MIT Open Access Articles

Spectroscopy of low-frequency noise and its temperature dependence in a superconducting qubit

Massachusetts Institute of Technology

ʕ

Spectroscopy of low-frequency noise and its temperature dependence in a superconducting qubit

Fei Yan,[1,*] Jonas Bylander,[2] Simon Gustavsson,[2] Fumiki Yoshihara,[3] Khalil Harrabi,[3,†] David G. Cory,[4,5] Terry P. Orlando,[2,6] Yasunobu Nakamura,[3,7,‡] Jaw-Shen Tsai,[3,7] and William D. Oliver[2,8]

[1]*Department of Nuclear Science and Engineering, Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts 02139, USA*
[2]*Research Laboratory of Electronics, MIT, Cambridge, Massachusetts 02139, USA*
[3]*Institute of Physical and Chemical Research (RIKEN), Wako, Saitama 351-0198, Japan*
[4]*Institute for Quantum Computing and Department of Chemistry, University of Waterloo, Ontario, Canada N2L 3G1*
[5]*Perimeter Institute for Theoretical Physics, Waterloo, Ontario, Canada N2J 2W9*
[6]*Department of Electrical Engineering and Computer Science, MIT, Cambridge, Massachusetts 02139, USA*
[7]*Green Innovation Research Laboratories, NEC Corporation, Tsukuba, Ibaraki 305-8501, Japan*
[8]*MIT Lincoln Laboratory, 244 Wood Street, Lexington, Massachusetts 02420, USA*

We report a direct measurement of the low-frequency noise spectrum in a superconducting flux qubit. Our method uses the noise sensitivity of a free-induction Ramsey interference experiment, comprising free evolution in the presence of noise for a fixed period of time followed by single-shot qubit-state measurement. Repeating this procedure enables Fourier-transform noise spectroscopy with access to frequencies up to the achievable repetition rate, a regime relevant to dephasing in ensemble-averaged time-domain measurements such as Ramsey interferometry. Rotating the qubit's quantization axis allows us to measure two types of noise: effective flux noise and effective critical-current or charge noise. For both noise sources, we observe that the very same $1/f$-type power laws measured at considerably higher frequencies ($0.2-20$ MHz) are consistent with the noise in the $0.01-100$-Hz range measured here. We find no evidence of temperature dependence of the noises over $65-200$ mK, and also no evidence of time-domain correlations between the two noises. These methods and results are pertinent to the dephasing of all superconducting qubits.

I. INTRODUCTION

A major remaining obstacle to implementing fault-tolerant quantum computation with superconducting qubits is the insufficient coherence time T_2 compared to the gate-operation time. In the Bloch-Redfield picture of two-level system dynamics, there are two mechanisms that limit T_2: energy relaxation T_1 due to noise at the transition frequency ν_{01}, and dephasing T_φ due to low-frequency fluctuations of ν_{01}. In the qubit's energy eigenbasis, these contributions are due to the transverse and longitudinal components of the fluctuations, respectively. In cases when both relaxation and dephasing exhibit exponential decay laws, their inverse times add to a decoherence rate $T_2^{-1} = (2T_1)^{-1} + T_\varphi^{-1}$. While T_1 can now exceed 10 μs in several superconducting qubit modalities,[1–3] further improvements are required to comfortably exceed even the most lenient error correction thresholds. In general, energy relaxation is irreversible, such that further control-based improvements require resource-intensive multiqubit quantum error-correction protocols. Dephasing, on the other hand, can be refocused by dynamical-decoupling techniques,[2] with only a modest amount of resource overhead. The ultimate goal is to mitigate and eliminate the noise leading to both types of decoherence. To this end, a more detailed understanding of the noise processes, such as magnetic flux, critical current, and charge fluctuations, would expedite materials science, device engineering, and the development of coherent-control methods.

Effective surface spins have recently been identified as one dominant source of low-frequency magnetic-flux noise,[4,5] detrimental to several types of superconducting qubits; however, open questions remain regarding the nature of these spins. Their noise is known to be due to local fluctuators[6–9] and the spectrum exhibits a $1/f^\alpha$ power-law dependence from hertz to tens of megahertz.[10–14] Its dependence on the device geometry[15,16] merits further study.

Similarly, for the flux qubit,[17,18] the noise in the tunnel coupling Δ between the persistent-current states shows a $1/f$-type spectrum from hertz to hundreds of kilohertz.[2] This noise may originate in the critical-current fluctuations of the Josephson junctions[19–22] and/or fluctuating offset charges due to, *e.g.*, charge traps located in the oxides of the junction, metal-insulator interfaces, or surfaces. Charge noise can lead to dephasing even in the flux qubit, even though the junctions have a relatively high ratio of Josephson-tunneling to Coulomb-charging energies ($E_J/E_C \approx 50$ in our device).

In this paper, we introduce a measurement technique for low-frequency noise. A distinguishing feature of our technique is that it enables the measurement of noise spectra up to frequencies limited only by the achievable measurement repetition rate. This is important because noise measured in this manner resides (at least in part) within the relevant measurement bandwidth of time-domain experiments, *e.g.*, Ramsey interferometry, that use the standard ensemble averaging (that is, the averaging of multiple trials acquired at the same repetition rate) to estimate the qubit-state occupation probability.

We report on a direct characterization of the $1/f$-noise power spectral densities (PSD) $S(f)$ in an aluminum superconducting flux qubit. (The Al/AlO$_x$/Al device was made by shadow evaporation at NEC; the experiments were performed at MIT.) We distinguish between the two noises $\delta\varepsilon$, which is effective flux noise, and $\delta\Delta$, which can be parametrized

as effective critical-current noise or effective charge noise. Interestingly, we find that the same $1/f^\alpha$ power laws, measured at much higher frequencies,[2] extend down to the 10^{-2}–10^{+2}-Hz range nearly unchanged. Over the temperature range 65–200 mK, both noises $\delta\varepsilon$ and $\delta\Delta$ are independent of temperature, and any $\delta\varepsilon$–$\delta\Delta$ noise correlations are very small or nonexistent.

II. EXPERIMENTAL METHODS, ANALYSIS, AND RESULTS

The flux qubit's two-level Hamiltonian is $\hat{\mathcal{H}} = -(h/2)(\varepsilon\hat{\sigma}_x + \Delta\hat{\sigma}_z)$. Here, $\varepsilon = 2I_p\Phi_b/h$ is the energy detuning between the diabatic states of classical circulating current $I_p = 0.18\ \mu$A, and ε is adjusted by the external magnetic flux Φ via $\Phi_b = \Phi - \Phi_0/2$ ($\Phi_0 = h/2e$ is the superconducting flux quantum); see the schematic in Figs. 1(a) and 1(b). $\Delta = 5.4$ GHz is the tunnel coupling that hybridizes the persistent-current states, and is established during fabrication by the E_J/E_C ratio. We write each of the parameters $\lambda = \varepsilon, \Delta$ as the sum of its nominal value and a time-dependent fluctuation $\lambda(t) = \lambda^{(0)} + \delta\lambda(t)$. We distinguish between the effects of $\delta\varepsilon$ and $\delta\Delta$ fluctuations by rotating the qubit's quantization axis (eigenbasis), thereby altering the sensitivity of the energy-level splitting $\nu_{01} = \sqrt{\varepsilon^2 + \Delta^2}$ to fluctuations $D_\lambda = \partial\nu_{01}/\partial\lambda = \lambda/\nu_{01}$ [see Fig. 1(b)]. The dominant contributor to longitudinal fluctuations in the qubit's energy eigenbasis is $\delta\Delta$ noise at $\varepsilon = 0$ (as the second-order contribution from $\delta\varepsilon$ noise is negligible[2]), and $\delta\varepsilon$ (flux) noise for $|\varepsilon| \gtrsim 0.1$ GHz.

