System characterization of dispersive readout in superconducting qubits

Daniel Sank,* Alex Opremcak, Andreas Bengtsson, Mostafa Khezri, Jimmy Chen, Ofer Naaman, and Alexander Korotkov
Google Quantum AI, Goleta, California 93111, USA

(Received 13 March 2024; revised 30 December 2024; accepted 31 January 2025; published 21 February 2025)

Designing quantum systems with the measurement speed and accuracy needed for quantum error correction using superconducting qubits requires iterative design and testing informed by accurate models and characterization tools. We introduce a single protocol, with few prerequisite calibrations, to measure the dispersive shift, resonator linewidth, and drive power used in the dispersive readout of superconducting qubits. We find that the resonator linewidth is poorly controlled, with a factor of 2 variation between the maximum and minimum measured values, and this variation is likely to require focused attention in future quantum error correction experiments. We also introduce a protocol for measuring the efficiency of the readout system using the same power levels as are used in typical qubit readout, and without the need to measure the qubit dephasing. We routinely run these protocols on chips with tens of qubits to a hundred qubits, driven by automation software with little human interaction. Using the extracted system parameters, we find that a model based on these parameters predicts the readout signal-to-noise ratio to within 10% over a device with 54 qubits.

DOI: 10.1103/PhysRevApplied.23.024055

I. INTRODUCTION

The scale and accuracy of quantum computing systems are increasing. Error correction protocols sit on top of an underlying physical level comprising individual qubits, each of which is an analog device with many physical parameters. Can these parameters be extracted reliably, and can the performance of the full system be predicted from a model based on measurements of these parameters? A demonstration of beyond-classical computation in superconducting qubits [1] has answered these questions affirmatively, showing that the cross-entropy in a random circuit sampling experiment was well predicted by a model based on single- and two-qubit component metrics. More recently, the same was shown in experiments with primordial surface codes, where logical error rates were well predicted by model-based component metrics [2,3]. Such precise agreement between prediction and experiment relies on accurate characterization. Furthermore, an accurate model of the readout system is an essential ingredient in tuning a quantum processor for best performance [4].

*Contact author: sank.daniel@gmail.com

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.

The value of accurate characterization extends beyond predicting the performance of and tuning existing devices. Fault-tolerant quantum computation requires faster, more accurate, and lower-leakage readout than has been reported to date in a system with sufficient multiplexing. Achieving the required speed-up while isolating the qubits from decay through the readout system (via e.g., the Purcell effect [5–8]), spectral crosstalk, and dephasing imparted by noise in the readout resonators [9–13] requires improved circuits. As we set ourselves to designing those new circuits, we want to know that the models in which we root our design are founded in agreement with experiment.

To these ends, we have developed a set of experiments that extract the parameters of the dispersive readout system in a superconducting qubit device. To satisfy the needs of increasingly larger devices, these experiments have been designed with automation in mind; they require few prerequisite calibrations and only simple model fitting. These experiments are run routinely on chips with up to 100 qubits via automation software with little human interaction [2,14].

The remainder of this paper is organized as follows. Section II reviews a simple and well-known model of dispersive measurement. Section III describes a single experiment which, using the qubit as a detector of the resonator energy, extracts the resonator dispersive shift, resonator linewidth, and resonator drive power. Section IV presents two additional ways to measure the resonator linewidths and gives an example wherein increased and

reliable measurement statistics revealed the root cause of a systematic design error. Section V introduces a protocol to measure the readout efficiency without measuring the qubit coherence. Section VI compares the measured readout performance with the performance predicted from a model based on our component metrics, finding ±10% accuracy in signal-to-noise ratio (SNR). Finally, Sec. VII considers the future outlook and presents conclusions. The experiments described here underpin the optimization procedure described in detail in Ref. [4] and used in Refs. [1,2,14].

II. MODEL

We study the system illustrated in Fig. 1(a), in which several qubit-resonator pairs are coupled to a common bandpass Purcell filter. As discussed in Ref. [7], the dynamics of each resonator can be accurately described by the simplified circuit shown in Fig. 1(b), with the resonator directly coupled to the input-output line. In this picture, the effect of the Purcell filter is to rescale the linewidth κ of the resonator. In a frame rotating with frequency ω_{frame}, and in the rotating-wave approximation, a single driven resonator is described by the pair of equations

$$\begin{aligned}\dot{\alpha}(t) &= -i\Delta\alpha(t) - \frac{\kappa}{2}\alpha(t) + \sqrt{\kappa}A(t),\\ B(t) &= -A(t) + \sqrt{\kappa}\alpha(t),\end{aligned} \tag{1}$$

where α is the complex resonator amplitude, ω_r is the resonator frequency, $\Delta \equiv \omega_r - \omega_{\text{frame}}$, is the detuning of the resonator from the frame, κ is the resonator's linewidth (assumed to be set by the coupling to the drive line (in this work $1/\kappa \approx 30$ ns), and $A(t)$ and $B(t)$ are the complex amplitudes of the incident (drive) and reflected (measured) waves at the point that the input-output transmission line couples to the resonator. Here, A and B are both in the rotating frame [see Fig. 1(b)] and are normalized such that $|A|^2$ and $|B|^2$ represent photon flux (with dimensions of inverse time), and $|\alpha|^2 = \bar{n}$ is the mean number of photons in the resonator, which is assumed to always be in a coherent state [15]. The resonator frequency depends on the state of the qubit; the qubit states $|0\rangle$ and $|1\rangle$ shift the resonator to frequencies $\omega_{r,|0\rangle}$ and $\omega_{r,|1\rangle}$ with difference $\omega_{r,|1\rangle} - \omega_{r,|0\rangle} = 2\chi < 0$ and "middle frequency" $\omega_{r,m} \equiv (\omega_{r,|0\rangle} + \omega_{r,|1\rangle})/2$.

The different values of resonator frequency for the two qubit states lead to different outgoing waves B, enabling qubit state measurement. For each measurement shot, we amplify and digitize the outgoing wave and compute a complex number

$$Z = C\int_0^T B(t)W(t)\,dt, \tag{2}$$

where $W(t)$ is a dimensionless weight function chosen to maximize the SNR [16,17], T is the total integration time,

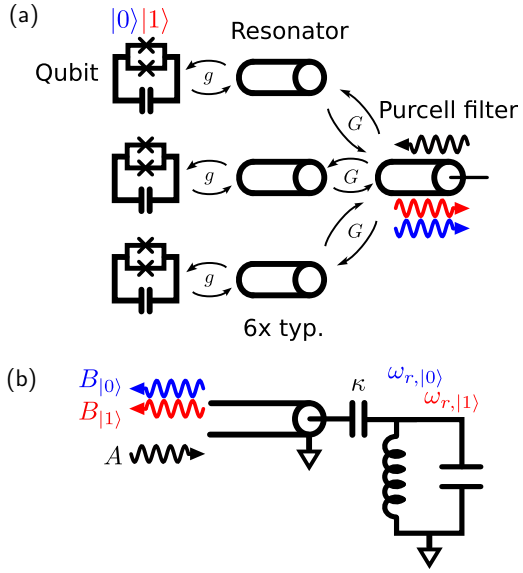


FIG. 1. Schematic diagram of the readout system. (a) The readout circuit includes a shared Purcell filter coupled to six readout resonators. Each readout resonator is coupled to a transmon qubit. The $|0\rangle$ (blue) and $|1\rangle$ (red) states of the transmon correspond to different inductances of the superconducting quantum interference device, which lead to different phases of the reflected signals. (b) The readout circuit can be approximated as a single resonator, with an effective κ, whose frequency depends on the qubit state.

and C is a constant capturing the conversion from the physical signal to digitized samples. The two qubit states $|0\rangle$ and $|1\rangle$ correspond to two different outgoing waves $B_{|i\rangle}$ with two associated points $Z_{|i\rangle}$. The distance in the complex plane between $Z_{|0\rangle}$ and $Z_{|1\rangle}$ is

$$\begin{aligned}\left|Z_{|0\rangle} - Z_{|1\rangle}\right| &= C\left|\int_0^T (B_{|0\rangle}(t) - B_{|1\rangle}(t))W(t)\,dt\right|\\ &= C\sqrt{\kappa}\left|\int_0^T (\alpha_{|0\rangle}(t) - \alpha_{|1\rangle}(t))W(t)\,dt\right|.\end{aligned} \tag{3}$$

Noise in the system adds random displacements to Z. Typically, these displacements are described by a two-dimensional Gaussian distribution. As discussed in detail in the Appendix, the one-dimensional width σ of this distribution is effectively given by $2\sigma^2 = C^2(1/2)\int_0^T |W(t)|^2\,dt$. The SNR is therefore

$$\begin{aligned}\text{SNR} &\equiv \frac{\left|Z_{|0\rangle} - Z_{|1\rangle}\right|^2}{2\sigma^2}\\ &= \eta\kappa\frac{\left|\int_0^T \left(\alpha_{|0\rangle}(t) - \alpha_{|1\rangle}(t)\right)W(t)\,dt\right|^2}{(1/2)\int_0^T |W(t)|^2\,dt},\end{aligned} \tag{4}$$

where the quantum efficiency $\eta < 1$ describes the reduction in SNR due to loss and added noise as the signal

travels from the readout resonator to the detector [18–20]; achieving $\eta = 1$ is possible only with a phase-sensitive amplifier and is limited to $\eta \leq 0.5$ with a phase-preserving amplifier, as is used in this experiment. The remainder of this paper describes measurements of the parameters appearing in Eq. (1) and comparison of Eq. (4) against experiment.

