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

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Spectroscopy of low-frequency noise and its temperature dependence in a superconducting qubit

Fei Yan, Jonas Bylander, Simón Gustavsson, Fumiki Yoshihara, Khalil Harrabi, David G. Cory, Terry P. Orlando, and William D. Oliver

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

(Received 26 January 2012; revised manuscript received 4 April 2012; published 22 May 2012)

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.

DOI: 10.1103/PhysRevB.85.174521 PACS number(s): 85.25.Cp, 74.40.−n, 03.67.Lx, 03.65.Yz

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 $\nu_{01}$, and dephasing $T_\phi$ due to low-frequency fluctuations of $\nu_{01}$. 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 $T_\text{dec} = 1/(2T_1) + 1/T_\phi$. While $T_1$ can now exceed 10 $\mu s$ in several superconducting qubit modalities,

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, 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, 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 and the spectrum exhibits a 1/f-type power-law dependence from hertz to tens of megahertz. Its dependence on the device geometry merits further study.

Similarly, for the flux qubit, the noise in the tunnel coupling $\Delta$ between the persistent-current states shows a 1/f-type spectrum from hertz to hundreds of kilohertz. This noise may originate in the critical-current fluctuations of the Josephson junctions 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 $\delta\epsilon$, 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 $1/f^2$ power laws, measured at much higher frequencies, 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 $\hat{H} = -(h/2)(\varepsilon \hat{\sigma}_x + \Delta \hat{\sigma}_z)$. Here, $\varepsilon = 2I_p\Phi_0/h$ is the energy detuning between the diabatic states of classical circulating current $I_p = 0.18 \mu A$, and $\varepsilon$ is adjusted by the external magnetic flux $\Phi$ via $\Phi_0 = \Phi - \Phi_0/2$ ($\Phi_0 = h/2e$ 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 $E_0/E_C$ 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 $\delta \nu_0 = \sqrt{\Delta^2 + \Delta^2} \rightarrow$ fluctuations $D_\nu = \partial \delta \nu_0/\partial \lambda = \lambda(0)/\nu_0$ [see Fig. 1(b)]. The dominant contributor to longitudinal fluctuations in the qubit’s energy eigenbasis is $\delta \nu_0$ noise at $\varepsilon = 0$ (as the second-order contribution from $\delta \varepsilon$ noise is negligible), 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 $I_0$ to zero. (This is close to its optimal value of $I_0^* = 50$ nA in this device, at which $\delta I_0$ fluctuations are decoupled from the qubit.)

To directly probe the fluctuations, we repeatedly let the qubit undergo Ramsey free induction [see Fig. 1(c)]. Instead of scanning the pulse separation $\tau$, we fix it at a value $\tau_0$, 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 $\nu(0) = \nu_0 - \nu_{\text{sw}} = 1/4\tau_0$ (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 $\delta \nu(t) = \nu(0) + \delta \nu(t)$, which in turn translate into fluctuations of the SQUID’s switching probability $p_{\text{sw}} = p_0 - a_0(T) \cos[2\pi \nu(t) \tau_0]$, 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 $p_0 = \text{cont}$ and obtain $\delta p_{\text{sw}}/\delta \nu = 2\pi a_0(T) \tau_0$. 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 $\tau_0$. The conversion factor from the noise $\delta \lambda$ to the switching probability $p_{\text{sw}}$ then becomes $p_{\text{sw}}(\tau_0, T) = 2\pi a(\tau_0, T) \tau_0 D_\nu$.