We use a hysteretic SQUID to read out the qubit's state. During qubit manipulation, we conveniently set the SQUID current I_b to zero. (This is close to its optimal value of $I_b^* = 50$ nA in this device, at which δI_b fluctuations are decoupled from the qubit.[23])

To directly probe the fluctuations, we repeatedly let the qubit undergo Ramsey free induction [see Fig. 1(c)]. Instead of scanning the pulse separation τ, we fix it at a value τ_0, chosen to maximize the qubit state's sensitivity to noise. We also fix the nominal detuning $\nu^{(0)}$ of the applied microwave frequency from the qubit's frequency to $\nu^{(0)} \equiv \nu_{01} - \nu_{\mu w} = 1/4\tau_0$ (a free-evolution $\pi/2$ rotation in the X–Y plane) [see Figs. 1(d) and 1(f)]. The $\delta\lambda(t)$ fluctuations translate into frequency fluctuations $\nu(t) = \nu^{(0)} + \delta\nu(t)$, which in turn translate into fluctuations of the SQUID's switching probability $p_{sw}(\nu, \tau_0) = p_0 - a_0(T)\cos[2\pi\nu(t)\tau_0]$, where $a_0(T)$ is the temperature-dependent read-out visibility ($2a_0 = 79\%$ at the refrigerator's $T = 12$-mK base temperature) and p_0 is the switching probability for the qubit's 50% superposition state. We linearize about a working point at $p_{sw} = p_0$ and obtain $\partial p_{sw}/\partial\nu = 2\pi a_0(T)\tau_0$. Due to this transfer function, a correction factor arises in the calculation of ensemble-averaged quantities, assuming Gaussian statistics (see Appendix Sec. B 2). This factor is $a(\tau_0, T)/a_0(T)$, where $a(\tau_0, T)$ is the amplitude of the fringe at pulse separation τ_0. The conversion factor from the noise $\delta\lambda$ to the switching probability p_{sw} then becomes $\eta_\lambda(\tau_0, T) = 2\pi a(\tau_0, T)\tau_0 D_\lambda$.

The single-shot read-out of the qubit, with a repetition time Δt, results in a binary time series $\{z_n\}$; see Fig. 1(g). (In this experiment, $\Delta t = 2$ ms to allow for the read-out induced
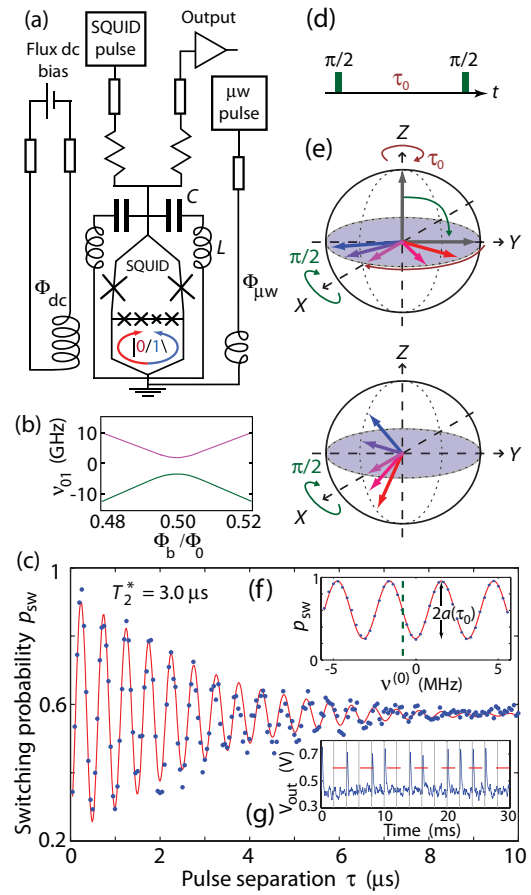

FIG. 1. (Color online) (a) Schematic of the qubit device with biasing and read-out circuitry. (b) Qubit energy-level diagram near $\Phi_b = \Phi_0/2$ ($\varepsilon = 0$). (c) Ramsey interference fringe measured at $\varepsilon = 0$ with 5000 averages per point. (d) Double-$\pi/2$ pulse (Ramsey) sequence. (e) Bloch sphere representation of dephasing and our measurement scheme: the initial $\pi/2$ pulse takes the Bloch vector from the ground state (along Z) to the equator (along Y); during the free-induction time $\tau = \tau_0$, the nominal detuning is $\nu^{(0)} = 1/4\tau_0$, but differs slightly at each trial (quasistatic noise), and the Bloch vectors therefore acquire different phases; the final $\pi/2$ pulse rotates the Bloch vectors out of the equatorial plane so that they can be distinguished by the SQUID read-out. (f) Calibration by scanning $\nu_{\mu w}$ with $\tau_0 = 0.3\ \mu$s, $\varepsilon = 0$. The red solid line is a sinusoidal fit used to determine the working point for maximal sensitivity (dashed line at $\nu^{(0)} = -0.8$ MHz). (g) Individual SQUID-switching events measured on an oscilloscope. The dashed horizontal line is a software threshold detector.

quasiparticles to relax between trials so that they contribute negligibly to heating.) Each element z_n represents the result of a Bernoulli trial with expectation value p_{sw}. The standard method to determine the noise PSD (Refs. 13, 15, and 24) is to ensemble average the switching events acquired during a gate time $t_{acq} = N_G\Delta t$, with typically $N_G \approx 1000$, to determine the average switching probability p_{sw} (binomially distributed), and then take the Fourier transform of the time series of switching probabilities p_{sw}; see Fig. 2(a). This approach estimates the noise spectrum at frequencies up to $1/2t_{acq}$. In contrast, in our approach, we calculate the bilateral PSD from the
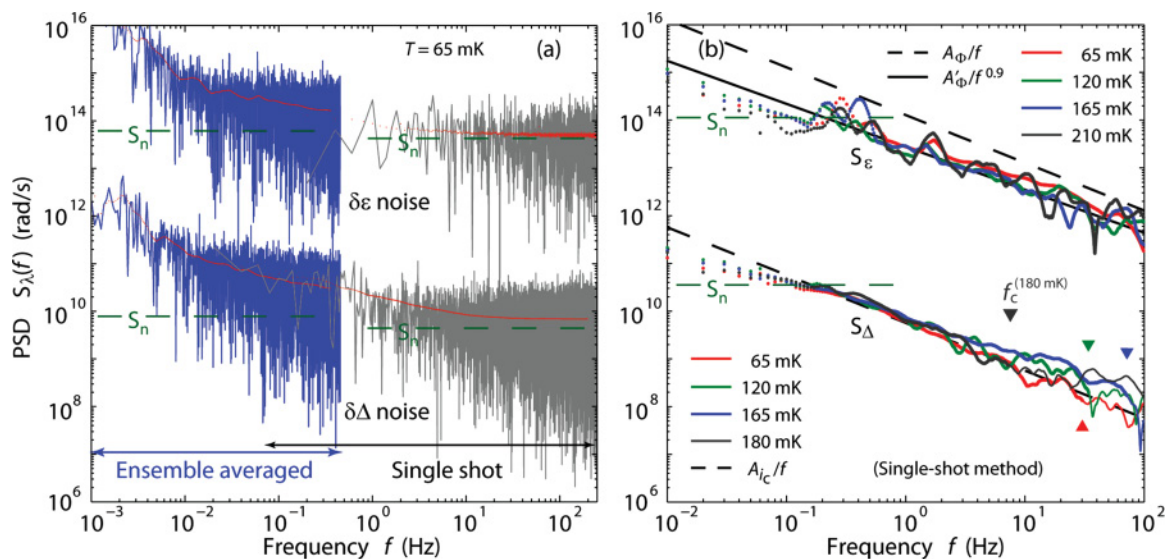
FIG. 2. (Color online) Bilateral noise PSDs. (a) Standard-method PSD of the ensemble-averaged time series of switching probabilities p_{sw} (blue data, on left) and of the binary time series of single-shot measurements $\{z_n\}$ (gray, on right), measured at base temperature. The red points (thin line) are averages of hundreds of traces. (S_ε measured at $\varepsilon = 450$ MHz.) (b) Cross PSDs of interleaved time series $S_\lambda(f)$ [Eq. (2)] vs device temperature (S_ε measured at $\varepsilon = 640$ MHz). The data were smoothed by a sliding average with a triangular weight function of width $\Delta f = f/4$. The dashed diagonal lines are the $1/f$ noises $S_\lambda(f) = (2\pi)^2 \kappa_\lambda^2 A_{\Phi,i_c}/|f|$, derived in Ref. 2; the sensitivities are $\kappa_{\Delta,i_c} \equiv \partial\Delta/\partial i_c = 3.2$ GHz (i_c is normalized critical current) and $\kappa_\varepsilon \equiv \partial\varepsilon/\partial\Phi = 1.1$ GHz/mΦ_0, and the noise strengths are $A_{i_c} = (4.0 \times 10^{-6})^2$ and $A_\Phi = (1.7\,\mu\Phi_0)^2$. The solid diagonal line is $S_\varepsilon(f) = (2\pi)^2 \kappa_\varepsilon^2 A'_\Phi/(|f/1\ \text{Hz}|^{0.9} \times 1\ \text{Hz})$, with $A'_\Phi = (0.8\,\mu\Phi_0)^2$. See Ref. 2 and its supplementary information for details. The horizontal green dashed lines are the sampling-noise levels $S_n(\tau_0, T)$ at low temperature; the triangles represent an upper cutoff frequency f_c for sufficient averaging, above which the data are not dependable.

recorded series of single-shot measurements (Bernoulli trials) $S(f_k) = |Z_k|^2/(N/\Delta t)$, where $f_k = k/N\Delta t, k = 0, \ldots, N/2$, and $\{Z_k\}$ is the discrete Fourier transform of $\{z_n\}$, with n typically ranging from 1 to $N = 5 \times 10^4$. [Contrary to Refs. 2, 11, and 12, we use the type-1 Fourier transform $S_x(f) = \int_{-\infty}^{\infty} dt\,\langle x(0)x(t)\rangle \exp(-i2\pi f t)$.] This method increases the upper cutoff frequency from $1/2N_G\Delta t$ to $1/2\Delta t$, which may approach $1/\tau_0$ and is limited only by the achievable repetition rate; see Figs. 2(a) and 2(b), Appendix A, and Fig. 5.