For some measurements, it is convenient to operate in the steady state, where the drive A is monochromatic, has constant magnitude, and has frequency equal to ω_{frame}, so that $\dot{\alpha} = 0$. Solving Eq. (1) in that case, we find

$$\alpha = \frac{2A}{\sqrt{\kappa}} \frac{1}{1 + i2\Delta/\kappa} \tag{5}$$

with photon occupation

$$\bar{n} = |\alpha|^2 = \frac{\kappa\,|A|^2}{\Delta^2 + (\kappa/2)^2}. \tag{6}$$

III. DISPERSIVE SHIFT, LINEWIDTH, AND POWER

Equation (4) for the SNR and related equations for the outgoing field B have four parameters: the dispersive shift χ, the resonator linewidth κ, the drive strength A, and the efficiency η. We measure the first three of these parameters in a single experiment based on the ac Stark effect, using the pulse sequence shown in Fig. 2(a). The qubit begins at the idle frequency $\omega_{q,0}$. We prepare the qubit in either $|0\rangle$ or $|1\rangle$ and then drive the resonator at a frequency $\omega_{d,r}$ ("omega drive resonator") for a duration much longer than $1/\kappa$ so that the resonator is driven into the steady state with a certain photon occupation $\bar{n}$. This resonator occupation causes the qubit frequency to shift (ac Stark shift) by $\delta\omega_q \approx 2\chi\bar{n}$ [21]. We then apply a pulse to the qubit at a frequency $\omega_{d,q}$ ("omega drive qubit"), which flips the qubit state if $\omega_{d,q} = \omega^*_{d,q} \approx \omega_{q,0} + \delta\omega_q = \omega_{q,0} + 2\chi\bar{n}$. Finally, we allow the resonator to ring down, and then we measure the state of the qubit using dispersive readout through that same resonator. The frequencies $\omega_{d,r}$ and $\omega_{d,q}$ are swept over a two-dimensional grid, and for each grid point, we record the probability that the final qubit state was different from the one prepared at the start of the pulse sequence. In other words, we record the probability that the qubit probe pulse frequency matched the ac-Stark-shifted frequency of the qubit in the presence of the resonator drive. Therefore, we are using the ac-Stark-shifted qubit frequency as a sensor of the resonator's internal energy. This protocol requires $1/\kappa \ll T_1$, where T_1 is the lifetime of the qubit state $|1\rangle$.

Results from an experiment are shown in Figs. 2(b) and 2(c). Typical values of $\omega_{d,r}/(2\pi)$ are $\omega_{r,m}/(2\pi) \pm 10$ MHz, and typical values of $\omega_{d,q}/(2\pi)$ are $\omega_{q,0}/(2\pi)^{+10\text{ MHz}}_{-30\text{ MHz}}$. Taking our value for $\bar{n}$ from Eq. (6) with $\Delta = (\omega_{r,m} \mp \chi) - \omega_{d,r}$, we find the qubit drive frequency

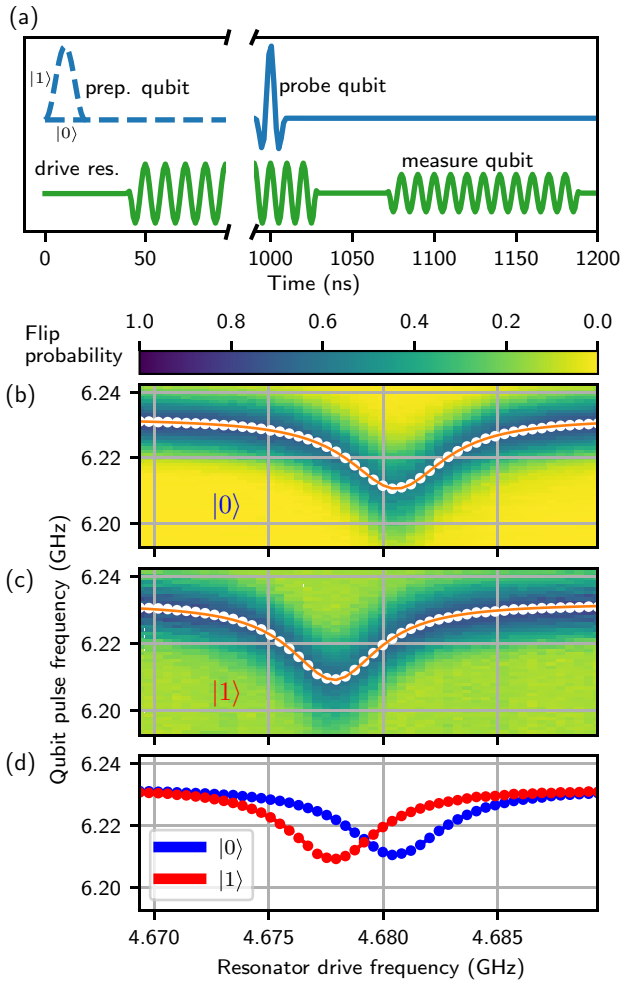


FIG. 2. The chi-kappa-power experiment. (a) Pulse sequence. The blue and green curves indicate the qubit and resonator drives, respectively. The dashed line indicates that the qubit preparation pulse is absent (present) for preparation of $|0\rangle$ ($|1\rangle$). The first resonator pulse at a variable frequency populates the resonator and ac Stark shifts the qubit, while the second resonator pulse at a fixed frequency reads out the final qubit state. (b) Probability of measuring the qubit in the opposite state from which it was prepared. The dark color bands, forming Lorentzian curves, indicate high probability. The white dots indicate the fitted minimum for each vertical slice of data, and the orange curves indicate the Lorentzian fits. (c) Lorentzian fits for $|0\rangle$ and $|1\rangle$.

where we expect to see the qubit state to flip,

$$\omega^*_{d,q} = \omega_{q,0} + 2\chi \frac{\kappa\,|A|^2}{(\omega_{d,r} - (\omega_{r,m} \mp \chi))^2 + (\kappa/2)^2}, \tag{7}$$

where the upper and lower signs are for the qubit in $|0\rangle$ and $|1\rangle$, respectively. Equation (7) describes a pair of Lorentzian curves with means $\omega_{r,m} \mp \chi$, depths $(2\chi)(4\,|A|^2/\kappa)$, and widths κ. We fit both curves, finding 2χ as the horizontal distance between their means and κ as their widths. Then, with χ and κ known, we find $|A|^2$

from their depths. We assume that $|A|^2$ is proportional to the pulse-generator power [22], so this calibration can be understood as a calibration between the generator power and the power incident on the resonator. The Lorentzian fits are shown overlaid with the raw data in Fig. 2(b), and the two fits are shown together in Fig. 2(c).

This experiment, which we call "chi-kappa-power" (CKP), is easy to calibrate because it requires only crude prior knowledge of the resonator and qubit frequencies and crude calibration of the qubit control pulses and readout. The two independent variables are both frequencies requiring no calibration, and we have no strong requirement on the contrast in the dependent variable, i.e., the qubit excitation probability indicated by the color scale in Fig. 2. Perhaps most importantly, the model in Eq. (7) is invariant across readout circuit designs and the CKP data is always fitted to a pair of Lorentzians with exactly five fit parameters.

This invariance in the analysis of the CKP experiment compares favorably against the traditional method of fitting microwave scattering parameters (S-parameters). In the traditional method, after the qubit state is prepared, a probe wave is applied to the resonator (the qubit is never measured). The scattered (transmitted or reflected) probe wave carries a phase and magnitude shift that depends on the probe frequency, the resonator frequency, and the resonator linewidth, i.e., the same parameters appearing in Eq. (7) except for $\omega_{q,0}$ and A. The probe frequency is swept, and the frequency-dependent S-parameters are fitted to a model. Unfortunately, that model is complicated by three factors. First, the microwave signal path typically involves several meters of cable, at least a dozen connectors, circulators and isolators, several amplifiers, and any number of stray couplings, e.g., between wire bond pads on the chip. Each of these components introduces its own contribution to the measured S-parameters. Second, in a multiplexed readout system, the frequency bandwidth is a limited resource, and the frequency spacing between resonators is only a few times larger than their linewidths. In this situation, each resonator's neighbors in frequency affect that resonator's S-parameters, so we have to fit the S-parameters for multiple resonators simultaneously. Third, real readout systems use Purcell filters [5,6,8], which introduce further modifications to the S-parameters; in particular, these modifications are *different* for each filter design. The complexity of measuring and fitting S-parameters is reflected throughout the quantum computing and microwave kinetic inductance detector literature in papers dedicated to handling nonidealities like those described above [23–27]. We find that these complexities in fitting S-parameters render the S-parameter method unsuitable for large-scale device characterization; however, because each resonator is coupled to exactly one qubit acting as an energy sensor for that resonator, the qubit ac Stark shifts measured in CKP show the simple Lorentzian profile seen in Fig. 2, making fitting easy and reliable. Moreover, CKP captures the resonator drive magnitude $|A|$, while the S-parameter method does not. Simulation of the CKP experiment shows that it extracts the correct values of χ, κ, and A to within a few percent.

