

# Enhancing Dispersive Readout of Superconducting Qubits through Dynamic Control of the Dispersive Shift: Experiment and Theory

François Swiadek<sup>1,2,\*</sup>, Ross Shillito,<sup>3</sup> Paul Magnard<sup>1</sup>, Ants Remm<sup>1,2</sup>, Christoph Hellings<sup>1,2</sup>, Nathan Lacroix<sup>1,2</sup>, Quentin Ficheux<sup>1</sup>, Dante Colao Zanuzzi<sup>1,2</sup>, Graham J. Norris,<sup>1,2</sup> Alexandre Blais,<sup>3,4</sup> Sebastian Krinner,<sup>1,2</sup> and Andreas Wallraff<sup>1,2,5</sup>

<sup>1</sup>Department of Physics, ETH Zurich, 8093 Zurich, Switzerland

<sup>2</sup>Quantum Center, ETH Zurich, 8093 Zurich, Switzerland

<sup>3</sup>Institut Quantique and Département de Physique, Université de Sherbrooke, Sherbrooke Quebec J1K 2R1, Canada

<sup>4</sup>Canadian Institute for Advanced Research, Toronto, Ontario M5G 1M1, Canada

<sup>5</sup>ETH Zurich—PSI Quantum Computing Hub, Paul Scherrer Institute, 5232 Villigen, Switzerland



(Received 15 August 2023; revised 15 July 2024; accepted 16 September 2024; published 20 November 2024)

The performance of a wide range of quantum computing algorithms and protocols depends critically on the fidelity and speed of the employed qubit readout. Examples include gate sequences benefiting from midcircuit real-time measurement-based feedback, such as qubit initialization, entanglement generation, teleportation, and, perhaps most importantly, quantum error correction. A prominent and widely used readout approach is based on the dispersive interaction of a superconducting qubit strongly coupled to a large-bandwidth readout resonator, frequently combined with a dedicated or shared Purcell filter protecting qubits from decay. By dynamically reducing the qubit-resonator detuning and thus increasing the dispersive shift, we demonstrate a beyond-state-of-the-art two-state-readout error of only 0.25% in 100-ns integration time. Maintaining low-readout-drive strength, we nearly quadruple the signal-to-noise ratio of the readout by doubling the readout-mode line width, which we quantify by considering the hybridization of the readout resonator and its dedicated Purcell filter. We find excellent agreement between our experimental data and our theoretical model. The presented results are expected to further boost the performance of new and existing algorithms and protocols critically depending on high-fidelity fast midcircuit measurements.

DOI: 10.1103/PRXQuantum.5.040326

Realizing high-fidelity and fast single-shot readout of a qubit [1–3] is essential for quantum error-correction protocols [4–8] in which qubit decoherence during readout and reset contributes significantly to the logical error. It is also key for algorithms requiring real-time feedback, such as teleportation [9–12], distillation [13,14], and initialization [15–18].

In superconducting circuits, the most commonly used readout architecture employs the state-dependent dispersive shift of the resonance frequency of a resonator coupled to the qubit to infer the qubit state [19–21]. While the frequency of the resonator is typically fixed, flux-tunable transmons allow us to control the

qubit-resonator detuning by modifying the transmon frequency [22], enabling high-fidelity fast entangling gates [23–25], and avoiding frequency collisions. Additionally, each qubit is often coupled to a microwave transmission line via a dedicated [6,7,26] or common Purcell filter [8,27,28] to protect the qubit from radiative decay [29–31]. Such measurements are usually performed with weak measurement tones to avoid nonlinearities and detrimental qubit state transitions, although high-power readout has been studied both theoretically [32] and experimentally [2].

In the past few years, significant improvements to the single-shot readout have been realized, reaching a two-level readout assignment error of  $4 \times 10^{-3}$  in 88 ns [3]. Faster readout protocols have been realized, with a  $8 \times 10^{-3}$  and  $9 \times 10^{-3}$  assignment error achieved in 40 ns by utilizing a nonlinear Purcell filter [33] and the distributed-element multimode nature of the readout resonator [34], respectively.

One of the critical parameters governing dispersive qubit readout is the detuning between the qubit and the

\*Contact author: francois.swiadek@phys.ethz.ch

readout resonator, which controls both the magnitude of the dispersive shift and the nonlinearities induced in the resonator. Different detuning regimes have been explored, including cases in which the resonator frequency is *lower* than the qubit [3,35]. Notably, the measurement fidelity has been shown to improve for smaller detunings [7,36], although these observations have not been fully explained.

In this work, we demonstrate an increase in the signal-to-noise ratio (SNR) and assignment fidelity by bringing the qubit frequency closer to the readout-resonator frequency using a flux pulse [see illustration in Figs. 1(a) and 1(b)], achieving a minimum two-level readout error of  $2.5 \times 10^{-3}$  in 100 ns. We credit this remarkable performance not only to an increase in the dispersive shift  $\chi$  imparted by the qubit on the cavity but also to an increase in the effective line width of the targeted normal mode response, caused by bringing the Lamb-shifted readout resonator closer to resonance with the Purcell filter [see Fig. 1(c)].

## I. READOUT-PARAMETER CHARACTERIZATION

We perform the experiment with a transmon qubit of transition frequency  $\omega_q/2\pi = 4.14$  GHz at the lower-flux

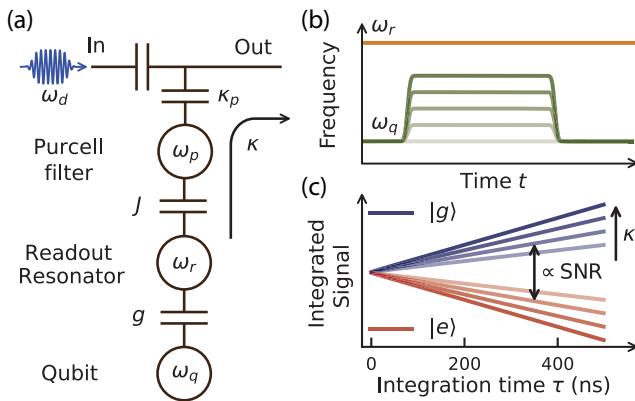


FIG. 1. (a) A schematic of a qubit coupled to a readout-resonator-Purcell-filter system. The qubit, of transition frequency  $\omega_q$ , is coupled capacitively at rate  $g$  to a readout resonator of frequency  $\omega_r$ . The readout resonator is, in turn, coupled at rate  $J$  to a Purcell filter of frequency  $\omega_p$ , which is coupled to a feed line at rate  $\kappa_p$ . We probe the system by measuring the transmission of a readout pulse at frequency  $\omega_d$  through the feed line. The effective decay rate of the readout resonator is indicated as  $\kappa$ . (b) A schematic of the time dependence of the qubit frequency  $\omega_q$  relative to the readout-resonator frequency  $\omega_r$ . The qubit, initially idling at the lower-flux sweet spot, is pulsed to a smaller detuning from the readout resonator using a fast Gaussian-filtered rectangular flux pulse. (c) An illustration of the rise of the SNR of the qubit readout with integration time  $\tau$ , parametrized by the effective line width of the readout resonator  $\kappa$  at approximately constant dispersive shift  $\chi$ . Increasing color saturation in (b) and (c) indicates increasing  $\kappa$  at reduced detuning between  $\omega_q$  and  $\omega_r$ .

sweet spot [37] and anharmonicity  $\alpha/2\pi = -181$  MHz. It has a lifetime  $T_1 = 30.4$   $\mu$ s and is capacitively coupled to a readout resonator with a coupling strength  $g/2\pi = 224$  MHz. The readout resonator is coupled to a feed line used for multiplexed readout [26] via a dedicated Purcell filter of line width  $\kappa_p$  and with a coupling strength  $J$  [see Fig. 1(a)]. The qubit is located on a device used to execute a distance-3 surface code (see Fig. 8 in Appendix A). Further information on the device properties and its fabrication can be found in Ref. [7].

To determine the readout parameters as a function of the frequency detuning between the qubit and its readout resonator, we perform pulsed-spectroscopy experiments. We first prepare the qubit in the ground state  $|g\rangle$  or excited state  $|e\rangle$ , pulse the qubit to a chosen readout frequency  $\omega_q$  using a baseband flux pulse, and probe the readout circuit using a 2.2- $\mu$ s-long microwave tone. This duration corresponds to the maximum integration time of our readout electronics (see Appendix A). The flux pulses are Gaussian-filtered ( $\sigma = 0.5$  ns) rectangular pulses with short rising and falling edges minimizing coupling to two-level systems [7] [see Fig. 1(b)].

We repeat the experiment for five different qubit-readout-resonator detunings  $\Delta_{qr}/2\pi$ , spanning  $-2.7$  GHz to  $-1.3$  GHz, where  $\Delta_{qr} = \omega_q - \omega_r^g$ . We denote  $\omega_r^{g/e}$  as the readout-resonator frequency with the qubit prepared in the ground or excited state. The measured (light-colored lines)

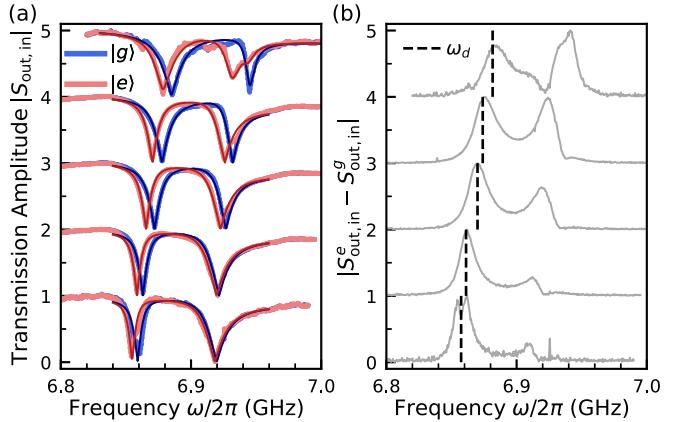


FIG. 2. (a) The readout circuitry transmission spectra measured for five qubit-readout-resonator detunings ( $\Delta_{qr}/2\pi \in [-2.7, -2.4, -1.9, -1.6, -1.3]$  GHz, from bottom to top, vertically shifted by increments of one). Spectra are shown for both the qubit prepared in the ground state  $|g\rangle$  (blue) and in the excited state  $|e\rangle$  (red). The solid lines are fits based on a coupled qubit-readout-resonator-Purcell-filter model (see Appendix A). (b) The corresponding differences of the complex-valued transmission responses  $|S_{\text{out,in}}^{(e)} - S_{\text{out,in}}^{(g)}|$  measured for the qubit prepared in the ground or excited state. The vertical dashed black lines indicate the selected readout frequency, chosen such that the response in transmission is maximum (gray) for the low-readout mode.

and calculated (dark-colored lines) transmission responses are shown in Fig. 2(a), with blue (red) lines corresponding to the qubit prepared in the ground (excited) state. From a fit to a coupled qubit-readout-resonator-Purcell-filter model [see Appendix A and the solid blue and red lines in Fig. 2(a)], we extract the relevant readout parameters for each value of  $\Delta_{qr}$ . The measured (dots) and calculated (lines) dressed-readout-resonator frequencies  $\omega_r^g$  and  $\omega_r^e = \omega_r^g + 2\chi$  are shown in Fig. 3(a) (blue and red circles) as a function of the detuning  $\Delta_{qr}$ , along with the Purcell-filter frequency  $\omega_p/2\pi = 6.900$  GHz, which remains constant. The variation in the resonator frequencies  $\omega_r^{g/e}$  is due to the Lamb shift  $g^2/\Delta_{qr}$  caused by the qubit [22]. Furthermore, we extract both a large intended Purcell-filter line width  $\kappa_p/2\pi = 34.5$  MHz and a large intended coupling strength between the readout resonator and the Purcell filter, of  $J/2\pi = 27.5$  MHz.