Both the PSDs originating from single-shot and from ensemble-averaged measurements in Fig. 2(a) exhibit statistical-sampling noise $S_n(\tau_0, T) = (2\pi)^2 \sigma_s^2 \Delta t/\eta_\lambda^2(\tau_0, T)$, where the variance is $\sigma_s^2 = p_{sw}(1 - p_{sw})$. We can eliminate this white background noise by calculating the bilateral cross PSD of the two interleaved (single-shot) time series $z'_n = z_{2n-1}$ and $z''_n = z_{2n}$:

$$S^{\text{cross}}(f_k) = \frac{Z'_k (Z''_k)^*}{N/\Delta t}, \quad f_k = \frac{k}{N\Delta t}, \quad (1)$$

where now $k = 0, \ldots, N/4$. Dividing Eq. (1) by the conversion factor, we obtain the spectral density of the fluctuation λ:

$$S_\lambda(f_k) = (2\pi)^2 S^{\text{cross}}(f_k)/\eta_\lambda^2(\tau_0, T). \quad (2)$$

We typically average the spectra of 500 time series to improve statistical accuracy without compromising bandwidth, and recalibrate the working point periodically (hourly). Note that both the $1/f$ noise and the sampling noise dominate all other background noise at the temperatures considered.

The $\delta\varepsilon$ and $\delta\Delta$ noise PSDs are plotted in Fig. 2(b) for several temperatures. There is striking agreement with the $1/f^\alpha$ power

laws inferred in Ref. 2, measured at considerably higher frequencies (0.02–20 MHz). Noise that is strictly $1/f^{\alpha=1}$ over the frequency range relevant to free induction (here $10^{-1} \sim 10^6$ Hz) gives a Gaussian decay function of the temporal Ramsey oscillations. Assuming that our $\delta\Delta$ noise satisfies these criteria, we use the approach of Ref. 25 to calculate the inhomogeneously broadened decay-time constant T_φ^*. With noise sensitivity κ_{Δ,i_c} and strength A_{i_c} as defined in the legend of Fig. 2, $T_1 = 12\ \mu$s, and typical parameters $\tau_0 = 1\ \mu$s and $t_{acq} = 10$ s, we obtain $T_\varphi^* = (2\pi\,\kappa_{\Delta,i_c} D_\Delta)^{-1} A_{i_c}^{-1/2} (\ln[t_{acq}/2\,\tau_0])^{-1/2} = 3.2\ \mu$s, in very good agreement with the observed T_φ^* in Fig. 1(c).

We now turn to possible $\delta\varepsilon - \delta\Delta$ noise correlations in the time domain to check that the two spectra in Fig. 2 are not due to one and the same mechanism. Figure 3 shows how we repeatedly measured the switching probability at alternating flux biases $p_{sw}(\pm\varepsilon^{(0)} + \delta\varepsilon, \Delta^{(0)} + \delta\Delta)$, with $\pm\varepsilon^{(0)}$ chosen such that the effects of the two noises on $\delta\nu$ were similar in magnitude, i.e., $D_\varepsilon(\varepsilon^{(0)})|\delta\varepsilon| \approx D_\Delta(\varepsilon^{(0)})|\delta\Delta|$. We set the pulse separation τ_0 and nominal frequency detuning $\nu^{(0)}$. With the energy-level splitting $\nu(\varepsilon, \Delta) = \nu^{(0)}(\varepsilon^{(0)}, \Delta^{(0)}) + \delta\nu(\varepsilon^{(0)}, \delta\varepsilon, \delta\Delta)$, we use the decay function of the Ramsey fringe to infer the noise correlations from the measured p_{sw}:

$$p_{sw}(\varepsilon, \Delta) = p_0 - a_0 \exp(-\tau_0/2T_1)$$
$$\times \exp(-[\tau_0/T_\varphi^*(\varepsilon)]^2) \cos(2\pi\,\nu(\varepsilon, \Delta)\,\tau_0), \quad (3)$$

where $[1/T_\varphi^*(\varepsilon)]^2 = [1/T_\varphi^*(0)]^2 + K A_\Phi D_\varepsilon^2$ and K is a constant that we have determined independently, along with the other parameters in the equation. At the bias points $\varepsilon = \pm\varepsilon^{(0)}$, $\delta\varepsilon$ fluctuations induce negatively correlated $\delta\nu$ fluctuations
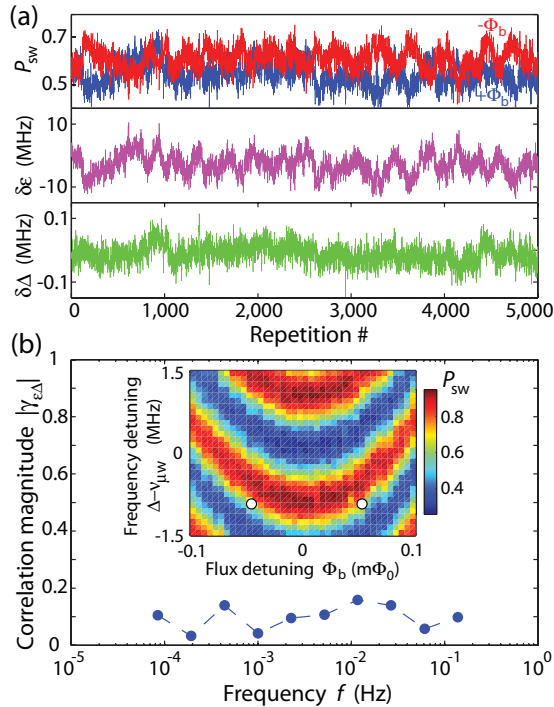
(a)

(b)

FIG. 3. (Color online) $\delta\varepsilon-\delta\Delta$ noise correlations measured at base temperature. (a) Top panel: Time series of p_{sw} with fixed $\tau_0 = 0.3\ \mu\text{s}$, nominal $\nu^{(0)} = -0.8$ MHz, and 1000 averages per measured point. The flux-bias polarity ($\pm\varepsilon^{(0)}$) was alternated between positive (blue) and negative (red) [$\pm\Phi_b = \pm0.058\ \text{m}\Phi_0$, see inset in (b)], with a 2-s repetition period. (a) Middle and bottom panels: Differential ($\delta\varepsilon$) and common-mode ($\delta\Delta$) noise inferred from the data in the top panel. (b) Magnitude of the correlation function [Eq. (4)]; the smaller the magnitude, the more indeterminate the phase, and therefore we cannot discriminate between possible correlation and anticorrelation. Inset: 30-min measurement of Ramsey fringes (frequency detuning referred to $\varepsilon = 0$, where $\nu_{01} = \Delta - \nu_{\mu w}$). White circles indicate the bias points used in the top panel of (a).