The CKP experiment does have some basic prerequisites. Before running it, the qubit pi pulses and basic readout are calibrated as described in Ref. [28], from which we learn the approximate values of $\omega_{q,0}$, $\omega_{r,|0\rangle}$, and the duration and magnitude of the pi pulse. For CKP, we fix the qubit pulse duration to be 100 ns and scale the pulse amplitude so that the pulse area is equal to $0.8\times$ the area in a pi pulse. To choose the resonator drive power A in the CKP experiment, we drive the resonator at $\omega_{r,|0\rangle}$ and find the power at which the observed qubit ac Stark shift is -20 MHz. Driving the resonator too hard leads to measurement-induced state transitions [29–31] and spoils the CKP data.

IV. OTHER MEASUREMENTS OF KAPPA

The CKP experiment measures κ in the frequency domain as resonator line widths, but κ is also the energy ringdown rate of the resonator in the time domain. We measure κ in the time domain with two different protocols: (1) observing the time-dependent ringdown of the qubit ac Stark shift, and (2) observing the time-dependent voltage leaked from the resonator itself.

A. ac Stark ringdown

In this method, we use the qubit's ac Stark shift as a probe of the resonator energy, just as in the CKP experiment, but now in the time domain. The pulse sequence is similar to the CKP sequence. The qubit is left in state $|0\rangle$ and the resonator is driven for a duration several times $1/\kappa$ to put it into the steady state. The resonator drive is turned off, and then a pulse of variable frequency is applied to the qubit at a variable time. In this way, we find the qubit frequency as a function of the time after the end of resonator drive. As the energy in the resonator decays as $\exp(-\kappa t)$, and because the qubit ac Stark shift is approximately linear in the resonator energy, we see the qubit frequency relax exponentially to its idle value. Fitting this exponential curve gives us κ. Figure 3(a) shows an example of the measured time-dependent qubit ac Stark shift with an exponential fit. We call this method "ac Stark ringdown." The ac Stark ringdown method has advantages similar to CKP: it avoids fitting an S-parameter model to the multiplexed readout system, and it requires only crude knowledge of the qubit and resonator frequencies.

B. Voltage ringdown

The voltage ringdown method is similar to the ac Stark ringdown method, except that we measure the

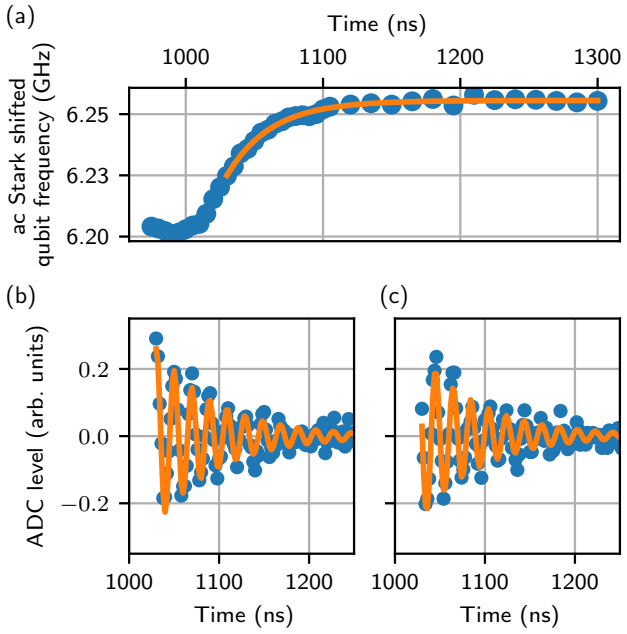


FIG. 3. Kappa measured in the time domain. The blue dots represent experimental measurements and the orange lines indicate fits. (a) Time-dependent ac Stark shift of the qubit. (b),(c) Decaying oscillations of the voltage arriving at the receiver's I and Q channel analog-to-digital converters.

time-dependent voltage arriving at the receiver and fit its exponential decay. The voltage decays with rate $\kappa/2$ instead of the energy decay rate κ. Figures 3(b) and 3(c) show examples of the oscillating and decaying voltages collected by the I and Q channels of the receiver.

C. Correlations between measurement methods

Figure 4 shows the correlations between our three methods for measuring κ. Evidently, the two methods based in the ac Stark effect agree the best, while the method using the ringdown of the received voltage agrees the least well with the other two.

D. Distribution

The designed value of κ for the device used in all measurements described so far was $(30\ \text{ns})^{-1}$, but as we can see in Fig. 4, the measured values range from $(40\ \text{ns})^{-1}$ to $(20\ \text{ns})^{-1}$, a factor of 2 variation from smallest to largest. This is a variation over resonator-qubit pairs on the chip, not a fluctuation in time. Such a large variation poses a problem for device design. Any given requirement on readout duration implies a lower bound on κ. Suppose we designed such that our smallest realized κ values were at the design value of $(30\ \text{ns})^{-1}$; then, with a $2\times$ variation, our largest κ would be $(15\ \text{ns})^{-1}$, which would incur too much damping on the qubits [5,6]. On the other hand, if we design defensively so that the fastest κ were $(30\ \text{ns})^{-1}$, then the slowest would be $(60\ \text{ns})^{-1}$, meaning our readout system would be approximately $2\times$ slower than intended. Because the readout is multiplexed, the slowest readout determines the repetition rate of cyclic protocols like the surface code [4], so even a few resonators in the low-κ tail of the distribution could be problematic.

We suspected that the root cause of the variance in κ might be an impedance mismatch in the receiver, specifically coming from a reflection at the custom infrared (IR) filter or circulator closest to the qubit chip. We constructed a "cryo-VNA" apparatus, similar to the one described in Ref. [25], enabling *in situ* microwave calibration near the plane of the device under test and used it to measure the reflection coefficients of the IR filter and circulator. The results are shown in Fig. 5. If the receiver impedance mismatch were the only source of variation in κ, then the ratio of the largest and smallest observed values of κ would be related to the receiver's reflection coefficient via the voltage standing wave ratio (VSWR):

$$\frac{\kappa_{\max}}{\kappa_{\min}} = \left(\frac{1+|S_{11}|}{1-|S_{11}|}\right)^2 = \text{VSWR}^2. \tag{8}$$

Near the center of our readout band at 4.8 GHz, the IR filter reflection coefficient is as high as -16 dB (corresponding to a wave impedance mismatch in the range $[36\ \Omega, 68\ \Omega]$ in a 50 Ω system) which, via Eq. (8), predicts $\kappa_{\max}/\kappa_{\min} \approx 1.9$, close to the observed value of 2; however, notice that the $|S_{11}|^2$ curves of the two IR filters shown in Fig. 5 differ by 4 dB in the readout band near 4.8 GHz, and note that the particular IR filters and circulators characterized in Fig. 5 are not the same physical units used in the experiments with qubits represented in Fig. 4. Therefore, we do not expect perfect agreement between the observed κ variation and the variation predicted by Eq. (8) and values of $|S_{11}|$ measured in the cryo-VNA. Indeed, in another setup, we found that removing the IR filter and using a circulator with a better impedance match reduced the spread in κ from $3\times$ to $2\times$ (data not shown).

These findings suggest that impedance matching is a critical aspect of a multiplexed superconducting qubit readout system, as it plays a role in determining the system's maximum speed. In these systems, signal losses must be minimized, particularly between the qubit chip and the first amplifier (usually a Josephson parametric amplifier), so using attenuators to improve the impedance match before the first amplifier is unlikely to be viable. These results highlight the importance of the microwave design of the sample package and readout receiver, which should be the subject of further study and development.

After initial submission of this work, Ref. [32] was published, which presents a similar discussion about the importance of off-chip impedance mismatches in controlling κ. That work provides a detailed investigation into

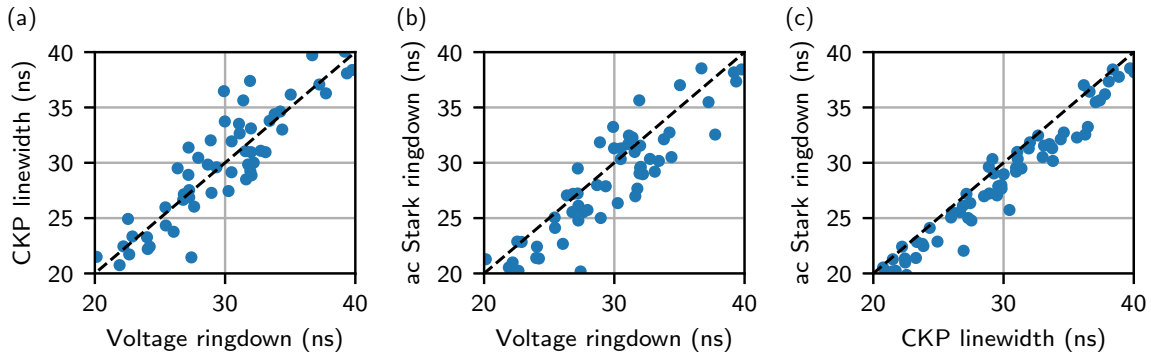


FIG. 4. Comparison of three methods of measuring $1/\kappa$. (a) Correlation between the CKP and voltage ringdown methods. (b) Correlation between the ac Stark ringdown and voltage ringdown methods. (c) Correlation between the ac Stark ringdown and CKP methods.

how intentional impedance mismatches in the readout circuit affect the sensitivity of κ to unwanted impedance mismatches.