We consider a standard circuit-QED approach to modeling the transmon-resonator-Purcell-filter system (for details, see Appendix B). In the case of a weak drive  $\mathcal{E}$  applied to the filter mode, the readout resonator and the Purcell-filter responses can be considered as linear. As such, the dynamics can be effectively mapped to the equations of motion [38],

$$\begin{bmatrix} \dot{\alpha}^{g/e} \\ \dot{\beta}^{g/e} \end{bmatrix} = -i \begin{bmatrix} \omega_r^{g/e} & J \\ J & \omega_p - i\kappa_p/2 \end{bmatrix} \begin{bmatrix} \alpha^{g/e} \\ \beta^{g/e} \end{bmatrix} + \begin{bmatrix} 0 \\ \mathcal{E} e^{-i\omega_d t} \end{bmatrix}, \quad (1)$$

where  $\alpha$  and  $\beta$  represent the coherent fields of the readout resonator and Purcell filter, respectively. In the regime

$J \approx \kappa_p$ , we observe two distinct hybridized readout-resonator-Purcell-filter modes [see Fig. 2(a)]. We denote these as the *low*- and *high*-readout modes, respectively, the lowest and highest of the two modes in the transmission spectrum. The frequency and line width of these modes can be determined, respectively, from the real and the imaginary parts of the eigenvalues of the equations of motion in Eq. (1) in the absence of a drive:

$$\begin{aligned} \omega_{l,h}^{g/e} &= \frac{\omega_r^{g/e} + \omega_p}{2} \pm \frac{1}{2} \text{Re} \sqrt{\left( \Delta_{rp}^{g/e} + \frac{i\kappa_p}{2} \right)^2 + 4J^2}, \\ \kappa_{l,h}^{g/e} &= \frac{\kappa_p}{2} \mp \text{Im} \sqrt{\left( \Delta_{rp}^{g/e} + \frac{i\kappa_p}{2} \right)^2 + 4J^2}. \end{aligned} \quad (2)$$

The readout-resonator-Purcell-filter hybridization leads to a *distribution* of the total qubit-induced dispersive shift  $\chi$  on the readout-resonator-Purcell-filter system. Using the model in Eq. (2), we can extract the dispersive shifts of the *low* and *high* modes, respectively,  $\chi_{l/h} = (\omega_{l/h}^e - \omega_{l/h}^g)/2$  [see purple (green) circles for the *low* (*high*) mode in Fig. 3(b)]. While the total dispersive shift  $\chi = \chi_l + \chi_h$  shows the expected  $\alpha g^2/\Delta_{qr}^2$  dependence in Fig. 3(b) (solid black line), with  $2\chi/2\pi \in [-5.67, -19.49]$  MHz, the *low*-mode dispersive shift only shows small variations in that range, staying between  $2\chi_l/2\pi \in [-4.17, -6.69]$  MHz [see the solid purple line in Fig. 3(b)]. In contrast, the *high*-mode dispersive shift shows a similar scaling with  $\Delta_{qr}$  as the total dispersive shift, with  $2\chi_h/2\pi \in [-1.5, -13.18]$  MHz [solid green line in Fig. 3(b)].

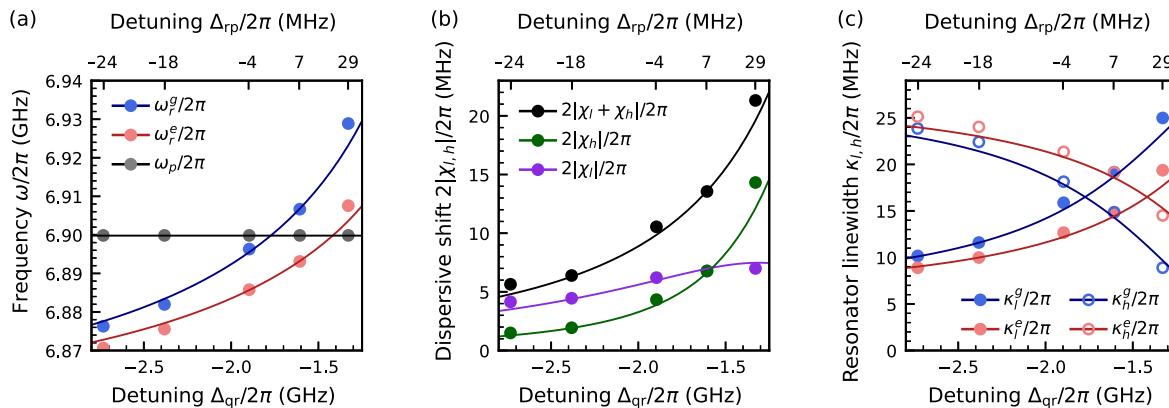


FIG. 3. (a) The readout-resonator frequency  $\omega_r^g$  ( $\omega_r^e$ ) conditioned on the qubit being prepared in the ground (excited) state and the Purcell-filter frequency  $\omega_p$  as a function of  $\Delta_{qr}$ . (b) The measured dispersive shift  $\chi_l$  ( $\chi_h$ ) of the lower- (higher-) frequency hybridized readout mode [see the green (purple) points], as a function of  $\Delta_{qr}$ . The two contributions sum up to the bare-readout-resonator-mode dispersive shift ( $\chi = \chi_l + \chi_h \approx \alpha g^2/\Delta_{qr}^2$ ). The solid lines (purple and green) are fits based on a qubit-readout-resonator-Purcell-filter model (see Appendix A). (c) The measured effective readout-resonator line width  $\kappa_l$  ( $\kappa_h$ ) for the lower- (higher-) frequency hybridized readout mode for the qubit prepared in the ground (excited) state [see blue (red)], as a function of  $\Delta_{qr}$ . The solid lines are fits based on a qubit-readout-resonator-Purcell-filter model (see Appendix A). In (a)–(c), the corresponding detunings  $\Delta_{rp} = \omega_r^g - \omega_p$  between the Purcell filter and the readout resonator, indicating the degree of hybridization of the two resonator modes, are shown on the top axis.

We observe that the dispersive shift of the *low* mode is dominant for qubit-readout-resonator detunings below  $-1.6$  GHz, after which the dispersive shift of the *high* mode becomes larger. The crossing point at which  $\chi_l = \chi_h$ , in the vicinity of the qubit-readout-resonator detuning  $\Delta_{qr}/2\pi = -1.6$  GHz, corresponds to an equal hybridization of the two readout modes. It coincides with  $\omega_p = \omega_r^g + \chi = \omega_r^e - \chi$  being equidistant to the ground- and excited-state responses of the readout resonator [see Fig. 3(a)].

Our model also gives us valuable information about the line width of the low and high modes, for the qubit prepared in the ground  $|g\rangle$  or excited state  $|e\rangle$ ; namely,  $\kappa_l^g$  ( $\kappa_h^g$ ) and  $\kappa_l^e$  ( $\kappa_h^e$ ), as a function of the qubit-readout-resonator detuning  $\Delta_{qr}$ . As shown in Fig. 3(c), while  $|\chi_l| > |\chi_h|$  for  $\Delta_{qr}/2\pi \leq -1.6$  GHz, we have  $\kappa_l^{g/e} < \kappa_h^{g/e}$ . This is expected, as for  $\Delta_{qr}/2\pi \leq -1.6$  GHz, the *low* mode has a larger weight in the readout resonator. The difference between  $\kappa_l^g$  and  $\kappa_l^e$  for the *low* mode derives from the frequency detuning  $\Delta_{rp}^{g/e} = \omega_r^{g/e} - \omega_p$ , between the readout-resonator frequency and the Purcell-filter frequency. In particular, for the *low* mode,  $\kappa_l^g > \kappa_l^e$  for all detunings, while for the *high* mode,  $\kappa_h^g < \kappa_h^e$ , which can be seen from the analysis of the normal-mode Hamiltonian in Appendix B.

In the vicinity of the detuning leading to an equal hybridization of the *low* and *high* modes,  $\Delta_{qr}/2\pi \approx -1.6$  GHz, we further note that all  $\kappa_h^e/2\pi \approx \kappa_l^g/2\pi \approx 19$  MHz and  $\kappa_h^g/2\pi \approx \kappa_l^e/2\pi \approx 14$  MHz. After this crossing point, we observe that while  $|\chi_l| < |\chi_h|$  for  $\Delta_{qr}/2\pi \geq -1.6$  GHz, we find  $\kappa_l^{g/e} > \kappa_h^{g/e}$ , which we exploit in Sec. II. The detailed parameters are summarized in Table I.

## II. SINGLE-SHOT READOUT

We perform single-shot readout for the qubit-readout-resonator detunings  $\Delta_{qr}/2\pi \in [-2.7, -2.4, -1.9, -1.6, -1.3]$  GHz (see Fig. 2) as a function of the readout-pulse power and integration time  $\tau \in [50, 100, 200, 300, 400]$  ns. Each experiment consists of  $10^4$  single-shot measurements with the qubit prepared in the ground or excited state. The detuning is varied by tuning the qubit to a chosen frequency  $\omega_q$ , using a flux pulse as described above. We use a rectangular readout pulse with a duration of 450 ns, convolved with a Gaussian filter of width  $\sigma = 0.5$  ns, and integrate the readout signal for a time  $\tau$  using mode-matched weights [39] to discriminate the ground- ( $|g\rangle$ ) and excited- ( $|e\rangle$ ) qubit-state responses. The flux pulse lasts longer than the readout pulse. In addition, we use a preselection readout to reduce the residual excited-state population of the qubit to below 0.1% [40].

We express the readout power as a function of the number of photons in the readout resonator  $n_g$  when the qubit is prepared in the ground state, relative to the critical number

of photons in the resonator  $n_{\text{crit}} = \Delta_{qr}^2/4g^2$  [19] at a given qubit-readout-resonator detuning  $\Delta_{qr}$ . By measuring the qubit-induced ac Stark shift  $\Delta_{ac} = 2g^2/\Delta_{qr}$  on the readout resonator at  $\Delta_{qr}/2\pi = -2.7$  GHz, we can infer the number of photons  $n_g$  in the resonator when the qubit is prepared in the ground state,  $n_g = \Delta_{ac}/2(\chi_l + \chi_h)$  (see Appendix D). The photon number  $n_g$  at other detunings and the photon number  $n_e$  when the qubit is prepared in the excited state for all detunings are inferred using semiclassical analysis (see Appendix B).