(and consequently p_{sw} fluctuations), whereas $\delta\Delta$ fluctuations induce positively correlated $\delta\nu$ fluctuations. At each time step, the measurement of p_{sw} for $\varepsilon = \pm\varepsilon_0$ yields a system of two nonlinear equations in the two unknowns $\delta\varepsilon$ and $\delta\Delta$. We solve this system numerically: Fig. 3(a) shows the raw p_{sw} data and extracted $\delta\varepsilon$ and $\delta\Delta$ versus time. We then calculate the cross PSD $S_{\varepsilon\Delta}(f)$ as the Fourier transform of the cross-correlation function, and obtain an upper bound on the magnitude of the correlation function, as shown in Fig. 3(b):

$$|\gamma_{\varepsilon\Delta}(f)| = \frac{|S_{\varepsilon\Delta}(f)|}{[S_\varepsilon(f)S_\Delta(f)]^{1/2}} < 0.2. \qquad (4)$$

Finally, we measured the temperature dependencies of the two types of noise. Figure 4 shows the integrated noise powers Π_λ versus temperature T in the $65-200$-mK range, where our read-out visibility is sufficient. We observe in essence temperature independence for both noises. For $\delta\varepsilon$ (flux) noise, this is consistent with previous observations in SQUIDs;[5,10] we discuss the $\delta\Delta$ noise in the following.
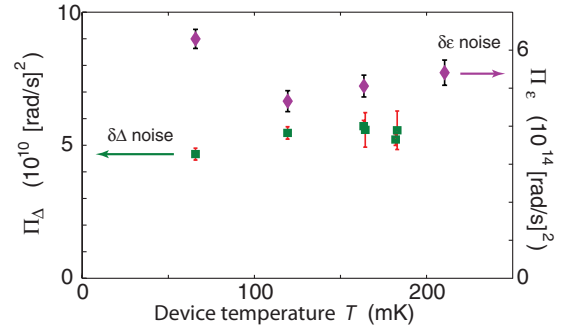

FIG. 4. (Color online) Temperature dependencies of the total noise powers $\Pi_\lambda(T) = (2\pi)^2\kappa_\lambda^2 \int_{F_\lambda} df\, S_\lambda(f; T)$ in the frequency intervals $F_\varepsilon = 0.02-50$ Hz and $F_\Delta = 0.02-2$ Hz [*cf.* Fig. 2(b)]. It is possible to measure the $\delta\varepsilon$ noise up to somewhat higher temperatures and frequencies than the $\delta\Delta$ noise. Note that the Π_λ values depend on the integration limits although the choice of F_λ does not make any significant difference in the trends. (The double data points for the 165- and 180-mK $\delta\Delta$ noise were measured with different pulse spacings τ_0.) We estimated the device temperature from switching-current measurements on the SQUID, which suggest saturation at dilution-refrigerator temperatures below about 65 mK. The error bars are derived only from the fit error of the read-out visibility $a(\tau_0, T)$, included in the p_{sw}-to-$\delta\lambda$ conversion factor $\eta_\lambda(\tau_0, T)$ [see Fig. 1(f)].

III. DISCUSSION

In order to analyze the $\delta\Delta$ noise, we parametrize it as an effective, normalized critical-current noise $i_c = \delta I_c/I_c$, with $I_c = 0.4\ \mu\text{A}$, in a Josephson junction with area $\mathcal{A} = (0.2\ \mu\text{m})^2$. Van Harlingen *et al.*[19] found a "canonical" value for the $1/f$ δI_c-noise power at 1 Hz and 4.2 K: $A_{I_c}^{\text{can}} \approx 144\ (\text{pA})^2(I_c/\mu\text{A})^2/(\mathcal{A}/\mu\text{m}^2)$ in several SQUIDs and qubits of various sizes, made of different materials. The authors hypothesized a quadratic temperature dependence, consistent with certain plausible models for the noise sources below 100 mK, while noting that other models suggest a linear dependence. The bilateral normalized noise PSD then becomes $S_{i_c}^{\text{can}}(f) = A_{I_c}^{\text{can}} I_c^{-2} (T/4.2\ \text{K})^2/|f|$, which, for $T = 65$ mK, is considerably lower (almost 20 times) than our measured value. On the other hand, Eroms *et al.*[22] measured resistance fluctuations in aluminum tunnel junctions: they found about 100 times lower noise power at 4.2 K, a linear temperature dependence, and saturation below 0.8 K, *i.e.*, $S_{i_c}(f) = (1/100) \times A_{I_c}^{\text{can}} I_c^{-2} (T/4.2\ \text{K})/|f|$. With $T = 0.8$ K, this gives a value about 2.5 times lower than what we observe. We also note that recently, contrary to these findings, Paik *et al.*[3] reported no evidence for $1/f$ i_c noise in a Josephson junction.

An alternative source of $\delta\Delta$ noise is the fluctuating offset charges δQ, known to exhibit $1/f$ noise;[26–28] these charges effectively supply a gate voltage to each island. The charge-noise power typically observed in single-electron tunneling (SET) devices is proportional to temperature[29] (although quadratic dependencies have also been observed[27]) and saturates below about 200 mK, due to self-heating of the SET, at a "canonical" value A_Q of about $(1\sim10\ \text{m}e)^2$ at 1 Hz. We estimate our qubit's maximum sensitivity to charge fluctuations $\kappa_{\Delta,Q} \equiv \partial\Delta/\partial Q$ to be in the range $0.1\sim1$ MHz/e. We can then parametrize

the $\delta\Delta$ noise as charge noise and estimate the dephasing time $T_\varphi^* = (2\pi\,\kappa_{\Delta,Q}D_\Delta)^{-1}A_Q^{-1/2}(\ln[t_{\mathrm{acq}}/2\,\tau_0])^{-1/2} \approx 4\sim 400\,\mu$s. The lower end of this range is not far from our observed value. Moreover, the tunneling of charged quasiparticles between the small islands constituting our device may displace offset charges and contribute to dephasing at $\varepsilon = 0$.

In conclusion, our spectroscopy of both $\delta\varepsilon$ noise (flux noise) and $\delta\Delta$ noise (effective critical current or charge noise), facilitated by single-shot measurements and thorough data analysis, shows that the very same $1/f^\alpha$ dependencies, measured at substantially higher frequencies, extend down to millihertz frequencies. This apparently indicates that the same noise mechanisms are active and dominant over some 10 orders of magnitude or more for $\delta\varepsilon$ noise and at least 8 orders of magnitude for $\delta\Delta$ noise. The $\delta\varepsilon$ noise may extend, with roughly constant slope (on a logarithmic scale), up to the qubit's transition frequency at several gigahertz[2]: there, this noise is nearly transverse to the flux qubit's energy eigenbasis, and would therefore also contribute to energy relaxation. The small, if not negligible, $\delta\varepsilon - \delta\Delta$ noise correlations (over $7 \times 10^{-4} - 2 \times 10^{-1}$ Hz) show that the noises are due to distinct underlying mechanisms. Moreover, both noises are temperature independent in the $65-200$-mK range, which suggests that the microscopic mechanisms are dominated by even lower energy scales than that. This is useful information for the development of noise models. It also calls for further studies of the reproducibility of the device properties, and, in particular, of the $\delta\Delta$ noise, as it limits the coherence time in superconducting flux and transmon qubits.

APPENDIX A: SPECTRAL DENSITY AND THE STATISTICAL NOISE FLOOR

Here, we describe how we calculate the noise power spectral density (PSD) from the noisy time series, and eliminate the statistical white-noise floor due to sampling.

1. PSD

The fluctuations of our qubit's transition frequency constitute a zero-mean, wide-sense stationary process $\delta\nu(t)$; at our chosen working point, $\delta\nu(t) = \nu(t) - 1/4\tau_0$. We seek its bilateral noise PSD (in units of rad/s, *i.e.*, we use the angular-frequency correlator)

$$S(f) = (2\pi)^2 \lim_{T \to \infty} \frac{\langle |V(f)|^2 \rangle}{T}. \tag{A1}$$

Our measurements' raw data, however, consists of a binary time series $\{z_n\}$ with elements of expectation value $y_n = \langle z_n \rangle = p_0 + a_0 \sin[2\pi\,\delta\nu(n\Delta t)\,\tau_0]$, where Δt is the time step. The statistical properties of $\{z_n\}$ represent those of the underlying process $\delta\nu(t)$, up to a conversion factor and a correction factor (explained in the next section). We can therefore take the discrete Fourier transform $\{Z_k\} = \mathcal{F}[\{z_n\}]$, identify $Z(f_k) = Z_k \times \Delta t$ for $f_k = k/N\Delta t$, and compute the discrete, bilateral noise PSD over the frequency range from $1/t_{\mathrm{tot}} = 1/N\Delta t = 10$ mHz to $1/2t_{\mathrm{acq}} = 1/2\Delta t = 250$ Hz:

$$S_{k=0} = (2\pi)^2 \frac{1}{2} \frac{Z_0^2(\Delta t)^2}{N\Delta t}, \quad S_{k \neq 0} = (2\pi)^2 \frac{Z_k^2(\Delta t)^2}{N\Delta t}, \quad \text{(A2)}$$

for $\quad k = 0, \ldots, (N-1)/2 \quad$ with $\quad f_k = \dfrac{k}{N\Delta t}$.

We then take the statistical average of M different PSDs obtained from different time series measured in succession,

$$\langle S_k \rangle_{\mathrm{stat}} = \frac{1}{M} \sum_{m=1}^{M} S_k^{(m)}, \tag{A3}$$

and finally smooth the result with a sliding average in the frequency domain.