The single-shot read-out of the qubit, with a repetition time $\Delta t$, 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 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 $p_{\text{sw}}$. 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 $t_{\text{acq}} = N G \Delta t$, with typically $N_G \approx 1000$, to determine the average switching probability $p_{\text{sw}}$ (binomially distributed), and then take the Fourier transform of the time series of switching probabilities $p_{\text{sw}}$; see Fig. 2(a). This approach estimates...
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 $N = 5 \times 10^5$. [Contrary to Refs. 2, 11, and 12, we use the type-1 Fourier transform $S_i(f_k) = \int_{-\infty}^{\infty} dt \langle x(0) \exp(-i 2 \pi f t) \rangle$.] This method increases the upper cutoff frequency from 1/2$N_0 \Delta t$ to 1/2$\Delta t$, which may approach 1/$r_0$ 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 $S_{S}(r_0, T) = (2\pi)^2 \sigma_n^2 \Delta t/\eta_0^2(\tau_0, T)$, where the variance is $\sigma_n^2 = p_{sw}(1 - p_{sw})$. We can eliminate this white background noise by calculating the bilateral cross PSD of the two interleaved (single-shot) time series $z_n^{(1)} = z_{2n-1}$ and $z_n^{(2)} = z_{2n}$:

$$S^{cross}(f_k) = \frac{Z_k^*(Z_k)}{N/\Delta t}, \quad f_k = \frac{k}{N \Delta t}, \quad \text{(1)}$$

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

$$S_i(f_k) = (2\pi)^2 \frac{S^{cross}(f_k)}{\eta_0^2(\tau_0, T)}, \quad \text{(2)}$$

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 $\delta \Lambda$ noise PSDs are plotted in Fig. 2(b) for several temperatures. There is striking agreement with the $1/f^a$ power laws inferred in Ref. 2, measured at considerably higher frequencies ($0.02–20$ MHz). Noise that is strictly $1/f^{a_2}$ over the frequency range relevant to free induction (here $10^{-1} \sim 10^0$ Hz) gives a Gaussian decay function of the temporal Ramsey oscillations. Assuming that our $\delta \Lambda$ noise satisfies these criteria, we use the approach of Ref. 25 to calculate the inhomogeneously broadened decay-time constant $T^*_\Lambda$. With noise sensitivity $\kappa_{\Lambda,i}$ and strength $\Lambda_i$ as defined in the legend of Fig. 2, $T^*_\Lambda = (2\pi \kappa_{\Lambda,i} D_\Lambda)^{-1} A_{\delta \Lambda}/(\ln(2\tau_0))^{-1/2} = (3.2 \mu s)$, in very good agreement with the observed $T^*_\Lambda$ in Fig. 1(c).

We now turn to possible $\delta \varepsilon - \delta \Lambda$ 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 $p_{sw}(\pm \delta \varepsilon, \Lambda) = p_{0} \varphi(\pm \delta \varepsilon, \Lambda + \delta \Lambda)$, with $\pm \delta \varepsilon$ chosen such that the effects of the two noises on $\delta \varepsilon$ were similar in magnitude, i.e., $D_{\Lambda}(\delta \varepsilon) \sim D_{\delta \varepsilon}(\delta \varepsilon)$. We set the pulse separation $r_0$ and nominal frequency detuning $\nu_0$. With the energy-level splitting $\nu(\varepsilon, \Lambda) = \nu_0(\varepsilon, \Lambda + \delta \varepsilon, \delta \varepsilon, \delta \Lambda)$, we use the decay function of the Ramsey fringe to infer the noise correlations from the measured $p_{sw}$:

$$p_{sw}(\varepsilon, \Delta) = p_0 - a_0 \exp(-\tau_0/2\tau_1) \times \exp(-[\varepsilon/T^*_\Lambda(\varepsilon/\lambda)^2] \cos(2\pi \nu(\varepsilon, \Lambda)(\tau_0)), \quad \text{(3)}$$

where $[1/T^*_\Lambda(\varepsilon)]^2 = [1/T^*_\Lambda(0)]^2 + K \Lambda_d D^2$ and $K$ is a constant that we have determined independently, along with the other parameters in the equation. At the bias points $\varepsilon = \pm \delta \varepsilon$, $\delta \varepsilon$ fluctuations induce negatively correlated $\delta \varepsilon$ fluctuations.
(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 $\nu = \pm \nu_0$ yields a system of two nonlinear equations in the two unknowns $\delta \nu$ and $\delta \Delta$. We solve this system numerically: Fig. 3(a) shows the raw $p_{sw}$ data and extracted $\delta \nu$ and $\delta \Delta$ versus time. We then calculate the cross PSD $S_{\nu \Delta}(f)$ 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):

$$|\gamma_{\nu \Delta}(f)| = \left| \frac{S_{\nu \Delta}(f)}{S_{\nu}(f) S_{\Delta}(f)} \right|^{1/2} < 0.2. \quad (4)$$