The readout drive frequency  $\omega_d$  is chosen such that the difference in the response in transmission when the qubit is prepared in the ground or excited state  $|S_{\text{out,in}}^{(e)} - S_{\text{out,in}}^{(g)}|$ , is maximum for the *low* mode [see the vertical black dashed lines in Fig. 2(b)]. Use of the *low* mode results in the largest signal on average and has a large value of  $\kappa_{l/h}^g$  relative to  $\chi_{l/h}$  over the range of qubit-readout-resonator detunings considered here, making it beneficial for fast readout. In the theoretical model (see Appendix C), this choice corresponds to selecting the drive frequency that leads to the largest steady-state displacement between the coherent *g*-state and *e*-state Purcell-filter-mode responses. This assumes a fixed weak drive power, such that the response is in the linear regime. We have found that this choice consistently leads to a stronger resonator response for the qubit being in the excited state ( $n_e > n_g$ ). We credit this to a smaller effective line width for the excited state for the lower mode,  $\kappa_l^e < \kappa_l^g$  [see Fig. 3(c)]. We find this to be an appropriate choice of drive frequency, as the Kerr nonlinearity imparted on the resonator is weaker for the excited state than for the ground state (see Appendix B).

From the acquired single-shot histograms (see example in Fig. 4), we extract the SNR in terms of power as [39]

$$\text{SNR} \equiv \left| \frac{\mu_g - \mu_e}{(\sigma_g + \sigma_e)/2} \right|^2, \quad (3)$$

where  $\mu_{g/e}$  and  $\sigma_{g/e}$  are, respectively, the mean and the standard deviation of the Gaussian distributions of the *g/e*-state responses. In Fig. 4, the solid black line indicates the distance between the means  $\mu_g$  and  $\mu_e$  and the radii of the black circles are given by the square root of the covariance-matrix diagonal elements of the bimodal Gaussian distribution.

We characterize the measurement by the average assignment error  $\varepsilon_a$  for two-state readout, limited by the overlap error between the Gaussian distributions and the qubit lifetime  $T_1$ , defined as [39]

$$\begin{aligned} \varepsilon_a &= 1 - \mathcal{F}_{g,e} \\ &= [P(e|g) + P(g|e)]/2 \\ &\gtrsim \frac{1}{2} \left[ 1 - \text{erf} \left( \sqrt{\text{SNR}/8} \right) \right] + \frac{\tau}{2T_1}, \end{aligned} \quad (4)$$

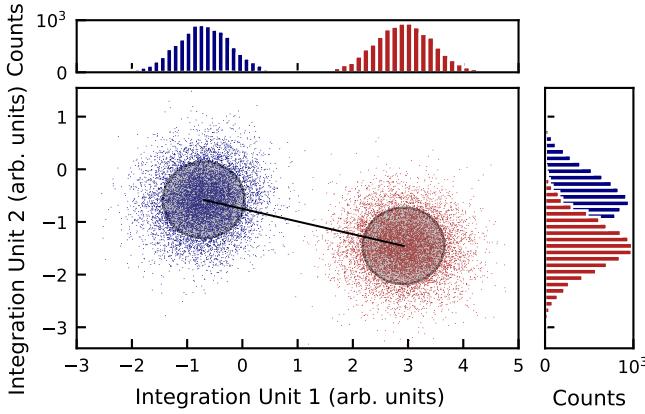


FIG. 4. The single-shot readout histogram for a qubit-readout-resonator detuning of  $\Delta_{qr}/2\pi = -1.3$  GHz, a readout integration time of  $\tau = 100$  ns, and  $n_g/n_{crit} = 0.93$ . We assign the measured state using a bimodal Gaussian-mixture model. The marginal distributions of this model along each axis are plotted along the corresponding axis. A solid black line indicates the distance between the means  $\mu_g$  and  $\mu_e$  of the Gaussian distributions of the ground- and excited-state responses. The square roots of the covariance-matrix diagonal elements of the Gaussian distributions  $\sigma_g$  and  $\sigma_e$  are used as the radii of the black circles.

where  $P(i|j)$  is the probability of measuring the state  $|i\rangle$  when having prepared the state  $|j\rangle$  and where the average two-state readout fidelity  $\mathcal{F}_{g,e}$  characterizes the quality of the readout. The factor of 2 present in the  $T_1$ -limit term arises from the fact that only  $P(g|e)$  is affected by loss events.

In Fig. 5(a), we present the lowest measured average assignment errors  $\varepsilon_a$  as a function of the qubit-readout-resonator detuning  $\Delta_{qr}/2\pi \in [-2.7, -2.4, -1.9, -1.6, -1.3]$  GHz and as a function of the readout integration times  $\tau \in [50, 100, 200, 300, 400]$  ns. We observe that  $\varepsilon_a < 1 \times 10^{-2}$  for  $\tau \geq 100$  ns. When  $\tau \geq 100$  ns and for all qubit-readout-resonator detunings, the variations in the average assignment error are small and stay between  $2.5 \times 10^{-3} \leq \varepsilon_a \leq 1 \times 10^{-2}$ , except for  $\Delta_{qr}/2\pi = -2.7$  GHz and  $\tau = 100$  ns. The best assignment error,  $\varepsilon_a = 2.5 \times 10^{-3}$ , is reached at  $\tau = 100$  ns and  $\Delta_{qr}/2\pi = -1.3$  GHz. This observation suggests that beyond this integration time, the assignment error is no longer limited by the SNR, which would continue to increase for longer readout times. This is further demonstrated in Fig. 5(b), where we indicate the measured SNR corresponding to each lowest measured average assignment error in Fig. 5(a) as a function of the same qubit-readout-resonator detuning and readout-integration-time range. We observe that a  $\text{SNR} \geq 30$  leads to  $2.5 \times 10^{-3} \leq \varepsilon_a \leq 1 \times 10^{-2}$  for  $\tau \geq 100$  ns and for all detunings  $\Delta_{qr}$ . On the other hand,  $\text{SNR} \leq 14$  leads to a larger assignment error,  $3.87 \times 10^{-2} \leq \varepsilon_a \leq 2.87 \times 10^{-1}$ . In particular, we find that the

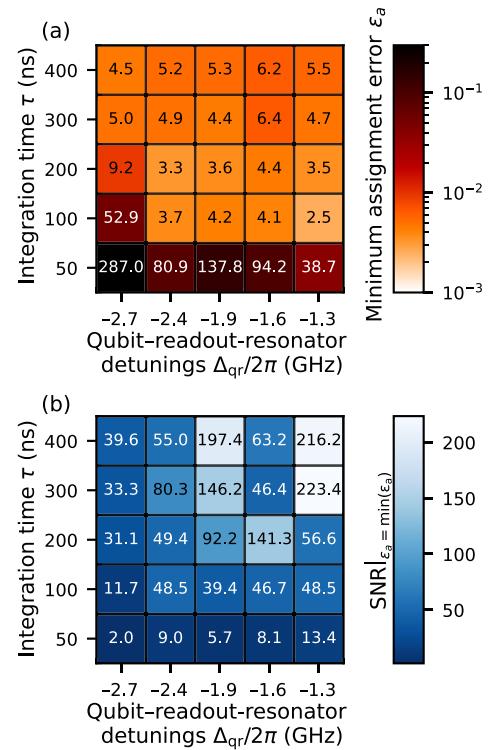


FIG. 5. (a) The minimum assignment error  $\varepsilon_a$  measured as a function of the qubit-readout-resonator detuning  $\Delta_{qr}/2\pi \in [-2.7, -2.4, -1.9, -1.6, -1.3]$  GHz and the readout integration time  $\tau \in [50, 100, 200, 300, 400]$  ns. The annotated values are in *per mille* units of probability: the lowest assignment error  $\varepsilon_a = 2.5 \times 10^{-3}$  is reached at a qubit-readout-resonator detuning of  $\Delta_{qr}/2\pi = -1.3$  GHz, for an integration time of  $\tau = 100$  ns. (b) The measured SNR corresponding to the minimum assignment error in (a).

best assignment error,  $\varepsilon_a = 2.5 \times 10^{-3}$ , is reached for  $\text{SNR} = 48.5$ . An  $\text{SNR} \geq 200$  leads to assignment errors on the same order as an  $\text{SNR} \approx 50$  (see, e.g., at  $\Delta_{qr}/2\pi = -1.6$  GHz compared to at  $\Delta_{qr}/2\pi = -1.3$  GHz, with  $\tau \in [300, 400]$  ns).

We next compare the readout performance in terms of the SNR at different qubit-readout-resonator detunings as a function of the readout power  $n_g/n_{crit}$ , for a fixed integration time  $\tau = 100$  ns [see Fig. 6(a)]. The shaded regions indicate the theoretical SNR prediction from the linear response to the readout drive power [41],

$$\text{SNR}(t) = 2\eta\kappa_p \int_0^t |\beta_e(t') - \beta_g(t')|^2 dt', \quad (5)$$

where  $\eta$  is the measurement efficiency. This expression is in line with Eq. (3). We note that the SNR for the smallest detunings,  $\Delta_{qr} \in [-1.6, -1.3]$  GHz, is significantly higher, which we credit to the increase in the line width of the targeted lower mode  $\kappa_l^{g/e}$  [see Fig. 3(c)]. This

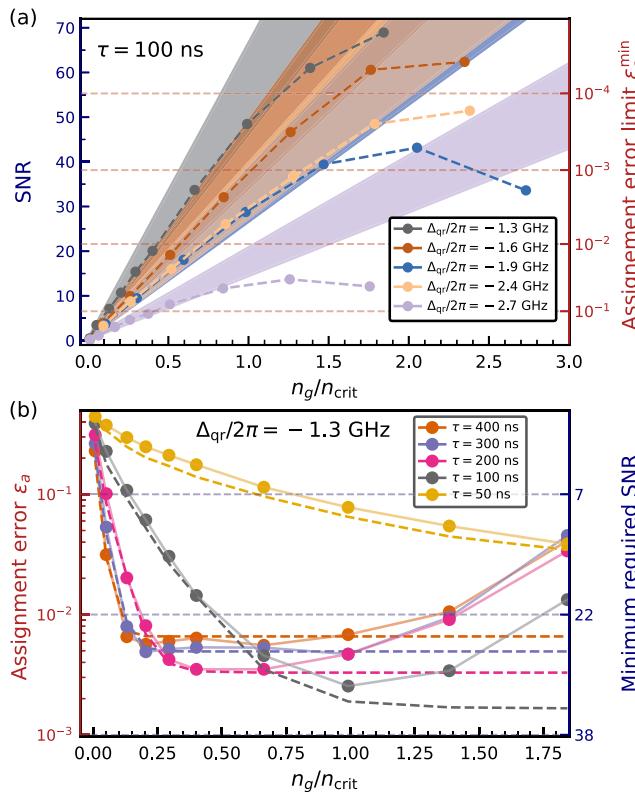


FIG. 6. (a) The SNR as a function of  $n_g/n_{\text{crit}}$  for the indicated qubit-readout-resonator detunings  $\Delta_{\text{qr}}$  and at a fixed readout-integration time of  $\tau = 100$  ns. The shaded regions provide estimates from the analytical solution in the linear regime, detailed in Appendix C. (b) The average assignment error  $\varepsilon_a$  as a function of  $n_g/n_{\text{crit}}$  for the indicated readout integration times  $\tau$  at a fixed qubit-readout-resonator detuning  $\Delta_{\text{qr}}/2\pi = -1.3$  GHz. The solid lines are plotted for ease of visualization. The dashed lines correspond to the theoretical limit imposed by the relaxation time of the qubit given by  $\varepsilon_a^{\min} = 0.5[1 - \text{erfc}(\sqrt{\text{SNR}/8})] + \tau/2T_1$ .

increased line width results in the pointer states  $\beta_{g/e}(t)$  reaching the steady state faster, thus maximizing the SNR rate. The shaded region contains the upper- and lower-bound estimates of the SNR based on uncertainties in the model parameters (see Appendix C).