With this method [Eqs. (A2) and (A3)], each element z_n is the result of a single-shot measurement; the sampling time step t_{acq} is the same as the pulse-sequence repetition time Δt. This sets it apart from the standard approach of first taking the ensemble average of typically $N_{\mathrm{G}} = 1000$ samples in the time domain, before calculating the PSD of the resulting N/N_{G} sampled points. The acquisition time is then $t_{\mathrm{acq}} = N_{\mathrm{G}}\Delta t$, and the upper cutoff frequency becomes only $1/2t_{\mathrm{acq}} \approx 0.25$ Hz.

2. White-noise floor

The PSD of the single-shot time sequence suffers from statistical sampling noise because each time step n constitutes a Bernoulli trial (b_n): the read-out SQUID switches ($b_n = 1$) with probability p and does not switch ($b_n = 0$) with probability $1 - p$. This statistical noise has a white spectrum; it dominates possible white background noise from other sources, and dominates also the $1/f$ noise at high frequencies. To estimate it, we can treat the stochastic variable b_n as independent and identically distributed (i.i.d.) with ensemble-averaged mean $\langle b \rangle = p$ and variance $\sigma_b^2 \equiv \langle (\Delta b)^2 \rangle \equiv \langle b^2 \rangle - \langle b \rangle^2 = p(1-p)$. By sampling at a fixed rate $1/\Delta t$, the white-noise floor of the bilateral PSD becomes

$$S_{\mathrm{n}}(f_k) = (2\pi)^2 \left(\sigma_b^2 + \langle b \rangle^2 \delta_{k,0} \right) \Delta t. \tag{A4}$$

Here, we use Kronecker's delta $\delta_{k,0}$ in the discrete PSD.

The same expression is valid for the PSD of the ensemble-averaged time series, the constituent elements of which have a binomially distributed switching probability (c for "counts") averaged over a gate time $t_{\mathrm{acq}} = N_{\mathrm{G}}\Delta t$: we obtain Eq. (A4) after substituting $\langle c \rangle = p$ and $\sigma_c^2 = p(1-p)/N_{\mathrm{G}}$ for $\langle b \rangle$ and σ_b^2, respectively.

Equation (A4) is, in fact, a modification of Carson's theorem, which is valid for temporally random pulse arrivals. There, one considers a random pulse train $w(t) = \sum_{l=1}^{L} b_l\, g(t - t_l)$, in which $g(t)$ is the pulse envelope, the

stochastic variable b_l is the (continuous) pulse height, and the stochastic variable t_l is the pulse-arrival time. The Fourier transform of $w(t)$ is $W(f) = G(f) \sum_{l=1}^{L} b_l \exp(-i2\pi f t_l)$, where $G(f) = \mathcal{F}[g(t)]$. Carson's theorem is then

$$S_{\mathrm{n}}^{\mathrm{Carson}}(f_k) = (2\pi)^2 [\langle 1/\Delta t \rangle \langle b^2 \rangle |G(f_k)|^2 + \langle b \rangle^2 \delta(f_k)]. \quad \text{(A5)}$$

In our case, the pulse height b_l is binary and the pulse-arrival rate is fixed at $1/\Delta t$; we can therefore write (with Kronecker's delta) $G = \mathcal{F}[\delta_{m,0}] = 1 \times \Delta t$. We just have to replace the mean square $\langle b^2 \rangle$ by the variance σ_b^2 and set $\delta(f_k) = \delta_{k,0} \Delta t$ to obtain Eq. (A4).

Parenthetically, one can also derive Eq. (A4) by using the Wiener-Khintchine theorem. The autocorrelation function is

$$R_{bb}^{(m)} = (1/N) \times \sum_{n=0}^{N-1} b_n b_{n-m} = \langle b \rangle^2 + \sigma_b^2 \delta_{m,0}, \quad \text{(A6)}$$

and $\mathcal{F}[1] = \delta_{k,0} \times \Delta t$, so that

$$S_{\mathrm{n}}(f_k) = (2\pi)^2 \mathcal{F}[R_{bb}^{(m)}] = \text{Eq. (A4)}. \quad \text{(A7)}$$

We find that the PSD resulting from a simulation of Bernoulli and binomially distributed noise agrees well with the measured data and with Eq. (A4): we therefore conclude that our experimental noise floor is due to the statistical sampling.

If the data consisted of a train of pulses of finite length in time, the PSD would have a rolloff near the Nyquist frequency $1/2\Delta t$. For example, the Fourier transform of a boxcar (square) pulse of length Δt is the function $\Delta t \, \mathrm{sinc}(\pi f \Delta t)$. In our case, after conversion of the SQUID's response (the presence or absence of a voltage pulse) to binary form, our data can be seen as represented by a train of delta functions, and their Fourier transform is frequency independent, *i.e.*, our white-noise floor has no rolloff.

3. Cross PSD: White-noise elimination

In order to eliminate the white-noise floor, at the expense of a halved Nyquist frequency, we calculate the discrete cross PSD of interleaved time series, *i.e.*, by setting $z_n' = z_{2n-1}$ and $z_n'' = z_{2n}$ (with $n = 1, \ldots, N/2$) and computing the cross spectrum of z_n' with z_n''. (We again assume that the stochastic switching process is uncorrelated from sample to sample, at frequency $1/\Delta t$.) The resulting PSD is

$$S_{k=0}^{\mathrm{cross}} = (2\pi)^2 \frac{1}{2} \frac{Z_k' (Z_k'')^*}{N/2\Delta t}, \quad S_{k\neq 0}^{\mathrm{cross}} = (2\pi)^2 \frac{Z_k' (Z_k'')^*}{N/2\Delta t}, \quad \text{(A8)}$$

where $k = 0, \ldots, N/4$ and $f_k = k/N\Delta t$. These expressions reproduce the noise spectrum, with the use of correction factors, as explained in the following section [Eqs. (B4)–(B10)].

Compared to the previous section, we have eliminated the white noise by circumventing the zero-delay autocorrelation term $R_{bb}^{(0)}$ in Eq. (A6), and are only left with the delta-function component

$$S_{\mathrm{n}}^{\mathrm{cross}}(f_k) = (2\pi)^2 \langle b \rangle^2 \delta_{k,0} \Delta t. \quad \text{(A9)}$$

In the same way as in Eqs. (A6) and (A7), this can be derived by applying the Wiener-Khintchine theorem

to the cross-correlation function, which this time simply gives $R_{b'b''}^{(m)} = \langle b \rangle^2$ (the subscript $b'b''$ indicates two interleaved subseries obtained from the original series b_n), and therefore

$$S_{\mathrm{n}}^{\mathrm{cross}}(f_k) = (2\pi)^2 \mathcal{F}[R_{bb'}^{(m)}] = \text{Eq. (A9)}. \quad \text{(A10)}$$

When calculating the PSD, we take the statistical average of S_k^{cross}, keeping the averaging coherent throughout (*i.e.*, retaining S_k^{cross} as a complex quantity) and, just as for S_k, smoothen it with a sliding average before plotting its magnitude $|\langle \langle S_k^{\mathrm{cross}} \rangle_{\mathrm{stat}} \rangle_{\mathrm{fq}}|$.

The result (A8) is equivalent to the explicit subtraction of the incoherent noise S_{n} from S_k [Eqs. (A2) and (A4)], with the advantage, however, of drastically reduced uncertainty, in particular at high frequencies where the ($1/f$-noise) signal is much smaller than the white noise. This method is appropriate for the analysis of, *e.g.*, $1/f$-type noise. However, it is not applicable in a predominantly white-noise environment: then, the noise under study would be eliminated along with the statistical white noise.

APPENDIX B: CORRECTION FACTORS: QUASISTATIC NOISE AND THE NONLINEAR TRANSFER FUNCTION

In this section, we treat the effects on the PSD caused by quasistatic noise, and by the sine nonlinearity in the conversion from the measured switching events to the variations of the qubit's transition frequency.

