidelity CZ and *ZZ*-free iSWAP gates with a tunable coupler, Phys. Rev. X **11**, 021058 (2021).

[33] See the Supplemental Material at http://link.aps.org/supplemental/10.1103/77qg-p68k for the experimental setup, noise power spectral density and numerical simulations of the frequency binary-search estimation scheme. The Supplemental Material includes Refs. [54–57].

[34] C. Bonato, M. S. Blok, H. T. Dinani, D. W. Berry, M. L. Markham, D. J. Twitchen, and R. Hanson, Optimized quantum sensing with a single electron spin using real-time adaptive measurements, Nat. Nanotechnol. **11**, 247 (2015).

[35] T. Joas, S. Schmitt, R. Santagati, A. A. Gentile, C. Bonato, A. Laing, L. P. McGuinness, and F. Jelezko, Online adaptive quantum characterization of a nuclear spin, npj Quantum Inf. **7**, 56 (2021).

[36] A. M. Childs, J. Preskill, and J. Renes, Quantum information and precision measurement, J. Mod. Opt. **47**, 155 (2000).

[37] R. S. Said, D. W. Berry, and J. Twamley, Nanoscale magnetometry using a single-spin system in diamond, Phys. Rev. B **83**, 125410 (2011).

[38] P. Cappellaro, Spin-bath narrowing with adaptive parameter estimation, Phys. Rev. A **85**, 030301 (2012).

[39] N. M. Nusran, M. U. Momeen, and M. V. G. Dutt, High-dynamic-range magnetometry with a single electronic spin in diamond, Nat. Nanotechnol. **7**, 109 (2012).

[40] G. Waldherr, J. Beck, P. Neumann, R. S. Said, M. Nitsche, M. L. Markham, D. J. Twitchen, J. Twamley, F. Jelezko, and J. Wrachtrup, High-dynamic-range magnetometry with a single nuclear spin in diamond, Nat. Nanotechnol. **7**, 105 (2012).

[41] We clarify that this is not a true binary search in the sense of gaining exactly one bit of information per measurement. The search does, however, follow a binary-search tree where the two options at each node provide the most information within the approximations that we use. A true *quantum* binary search could be implemented for a decoherence-free qubit with ideal initialization and readout [36]. The focus here is on estimating noise in a physical qubit.

[42] In the limit $\mu/\sigma \to 0$, for $m = +1$, the likelihood function with zero phase (a squared cosine) has a single maximum at $\mu = 0$ within the 95% credible interval of the prior distribution and the corresponding posterior distribution can be approximated by a Gaussian. In contrast, for $m = -1$ in the same limit, the zero-phase likelihood function (a squared sine) exhibits two global maxima within the 95% credible interval and a minimum at the maximum prior value, resulting in a bimodal posterior distribution that cannot be adequately approximated by a single Gaussian.

[43] J. Benestad, J. A. Krzywda, E. van Nieuwenburg, and J. Danon, Efficient adaptive Bayesian estimation of a slowly fluctuating Overhauser field gradient, SciPost Phys. **17**, 014 (2024).

[44] Note that we do not claim exponential scaling with sensing time, which is fundamentally bounded by the Heisenberg limit $\sigma^2(\varepsilon) \sim 1/\tau^2$. Instead, our protocol achieves exponential scaling with the number of single-shot measurements N.

[45] 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).

[46] In the RB experiment, μ_0 is reset to zero before each estimation sequence, which explains the increased σ_0 and N compared to the Ramsey experiment of the previous section, where we do not reset μ_0 before each estimation sequence. We believe that further improvements in the single-qubit gate fidelity could be achieved by using the same settings as in the Ramsey experiment. Still, this work focuses on demonstrating the enhancement provided by FBS, rather than achieving the highest possible fidelity for this particular setup.

[47] P. J. J. O'Malley, *et al.*, Qubit metrology of ultralow phase noise using randomized benchmarking, Phys. Rev. Appl. **3**, 044009 (2015).

[48] E. Nielsen, K. Rudinger, T. Proctor, A. Russo, K. Young, and R. Blume-Kohout, Probing quantum processor performance with pyGSTi, Quantum Sci. Technol. **5**, 044002 (2020).

[49] A. Hashim, L. B. Nguyen, N. Goss, B. Marinelli, R. K. Naik, T. Chistolini, J. Hines, J. Marceaux, Y. Kim, P. Gokhale, *et al.*, A practical introduction to benchmarking and characterization of quantum computers, ArXiv:2408.12064.

[50] Y. R. Sanders, J. J. Wallman, and B. C. Sanders, Bounding quantum gate error rate based on reported average fidelity, New J. Phys. **18**, 012002 (2015).

[51] T. Nakajima, A. Noiri, K. Kawasaki, J. Yoneda, P. Stano, S. Amaha, T. Otsuka, K. Takeda, M. R. Delbecq, G. Allison, *et al.*, Coherence of a driven electron spin qubit actively decoupled from quasistatic noise, Phys. Rev. X **10**, 011060 (2020).

[52] B.-J. Liu, Y.-Y. Wang, T. Sheffer, and C. Wang, Observation of discrete charge states of a coherent two-level system in a superconducting qubit, Phys. Rev. Lett. **133**, 160602 (2024).

[53] F. Ye, A. Ellaboudy, and J. M. Nichol, Stabilizing an individual charge fluctuator in a Si/SiGe quantum dot, Phys. Rev. Appl. **23**, 044063 (2025).

[54] J. S. Rojas-Arias, Y. Kojima, K. Takeda, P. Stano, T. Nakajima, J. Yoneda, A. Noiri, T. Kobayashi, D. Loss, and S. Tarucha, The origins of noise in the Zeeman splitting of spin qubits in natural-silicon devices, ArXiv:2408.13707.

[55] B. Paquelet Wuetz, D. Degli Esposti, A.-M. J. Zwerver, S. V. Amitonov, M. Botifoll, J. Arbiol, A. Sammak, L. M. Vandersypen, M. Russ, and G. Scappucci, Reducing charge noise in quantum dots by using thin silicon quantum wells, Nat. Commun. **14**, 1385 (2023).

[56] E. Magesan, J. M. Gambetta, B. R. Johnson, C. A. Ryan, J. M. Chow, S. T. Merkel, M. P. da Silva, G. A. Keefe, M. B. Rothwell, T. A. Ohki, M. B. Ketchen, and M. Steffen, Efficient measurement of quantum gate error by interleaved randomized benchmarking, Phys. Rev. Lett. **109**, 080505 (2012).

[57] P. J. Huber, *Robust Statistics*, Wiley Series in Probability and Mathematical Statistics (J. Wiley & Sons, New York, 1981).