In all instances, we observe a saturation of the SNR at a readout power  $n_g \gtrsim n_{\text{crit}}$ , where the dispersive approximation is known to break down [32,42,43]. This is in part due to the broadening of the pointer states caused by the qubit-induced Kerr nonlinearity of the resonator (see Fig. 10), measurement-induced state transitions [44], and ionization [45–47].

Finally, in Fig. 6(b), we compare the average assignment error  $\varepsilon_a$  at different readout integration times  $\tau \in [50, 100, 200, 300, 400]$  ns for a fixed qubit-readout-resonator detuning  $\Delta_{\text{qr}}/2\pi = -1.3$  GHz as a function of

the readout power  $n_g/n_{\text{crit}}$ . For  $n_g/n_{\text{crit}} < 1$ , we find excellent agreement between the experimental data (dots) and the approximate theoretical limit (dashed lines) in Eq. (4). Here, we note that the 50-ns measurement is clearly limited by the SNR, which consistently improves at higher drive powers. For 100 ns, the minimum assignment error  $\varepsilon_a^{\min} = 2.5 \times 10^{-3}$  is limited by the intrinsic lifetime of the qubit  $T_1$  rather than the SNR, which can be seen by comparison with the calculated assignment error (gray dashed line), which plateaus at higher readout powers. We notice a distinct upturn in the 100-ns measurement at higher drive powers, which we attribute to nonlinearities and measurement-induced transitions.

For the longer readout times  $\tau \in [200, 300, 400]$  ns, the minimum assignment error is obtained at *lower* drive powers, since these measurements reach a larger SNR. Given that the assignment error is limited by the qubit lifetime, the increase in SNR by increasing the drive power has little impact on the final assignment error  $\varepsilon_a$ , as indicated by the plateaus (dashed lines).

### III. CONCLUSIONS

We have demonstrated beyond-state-of-the-art single-shot readout reaching a minimum assignment error of  $2.5 \times 10^{-3}$  in only 100 ns when reducing the qubit detuning from the resonator by applying a flux pulse (see our work in perspective with other techniques in Fig. 7). We have provided new insights on dispersive readout for a qubit-readout-resonator-Purcell-filter system, in a strongly hybridized regime in which the coupling strength between the readout resonator and the Purcell filter,  $J$ , is comparable to the Purcell-filter line width  $\kappa_p$ . We have shown that by probing the dispersive regime via flux pulses, we can increase the effective decay rate of the targeted readout mode, thus allowing us to reach a larger SNR in a shorter integration time.

We have studied the dispersive readout of a superconducting qubit using a readout resonator coupled to a dedicated Purcell filter. The ability to adjust the detuning between the qubit and the readout-resonator-Purcell-filter combination has allowed us to investigate in detail the model describing the system and the dependence of important parameters, such as the dispersive shift and the effective readout bandwidth, on said detuning. In addition, adjusting the detuning dynamically, as demonstrated in our experiments, allows us to improve qubit readout while fulfilling other constraints, such as being able to bias the qubit at a magnetic-flux-insensitive bias point or at another bias point that may be useful for implementing other functionalities such as magnetic-flux-controlled two-qubit gate operations. While directly applicable to the readout optimization of flux-tunable qubits, we expect that

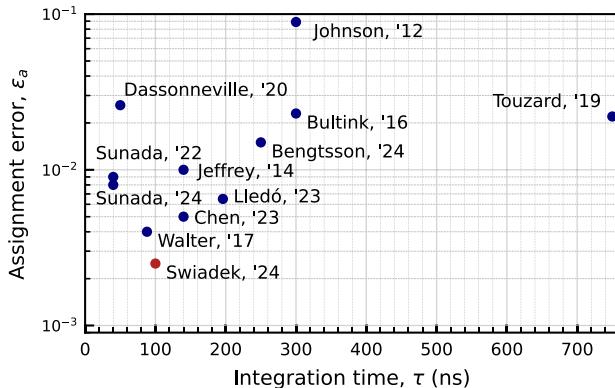


FIG. 7. The two-level average assignment error reported in Johnson *et al.* [15], Jeffrey *et al.* [30], Bultink *et al.* [48], Walter *et al.* [3], Dassonneville *et al.* [49], Touzard *et al.* [50], Chen *et al.* [51], Sunada *et al.* [33,34], Llédó *et al.* [52], Bengtsson *et al.* [53] (blue circles), and this work (red circle) as a function of the readout integration time. Jurcevic *et al.* [54] have reached a two-level assignment error of  $3.5 \times 10^{-2}$  using the excited-state promotion technique [55], which is not plotted here.

other systems, including fixed-frequency qubits, will benefit from the theory framework that we have provided to optimize readout circuits [53,56].

The research presented here significantly improves the understanding of the readout process on short time scales, specifically elucidating the dependence of the SNR on the dispersive shifts, couplings, and frequencies of the readout-circuitry components. It provides theoretical and experimental insights into how one may optimize the system parameters to achieve an assignment error below  $10^{-3}$  in 100 ns. In addition, we expect our work to contribute to reducing the duration of midcircuit measurements for quantum error correction [7], without compromising on the readout fidelity. The combination of techniques discussed in this paper with machine-learning methods [57] for the optimization of pulse shapes, is likely to further decrease readout times while maintaining low readout errors.

## ACKNOWLEDGMENTS

The team in Zurich thanks Johannes Herrmann and Stefania Lažar for contributions to the experimental setup. The team in Sherbrooke thanks Cristóbal Lledó and Catherine Leroux for insightful discussions. The team in Zurich acknowledges that research was sponsored by IARPA and the Army Research Office, under the Entangled Logical Qubits program, and was accomplished under Cooperative Agreement Number W911NF-23-2-0212, by the European Union (EU) Flagship on Quantum Technology H2020-FETFLAG-2018-03 project 820363 OpenSuperQ, by the National Center of Competence in Research “Quantum Science and Technology” (NCCR

QSIT), a research instrument of the Swiss National Science Foundation (SNSF), Grant No. 51NF40-185902), by the SNSF R’Equip Grant No. 206021-170731, by the EU program H2020-FETOPEN project 828826 Quromorphic, and by ETH Zurich. S.K. acknowledges financial support from Fondation Jean-Jacques et Félicia Lopez-Loreta and the ETH Zurich Foundation. The team in Sherbrooke acknowledges the financial support by NSERC, the Canada First Research Excellence Fund, and the Ministère de l’Économie et de l’Innovation du Québec. Support is also acknowledged from the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Quantum Systems Accelerator. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of IARPA, the Army Research Office, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

S.K., P.M., and F.S. planned the experiments with support from all coauthors and S.K. and F.S. performed the experiments. F.S., R.S., and S.K. analyzed the data. R.S. and F.S. worked on the theory. F.S. and A.R. designed the device and A.R., S.K., D.C.Z., and G.J.N. fabricated the device. C.H., N.L., and A.R. developed the experimental software framework. S.K., A.R., F.S., C.H., and N.L. contributed to the experimental setup and maintained it. F.S., R.S., and S.K. prepared the figures for the manuscript and S.K., A.W., A.B., and Q.F. provided feedback. F.S. and R.S. wrote the manuscript, with inputs from all coauthors. A.W., S.K., and A.B. supervised the work.

The authors declare no competing interests.

## APPENDIX A: EXPERIMENTAL SETUP AND DEVICE CHARACTERIZATION

We used a qubit of a 17-qubit quantum device, shown in Fig. 8, to perform the experiment. We fabricated the 17-qubit quantum processor by sputtering a niobium 150-nm thin film onto a high-resistivity intrinsic silicon substrate. All coplanar waveguides, capacitors, and qubit islands were patterned using photolithography and reactive-ion etching. The aluminum-titanium-aluminum trilayer air bridges establish a well-connected ground plane on the device and connect signal lines split by crossings. We fabricated aluminum-based Josephson junctions using shadow evaporation of aluminum through a resist mask defined by electron-beam lithography.

We characterized the properties of the qubit using spectroscopy and standard time-domain measurements. The qubit has an idling frequency  $\omega_q/2\pi = 4.144$  GHz with an upper-flux-bias-insensitive frequency  $\omega_q/2\pi = 5.842$  GHz,

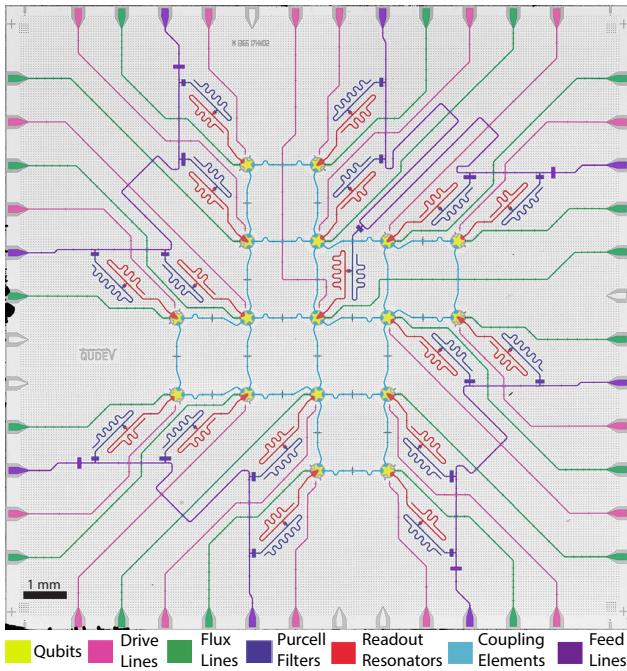


FIG. 8. A false-colour micrograph of the 17-qubit device used for the experiment, adapted from Ref. [7]; the scale bar denotes 1 mm. The experiment was realized using the qubit in the feed line at the bottom left corner, as presented in Fig. 1.

an anharmonicity  $\alpha/2\pi = -181$  MHz, a lifetime  $T_1 = 30.4$   $\mu$ s, a Ramsey-decay time  $T_2^* = 29.2$   $\mu$ s, and an echo-decay time  $T_2^e = 33.9$   $\mu$ s.