1. Decay of the Ramsey fringe: Quasistatic noise

Noise in the effective longitudinal field coupled to the qubit results in decoherence of the quantum superposition. We denote a fluctuation as "quasistatic" or "incoherent" noise, when it can be considered as static during each free-induction period, but varying over the longer time span between experimental realizations. Dephasing results from such uncorrelated fluctuations of the Larmor frequency ν_{01}, and therefore of the accrued phase of the superposition state $\varphi(\tau) = 2\pi \int_0^{\tau} dt \, \nu_{01}(t)$. It leads to decay of the Ramsey free-induction signal, as each measured point is the incoherent average of many experimental realizations. We describe this fluctuation by a standard deviation

$$\sigma^2 = 2 \int_{1/t_{\mathrm{acq}}}^{1/\tau} df \, S(f). \quad \text{(B1)}$$

The higher integration limit is here the inverse of the free-induction time $1/\tau \approx 0.1-100$ MHz; fluctuations at even higher frequencies are effectively canceling out. The lower limit is given by the total acquisition time t_{acq} used to infer the qubit's population at each fixed free-induction time span τ. Typically, averaging over $N_{\mathrm{avg}} = 5000$ measurements with a repetition time $\Delta t = 2$ ms, we obtain $t_{\mathrm{acq}} = N_{\mathrm{avg}} \Delta t = 10$ s. (If instead the measurements were done in the opposite order, stepping over τ in the inner loop, with $N_{\mathrm{pts}} \approx 100$ steps, and averaging over N_{avg} in the outer loop, the total acquisition time

would be $N_{pts}N_{avg}\Delta t = 1000$ s, and the lower cutoff frequency would be correspondingly lower.)

Ensemble averaging over all realizations of $\delta\varphi(\tau)$ and assuming Gaussian fluctuations resulting from numerous fluctuators, we obtain the dephasing envelope

$$h(\tau) = \langle \exp[i\delta\varphi(\tau)]\rangle = \exp[-\langle(\delta\varphi)^2\rangle/2]$$
$$= \exp\left(-\tau^2 \int_{f_{low}}^{f_{high}} df\, S(f)\, \mathrm{sinc}^2(\pi f\tau)\right)$$
$$\approx \exp(-\sigma^2\tau^2/2), \tag{B2}$$

where the sinc-squared function is due to the square time window of the Ramsey pulse sequence, and we can approximate it by unity for $f < 1/2\tau$.

As an illustration, we now evaluate $h(\tau)$ for the two cases of $1/f$ noise and white noise. For $1/f$ noise, $S(f) = A/f$, Eq. (B1) becomes $\sigma^2 = 2A\ln[(1/2\tau)/(1/t_{acq})]$. The weak, logarithmic sensitivity to the cutoff frequencies effectively allows us to treat it as a time-independent constant $\sigma^2 \approx 2AC$, giving Gaussian decay $p_{1/f}(\tau) = \exp(-\sigma^2\tau^2/2)$. For white noise $S(f) = S_w$, on the other hand, the integral is linearly sensitive to the upper cutoff frequency, so that $\sigma^2 = S_w/\tau$, yielding an exponential decay $p_w(\tau) = \exp(-S_w\tau/2)$. Here, the exponent is proportional to time; we can therefore identify $1/T_\varphi^* = S_w/2$ as the dephasing rate.

2. Repeated fixed-time free induction

The previous section described how quasistatic noise determines the dephasing of the Ramsey fringe. Now we turn to its effect on Ramsey interference with a fixed free-induction time τ_0, repeated numerous times.

With our single-shot measurements, each element of the binary time series $\{z_n\}$ is a Bernoulli random variable z_n with expectation value given by the switching probability p_{sw}, which we now denote as

$$y_n = p_0 + a_0 \sin x_n. \tag{B3}$$

This function has a nonlinear dependence on $x_n = 2\pi\delta\nu_n\tau_0$, the phase accrued during τ_0, where $\delta\nu_n$ is the average fluctuation of the transition frequency at time step n. This phase x_n, in turn, has noise contributions from two distinct frequency intervals, "1" and "2."

We denote as "interval 1" the frequencies which we can resolve by taking the Fourier transform of the series $\{z_n\}$, of total length $N\Delta t$ and step size Δt, i.e., from $1/t_{tot} = 1/N\Delta t \approx 10^{-2}$ Hz to $1/2t_{acq} = 1/2\Delta t \approx 250$ Hz (or with the interleaving method up to $1/4\Delta t \approx 125$ Hz). The noise within this interval has zero mean and variance σ_1^2 [Eq. (B1)].

In addition, there is a contribution from the quasistatic noise in "interval 2," which is the range from $1/2t_{acq}$ to $1/\tau_0$; see Fig. 5. This noise cannot be resolved, but acts in aggregate and leads to dephasing, e.g., in a Ramsey-fringe experiment. It has zero mean and variance σ_2^2 [Eq. (B1)]. Noise at even higher frequencies than $1/\tau_0$ averages out during free induction.

At each time step n, the element x_n is subject to noise contributions from both intervals, and their variances add up, $\sigma^2 = \sigma_1^2 + \sigma_2^2$. We write $x_n = u_n + v_n$, where u and v refer to
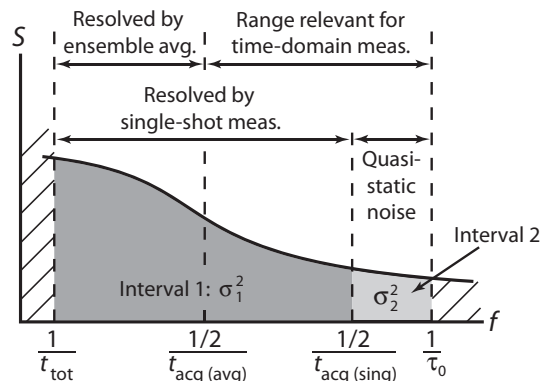

FIG. 5. Sketch of the PSD, indicating the frequency intervals resolved by the ensemble-averaging and single-shot sampling methods. Also indicated are the Gaussian, quasistatic noise, and the variances $\sigma_{1,2}^2$ [Eq. (B1)]. Here, t_{tot} is the total length of the time trace (can be several minutes to hours); $t_{acq(avg)} = N_G\Delta t = 1 \sim 10$ s is the acquisition time per measured point in time-domain experiments such as Ramsey and spin-echo decay; $t_{acq(sing)} = \Delta t = 2$ ms is the repetition time (acquisition time of the single-shot samples); and $\tau_0 \approx 1\,\mu$s is the pulse spacing.

the noise originating in intervals 1 and 2, respectively. Here, u_n has correlations between the different time steps n due to the memory effect of the $1/f$ noise; on the other hand, v_n is incoherent and can be taken as a Gaussian i.i.d. random variable.

While it is impossible to unequivocally infer x_n from the measured z_n at each instance n, we can infer statistical properties of $\{x_n\}$, such as its correlations and spectral density, up to the frequency $1/2t_{acq} = 1/2\Delta t$, which can approach $1/\tau_0$. This is advantageous compared to the ensemble-averaging method, which has a longer acquisition time $t_{acq} = N_G\Delta t$.

We can write the $m \neq n$ autocovariance function for $\Delta z_n = z_n - \langle z_n\rangle$ as

$$\langle\Delta z_m \Delta z_n\rangle = \langle\Delta y_m \Delta y_n\rangle = a_0^2\langle\sin x_m \sin x_n\rangle$$
$$\simeq a_0^2\langle x_m x_n\rangle = a_0^2\langle u_m u_n\rangle. \tag{B4}$$

The first equality holds because the Bernoulli trials are independent, and the last equality is the consequence of v_n being i.i.d., which implies $\langle u_m v_n\rangle = \langle v_m v_n\rangle = 0$. The third step is an equality only when $|x_n| \ll 1$; when σ_2 is large, e.g., at higher temperatures, or when we use a larger free-induction time τ_0 to decrease the statistical noise level, the variation of x_n can be large, and then this is not a good approximation. Instead of approximating, however, we can compensate the result for the sine nonlinearity. Expanding the correlator $\langle\Delta y_m\Delta y_n\rangle$, we

TABLE I. $\delta\varepsilon$ noise ($\varepsilon = 640$ MHz). Data in Figs. 2 and 4.

Temp. (mK)	τ_0 (ns)	$\exp(\sigma^2\tau_0^2)$
65	50	1.5
120	50	1.6
165	50	1.6
210	50	1.8

TABLE II. $\delta\Delta$ noise ($\varepsilon = 0$). Data in Figs. 2 and 4.