Finally, we measured the temperature dependencies of the two types of noise. Figure 4 shows the integrated noise powers $\Pi_\nu(T)$ in the frequency intervals $F_\nu = 0.02$–50 Hz and $F_\Delta = 0.02$–2 Hz [cf. Fig. 2(b)]. It is possible to measure the $\delta \nu$ noise up to somewhat higher temperatures and frequencies than the $\delta \Delta$ noise. Note that the $\Pi_\nu$ values depend on the integration limits although the choice of $F_\nu$ 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 $\tau_{\nu}$.) 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 $\alpha(\nu_0, T)$, included in the $p_{sw}$-to-$\delta \nu$ conversion factor $\eta_\nu(\nu_0, T)$ [see Fig. 1(a)].

**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 $I_c = 0.4 \mu A$, in a Josephson junction with area $A = (0.2 \mu m)^2$. Van Harlingen et al.\textsuperscript{19} found a “canonical” value for the $1/f$ $\delta I_c$-noise power at 1 Hz and 4.2 K: $A_{\nu \Delta}^{\text{can}} \approx 144 (\text{pA})^2 (I_c/\mu A)^2 (A/\mu m^2)$ 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 $S_{\nu \Delta}(f) = A_{\nu \Delta}^{\text{can}} I_c^{-2} (T/2.4 K)^2 / |f|^{1/2}$, which, for $T = 65$ mK, is considerably lower (almost 20 times) than our measured value. On the other hand, Eroms et al.\textsuperscript{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., $S_{\nu \Delta}(f) = (1/100) \times A_{\nu \Delta}^{\text{can}} I_c^{-2} (T/4.2 K)^2 / |f|$. 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.\textsuperscript{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 $\delta Q$, known to exhibit $1/f$ noise;\textsuperscript{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\textsuperscript{29} (although quadratic dependencies have also been observed\textsuperscript{27}) and saturates below about 200 mK, due to self-heating of the SET, at a “canonical” value $A_Q$ of about $(1 \sim 10 \mu e)^2$ at 1 Hz. We estimate our qubit’s maximum sensitivity to charge fluctuations $\kappa_{\Delta,Q} \equiv \delta \Delta / \delta Q$ to be in the range $0.1 \sim 1 \text{ MHz/e}$. We can then parametrize
the \( \delta \Delta \) noise as charge noise and estimate the dephasing time \( T_\phi = (2\pi \kappa_\Delta \zeta_\Delta D_\Delta)^{-1} A^1_{\Delta} / (\ln(\tau_{acq}/2 \tau_0)) \approx 4 \sim 400 \mu s \). 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 \( \epsilon = 0 \).

In conclusion, our spectroscopy of both \( \delta \epsilon \) 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 \( \delta \epsilon \) noise and at least 8 orders of magnitude for \( \delta \Delta \) noise. The \( \delta \epsilon \) 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, \( \delta \epsilon - \delta \Delta \) noise correlations (over \( 7 \times 10^{-4} - 2 \times 10^{-1} \text{Hz} \)) 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.

ACKNOWLEDGMENTS

We acknowledge discussions with G. Chen, L. DiCarlo, M. Gustafsson, X. Jin, and L. Wang. We thank the LTSE team at MIT Lincoln Laboratory for technical assistance. This work was sponsored in part by the US Government, the Laboratory for Physical Sciences, the US Army Research Office (W911NF-12-1-0036), the National Science Foundation (PHY-1005373), and the Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST), NICT Commissioned Research, MEXT kakenhi “Quantum Cybernetics.”