Following the method in Ref. [26], we fitted the transmission amplitude of the readout signal through a feed line

to the function

$$|S_{\text{out,in}}|(\omega) = (A + k(\omega - \omega_0)) \times \left| \cos(\phi) - e^{i\phi} \frac{\kappa_p(-2i\Delta_r^{g/e})}{4J^2 + (\kappa_p - 2i\Delta_p)(-2i\Delta_r^{g/e})} \right|, \quad (\text{A1})$$

where  $A$  is the amplitude,  $k$  describes a tilt in the spectrum centered at  $\omega_0$ ,  $\phi$  is a phase rotation induced by the capacitive couplings to other lines,  $\kappa_p$  is the external coupling rate of the Purcell filter,  $\Delta_p = \omega - \omega_p$  is the detuning between the drive frequency  $\omega$  and the Purcell-filter frequency  $\omega_p$ , and  $\Delta_r^{g/e} = \omega - \omega_r^{g/e}$  is the detuning between the drive frequency and the resonator frequency conditioned on the state of the qubit. The relevant parameters for the studied qubit at the indicated qubit-resonator detunings are provided in Table I.

We installed the device in a magnetically shielded sample holder mounted at the base plate (9 mK) of a cryogenic measurement setup [58] and connected it to the control and measurement electronics as shown in Fig. 9. We use a dc signal to generate a current inducing a magnetic flux in the superconducting quantum interference device (SQUID) loop of the transmon qubit, to control its idle frequency. We use arbitrary waveform generators (AWGs) to apply voltage pulses (2.4 GSa/s sampling rate) to the qubit to tune its frequency for readout. The dc and AWG signals are combined using a bias tee. A precompensation of distortions in the flux line is applied, as in Ref. [7].

We perform the single-shot readout experiments with an ultrahigh-frequency quantum analyzer (UHFQA) by

TABLE I. A list of the readout parameters extracted for the qubit-readout-resonator detunings  $\Delta_{qr}/2\pi$  spanning  $-2.7$  GHz to  $-1.3$  GHz using pulsed-spectroscopy measurements.

| Qubit-readout-resonator detuning                 | $\Delta_{qr}/2\pi$ (GHz)  | -2.7    | -2.4    | -1.9    | -1.6    | -1.3    |
|--|---------------------------|---------|---------|---------|---------|---------|
| Qubit frequency during readout                   | $\omega_q/2\pi$ (MHz)     | 4144    | 4500    | 5000    | 5300    | 5600    |
| Bare-readout-resonator frequency                 | $\omega_{r,b}/2\pi$ (MHz) | 6854.63 | 6858.02 | 6857.98 | 6859.74 | 6864.86 |
| Dressed-readout-resonator frequency              | $\omega_r^g/2\pi$ (MHz)   | 6876.27 | 6881.96 | 6896.31 | 6906.66 | 6928.88 |
| Purcell-filter frequency                         | $\omega_p/2\pi$ (MHz)     | 6899.86 | 6899.86 | 6899.86 | 6899.86 | 6899.86 |
| Readout drive frequency                          | $\omega_d/2\pi$ (MHz)     | 6857.4  | 6861.2  | 6870.0  | 6874.0  | 6881.6  |
| Qubit-readout-resonator coupling                 | $g_b/2\pi$ (MHz)          | 224.32  | 205.61  | 211.49  | 204.2   | 205.53  |
| Qubit-charge-readout-resonator coupling          | $g/2\pi$ (MHz)            | 284.01  | 271.40  | 293.71  | 292.27  | 302.34  |
| Readout-resonator-Purcell-filter coupling        | $J/2\pi$ (MHz)            | 27.9    | 27.9    | 27.9    | 27.9    | 27.9    |
| Low-mode line width, qubit in $ g\rangle$ state  | $\kappa_l^g/2\pi$ (MHz)   | 10.16   | 11.61   | 15.88   | 19.17   | 25.00   |
| Low-mode line width, qubit in $ e\rangle$ state  | $\kappa_l^e/2\pi$ (MHz)   | 8.89    | 9.99    | 12.67   | 14.87   | 19.37   |
| High-mode line width, qubit in $ g\rangle$ state | $\kappa_h^g/2\pi$ (MHz)   | 23.86   | 22.41   | 18.14   | 14.85   | 8.90    |
| High-mode line width, qubit in $ e\rangle$ state | $\kappa_h^e/2\pi$ (MHz)   | 25.13   | 24.03   | 23.35   | 19.15   | 14.52   |
| Low-mode dispersive shift                        | $2\chi_l/2\pi$ (MHz)      | -4.15   | -4.45   | -6.21   | -6.77   | -6.99   |
| High-mode dispersive shift                       | $2\chi_h/2\pi$ (MHz)      | -1.50   | -1.94   | -4.33   | -6.78   | -14.32  |
| Critical readout-resonator photon number         | $n_{\text{crit}}$         | 23.14   | 19.26   | 10.42   | 7.55    | 4.83    |

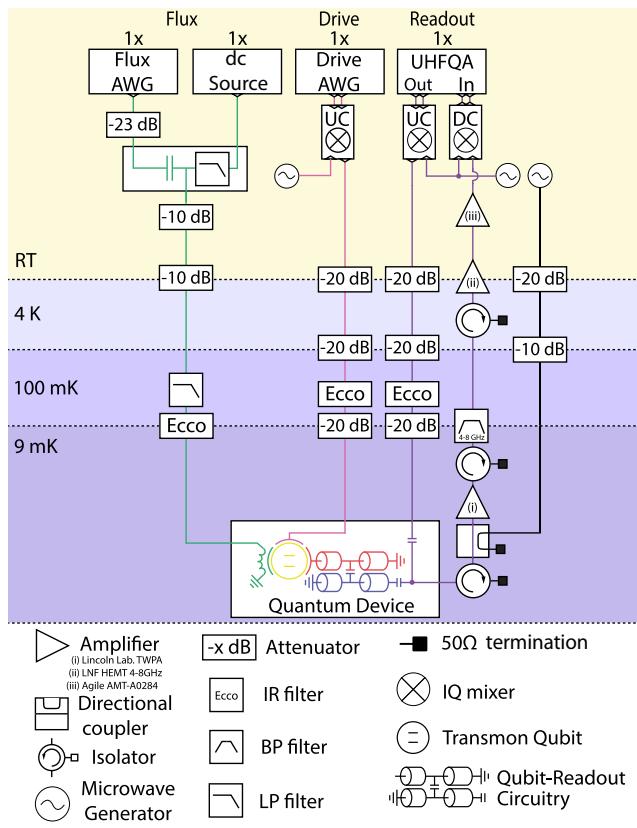


FIG. 9. A schematic of the wiring and control electronics. The qubit (yellow) on the quantum device is connected to the room-temperature electronics via flux lines (green), drive lines (pink), and readout lines (purple) through its readout resonator (red) and Purcell filter (blue). The background colors indicate the temperature stages of the experimental setup.

using an in-phase and quadrature (IQ) mixer to up-convert the frequency-multiplexed readout pulses from an intermediate-frequency signal sampled at 1.8 GSa/s to the gigahertz frequency range of the readout circuitry. At the output of the device feed line, the readout signal passes through a wide-bandwidth near-quantum limited traveling-wave parametric amplifier (TWPA) [59], a high-electron-mobility transistor (HEMT) amplifier, and room-temperature amplifiers. It is then down-converted with an IQ mixer and digitally demodulated and integrated in the UHFQA. Further details on the device fabrication, characterization, and the experimental setup can be found in Ref. [7].

## APPENDIX B: MODEL

To model the system, we use the Hamiltonian

$$\begin{aligned} \hat{H}_{hp} = & 4E_c\hat{n}_t^2 - E_J(\Phi)\cos\hat{\phi}_t \\ & + \omega_{r,b}\hat{a}^\dagger\hat{a} - ig(\hat{n}_t - n_g)(\hat{a} - \hat{a}^\dagger) \\ & + \omega_p\hat{f}^\dagger\hat{f} + J(\hat{f}^\dagger - \hat{f})(\hat{a}^\dagger - \hat{a}) \\ & + 2i\mathcal{E}\sin(\omega_d t)(\hat{f}^\dagger - \hat{f}), \end{aligned} \quad (B1)$$

where  $\hat{n}_t$  is the charge operator of the transmon,  $\hat{a}$  the readout-resonator mode-creation operator and  $\hat{f}$  the Purcell-filter mode-creation operator.  $E_c$  is the charging energy of the transmon,  $E_J(\Phi)$  the flux-tunable Josephson energy of the transmon,  $\omega_{r,b}$  and  $\omega_p$  are the bare-resonator and Purcell-filter frequencies, and  $g$  and  $J$  are the transmon-resonator and resonator–Purcell-filter coupling rates, respectively.  $\mathcal{E}$  and  $\omega_d$  are the drive amplitude and the drive frequency.

Further, we consider a master equation

$$\dot{\rho} = -i[\hat{H}_{hp}, \hat{\rho}] + \kappa_p \mathcal{D}[\hat{f}], \quad (B2)$$

where  $\kappa_p$  is the coupling rate between the Purcell filter and the feed line.