Temp. (mK)	τ_0 (ns)	$\exp\left(\sigma^2\tau_0^2\right)$
65	300	1.2
120	300	1.3
165	1200	7.4
180	1000	5.6

obtain

$$\langle \sin x_m \sin x_n \rangle = \langle \sin(u_m + v_m) \sin(u_n + v_n) \rangle$$
$$= \langle (\sin u_m \cos v_m + \cos u_m \sin v_m)(m \to n) \rangle. \tag{B5}$$

Since sine is an odd function and v_n is a zero-mean, Gaussian i.i.d. variable, $\langle \sin v_{m,n} \rangle = 0$, and (B5) becomes

$$\langle \sin u_m \cos v_m \sin u_n \cos v_n \rangle = \langle \cos v_m \rangle \langle \cos v_n \rangle \langle \sin u_m \sin u_n \rangle. \tag{B6}$$

The cosine factors depend on noise in interval 2, *i.e.*, above the sampling frequency. This is similar to dephasing due to quasistatic noise, which acts uniformly on all the samples in time (incoherent averaging over a distribution of the noise), and leads to Gaussian decay functions

$$\langle \cos v_{m,n} \rangle = \exp\left(-\sigma_2^2 \tau_0^2/2\right). \tag{B7}$$

For the sine factor, the noise is from interval 1, *i.e.*, it is resolved by our sampling, and therefore is not uniform. The process is a combination of ensemble-averaged incoherent noise and a frequency-dependent filtering due to the $(m - n)\Delta t$ time difference in the correlator. Evaluating this factor, we obtain Gaussian damping of a hyperbolic-sine function of the correlator

$$\langle \sin u_m \sin u_n \rangle = \iint du_m du_n \, \sin u_m \sin u_n \, N(0,\sigma)$$
$$= \exp\left(-\sigma_1^2 \tau_0^2\right) \sinh(\langle u_m u_n \rangle), \tag{B8}$$

where the integral is taken over a two-dimensional normal distribution $N(0,\sigma)$ with zero mean and correlation matrix $\sigma = \{\sigma_m, \sigma_n, \sigma_{mn}\}$. (The distribution widths are equal, $\sigma_m = \sigma_n$, and $\sigma_{mn} = \langle u_m u_n \rangle$ is the correlation function.)

The correlator (B4) finally becomes

$$\langle \Delta z_m \Delta z_n \rangle = a_0^2 \exp\left(-\sigma_1^2 \tau_0^2\right) \exp\left(-\sigma_2^2 \tau_0^2\right) \sinh(\langle u_m u_n \rangle). \tag{B9}$$

Note that no approximation has been made so far [*cf.* Eq. (A6)]. If the noise correlation due to $1/f$-type noise is small, as in our case where $\exp(\sigma_1^2 \tau_0^2) < 10$, we can neglect the frequency-dependent filtering effect and approximate $\sinh(\langle u_m u_n \rangle) \approx \langle u_m u_n \rangle$.

Now remains only the determination of the correction factors, which we know from the calibration measurement

TABLE III. $\delta\Delta$ noise ($\varepsilon = 0$). Data in Fig. 4 (but not in Fig. 2).

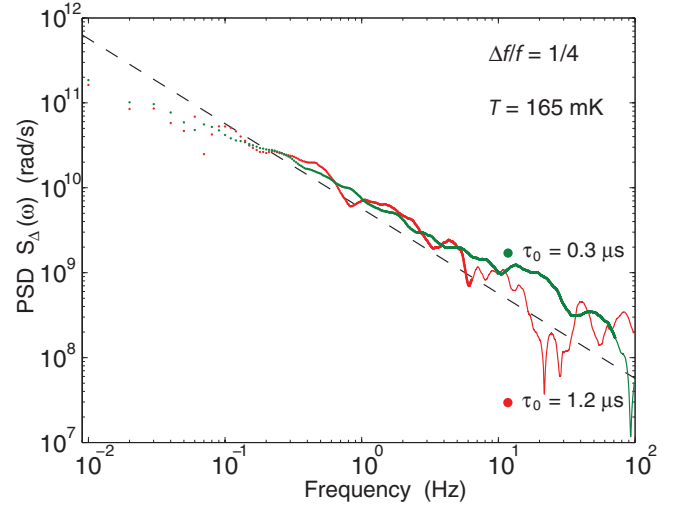| Temp. (mK)	τ_0 (ns)	$\exp\left(\sigma^2\tau_0^2\right)$
165	300	1.3
180	300	1.6

FIG. 6. (Color online) $\delta\Delta$ noise at 165 mK with two different pulse spacings τ_0, showing reproducibility of the noise PSD; *cf.* Tables II and III.

$\exp[(\sigma_1^2 + \sigma_2^2)\tau_0^2] = [a_0/a(\tau_0)]^2$, where we identify $a(\tau_0)/a_0 = h(\tau_0)$ [Eq. (B2)], so that, finally,

$$\langle u_m u_n \rangle \approx \langle \Delta z_m \Delta z_n \rangle/[a(\tau_0)]^2. \tag{B10}$$

We note that it resembles the signal damping due to dephasing in a Ramsey fringe. The actual numbers used in our analysis of the data in Figs. 2 and 4 are presented in Tables I–III.

APPENDIX C: DATA SMOOTHING AND REPRODUCIBILITY OF THE PSD

Figures 6 and 7 show the reproducibility of our results, with sets of data taken on different days. Figures 8 and 9 show that our PSD's power laws are independent of the choice of smoothing windows.
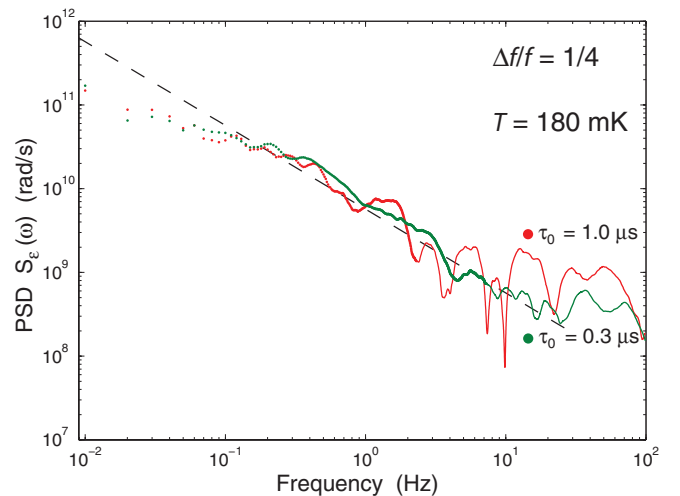

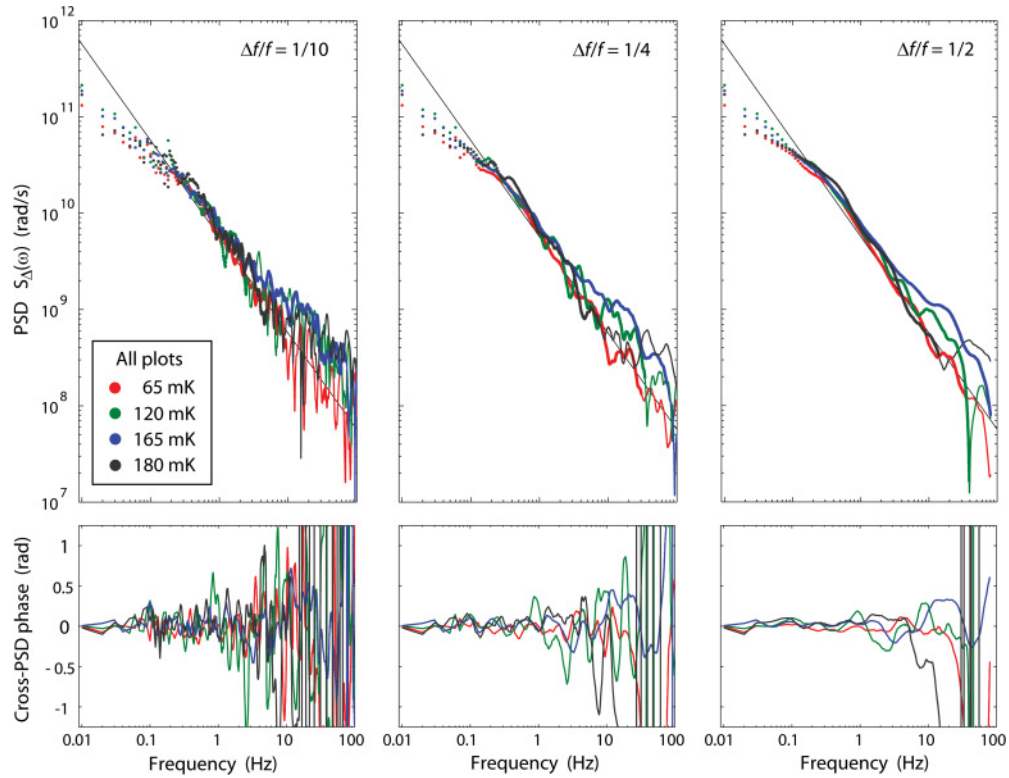FIG. 7. (Color online) $\delta\Delta$ noise at 180 mK, otherwise like Fig. 6.