V. READOUT EFFICIENCY

With χ, κ, and the resonator drive power measured in the CKP experiment, we turn to the readout quantum efficiency η. The efficiency is the fraction of available information from the readout that is actually collected and contributes to the SNR. A common method for measuring quantum efficiency uses the fact that measuring the qubit energy necessarily causes the qubit to dephase [33,34]. The magnitude of the off-diagonal element of the qubit density matrix ρ_{10} is bounded by $|\rho_{10}| \leq (1/2)\exp(-\mathrm{SNR}/4)$. For a given resonator drive pulse and measured $|\rho_{10}|$, the efficiency is the ratio of the measured SNR to the maximum SNR allowed by the inequality. This method has a few limitations. First, the coherence ρ_{10} typically decays over time due to frequency noise sources independent of the readout signal, so the measured value of $|\rho_{10}|$ does not directly indicate the dephasing due to the resonator drive. Second, by construction, the SNR levels of interest for projective measurement correspond to tiny values of $|\rho_{10}|$ because projective measurement means that the qubit state is projected to one of the poles of the Bloch sphere were the phase coherence is zero. Therefore, it is awkward to use this strategy for measuring efficiency with the same resonator drive power levels as would actually be used in, e.g., a quantum error correction code based on projective measurements.

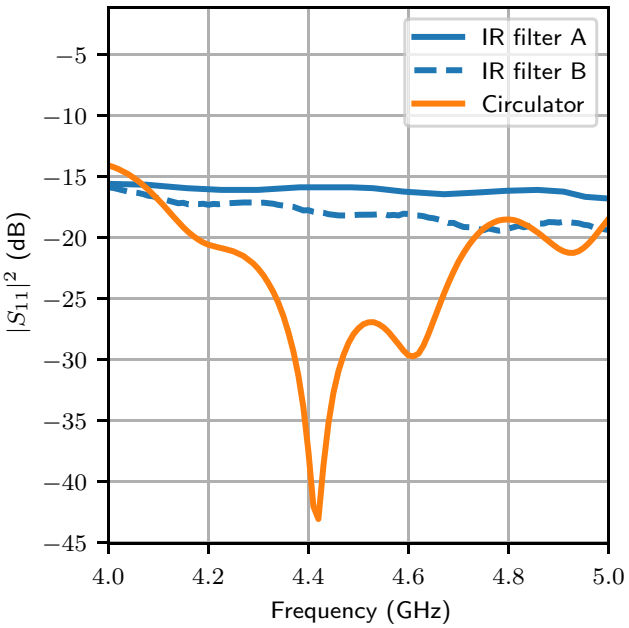


FIG. 5. Reflection coefficient of the IR filter and circulator.

Our approach is based on the fact that with χ, κ, and power calibrated from the CKP experiment, we can predict the rate of increase of SNR with increasing resonator drive pulse length. With the rotating frame frequency at $\omega_{r,m}$ so that $\Delta = \mp\chi$, the steady-state resonator fields for the qubit in states $|0\rangle$ and $|1\rangle$ are

$$\alpha_{|0\rangle,|1\rangle} = \frac{2A}{\sqrt{\kappa}} \frac{1}{1 \pm ix}, \tag{9}$$

where $x = -2\chi/\kappa > 0$ and the upper and lower signs are for $|0\rangle$ and $|1\rangle$. The distance between the fields is

$$\left|\alpha_{|0\rangle} - \alpha_{|1\rangle}\right| = 4\frac{|A|}{\sqrt{\kappa}} \frac{x}{1+x^2}, \tag{10}$$

and the mean number of photons in either state is

$$\bar{n} = |\alpha|^2 = \frac{4\,|A|^2}{\kappa} \frac{1}{1+x^2} = \bar{n}^* \frac{1}{1+x^2}, \tag{11}$$

where $\bar{n}^* = 4\,|A|^2/\kappa$ is the number of photons that would be in the resonator if the resonator were driven on resonance. Combining Eqs. (4), (10), and (11), we find the rate

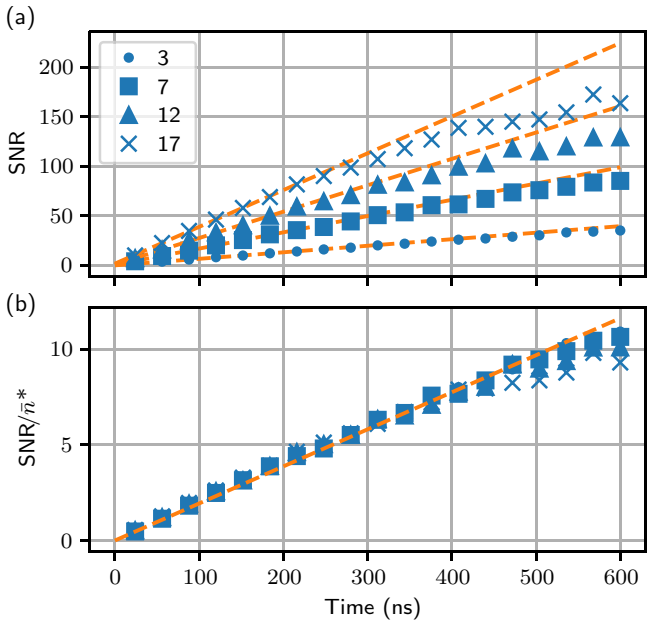


FIG. 6. (a) SNR versus integration time for several values of $\bar{n}^*$. The dashed lines are linear fits to the first 200 ns of data. (b) The same data but normalized by $\bar{n}^*$. The dashed line is a guide to highlight the curve bending.

of increase in SNR per time when the resonator is in the steady state as

$$\frac{d\mathrm{SNR}}{dt} = \eta\,\bar{n}\,\kappa\left(\frac{1}{1/2}\right)\left(\frac{4x^2}{1+x^2}\right) \tag{12}$$

$$= \eta\,\bar{n}^*\,\kappa\left(\frac{1}{1/2}\right)\left(\frac{2x}{1+x^2}\right)^2. \tag{13}$$

The three factors in Eq. (12) can be roughly interpreted as the leaked photon flux, the noise, and the information per photon. With χ, κ, and $\bar{n}$ known, a measurement of the "SNR flux" $d\mathrm{SNR}/dt$ leaves η as the only free parameter, which we find by fitting a line to the SNR flux versus time. To measure $d\mathrm{SNR}/dt$, we prepare the qubit in either $|0\rangle$ or $|1\rangle$ and drive the resonator for several times $1/\kappa$ to put it into the steady state; we then continue driving the resonator while collecting the dispersed signal for a variable time. The results are shown in Fig. 6.

We repeat the experiment for a few different values of the drive power, observing that the SNR flux depends linearly on the drive power $\bar{n}^*$ as expected, i.e., that we are not saturating the parametric amplifier. This approach was used in Ref. [35] to characterize the readout efficiency with SNake Impedance Matched Parametric Amplifier (SNIMPA) parametric amplifiers. Note that with phase-preserving amplifiers, as used here, the maximum attainable efficiency is $\eta = 0.5$.

The data in Fig. 6 do not fall on completely straight lines. In some experiments, this happens because the qubit energy decay time (T_1) is too short, causing the cloud of points for $|1\rangle$ to distort toward the cloud for $|0\rangle$, reducing the extracted SNR. Interestingly, in some experiments, we notice the curves bend over at a particular value of SNR regardless of drive power. The origin of this bending at a certain value of SNR is not understood.

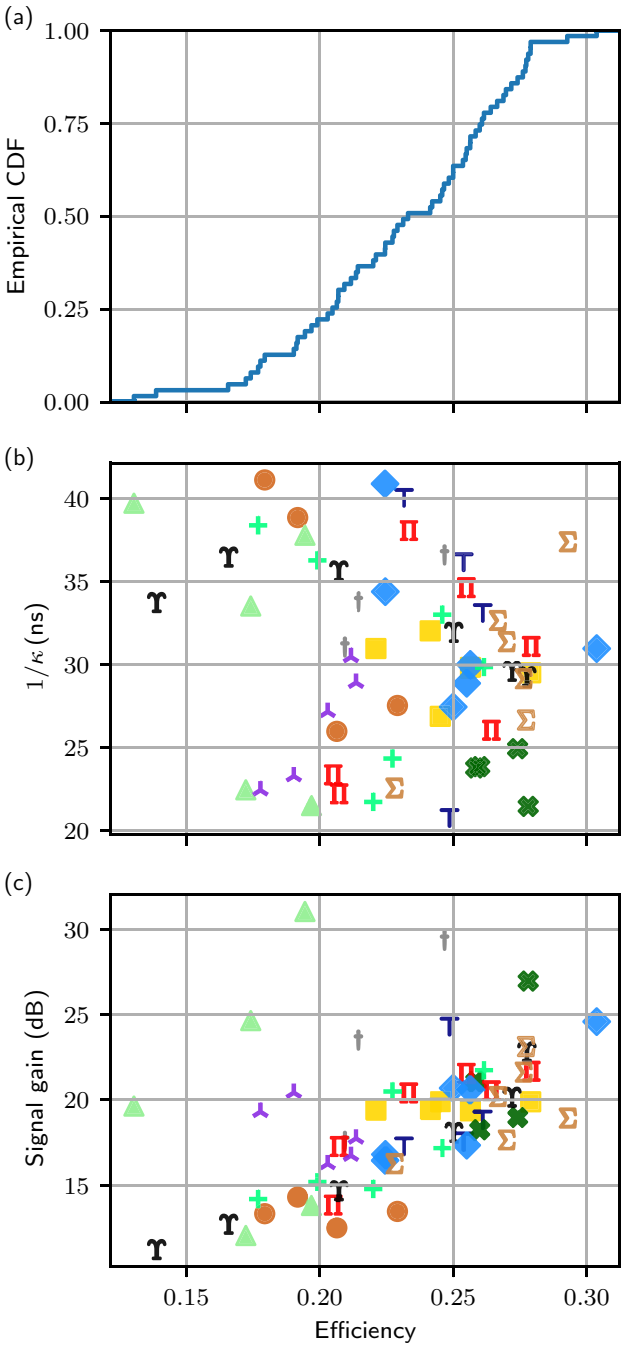


FIG. 7. (a) Cumulative distribution function of the readout efficiency. (b) Readout efficiency versus $1/\kappa$. (c) Readout efficiency versus parametric amplifier gain. In (b),(c), each marker symbol corresponds to a different multiplexed readout line on the chip.