We first diagonalize the transmon-resonator subsystem. We follow the notation of Ref. [21] and assume a Kerr-nonlinear oscillator model for the transmon, valid in the low-readout-power regime. A Schrieffer-Wolff transformation yields an effective Hamiltonian

$$\begin{aligned} \hat{H} = & \bar{\omega}_q\hat{b}^\dagger\hat{b} + \omega_p\hat{f}^\dagger\hat{f} + \omega_r^g\hat{a}^\dagger\hat{a} + 2\chi\hat{a}^\dagger\hat{a}\hat{b}^\dagger\hat{b} \\ & - 2\lambda'\lambda^3 E_C\hat{a}^\dagger\hat{a}^2\hat{b}^\dagger\hat{b} - \frac{\alpha}{2}\hat{b}^\dagger\hat{b}^2 - \frac{E_c}{2}\lambda^4\hat{a}^2\hat{a}^2 \\ & + J\left(\left[1 - 2\lambda\lambda'\hat{b}^\dagger\hat{b}\right]\hat{a}^\dagger\hat{f} + \lambda\hat{b}^\dagger\hat{f} + \text{H.c.}\right), \end{aligned} \quad (B3)$$

where

$$\chi = -g^2 E_C / (\Delta_{qr}(\Delta_{qr} - E_c)), \quad (B4)$$

$$\lambda' = \lambda E_c / [\Delta_{qr} + E_c(1 - 2\lambda^2)], \quad (B5)$$

with  $\lambda = g/\Delta_{qr}$  and  $\Delta_{qr} = \omega_q - \omega_{r,b}$ . Further, the qubit and resonator frequencies become Lamb shifted, with  $\bar{\omega}_q \approx \omega_q + g^2/\Delta_{qr}$ ,  $\omega_r^g \approx \omega_{r,b} - g^2/\Delta_{qr}$ . The contribution  $-2\lambda'\lambda^3 E_C\hat{a}^\dagger\hat{a}^2\hat{b}^\dagger\hat{b}$  normalizes down the effective Kerr nonlinearity when the qubit is in the excited state and we note that  $K_e/K_g \approx 1 + 4\lambda'/\lambda$ , where  $4\lambda'/\lambda < 0$  for  $\Delta_{qr} < 0$ . This nonlinearity leads to a significantly larger increase in the Gaussian width of the ground-state response than the excited-state response, as seen in Figs. 10(c)–10(f) [60,61]. For this reason, we quote the drive power in the main text as a function of  $n_g/n_{\text{crit}}$  as opposed to  $n_e/n_{\text{crit}}$ .

We also note a correlation between the broadening of the ground-state response and an increase in the overlap errors  $P(g|e)$  and  $P(e|g)$  for  $n_g/n_{\text{crit}} \gtrsim 1$  where nonlinear effects are expected to be more important [see Figs. 10(a) and 10(b)]. The frequency renormalizations from the effective coupling of the filter to the qubit are on the order of  $J^2\lambda^2/\Delta_{qr}^2$  and can be safely ignored in this regime; this has been corroborated by numerical diagonalization of Eq. (B1) and simulation of the master equation. More

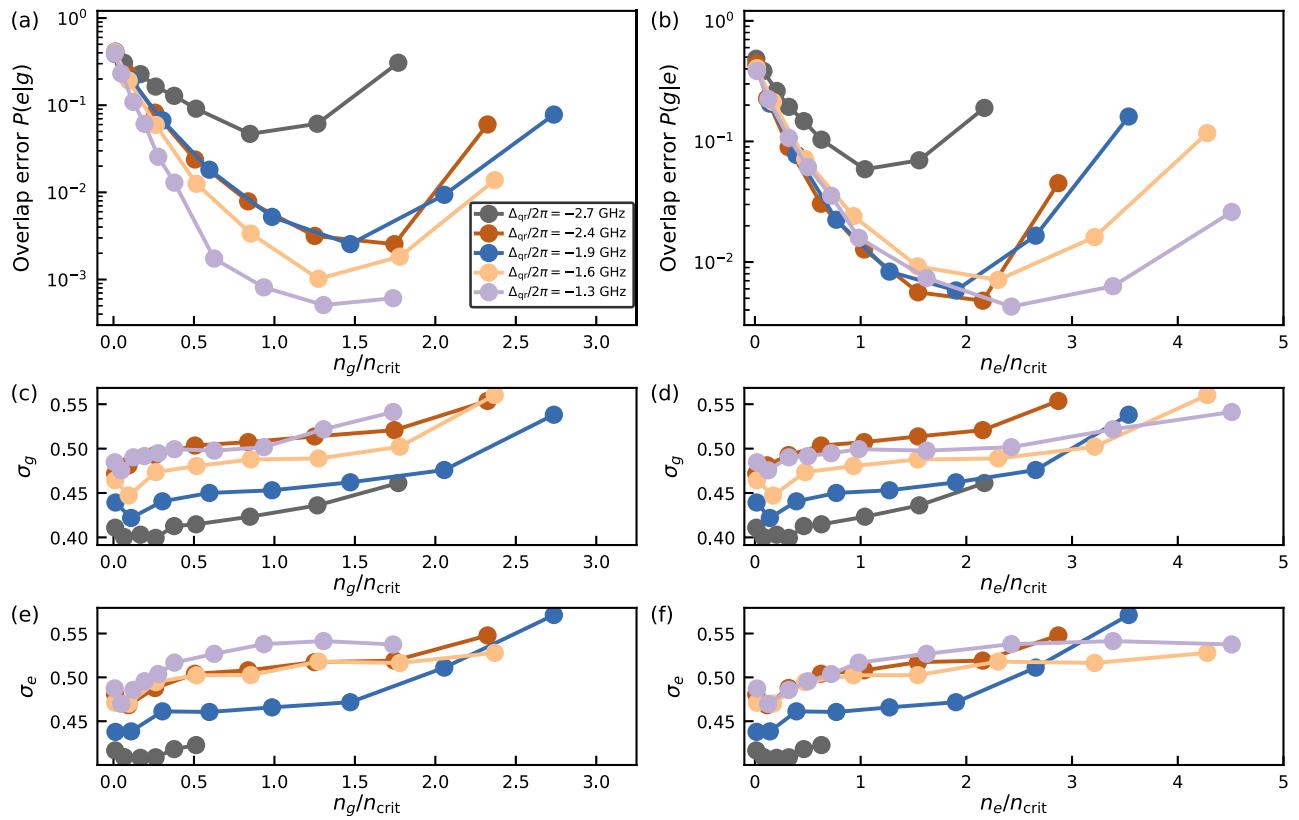


FIG. 10. (a),(b) The overlap error (a)  $P(e|g)$  and (b)  $P(g|e)$  for the indicated qubit-readout-resonator frequency detunings  $\Delta_{qr}$  at a fixed integration time of  $\tau = 100$  ns. (c),(d) The Gaussian width  $\sigma_g$  of the ground-state single-shot histogram as a function of (c)  $n_g/n_{\text{crit}}$  and (d)  $n_e/n_{\text{crit}}$ . (e),(f) The Gaussian width  $\sigma_e$  of the ground-state single-shot histogram as a function of (e)  $n_g/n_{\text{crit}}$  and (f)  $n_e/n_{\text{crit}}$ .

importantly, we note that the effective coupling strength between the resonator and Purcell filter,  $J[1 - 2\lambda\lambda' \hat{b}^\dagger \hat{b}]$ , only weakly depends on the qubit state.

### APPENDIX C: LINEAR RESPONSE

For sufficiently small drive amplitudes, we can assume negligible impact from the Kerr nonlinearity and take the resonator and filter responses to be linear. As such, we can use the relation

$$\begin{bmatrix} \dot{\alpha}^{g/e} \\ \dot{\beta}^{g/e} \end{bmatrix} = -i \begin{bmatrix} \omega_r^{g/e} - \omega_d & J^{g/e} \\ J^{g/e} & \omega_p - \omega_d - i\frac{\kappa_p}{2} \end{bmatrix} \begin{bmatrix} \alpha^{g/e} \\ \beta^{g/e} \end{bmatrix} + \begin{bmatrix} 0 \\ \mathcal{E} \end{bmatrix}, \quad (\text{C1})$$

where  $\alpha$  and  $\beta$  represent the coherent fields of the readout resonator and Purcell filter, respectively,  $\mathcal{E}$  is the drive amplitude,  $J^{g/e} = J[1 - \lambda\lambda'(\langle \hat{\sigma}_z \rangle + 1)]$  is the effective readout-resonator-Purcell-filter coupling, and  $\omega_r^{g/e} = \omega_r^g + 2\chi$ . Diagonalizing the equation of motion in the

absence of a drive ( $\mathcal{E} = \omega_d = 0$ ) yields eigenvalues

$$\lambda_{l,h}^{g/e} = \frac{\omega_r^{g/e} + \omega_p - i\kappa_p/2}{2} \pm \frac{1}{2} \sqrt{\left( \Delta_{rp}^{g/e} + \frac{i\kappa_p}{2} \right)^2 + 4J^{2,g/e}}, \quad (\text{C2})$$

where  $\Delta_{rp}^{ge} = \omega_r^{ge} - \omega_p$ . For  $4J \gg \kappa_p$ , the eigenvalues approximately correspond to a normal mode, where the indices  $l$  and  $h$  correspond to the *low* and *high* mode, respectively. In this fashion, the real and imaginary components of  $\lambda_{l,h}^{g/e}$  correspond to the frequency and line width of these *low* and *high* modes:

$$\omega_{l,h}^{g/e} = \text{Re}[\lambda_{l,h}^{g/e}], \quad \kappa_{l,h}^{g/e} = -2 \text{Im}[\lambda_{l,h}^{g/e}]. \quad (\text{C3})$$

To obtain a qualitative understanding of the eigenvalues, we perform an expansion of the square root in

$(\Delta_{\text{rp}}^{g/e} + i\kappa_p/2)$ . Assuming  $J^{g/e} \approx J$ , this yields

$$\lambda_{l,h}^{g/e} = \frac{\omega_r^{g/e} + \omega_p - i\kappa_p/2}{2} \pm \left( J + \frac{\Delta_{\text{rp}}^{2,g/e} - i\Delta_{\text{rp}}^{g/e}\kappa_p - \kappa_p^2/4}{8J} \right). \quad (\text{C4})$$

Consequently, we see that the frequencies  $\omega_{l,h}^{g/e}$  of the two sets of normal modes are approximately separated by  $2J$ , with the relative dispersive shift of each mode  $\chi_{l,h}$  between the low and high mode being as follows:

$$\begin{aligned} 2\chi_l &= \omega_l^e - \omega_l^g \approx \chi - \frac{\Delta_{\text{rp}}^g \chi + \chi^2}{2J}, \\ 2\chi_h &= \omega_h^e - \omega_h^g \approx \chi + \frac{\Delta_{\text{rp}}^g \chi + \chi^2}{2J}, \\ \kappa_l^{g/e} &\approx \frac{\kappa_p}{2} + \frac{\Delta_{\text{rp}}^{g/e} \kappa_p}{4J}, \\ \kappa_h^{g/e} &\approx \frac{\kappa_p}{2} - \frac{\Delta_{\text{rp}}^{g/e} \kappa_p}{4J}. \end{aligned} \quad (\text{C5})$$

Noting that  $\Delta_{\text{rp}}^e < \Delta_{\text{rp}}^g$ , we see that  $\kappa_h^g < \kappa_h^e$  and  $\kappa_l^g > \kappa_l^e$  for  $\Delta_{\text{rp}}^g < 0$  and vice versa for  $\Delta_{\text{rp}}^g > 0$ .

The relevant frequencies and line widths extracted from the normal-mode Hamiltonian in Eq. (C1) are plotted in Fig. 3.