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)

\[
S(\nu) = (2\pi)^2 \lim_{T \to \infty} \frac{\langle |V(\nu)|^2 \rangle}{T}.
\]

Our measurements’ raw data, however, consists of a binary time series \( \{z_n\} \) with elements of expectation value \( y_n = \langle z_n \rangle = p_0 + a_0 \sin(2\pi \nu (n \Delta t) \tau_0) \), where \( \Delta t \) 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 \( Z_k = Z_k \times \Delta t \) for \( k = k/N \Delta t \), and compute the discrete, bilateral noise PSD over the frequency range from \( 1/\tau_{tot} = 1/N \Delta t = 10 \text{ mHz} \) to \( 1/2\tau_{acq} = 1/2 \Delta t = 250 \text{ Hz} \):

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

for \( k = 0, \ldots, (N - 1)/2 \) with \( f_k = k/N \Delta t \).

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

\[
\langle S_k \rangle_{\text{stat}} = \frac{1}{M} \sum_{m=1}^{M} S_k^{(m)}, \quad (A3)
\]

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 \( \tau_{acq} \) is the same as the pulse-sequence repetition time \( \Delta t \). This sets it apart from the standard approach of first taking the ensemble average of typically \( N_G = 1000 \) samples in the time domain, before calculating the PSD of the resulting \( N/N_G \) sampled points. The acquisition time is then \( \tau_{acq} = N_G \Delta t \), and the upper cutoff frequency becomes only \( 1/2 \tau_{acq} \approx 0.25 \text{ Hz} \).

2. White-noise floor

The PSD of the single-shot time sequence suffers from statistical sampling noise because each time step \( \tau_n \) 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 \( \sigma_b^2 = \langle (\Delta b)^2 \rangle = \langle b^2 \rangle - \langle b \rangle^2 = p(1 - p) \). By sampling at a fixed rate \( 1/\Delta t \), the white-noise floor of the bilateral PSD becomes

\[
S_b(f_k) = (2\pi)^2 \left( \sigma_b^2 + \langle b \rangle^2 \delta_{k,0} \right) \Delta t. \quad (A4)
\]

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 \( \tau_{acq} = N_G \Delta t \) : we obtain Eq. (A4) after substituting \( \langle c \rangle = p \) and \( \sigma_c^2 = p(1 - p)/N_G \) for \( \langle b \rangle \) and \( \sigma_b^2 \), 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 \( w(t) = \sum_{i=1}^{\infty} b_i g(t - t_i) \), in which \( g(t) \) is the pulse envelope, the
stochastic variable $b_t$ is the (continuous) pulse height, and the stochastic variable $t_i$ is the pulse-arrival time. The Fourier transform of $w(t)$ is $W(f) = G(f) \sum_{l=0}^{L} b_l \exp(-i2\pi ft_i)$, where $G(f) = \mathcal{F}[g(t)]$. Carson’s theorem is then

$$S_n^{\text{Carson}}(f_k) = (2\pi)^2 \left[ (1/\Delta t) |b|^2 \right] |G(f_k)|^2 + (b)^2 \delta(f_k). \quad (A5)$$

In our case, the pulse height $b_t$ is binary and the pulse-arrival rate is fixed at $1/\Delta t$; we can therefore write (with Kronecker’s delta) $G(f) = \mathcal{F}[\delta_{n,0}]$. We just have to replace the mean square $\langle b^2 \rangle$ by the variance $\sigma_n^2$ 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

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

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

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

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 $\Delta t$ is the function $\Delta t \text{sinc}(\pi f/\Delta t)$. 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 correlations of the Larmor frequency $\nu_0(t)$, and therefore of the accrued phase of the superposition state $\varphi(t) = 2\pi \int_0^t dt \nu_0(t)$. 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

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

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 $\tau_{\text{acq}}$ used to infer the qubit’s population at each fixed free-induction time span $\tau$. Typically, averaging over $N_{\text{avg}} = 5000$ measurements with a repetition time $\Delta t = 2$ ms, we obtain $\tau_{\text{acq}} = N_{\text{avg}} \Delta t = 10$ s. (If instead the measurements were done in the opposite order, stepping over $\tau$ in the inner loop, with $N_{\text{pts}} \approx 100$ steps, and averaging over $N_{\text{avg}}$ in the outer loop, the total acquisition time