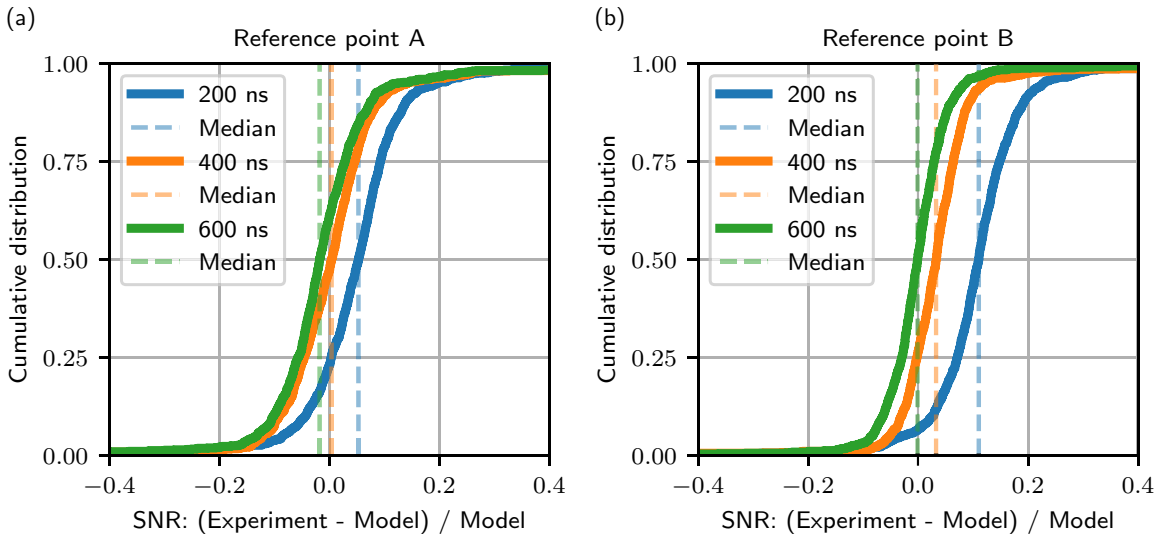


FIG. 8. Cumulative distribution of the relative error in measured versus predicted readout SNR. (a) Qubits calibrated at reference point A. (b) Qubits calibrated at reference point B.

We also study the statistical distribution of the measured efficiencies. Figure 7 shows the cumulative distribution function of η along with scatter plots of η versus $1/\kappa$ and versus the parametric amplifier gain (gain is measured by observing the signal level as we turn the parametric amplifier pump on and off). Evidently, η is uncorrelated with $1/\kappa$ as expected, having a Pearson correlation of 5%; however, we see increasing η with increasing gain (in dB) with a Pearson correlation of 58%. This correlation of efficiency with gain is unsurprising as the receiver hardware (package, cabling, circulators) used here are similar to that used in Ref. [35]. Using Eq. (3) in that reference with $\alpha = 0.5$, $T_Q = T_P = 112$ mK, and $T_H = 2.5$ K, we estimate that as the parametric amplifier gain ranges from 10 to 20 dB, the efficiency should range approximately from 12% to 23%, in line with our observations here.

VI. MODEL AND MEASUREMENTS OF READOUT SNR

With our system parameters measured, we come to comparison of predicted and measured SNR. We measure the SNR in single-qubit readout on 54 qubits in a Sycamore-style device. Whereas the system-characterization experiments discussed above are conducted with the resonator in the steady state, here, we use readout pulses. For each qubit, we measure the system parameters described above with the qubit at two different frequencies, which we refer to as reference points A and B. We then vary the qubits' frequencies in a 360 MHz band around these reference points. At each frequency, we tune the readout using Eq. (4) and the system parameters from the reference point, targeting SNR = 15 (corresponding to a separation error [6] of 0.3%). We repeat this procedure for three different pulse durations. For each qubit frequency and pulse duration, we model the expected SNR via Eq. (4). We plot the disagreement between the measured and modeled SNR, marginalized over the qubit frequencies, in Fig. 8, finding most of the data within $\pm 10\%$. Interestingly, the prediction seems to be the least accurate for the shortest pulse duration (200 ns), where the measured SNR is somewhat *larger* than predicted by the model. In any case, the agreement between the experiment and the model is good enough that we can continue to use it as the basis for development of more advanced designs.

VII. CONCLUSIONS

We have shown methods for measuring the four most important parameters in a dispersive readout system. The dispersive shift χ, resonator line width κ, and drive power at the readout resonator, were measured in a single self-calibrating procedure (CKP). Three different measurements of κ—voltage ringdown, ac Stark ringdown, and CKP—gave similar results, with CKP and ac Stark ringdown agreeing the best. The efficiency can be measured by comparing the predicted and observed SNR. Combining the measured parameters into a simple model predicted the SNR in single-qubit readout on a Sycamore processor with 54 qubits to $\pm 10\%$. We found large variation in κ and that attention to impedance matching in the readout receiver reduces this variation. Further improvements to the variation in κ are required for future fault-tolerant quantum computing systems.

Two future directions suggested by this work are (1) improved control over κ and (2) accurate measurement and modeling of multiplexed readout with sufficient speed for fault-tolerant quantum computing.

ACKNOWLEDGMENTS

We thank Ebrahim Forati, Reza Fatemi, and Juan Atalaya for critical readings of the manuscript. The Google Quantum AI team fabricated the processor, built the cryogenic and control systems, optimized the processor performance, and provided the tools that enabled execution of this experiment.

APPENDIX: SNR MODEL

This appendix presents the details of the readout receiver, including analog mixing and filtering, and digital signal processing. The discussion centers around derivation of Eq. (4) via frequency-domain analysis, with attention drawn to the various assumptions and limits under which that equation is correct. While the analysis includes consideration of discrete time sampling, we will see that the analytic result is valid precisely when the filtering and sampling are designed to preserve the properties of the incoming analog signals and noise. All integrals in this appendix are taken over the interval $[-\infty, \infty]$ unless otherwise noted.

A schematic of the receiver is shown in Fig. 9. The analog signal (and noise) come in on a transmission line with characteristic impedance Z_0. The signal is mixed to intermediate frequencies in an IQ mixer, filtered in lowpass antialiasing filters with frequency response $\tilde{h}$, sampled at discrete times spaced by δt, and then digitally processed to produce a single point Z in the complex IQ plane.

1. Signal

Let the incoming microwave voltage pulse be denoted $V_{\mu w}$. Because $V_{\mu w}$ is real, it can be written as a Fourier transform over positive frequencies:

$$V_{\mu w}(t) = \int_0^\infty \frac{d\omega}{2\pi} \tilde{V}_{\mu w}(\omega) e^{i\omega t} + \int_0^\infty \frac{d\omega}{2\pi} \tilde{V}_{\mu w}(\omega)^* e^{-i\omega t}. \tag{A1}$$

If $V_{\mu w}$ has nonzero content only in a band $[\Omega - B, \Omega + B]$, then its Fourier-integral representation can be rewritten by a change of variables $\omega = \Omega + \delta\omega$ (where $\delta\omega$ is the detuning away from the local oscillator) as

$$V_{\mu w}(t) = \int_{-B}^{B} \frac{d\delta\omega}{2\pi} \left(\tilde{V}_{\mu w}(\Omega + \delta\omega) e^{i(\Omega+\delta\omega)t} + \tilde{V}_{\mu w}(\Omega + \delta\omega)^* e^{-i(\Omega+\delta\omega)t} \right) \tag{A2}$$

$$= 2\mathrm{Re}\left[\int_{-B}^{B} \frac{d\delta\omega}{2\pi} \tilde{V}_{\mu w}(\Omega + \delta\omega) e^{i(\Omega+\delta\omega)t} \right]. \tag{A3}$$