The steady-state expressions are found to be

$$\begin{bmatrix} \alpha_{\text{ss}}^{g/e} \\ \beta_{\text{ss}}^{g/e} \end{bmatrix} = \frac{i\mathcal{E}}{\Delta_{\text{rd}}^{g/e}(\Delta_{\text{pd}} - i\kappa_p/2) - J^2} \begin{bmatrix} J \\ -\Delta_{\text{rd}}^{g/e} \end{bmatrix}, \quad (\text{C6})$$

where  $\Delta_{\text{rd}}^{g/e} = \omega_r^{g/e} - \omega_d$  and  $\Delta_{\text{pd}} = \omega_p - \omega_d$ . Using this expression, the full time-dependent response takes the form

$$\begin{aligned} \beta^{g/e}(t) &= \beta_{\text{ss}}^{g/e} + i\mathcal{E} \frac{\lambda_h^{g/e} - \omega_r^{g/e}}{d} \frac{e^{-i(\lambda_h^{g/e} - \omega_d)t}}{(\lambda_h^{g/e} - \omega_d)} \\ &\quad - i\mathcal{E} \frac{\lambda_l^{g/e} - \omega_r^{g/e}}{d} \frac{e^{-i(\lambda_l^{g/e} - \omega_d)t}}{(\lambda_l^{g/e} - \omega_d)}, \end{aligned} \quad (\text{C7})$$

where  $d = \sqrt{(\Delta_{\text{rp}}^{g/e} + i\kappa_p/2)^2 + 4J^2}$ .

We then use Eq. (C7) to express the SNR as [41]

$$\text{SNR}(t) = 2\eta\kappa_p \int_0^t |\beta_e(t') - \beta_g(t')|^2 dt', \quad (\text{C8})$$

where  $\eta$  is the measurement efficiency. We note that this expression is the square of the often-used expression but is in line with Eq. (3). We then plot these results in

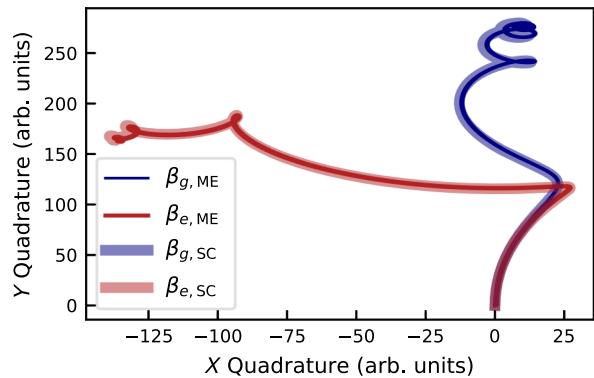


FIG. 11. An example comparison of the master equation  $\beta_{g/e,\text{ME}}$ , against the semiclassical trajectories  $\beta_{g/e,\text{SC}}$ , plotted in the phase space of the Purcell-filter mode for  $\Delta_{\text{qr}}/2\pi = -1.3$  GHz and  $\mathcal{E}/2\pi = 10$  MHz. The solid lines correspond to the master-equation expectations of the resonator and filter modes,  $\langle a \rangle$  and  $\langle f \rangle$ , respectively, where  $\langle X \rangle = \text{Re}\{a,f\}$ ,  $\langle Y \rangle = \text{Im}\{a,f\}$ . The transparent lines correspond to the semiclassical solution.

Fig. 6(a) with a  $\pm 1$ -MHz deviation in the calculated values of  $g$ ,  $J$ ,  $\omega_r$ , and  $\kappa_p$  to allow for uncertainties in the fitted parameters and nonidealities caused by spurious couplings to two-level systems, gain ripples of the TWPA, or frequency-dependent impedance mismatches of components in the readout lines, alongside a variation of up to 5% in the measurement efficiency at different frequencies. The shaded region contains the upper- and lower-bound estimates of the SNR based on these uncertainties.

Finally, we verify the validity of the semiclassical approximation. A negligible difference was noted in the trajectories in phase space between the expected internal coherent fields  $\langle \hat{a} \rangle$  and  $\langle \hat{f} \rangle$ , calculated by solving the master equation (B2), and the corresponding semiclassical predictions  $\alpha$  and  $\beta$  of the resonator and Purcell filter, respectively, confirming that the semiclassical model (in Appendix C) captures the state separation at low powers. Example trajectories at low power for the Purcell-filter mode are plotted for the  $\Delta_{\text{qr}}/2\pi = -1.3$  GHz case in Fig. 11.

## APPENDIX D: PHOTON-NUMBER AND DRIVE-POWER CALIBRATION

We measure the ac Stark shift  $\Delta_{\text{ac}}$  caused on the qubit prepared in the ground state by the readout resonator as a function of the power (see Fig. 12). For this purpose, we simultaneously apply a readout tone with a length of  $0.6 \mu\text{s}$  of variable power and a  $\pi$  pulse of variable frequency. We measure the excited-state population as a function of the drive-pulse frequency for each readout power for the  $\omega_q/2\pi = 4.14$  GHz qubit frequency. The frequency at which the excited state population is maximum corresponds to the instantaneous and ac-Stark-shifted qubit frequency.

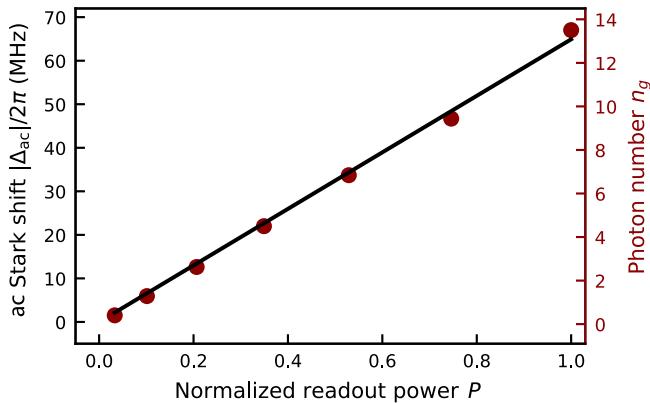


FIG. 12. The measured (dots) ac Stark shift  $\Delta_{\text{ac}}$  of the qubit prepared in the ground state at  $\omega_q/2\pi = 4.14$  GHz as a function of the normalized readout power. The corresponding inferred resonator photon number  $n_g = \Delta_{\text{ac}}/2(\chi_l + \chi_h)$  (solid line) is shown on the right axis.

We determine this frequency using a Gaussian fit. We infer and then calibrate the steady-state readout-resonator photon number  $n_g$  with the qubit prepared in the ground state from  $n_g = \Delta_{\text{ac}}/2(\chi_l + \chi_h)$  for the chosen drive powers [62].

Then, using the steady-state resonator response from Eq. (C1), this allows us to extract the effective drive amplitudes  $\mathcal{E}$ . The steady-state resonator responses for the qubit-resonator detunings  $\Delta_{\text{qr}}/2\pi \in [-2.7, -2.4, -1.9, -1.6, -1.3]$  GHz are then inferred from Eq. (C6) at the same drive powers.

Eichler, and A. Wallraff, Realizing repeated quantum error correction in a distance-three surface code, *Nature* **605**, 669 (2022).