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 fringe 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 $\nu_0(t)$, and therefore of the accrued phase of the superposition state $\varphi(t) = 2\pi \int_0^t dt \nu_0(t)$. 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

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

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 $\tau_{\text{acq}}$ used to infer the qubit’s population at each fixed free-induction time span $\tau$. Typically, averaging over $N_{\text{avg}} = 5000$ measurements with a repetition time $\Delta t = 2$ ms, we obtain $\tau_{\text{acq}} = N_{\text{avg}} \Delta t = 10$ s. (If instead the measurements were done in the opposite order, stepping over $\tau$ in the inner loop, with $N_{\text{pts}} \approx 100$ steps, and averaging over $N_{\text{avg}}$ in the outer loop, the total acquisition time
would be $N_{\text{pp}}N_{\text{avg}} \Delta t = 1000 \text{ s}$, and the lower cutoff frequency would be correspondingly lower.)

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

$$h(\tau) = \langle \exp[\delta \varphi(t)] \rangle = \exp[-((\delta \varphi)^2)/2]$$

where the sinc-squared function is due to the square time window.

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 $\sigma^2 = 2A \ln((1/\tau)/(1/\Delta f))$. The weak, logarithmic sensitivity to the cutoff frequencies effectively allows us to treat it as a time-independent constant $\sigma^2 \approx 2A/C$, 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 $1/T_w = S_w/2$ 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 $\tau_0$, 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 $p_w$, which we now denote as

$$y_n = p_0 + a_0 \sin x_n, \quad (B3)$$

This function has a nonlinear dependence on $x_n = 2\pi \delta y_n \tau_0$, the phase accrued during $\tau_0$, where $\delta y_n$ 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 $\Delta t$, i.e., from $1/\Delta f = 1/N \Delta t \approx 10^{-2} \text{ Hz}$ to $1/2\Delta f = 1/2 \Delta ft \approx 250 \text{ Hz}$ (or with the interleaving method up to $1/4 \Delta ft \approx 125 \text{ Hz}$). The noise within this interval has zero mean and variance $\sigma_{1z}^2$ [Eq. (B1)].

In addition, there is a contribution from the quasistatic noise in “interval 2,” which is the range from $1/2f_{\text{acq}}$ 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 $\sigma_{2z}^2$ [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, $\sigma^2 = \sigma_{1z}^2 + \sigma_{2z}^2$. 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 $\sigma_{1z}^2$ [Eq. (B1)]. Here, $t_{\text{acq}}$ is the total length of the time trace (can be several minutes to hours); $t_{\text{acq(avg)}} = N \Delta t = 1 \sim 10 \text{ s}$ is the acquisition time per measured point in time-domain experiments such as Ramsey and spin-echo decay; $t_{\text{acq(sing)}} = \Delta t = 2 \text{ ms}$ is the repetition time (acquisition time of the single-shot samples); and $\tau_0 \approx 1 \mu\text{s}$ is the pulse spacing.

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

<table>
<thead>
<tr>
<th>Temp. (mK)</th>
<th>$\tau_0$ (ns)</th>
<th>exp$(\sigma^2 T_0^2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>50</td>
<td>1.5</td>
</tr>
<tr>
<td>120</td>
<td>50</td>
<td>1.6</td>
</tr>
<tr>
<td>165</td>
<td>50</td>
<td>1.6</td>
</tr>
<tr>
<td>210</td>
<td>50</td>
<td>1.8</td>
</tr>
</tbody>
</table>

The first equality holds because the Bernoulli trials are independent, and the last equality is the consequence of $\tau_0$ 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 $\langle x_n \rangle \ll 1$; when $\sigma_2$ is large, e.g., at higher temperatures, or when we use a larger free-induction time $\tau_0$ 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

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 inf er statistical properties of $\{x_n\}$, such as its correlations and spectral density, up to the frequency $1/2f_{\text{acq}} = 1/2 \Delta f$, which can approach $1/\tau_0$. This is advantageous compared to the ensemble-averaging method, which has a longer acquisition time $t_{\text{acq}} = N \Delta t$.