The microwave signal is passed through the IQ mixer, resulting in two IF signals

$$I_{\mathrm{mixer}}(t) = A\cos(\Omega t) V_{\mu w}(t) = A\,\mathrm{Re}\left[\int_{-B}^{B} \frac{d\delta\omega}{2\pi} \tilde{V}_{\mu w}(\Omega + \delta\omega) e^{i\delta\omega t} \right], \tag{A4}$$

$$Q_{\mathrm{mixer}}(t) = -A\sin(\Omega t) V_{\mu w}(t) = A\,\mathrm{Im}\left[\int_{-B}^{B} \frac{d\delta\omega}{2\pi} \tilde{V}_{\mu w}(\Omega + \delta\omega) e^{i\delta\omega t} \right], \tag{A5}$$

where the factor A accounts for conversion loss in the mixer and we have dropped the higher-frequency (2Ω) mixing products because they are filtered away in the next step. The IF signals are passed through lowpass (antialiasing) filters, resulting in

$$I(t) = A\,\mathrm{Re}\left[\int_{-B}^{B} \frac{d\delta\omega}{2\pi} \tilde{V}_{\mu w}(\Omega + \delta\omega) \tilde{h}(\delta\omega) e^{i\delta\omega t} \right], \tag{A6}$$

$$Q(t) = A\,\mathrm{Im}\left[\int_{-B}^{B} \frac{d\delta\omega}{2\pi} \tilde{V}_{\mu w}(\Omega + \delta\omega) \tilde{h}(\delta\omega) e^{i\delta\omega t} \right]. \tag{A7}$$

Note that $\tilde{h}$ is dimensionless. Because $I(t)$ and $Q(t)$ are the real and imaginary parts of the same quantity, it is mechanically and conceptually convenient to define $Z(t) = I(t) + iQ(t)$, or in the frequency domain, $\tilde{Z}(\omega) = A\tilde{V}_{\mu w}(\Omega + \omega)\tilde{h}(\omega)$.

Given a frame rotating with frequency Ω, we define the rotating-frame representation $\bar{V}_{\mu w}$ of the microwave signal

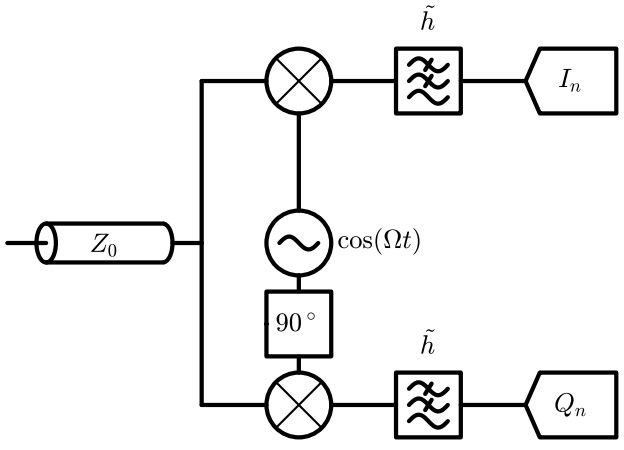


FIG. 9. The readout receiver. The signal and noise enter from the left along a transmission line. They are mixed to an intermediate frequency (IF) in an IQ mixer, here modeled as a local oscillator driving two mixers with a 90° phase shift. The IF signals are lowpass filtered and then digitized.

using the equation $V_{\mu w}(t) = \mathrm{Re}\left[e^{i\Omega t}\bar{V}_{\mu w}(t)\right]$, from which we find via Eq. (A3)

$$\bar{V}_{\mu w}(t) = 2\int_{-B}^{B}\frac{d\delta\omega}{2\pi}\tilde{V}_{\mu w}(\Omega+\delta\omega)e^{i\delta\omega t} \tag{A8}$$

or $\tilde{\bar{V}}_{\mu w}(\omega) = 2\tilde{V}_{\mu w}(\Omega+\omega)$. The postfilter IF signal can then be written as

$$\tilde{Z}(\omega) = \frac{A}{2}\tilde{\bar{V}}_{\mu w}(\omega)\tilde{h}(\omega). \tag{A9}$$

The IF signal is digitally sampled, forming the sequence

$$Z_n = D\int\frac{d\omega}{2\pi}\tilde{Z}(\omega)e^{i\omega n\delta t}, \tag{A10}$$

where D is a conversion factor from voltage to digital amplitude (D has dimensions of 1/voltage) and δt is the sampling interval. In the work described here, $\delta t = 1$ ns. The demodulated IQ point Z is computed digitally as

$$\begin{aligned} Z &= \frac{1}{N}\sum_{n=0}^{N-1} Z_n e^{-i\,\omega_{\mathrm{sb}}\,n\delta t}W_n \\ &= \frac{AD}{2N}\sum_{n=0}^{N-1}\int\frac{d\omega}{2\pi}\tilde{\bar{V}}_{\mu w}(\omega)\tilde{h}(\omega)e^{i\omega n\delta t}e^{-i\omega_{\mathrm{sb}}\,n\delta t}W_n, \end{aligned} \tag{A11}$$

where ω_{sb} is the sideband frequency of the channel and W_n defines a "window function" used to optimize the SNR [16, 17]. We define a dimensionless variable $x = (\omega-\omega_{\mathrm{sb}})\delta t$ and move the sum inside the integrals to get

$$Z = \frac{AD}{2N\delta t}\int\frac{dx}{2\pi}\tilde{\bar{V}}_{\mu w}(x/\delta t+\omega_{\mathrm{sb}})\tilde{h}(x/\delta t+\omega_{\mathrm{sb}})\sum_{n=0}^{N-1}e^{inx}W_n. \tag{A12}$$

We now assume that W_n can be interpolated as a smooth function $W:\mathbb{R}\to\mathbb{R}$. The discrete samples W_n can then be written in the frequency domain as

$$W_n = \frac{1}{\delta t}\int\frac{dy}{2\pi}\tilde{W}(y/\delta t)e^{iyn}. \tag{A13}$$

Putting everything together, we have

$$\begin{aligned} Z = \frac{AD}{2N\delta t^2}\int\frac{dx}{2\pi}&\tilde{\bar{V}}_{\mu w}(x/\delta t+\omega_{\mathrm{sb}})\tilde{h}(x/\delta t+\omega_{\mathrm{sb}}) \\ &\times\int\frac{dy}{2\pi}\tilde{W}(y/\delta t)\sum_{n=0}^{N-1}e^{i(x+y)n}. \end{aligned} \tag{A14}$$

The sum is the known Dirichlet kernel K, defined as $K(N,u) = \sin(Nu/2)/\sin(u/2)$. In the limit $N\to\infty$, the Dirichlet kernel approaches a comb of Dirac delta functions $K(N\to\infty,u) = \sum_{k=-\infty}^{\infty}(2\pi)\delta(u-2\pi k)$. Our first assumption is that N is large enough to use this Dirac comb approximation. We can now take the integral over y to find

$$\begin{aligned} Z = \frac{AD}{2N\delta t^2}\int\frac{dx}{2\pi}&\sum_{k=-\infty}^{\infty}\tilde{\bar{V}}_{\mu w}(x/\delta t+\omega_{\mathrm{sb}}) \\ &\times\tilde{h}(x/\delta t+\omega_{\mathrm{sb}})\tilde{W}((2\pi k-x)/\delta t) \end{aligned} \tag{A15}$$

and, having handled the sum over n, we switch back to $\omega = x/\delta t$:

$$\begin{aligned} Z = \frac{AD}{2N\delta t}\int\frac{d\omega}{2\pi}&\sum_{k=-\infty}^{\infty}\tilde{\bar{V}}_{\mu w}(\omega+\omega_{\mathrm{sb}}) \\ &\times\tilde{h}(\omega+\omega_{\mathrm{sb}})\tilde{W}(2\pi k/\delta t-\omega). \end{aligned} \tag{A16}$$

In a frequency-multiplexed system, the microwave signal in the rotating frame has several components

$$\bar{V}_{\mu w}(t) = \sum_{\text{channel }c} e^{i\omega_{\mathrm{sb},c}t}\bar{V}^{c}_{\mu w}(t), \tag{A17}$$

where each $\bar{V}^{c}_{\mu w}$ is a slowly varying envelope and each component in the sum is meant to probe one qubit's readout resonator. In the frequency domain,

$$\tilde{\bar{V}}_{\mu w}(\omega) = \sum_{\text{channel }c}\tilde{\bar{V}}^{c}_{\mu w}(\omega-\omega_{\mathrm{sb},c}). \tag{A18}$$

In this case, we find

$$\begin{aligned} Z = \frac{AD}{2N\delta t}\int\frac{d\omega}{2\pi}&\sum_{k=-\infty}^{\infty}\sum_{\text{channel }c}\tilde{\bar{V}}^{c}_{\mu w}(\omega+\omega_{\mathrm{sb}}-\omega_{\mathrm{sb},c}) \\ &\times\tilde{h}(\omega+\omega_{\mathrm{sb}})\tilde{W}(2\pi k/\delta t-\omega). \end{aligned} \tag{A19}$$

We now make five further assumptions. First, we assume that the various $\tilde{\bar{V}}^{c}_{\mu w}$ do not appreciably overlap. The degree to which this assumption is valid is a critical aspect of readout design. As multiplexing factors are increased and readout pulse duration is decreased, keeping the various frequency channels from overlapping requires separating them further in frequency space, which requires higher-bandwidth amplifiers, sharp-cutoff Purcell filters, larger qubit-resonator detuning, and consequently larger qubit-resonator coupling. While we do not study this issue in detail here, we mention that we have studied this issue analytically, numerically, and experimentally, finding that for realistically achievable parameters, multiplexing beyond ten qubits with readout fast enough for error correction will be a challenge. In short, frequency crowding is likely to be one of the most important challenges in superconducting qubit readout, and this should be investigated in depth.