FIG. 8. (Color online) $\delta\Delta$ noise with different smoothing windows $\Delta f/f$. We choose the upper cutoff frequency f_c for Fig. 2 as the lowest frequency for which the phase of the cross PSD deviates from zero by more than 1 rad. In that figure, we use the smoothing window $\Delta f/f = 1/4$. The spectrum displays no significant difference depending on $\Delta f/f$, and the structure can be attributed to insufficient averaging. The phase deviation is, also, due to insufficient averaging, and becomes larger for increasing temperature, for a fixed pulse separation τ_0.
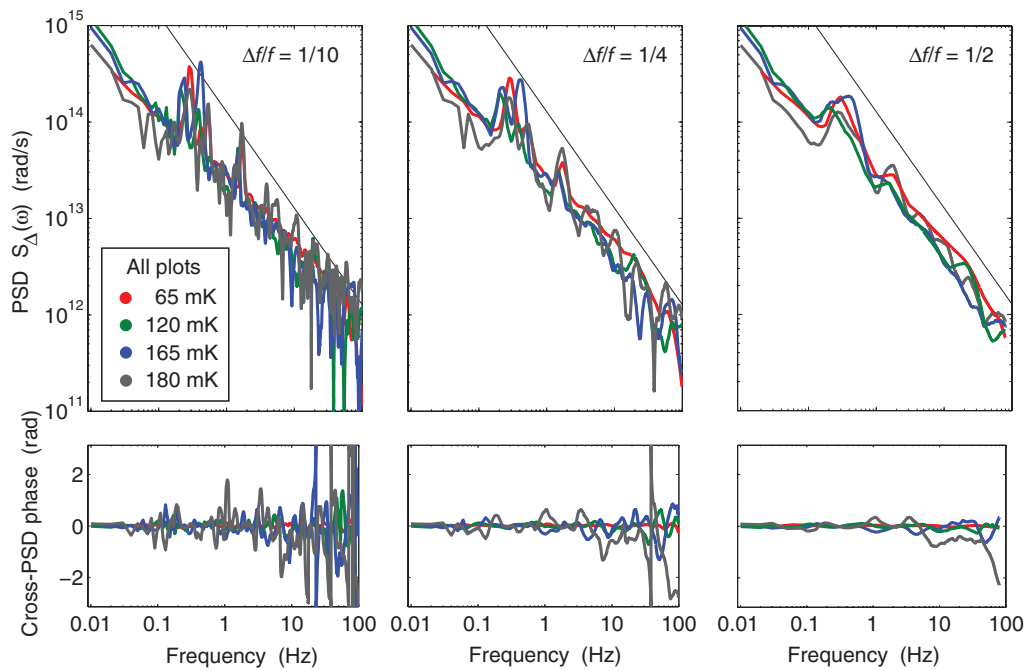

FIG. 9. (Color online) $\delta\varepsilon$ noise with different smoothing windows $\Delta f/f$ (*cf*. Fig. 8).

*fyan@mit.edu

†Present address: Physics Department, King Fahd University of Petroleum, and Minerals, Dhahran 31261, Saudi Arabia.

‡Present address: Research Center for Advanced Science and Technology (RCAST), University of Tokyo, Komaba, Meguro-ku, Tokyo 153-8904, Japan.

[1] Z. Kim, B. Suri, V. Zaretskey, S. Novikov, K. D. Osborn, A. Mizel, F. C. Wellstood, and B. S. Palmer, Phys. Rev. Lett. **106**, 120501 (2011).

[2] J. Bylander, S. Gustavsson, F. Yan, F. Yoshihara, K. Harrabi, G. Fitch, D. G. Cory, Y. Nakamura, J.-S. Tsai, and W. D. Oliver, Nat. Phys. **7**, 565 (2011).

[3] H. Paik, D. I. Schuster, L. S. Bishop, G. Kirchmair, G. Catelani, A. P. Sears, B. R. Johnson, M. J. Reagor, L. Frunzio, L. I. Glazman *et al.* , Phys. Rev. Lett. **107**, 240501 (2011).

[4] S. Sendelbach, D. Hover, A. Kittel, M. Mück, J. M. Martinis, and R. McDermott, Phys. Rev. Lett. **100**, 227006 (2008).

[5] S. Sendelbach, D. Hover, M. Mück, and R. McDermott, Phys. Rev. Lett. **103**, 117001 (2009).

[6] R. H. Koch, D. P. DiVincenzo, and J. Clarke, Phys. Rev. Lett. **98**, 267003 (2007).

[7] S. K. Choi, D.-H. Lee, S. G. Louie, and J. Clarke, Phys. Rev. Lett. **103**, 197001 (2009).

[8] F. Yoshihara, Y. Nakamura, and J. S. Tsai, Phys. Rev. B **81**, 132502 (2010).

[9] S. Gustavsson, J. Bylander, F. Yan, W. D. Oliver, F. Yoshihara, and Y. Nakamura, Phys. Rev. B **84**, 014525 (2011).

[10] F. C. Wellstood, C. Urbina, and J. Clarke, Appl. Phys. Lett. **50**, 772 (1987).

[11] G. Ithier, E. Collin, P. Joyez, P. J. Meeson, D. Vion, D. Esteve, F. Chiarello, A. Shnirman, Y. Makhlin, J. Schriefl *et al.* , Phys. Rev. B **72**, 134519 (2005).

[12] F. Yoshihara, K. Harrabi, A. O. Niskanen, Y. Nakamura, and J. S. Tsai, Phys. Rev. Lett. **97**, 167001 (2006).

[13] R. C. Bialczak, R. McDermott, M. Ansmann, M. Hofheinz, N. Katz, E. Lucero, M. Neeley, A. D. O'Connell, H. Wang, A. N. Cleland *et al.*, Phys. Rev. Lett. **99**, 187006 (2007).

[14] L. Faoro and L. B. Ioffe, Phys. Rev. Lett. **100**, 227005 (2008).

[15] T. Lanting, A. J. Berkley, B. Bumble, P. Bunyk, A. Fung, J. Johansson, A. Kaul, A. Kleinsasser, E. Ladizinsky, F. Maibaum *et al.*, Phys. Rev. B **79**, 060509 (2009).

[16] F. C. Wellstood, C. Urbina, and J. Clarke, IEEE Trans. Appl. Supercond. **21**, 856 (2011).

[17] T. P. Orlando, J. E. Mooij, L. Tian, C. H.van der Wal, L. L. Levitov, S. Lloyd, and J. J. Mazo, Phys. Rev. B **60**, 15398 (1999).

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

[19] D. J. Van Harlingen, T. L. Robertson, B. L. T. Plourde, P. A. Reichardt, T. A. Crane, and J. Clarke, Phys. Rev. B **70**, 064517 (2004).

[20] F. C. Wellstood, C. Urbina, and J. Clarke, Appl. Phys. Lett. **85**, 5296 (2004).

[21] M. Mück, M. Korn, C. G. A. Mugford, J. B. Kycia, and J. Clarke, Appl. Phys. Lett. **86**, 012510 (2005).

[22] J. Eroms, L. C. van Schaarenburg, E. F. C. Driessen, J. H. Plantenberg, K. Huizinga, R. N. Schouten, A. H. Verbruggen, C. J. P. M. Harmans, and J. E. Mooij, Appl. Phys. Lett. **89**, 122516 (2006).

[23] P. Bertet, I. Chiorescu, G. Burkard, K. Semba, C. J. P. M. Harmans, D. P. DiVincenzo, and J. E. Mooij, Phys. Rev. Lett. **95**, 257002 (2005).

[24] D. Sank, R. Barends, R. C. Bialczak, Y. Chen, J. Kelly, M. Lenander, E. Lucero, M. Mariantoni, M. Neeley, P. J. J. O'Malley *et al.*, arXiv:1111.2890.

[25] J. M. Martinis, S. Nam, J. Aumentado, K. M. Lang, and C. Urbina, Phys. Rev. B **67**, 094510 (2003).

[26] K. Bladh, D. Gunnarsson, A. Aassime, M. Taslakov, R. Schoelkopf, and P. Delsing, Phys. E (Amsterdam) **18**, 91 (2003).

[27] O. Astafiev, Y. A. Pashkin, Y. Nakamura, T. Yamamoto, and J. S. Tsai, Phys. Rev. Lett. **93**, 267007 (2004).

[28] N. M. Zimmerman and W. H. Huber, Phys. Rev. B **80**, 195304 (2009).

[29] M. Gustafsson, A. Pourkabirian, G. Johansson, J. Clarke, and P. Delsing, arXiv:1202.5350.