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

$$\langle \Delta z_m \Delta z_n \rangle = \langle \Delta y_m \Delta y_n \rangle = a_0^2 \langle \sin x_m \sin x_n \rangle \approx a_0^2 \langle x_m x_n \rangle = a_0^2 \langle u_m u_n \rangle. \quad (B4)$$

The first equality holds because the Bernoulli trials are independent, and the last equality is the consequence of $\tau_0$ 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 $\langle x_n \rangle \ll 1$; when $\sigma_2$ is large, e.g., at higher temperatures, or when we use a larger free-induction time $\tau_0$ 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 II. $\delta \Delta$ noise ($\varepsilon = 0$). Data in Figs. 2 and 4.

<table>
<thead>
<tr>
<th>Temp. (mK)</th>
<th>$\tau_0$ (ns)</th>
<th>$\exp(\sigma_0^2 \tau_0^2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>300</td>
<td>1.2</td>
</tr>
<tr>
<td>120</td>
<td>300</td>
<td>1.3</td>
</tr>
<tr>
<td>165</td>
<td>1200</td>
<td>7.4</td>
</tr>
<tr>
<td>180</td>
<td>1000</td>
<td>5.6</td>
</tr>
</tbody>
</table>

obtain

$$\langle \sin x_m \sin x_n \rangle = \langle (\sin u_m + v_m) \sin (u_n + v_n) \rangle = \langle (\sin u_m \cos v_m + \cos u_m \sin v_m)(m \to n) \rangle.$$ \hspace{1cm} (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

$$\langle \sin u_m \cos v_m \sin u_n \cos v_n \rangle = \langle \cos v_m \rangle \langle \cos v_n \rangle \langle \sin u_m \sin u_n \rangle.$$ \hspace{1cm} (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

$$\langle \cos v_{m,n} \rangle = \exp\left(-\sigma_0^2 \tau_0^2/2\right).$$ \hspace{1cm} (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

$$\langle \sin u_m \sin u_n \rangle = \int \int du_m du_n \sin u_m \sin u_n N(0,\sigma)$$

$$= \exp\left(-\sigma_0^2 \tau_0^2\right) \sinh(u_m u_n),$$ \hspace{1cm} (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

$$\langle \Delta z_m \Delta z_n \rangle = a_0^2 \exp\left(-\sigma_0^2 \tau_0^2\right) \exp\left(-\sigma_0^2 \tau_0^2\right) \sinh(u_m u_n).$$ \hspace{1cm} (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 $\exp(\sigma_0^2 \tau_0^2) < 10$, we can neglect the frequency-dependent filtering effect and approximate $\sinh(u_m u_n) \approx \langle u_m u_n \rangle$.

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).

<table>
<thead>
<tr>
<th>Temp. (mK)</th>
<th>$\tau_0$ (ns)</th>
<th>$\exp(\sigma_0^2 \tau_0^2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>165</td>
<td>300</td>
<td>1.3</td>
</tr>
<tr>
<td>180</td>
<td>300</td>
<td>1.6</td>
</tr>
</tbody>
</table>

FIG. 6. (Color online) $\delta \Delta$ noise at 165 mK with two different pulse spacings $\tau_0$, showing reproducibility of the noise PSD; cf. Tables II and III.

exp[$(\sigma_1^2 + \sigma_2^2) \tau_0^2] = \frac{a_0}{a_0(t_0)^2}$, where we identify $a(t_0)/a_0 = h(t_0)$ [Eq. (B2)], so that, finally,

$$\langle u_m u_n \rangle \approx \langle \Delta z_m \Delta z_n \rangle /\langle a(t_0)^2 \rangle.$$ \hspace{1cm} (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 $\tau_0$.

FIG. 9. (Color online) $\delta \varepsilon$ noise with different smoothing windows $\Delta f/f$ (cf. Fig. 8).
† fyan@mit.edu

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

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