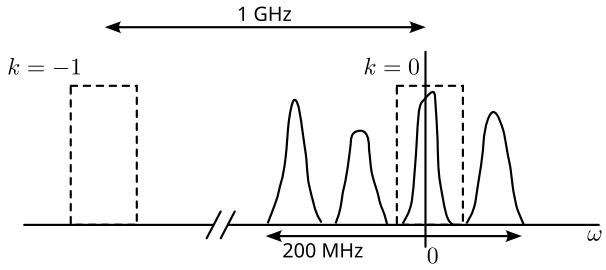


FIG. 10. Diagram of the frequency-domain integrand in Eq. (A19). The peaks drawn in solid lines represent the frequency components $\tilde{\tilde{V}}^c_{\mu w}$, and the dashed-line boxes represent $\tilde{W}(2\pi k/\delta t - \omega)$ for $k = 0$ and -1.

Second, we assume that the total bandwidth occupied by all $\tilde{\tilde{V}}^c_{\mu w}$ is smaller than $(2\pi)/\delta t$. Third, we choose ω_{sb} to be equal to one of the $\omega_{sb,c}$, which we denote ω_{sb,c^*}. This is just saying that the sideband frequency used in digital mixing is chosen to center one of the frequency channels at dc. Fourth, we assume that W is chosen to "window" just that centered frequency channel. By construction, $\tilde{h}(x/\delta t)$ is chosen to be nonzero only for $|x| \lesssim \pi/2$ [36]. Under these conditions, the only nonzero term in the integrand is for $k = 0$ (see Fig. 10). Fifth and finally, we assume that $\tilde{h}$ varies sufficiently slowly over the width of $\tilde{W}$ such that we can replace $\tilde{h}(\omega + \omega_{sb})$ with $\tilde{h}(\omega_{sb,c^*})$. With these assumptions in place, we have

$$\begin{aligned} Z &= \frac{AD}{2N\delta t}\tilde{h}(\omega_{sb,c^*})\int \frac{d\omega}{2\pi}\tilde{\tilde{V}}^{c*}_{\mu w}(\omega)\tilde{W}(-\omega) \\ &= \frac{AD}{2N\delta t}\tilde{h}(\omega_{sb,c^*})\int dt\, \bar{V}_{\mu w}(t)W(t). \end{aligned} \tag{A20}$$

To make a connection with the main text, note that the power flowing into the receiver in terms of the voltage is $\bar{V}^2_{\mu w}/2Z_0$, while in terms of the amplitude wave $B(t)$ from the main text, the power is $\hbar\omega\,|B|^2$. Therefore, $\bar{V}_{\mu w}(t) = \sqrt{2Z_0\hbar\omega}B(t)$ and

$$\begin{aligned} \left|Z_{|0\rangle} - Z_{|1\rangle}\right| &= \frac{AD}{2N\delta t}\sqrt{2Z_0\hbar\omega} \\ &\times \left|\int dt\, \left(B_{|0\rangle}(t) - B_{|1\rangle}(t)\right) W(t)\right|. \end{aligned} \tag{A21}$$

Evidently, the constant C from the main text is $C = (AD/2N\delta t)\sqrt{2Z_0\hbar\omega}$.

2. Noise

To complete our formula for the SNR, we need to find the width of the cloud of points in the IQ plane. Our approach is to write out the IF waveforms in the case where incident on the receiver is noise with a known spectral density. We then explicitly compute the average of the square modulus of the demodulated IQ points. The purpose of this approach is to highlight the assumptions and associated approximations.

Suppose that, along with our microwave voltage signal, there is voltage noise denoted V_ξ. The IF noise waveforms after the filter are

$$I(t) = A\int dt'\, V_\xi(t')\cos(\Omega t')h(t-t'), \tag{A22}$$

$$Q(t) = -A\int dt'\, V_\xi(t')\sin(\Omega t')h(t-t'), \tag{A23}$$

and the digital complex signal is

$$Z_n = AD\int dt'\, V_\xi(t')e^{-i\Omega t'}h(n\delta t - t'). \tag{A24}$$

The demodulated complex point Z is a weighted sum of random variables

$$\begin{aligned} Z &= \frac{1}{N}\sum_{n=0}^{N-1} Z_n e^{-i\omega_{sb}n\delta t}W^c_n \\ &= \frac{AD}{N}\sum_{n=0}^{N-1}\int dt' V_\xi(t')e^{-i\Omega t'}e^{-i\omega_{sb}n\delta t}h(n\delta t - t')W^c_n \\ &= \frac{AD}{N}\sum_{n=0}^{N-1}\int dt' V_\xi(n\delta t - t')e^{-i\Omega(n\delta t - t')}e^{-i\omega_{sb}n\delta t}h(t')W^c_n. \end{aligned} \tag{A25}$$

The mod square of Z is

$$\begin{aligned} |Z|^2 = \left(\frac{AD}{N}\right)^2 &\sum_{n,m=0}^{N-1}\int dt'\, dt'' \\ &\times V_\xi(n\delta t - t')V_\xi(m\delta t - t'')h(t')h(t'') \\ &\times e^{i\Omega(t'-t'')}e^{-i(\Omega+\omega_{sb})(n-m)\delta t}W^c_n W^{c\,*}_m. \end{aligned} \tag{A26}$$

Taking the statistical average, we now have

$$\begin{aligned} \left\langle|Z|^2\right\rangle = \left(\frac{AD}{N}\right)^2 &\sum_{n,m=0}^{N-1}\int dt'\, dt'' \\ &\times \left\langle V_\xi(n\delta t - t')V_\xi(m\delta t - t'')\right\rangle h(t')h(t'') \\ &\times e^{i\Omega(t'-t'')}e^{-i(\Omega+\omega_{sb})(n-m)\delta t}W^c_n W^{c\,*}_m. \end{aligned} \tag{A27}$$

The Wiener-Khinchin theorem relates the two-point average noise to the single-sided noise spectral density

$$\left\langle V_\xi(t)V_\xi(0)\right\rangle = \frac{1}{2}\int\frac{d\omega}{2\pi}S_{V_\xi}(\omega)e^{i\omega t}. \tag{A28}$$

Plugging this two-point average into our formula for $\langle|Z|^2\rangle$, we have

$$\langle|Z|^2\rangle = \frac{1}{2}\left(\frac{AD}{N}\right)^2 \sum_{n,m=0}^{N-1} \int dt'\, dt''\, \frac{d\omega}{2\pi} \times S_{V_\xi}(\omega) e^{i\omega((n-m)\delta t-(t'-t''))} h(t') h(t'') \times e^{i\Omega(t'-t'')} e^{-i(\Omega+\omega_{\rm sb})(n-m)\delta t} W_n^c W_m^{c\,*}. \quad \text{(A29)}$$

The integrals over t and t' can be evaluated as Fourier transforms of the filter h, leaving

$$\langle|Z|^2\rangle = \frac{1}{2}\left(\frac{AD}{N}\right)^2 \sum_{n,m=0}^{N-1} \int \frac{d\omega}{2\pi} \times S_{V_\xi}(\omega) e^{i\omega(n-m)\delta t} \left|\tilde{h}(\omega-\Omega)\right|^2 \times e^{-i(\Omega+\omega_{\rm sb})(n-m)\delta t} W_n^c W_m^{c\,*}. \quad \text{(A30)}$$

Now we push the sum inside and define $x = \delta t(\omega - \Omega - \omega_{\rm sb})$, giving

$$\langle|Z|^2\rangle = \frac{(AD)^2}{2N^2\delta t} \int \frac{dx}{2\pi} S_{V_\xi}(\Omega+\omega_{\rm sb}+x/\delta t) \times \left|\tilde{h}(x/\delta t+\omega_{\rm sb})\right|^2 \left|\sum_{n=0}^{N-1} e^{inx} W_n\right|^2. \quad \text{(A31)}$$

As we did for the signal, we assume that W_n can be extrapolated by a continuous function, express it in the frequency domain, and approximate the sum (the Dirichlet kernel) as a comb of delta functions. With the same assumptions we made for the signal, we find

$$\langle|Z|^2\rangle = \frac{(AD)^2}{2N^2\delta t^3} \int \frac{dx}{2\pi} S_{V_\xi}(\Omega+\omega_{\rm sb}+x/\delta t) \times \left|\tilde{h}(x/\delta t+\omega_{\rm sb})\right|^2 \left|\tilde{W}(-x/\delta t)\right|^2. \quad \text{(A32)}$$

Here, $\tilde{W}(x)$ limits the integration over x to a narrow range, so we approximate $\tilde{h}$ as a constant over that range, and similarly for S_{V_ξ}, leaving

$$\langle|Z|^2\rangle = S_{V_\xi}(\Omega+\omega_{\rm sb}) \left|\tilde{h}(\omega_{\rm sb})\right|^2 \frac{(AD)^2}{2N^2\delta t^3} \times \int \frac{dx}{2\pi} \left|\tilde{W}(-x/\delta t)\right|^2 = S_{V_\xi}(\Omega+\omega_{\rm sb}) \left|\tilde{h}(\omega_{\rm sb})\right|^2 \frac{(AD)^2}{2(N\delta t)^2} \times \int dt\, |W(t)|^2. \quad \text{(A33)}$$