- [8] R. Acharya, *et al.*, Suppressing quantum errors by scaling a surface code logical qubit, *Nature* **614**, 676 (2023).
- [9] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels, *Phys. Rev. Lett.* **70**, 1895 (1993).
- [10] D. Gottesman and I. L. Chuang, Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations, *Nature* **402**, 390 (1999).
- [11] L. Steffen, Y. Salathe, M. Oppliger, P. Kurpiers, M. Baur, C. Lang, C. Eichler, G. Puebla-Hellmann, A. Fedorov, and A. Wallraff, Deterministic quantum teleportation with feed-forward in a solid state system, *Nature* **500**, 319 (2013).
- [12] J. Qiu, Y. Liu, J. Niu, L. Hu, Y. Wu, L. Zhang, W. Huang, Y. Chen, J. Li, S. Liu, Y. Zhong, L. Duan, and D. Yu, Deterministic quantum teleportation between distant superconducting chips, [arXiv:2302.08756](https://arxiv.org/abs/2302.08756).
- [13] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Concentrating partial entanglement by local operations, *Phys. Rev. A* **53**, 2046 (1996).
- [14] S. Bravyi and A. Kitaev, Universal quantum computation with ideal Clifford gates and noisy ancillas, *Phys. Rev. A* **71**, 022316 (2005).
- [15] J. E. Johnson, C. Macklin, D. H. Slichter, R. Vijay, E. B. Weingarten, J. Clarke, and I. Siddiqi, Heralded state preparation in a superconducting qubit, *Phys. Rev. Lett.* **109**, 050506 (2012).
- [16] D. Ristè, J. G. van Leeuwen, H.-S. Ku, K. W. Lehnert, and L. DiCarlo, Initialization by measurement of a superconducting quantum bit circuit, *Phys. Rev. Lett.* **109**, 050507 (2012).
- [17] Y. Salathé, M. Mondal, M. Oppliger, J. Heinsoo, P. Kurpiers, A. Potočnik, A. Mezzacapo, U. Las Heras, L. Lamata, E. Solano, S. Filipp, and A. Wallraff, Digital quantum simulation of spin models with circuit quantum electrodynamics, *Phys. Rev. X* **5**, 021027 (2015).
- [18] J. Herrmann, S. M. Lima, A. Remm, P. Zapletal, N. A. McMahon, C. Scarato, F. Swiadek, C. K. Andersen, C. Hellings, S. Krinner, N. Lacroix, S. Lazar, M. Kerschbaum, D. C. Zanuz, G. J. Norris, M. J. Hartmann, A. Wallraff, and C. Eichler, Realizing quantum convolutional neural networks on a superconducting quantum processor to recognize quantum phases, *Nat. Commun.* **13**, 4144 (2022).
- [19] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation, *Phys. Rev. A* **69**, 062320 (2004).
- [20] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Approaching unit visibility for control of a superconducting qubit with dispersive readout, *Phys. Rev. Lett.* **95**, 060501 (2005).
- [21] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, Circuit quantum electrodynamics, *Rev. Mod. Phys.* **93**, 025005 (2021).
- [22] 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).
- [23] V. Negirneac, H. Ali, N. Muthusubramanian, F. Battistel, R. Sagastizabal, M. S. Moreira, J. F. Marques, W. J. Vlothuizen, M. Beekman, C. Zachariadis, N. Haider, A. Bruno, and L. DiCarlo, High-fidelity controlled-Z gate with maximal intermediate leakage operating at the speed limit in a superconducting quantum processor, *Phys. Rev. Lett.* **126**, 220502 (2021).
- [24] L. DiCarlo, M. D. Reed, L. Sun, B. R. Johnson, J. M. Chow, J. M. Gambetta, L. Frunzio, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf, Preparation and measurement of three-qubit entanglement in a superconducting circuit, *Nature* **467**, 574 (2010).
- [25] F. W. Strauch, P. R. Johnson, A. J. Dragt, C. J. Lobb, J. R. Anderson, and F. C. Wellstood, Quantum logic gates for coupled superconducting phase qubits, *Phys. Rev. Lett.* **91**, 167005 (2003).
- [26] J. Heinsoo, C. K. Andersen, A. Remm, S. Krinner, T. Walter, Y. Salathé, S. Gasparinetti, J.-C. Besse, A. Potočnik, A. Wallraff, and C. Eichler, Rapid high-fidelity multiplexed readout of superconducting qubits, *Phys. Rev. Appl.* **10**, 034040 (2018).
- [27] F. Arute, *et al.*, Quantum supremacy using a programmable superconducting processor, *Nature* **574**, 505 (2019).
- [28] Z. Chen, *et al.*, Exponential suppression of bit or phase errors with cyclic error correction, *Nature* **595**, 383 (2021).
- [29] M. D. Reed, B. R. Johnson, A. A. Houck, L. DiCarlo, J. M. Chow, D. I. Schuster, L. Frunzio, and R. J. Schoelkopf, Fast reset and suppressing spontaneous emission of a superconducting qubit, *Appl. Phys. Lett.* **96**, 203110 (2010).
- [30] E. Jeffrey, D. Sank, J. Y. Mutus, T. C. White, J. Kelly, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Megrant, P. J. J. O’Malley, C. Neill, P. Roushan, A. Vainsencher, J. Wenner, A. N. Cleland, and J. M. Martinis, Fast accurate state measurement with superconducting qubits, *Phys. Rev. Lett.* **112**, 190504 (2014).
- [31] N. T. Bronn, Y. Liu, J. B. Hertzberg, A. D. Córcoles, A. A. Houck, J. M. Gambetta, and J. M. Chow, Broadband filters for abatement of spontaneous emission in circuit quantum electrodynamics, *Appl. Phys. Lett.* **107**, 172601 (2015).
- [32] M. Boissonneault, J. M. Gambetta, and A. Blais, Improved superconducting qubit readout by qubit-induced nonlinearities, *Phys. Rev. Lett.* **105**, 100504 (2010).
- [33] Y. Sunada, K. Yuki, Z. Wang, T. Miyamura, J. Ilves, K. Matsuura, P. A. Spring, S. Tamate, S. Kono, and Y. Nakamura, Photon-noise-tolerant dispersive readout of a superconducting qubit using a nonlinear Purcell filter, *PRX Quantum* **5**, 010307 (2024).
- [34] Y. Sunada, S. Kono, J. Ilves, S. Tamate, T. Sugiyama, Y. Tabuchi, and Y. Nakamura, Fast readout and reset of a superconducting qubit coupled to a resonator with an intrinsic Purcell filter, *Phys. Rev. Appl.* **17**, 044016 (2022).
- [35] M. Khezri, A. Opremcak, Z. Chen, K. C. Miao, M. McEwen, A. Bengtsson, T. White, O. Naaman, D. Sank, A. N. Korotkov, Y. Chen, and V. Smelyanskiy, Measurement-induced state transitions in a superconducting qubit: Within the rotating-wave approximation, *Phys. Rev. Appl.* **20**, 054008 (2023).
- [36] M. Malekakhlagh, W. Shanks, and H. Paik, Optimization of the resonator-induced phase gate for superconducting qubits, *Phys. Rev. A* **105**, 022607 (2022).
- [37] M. D. Hutchings, J. B. Hertzberg, Y. Liu, N. T. Bronn, G. A. Keefe, M. Brink, J. M. Chow, and B. L. T. Plourde, Tunable superconducting qubits with flux-independent coherence, *Phys. Rev. Appl.* **8**, 044003 (2017).
- [38] C. W. Gardiner and M. J. Collett, Input, and output in damped quantum systems: Quantum stochastic differential equations and the master equation, *Phys. Rev. A* **31**, 3761 (1985).
- [39] J. Gambetta, W. A. Braff, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, Protocols for optimal readout of qubits using a continuous quantum nondemolition measurement, *Phys. Rev. A* **76**, 012325 (2007).
- [40] P. Magnard, P. Kurpiers, B. Royer, T. Walter, J.-C. Besse, S. Gasparinetti, M. Pechal, J. Heinsoo, S. Storz, A. Blais, and A. Wallraff, Fast and unconditional all-microwave reset of a superconducting qubit, *Phys. Rev. Lett.* **121**, 060502 (2018).
- [41] C. C. Bultink, B. Tarasinski, N. Haandbæk, S. Poletto, N. Haider, D. J. Michalak, A. Bruno, and L. DiCarlo, General method for extracting the quantum efficiency of dispersive qubit readout in circuit QED, *Appl. Phys. Lett.* **112**, 092601 (2018).
- [42] J. E. Johnson, E. M. Hoskinson, C. Macklin, D. H. Slichter, I. Siddiqi, and J. Clarke, Dispersive readout of a flux qubit at the single-photon level, *Phys. Rev. B* **84**, 220503 (2011).
- [43] 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).
- [44] D. Sank, Z. Chen, M. Khezri, J. Kelly, R. Barends, B. Campbell, Y. Chen, B. Chiaro, A. Dunsworth, A. Fowler, *et al.*, Measurement-induced state transitions in a superconducting qubit: Beyond the rotating wave approximation, *Phys. Rev. Lett.* **117**, 190503 (2016).
- [45] R. Shillito, A. Petrescu, J. Cohen, J. Beall, M. Hauru, M. Ganahl, A. G. Lewis, G. Vidal, and A. Blais, Dynamics of transmon ionization, *Phys. Rev. Appl.* **18**, 034031 (2022).
- [46] K. N. Nesterov and I. V. Pechenezhskiy, Measurement-induced state transitions in dispersive qubit readout schemes, [arXiv:2402.07360](https://arxiv.org/abs/2402.07360).
- [47] M. F. Dumas, B. Groleau-Paré, A. McDonald, M. H. Muñoz-Arias, C. Lledó, B. D’Anjou, and A. Blais, Unified picture of measurement-induced ionization in the transmon, [arXiv:2402.06615](https://arxiv.org/abs/2402.06615).
- [48] C. C. Bultink, M. A. Rol, T. E. O’Brien, X. Fu, B. C. S. Dikken, C. Dickel, R. F. L. Vermeulen, J. C. de Sterke, A. Bruno, R. N. Schouten, and L. DiCarlo, Active resonator reset in the nonlinear dispersive regime of circuit QED, *Phys. Rev. Appl.* **6**, 034008 (2016).
- [49] R. Dassonneville, T. Ramos, V. Milchakov, L. Planat, E. Dumur, F. Foroughi, J. Puertas, S. Leger, K. Bharadwaj, J. Delaforce, C. Naud, W. Hasch-Guichard, J. J. García-Ripoll, N. Roch, and O. Buisson, Fast high-fidelity quantum nondemolition qubit readout via a nonperturbative cross-Kerr coupling, *Phys. Rev. X* **10**, 011045 (2020).
- [50] S. Touzard, A. Kou, N. E. Frattini, V. V. Sivak, S. Puri, A. Grimm, L. Frunzio, S. Shankar, and M. H. Devoret,

- Gated conditional displacement readout of superconducting qubits, [Phys. Rev. Lett. \*\*122\*\*, 080502 \(2019\)](#).
- [51] L. Chen, H.-X. Li, Y. Lu, C. W. Warren, C. J. Križan, S. Kosen, M. Rommel, S. Ahmed, A. Osman, and J. Biznárová, *et al.*, Transmon qubit readout fidelity at the threshold for quantum error correction without a quantum-limited amplifier, [npj Quantum Inf. \*\*9\*\*, 26 \(2023\)](#).
- [52] C. Lledó, R. Dassonneville, A. Moulinas, J. Cohen, R. Shillito, A. Bienfait, B. Huard, and A. Blais, Cloaking a qubit in a cavity, [Nat. Commun. \*\*14\*\*, 6313 \(2023\)](#).
- [53] A. Bengtsson, A. Opremcak, M. Khezri, D. Sank, A. Bourassa, K. J. Satzinger, S. Hong, C. Erickson, B. J. Lester, K. C. Miao, A. N. Korotkov, J. Kelly, Z. Chen, and P. V. Klimov, Model-based optimization of superconducting qubit readout, [Phys. Rev. Lett. \*\*132\*\*, 100603 \(2024\)](#).
- [54] P. Jurcevic, *et al.*, Demonstration of quantum volume 64 on a superconducting quantum computing system, [Quantum Sci. Technol. \*\*6\*\*, 025020 \(2021\)](#).
- [55] S. S. Elder, C. S. Wang, P. Reinhold, C. T. Hann, K. S. Chou, B. J. Lester, S. Rosenblum, L. Frunzio, L. Jiang, and R. J. Schoelkopf, High-fidelity measurement of qubits encoded in multilevel superconducting circuits, [Phys. Rev. X \*\*10\*\*, 011001 \(2020\)](#).
- [56] D. Sank, A. Opremcak, A. Bengtsson, M. Khezri, Z. Chen, O. Naaman, and A. Korotkov, System characterization of dispersive readout in superconducting qubits, [arXiv:2402.00413](#).
- [57] B. Lienhard, A. Vepsäläinen, L. C. Govia, C. R. Hoffer, J. Y. Qiu, D. Ristè, M. Ware, D. Kim, R. Winik, A. Melville, B. Niedzielski, J. Yoder, G. J. Ribeill, T. A. Ohki, H. K. Krovi, T. P. Orlando, S. Gustavsson, and W. D. Oliver, Deep-neural-network discrimination of multiplexed superconducting-qubit states, [Phys. Rev. Appl. \*\*17\*\*, 014024 \(2022\)](#).
- [58] S. Krinner, S. Storz, P. Kurpiers, P. Magnard, J. Heinsoo, R. Keller, J. Lütolf, C. Eichler, and A. Wallraff, Engineering cryogenic setups for 100-qubit scale superconducting circuit systems, [EPJ Quantum Technol. \*\*6\*\*, 2 \(2019\)](#).
- [59] C. Macklin, K. O'Brien, D. Hover, M. E. Schwartz, V. Bolkhovsky, X. Zhang, W. D. Oliver, and I. Siddiqi, A near-quantum-limited Josephson traveling-wave parametric amplifier, [Science \*\*350\*\*, 307 \(2015\)](#).
- [60] N. Bartolo, F. Minganti, W. Casteels, and C. Ciuti, Exact steady state of a Kerr resonator with one- and two-photon driving and dissipation: Controllable Wigner-function multimodality and dissipative phase transitions, [Phys. Rev. A \*\*94\*\*, 033841 \(2016\)](#).
- [61] D. Roberts and A. A. Clerk, Driven-dissipative quantum Kerr resonators: New exact solutions, photon blockade and quantum bistability, [Phys. Rev. X \*\*10\*\*, 021022 \(2020\)](#).
- [62] D. I. Schuster, A. Wallraff, A. Blais, L. Frunzio, R.-S. Huang, J. Majer, S. M. Girvin, and R. J. Schoelkopf, ac Stark shift and dephasing of a superconducting qubit strongly coupled to a cavity field, [Phys. Rev. Lett. \*\*94\*\*, 123602 \(2005\)](#).