The quantum limit of voltage noise spectral density at frequency ω is $S_{V_\xi} = Z_0\hbar\omega/2$. Note also that while we computed $\langle|Z|^2\rangle$, the width σ in Eq. (4) is the width of a one-dimensional projection (marginal distribution) of the two-dimensional distribution of Z. If Z is Gaussian, then $\sigma^2 = \langle|Z|^2\rangle/2$. Collecting these facts, we have

$$2\sigma^2 = \frac{1}{4}\frac{(AD)^2 Z_0\hbar\Omega}{(N\delta t)^2} \left|\tilde{h}(\omega_{\rm sb})\right|^2 \int |W(t)|^2\, dt. \quad \text{(A34)}$$

3. Signal-to-noise ratio

Combining Eqs. (A21) and (A34), the signal-to-noise ratio is

$$\text{SNR} = \frac{\left|Z_{|0\rangle} - Z_{|1\rangle}\right|^2}{2\sigma^2} = \kappa \frac{\left|\int dt\, \left(\alpha_{|0\rangle}(t) - \alpha_{|1\rangle}(t)\right) W(t)\right|^2}{(1/2)\int dt\, |W(t)|^2}, \quad \text{(A35)}$$

where we used $B(t) = \sqrt{\kappa}\alpha(t)$. This result agrees with Eq. (4) up to the factor of η, which can be inserted by hand to represent either signal loss or excess noise above the quantum limit.

[1] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, *et al.*, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).

[2] Google Quantum AI and collaborators, Suppressing quantum errors by scaling a surface code logical qubit, Nature **614**, 676 (2023).

[3] S. Krinner, N. Lacroix, A. Remm, A. Di Paolo, E. Genois, C. Leroux, C. Hellings, S. Lazar, F. Swiadek, J. Herrmann, *et al.*, Realizing repeated quantum error correction in a distance-three surface code, Nature **605**, 669 (2022).

[4] 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).

[5] 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).

[6] 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).

[7] E. A. Sete, J. M. Martinis, and A. N. Korotkov, Quantum theory of a bandpass Purcell filter for qubit readout, Phys. Rev. A **92**, 012325 (2015).

[8] 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).

[9] P. Bertet, I. Chiorescu, G. Burkard, K. Semba, C. J. P. M. Harmans, D. P. DiVincenzo, and J. E. Mooij, Dephasing of a superconducting qubit induced by photon noise, Phys. Rev. Lett. **95**, 257002 (2005).

[10] A. P. Sears, A. Petrenko, G. Catelani, L. Sun, H. Paik, G. Kirchmair, L. Frunzio, L. I. Glazman, S. M. Girvin, and R. J. Schoelkopf, Photon shot noise dephasing in the strong-dispersive limit of circuit QED, Phys. Rev. B **86**, 180504 (2012).

[11] C. Rigetti, J. M. Gambetta, S. Poletto, B. L. T. Plourde, J. M. Chow, A. D. Córcoles, J. A. Smolin, S. T. Merkel, J. R. Rozen, G. A. Keefe, M. B. Rothwell, M. B. Ketchen, and M. Steffen, Superconducting qubit in a waveguide cavity with a coherence time approaching 0.1 ms, Phys. Rev. B **86**, 100506 (2012).

[12] F. Yan, S. Gustavsson, A. Kamal, J. Birenbaum, A. P. Sears, D. Hover, T. J. Gudmundsen, D. Rosenberg, G. Samach, S. Weber, *et al.*, The flux qubit revisited to enhance coherence and reproducibility, Nat. Commun. **7**, 12964 (2016).

[13] F. Yan, D. Campbell, P. Krantz, M. Kjaergaard, D. Kim, J. L. Yoder, D. Hover, A. Sears, A. J. Kerman, T. P. Orlando, S. Gustavsson, and W. D. Oliver, Distinguishing coherent and thermal photon noise in a circuit quantum electrodynamical system, Phys. Rev. Lett. **120**, 260504 (2018).

[14] Google Quantum AI and collaborators, Quantum error correction below the surface code threshold, Nature (2024).

[15] This assumption is valid because Heisenberg's equation of motion for the a operator is identical to the classical equation of motion for the field α, so long as the drive is linearly coupled (i.e., inductor-capacitor coupling).

[16] C. A. Ryan, B. R. Johnson, J. M. Gambetta, J. M. Chow, M. P. da Silva, O. E. Dial, and T. A. Ohki, Tomography via correlation of noisy measurement records, Phys. Rev. A **91**, 022118 (2015).

[17] D. Sank, *Fast, Accurate State Measurement in Superconducting Qubits*, Ph.D. thesis, University of California Santa Barbara, 2014. https://escholarship.org/uc/item/73m3g1h0.

[18] A. N. Korotkov, Selective quantum evolution of a qubit state due to continuous measurement, Phys. Rev. B **63**, 115403 (2001).

[19] A. A. Clerk, S. M. Girvin, and A. D. Stone, Quantum-limited measurement and information in mesoscopic detectors, Phys. Rev. B **67**, 165324 (2003).

[20] A. A. Clerk, M. H. Devoret, S. M. Girvin, F. Marquardt, and R. J. Schoelkopf, Introduction to quantum noise, measurement, and amplification, Rev. Mod. Phys. **82**, 1155 (2010).

[21] This linear approximation for the qubit ac Stark shift is valid when $\bar{n} \lesssim n_{\text{crit}} \equiv (1/4)(\Delta/g)^2$, where Δ is the detuning between the qubit and the resonator, and g is the qubit-resonator coupling strength.

[22] This assumption is substantiated by the observation that the SNR scales linearly with power, as can be seen in Sec. V.

[23] N. Zobrist, N. Klimovich, B. Eom, G. Coiffard, M. Daal, N. J. Swimmer, S. Steiger, B. Bumble, H. G. LeDuc, P. K. Day, and B. A. Mazin, Improving the dynamic range of single photon counting kinetic inductance detectors, J. Astron. Telesc. Instrum. Syst. **7**, 010501 (2021).

[24] M. Khalil, M. Stoutimore, F. Wellstood, and K. Osborn, An analysis method for asymmetric resonator transmission applied to superconducting devices, J. Appl. Phys. **111**, 0 (2012).

[25] H. Wang, S. Singh, C. R. H. McRae, J. C. Bardin, S.-X. Lin, N. Messaoudi, A. R. Castelli, Y. J. Rosen, E. T. Holland, D. P. Pappas, and J. Y. Mutus, Cryogenic single-port calibration for superconducting microwave resonator measurements, Quantum Sci. Technol. **6**, 035015 (2021).

[26] F. W. Carter, T. S. Khaire, V. Novosad, and C. L. Chang, scraps: An open-source python-based analysis package for analyzing and plotting superconducting resonator data, IEEE Trans. Appl. Supercond. **27**, 1 (2017).

[27] A. Megrant, C. Neill, R. Barends, B. Chiaro, Y. Chen, L. Feigl, J. Kelly, E. Lucero, M. Mariantoni, P. J. O'Malley, *et al.*, Planar superconducting resonators with internal quality factors above one million, Appl. Phys. Lett. **100**, 113510 (2012).

[28] Z. Chen, *Metrology of Quantum Control and Measurement in Superconducting Qubits*, Ph.D. thesis, University of California Santa Barbara, 2018. https://escholarship.org/uc/item/0g29b4p0.

[29] D. Sank, Z. Chen, M. Khezri, J. Kelly, R. Barends, B. Campbell, Y. Chen, B. Chiaro, A. Dunsworth, A. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Mutus, M. Neeley, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, A. Vainsencher, T. White, J. Wenner, A. N. Korotkov, and J. M. Martinis, Measurement-induced state transitions in a superconducting qubit: Beyond the rotating wave approximation, Phys. Rev. Lett. **117**, 190503 (2016).

[30] 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).

[31] M. F. Dumas, B. Groleau-Paré, A. McDonald, M. H. Muñoz Arias, C. Lledó, B. D'Anjou, and A. Blais, Measurement-induced transmon ionization, Phys. Rev. X **14**, 041023 (2024).

[32] A. Yen, Y. Ye, K. Peng, J. Wang, G. Cunningham, M. Gingras, B. M. Niedzielski, H. Stickler, K. Serniak, M. E. Schwartz, and K. P. O'Brien, Directional emission of a readout resonator for qubit measurement, Phys. Rev. Appl. **22**, 034035 (2024).

[33] 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).

[34] A. Eddins, J. M. Kreikebaum, D. M. Toyli, E. M. Levenson-Falk, A. Dove, W. P. Livingston, B. A. Levitan, L. C. G. Govia, A. A. Clerk, and I. Siddiqi, High-efficiency measurement of an artificial atom embedded in a parametric amplifier, Phys. Rev. X **9**, 011004 (2019).

[35] T. White, *et al.*, Readout of a quantum processor with high dynamic range Josephson parametric amplifiers, Appl. Phys. Lett. **122**, 014001 (2023).

[36] The Nyquist antialiasing criterion demands $|x| < \pi$, corresponding to $|\omega| < (2\pi)$ 500 MHz for our 1-Gs/s ADC, but our receiver is intentionally designed with a more conservative filter with a cutoff near 200 MHz.