PHYSICAL REVIEW X **13,** 041005 (2023)

Disentangling Losses in Tantalum Superconducting Circuits

Kevin D. Crowley,[1,†] Russell A. McLellan,[2,†] Aveek Dutta,[2,†] Nana Shumiya,[2] Alexander P. M. Place,[2] Xuan Hoang Le,[2]
Youqi Gang,[2] Trisha Madhavan,[2] Matthew P. Bland,[2] Ray Chang,[2] Nishaad Khedkar,[2] Yiming Cady Feng,[2]
Esha A. Umbarkar,[1] Xin Gui,[3] Lila V. H. Rodgers,[2] Yichen Jia,[4] Mayer M. Feldman,[1] Stephen A. Lyon,[2]
Mingzhao Liu,[4] Robert J. Cava,[3] Andrew A. Houck,[2] and Nathalie P. de Leon[2,*]

[1]*Department of Physics, Princeton University, Princeton, New Jersey 08540, USA*
[2]*Department of Electrical and Computer Engineering,*
Princeton University, Princeton, New Jersey 08540, USA
[3]*Department of Chemistry, Princeton University, Princeton, New Jersey 08540, USA*
[4]*Center for Functional Nanomaterials, Brookhaven National Laboratory, Upton, New York 11973, USA*

(Received 8 February 2023; revised 29 July 2023; accepted 29 August 2023; published 6 October 2023)

Superconducting qubits are a leading system for realizing large-scale quantum processors, but overall gate fidelities suffer from coherence times limited by microwave dielectric loss. Recently discovered tantalum-based qubits exhibit record lifetimes exceeding 0.3 ms. Here, we perform systematic, detailed measurements of superconducting tantalum resonators in order to disentangle sources of loss that limit state-of-the-art tantalum devices. By studying the dependence of loss on temperature, microwave photon number, and device geometry, we quantify materials-related losses and observe that the losses are dominated by several types of saturable two-level systems (TLSs), with evidence that both surface and bulk related TLSs contribute to loss. Moreover, we show that surface TLSs can be altered with chemical processing. With four different surface conditions, we quantitatively extract the linear absorption associated with different surface TLS sources. Finally, we quantify the impact of the chemical processing at single-photon powers, the relevant conditions for qubit device performance. In this regime, we measure resonators with internal quality factors ranging from 5 to 15×10^6, comparable to the best qubits reported. In these devices, the surface and bulk TLS contributions to loss are comparable, showing that systematic improvements in materials on both fronts are necessary to improve qubit coherence further.

DOI: 10.1103/PhysRevX.13.041005
Subject Areas: Materials Science, Quantum Information, Superconductivity

I. INTRODUCTION

Superconducting qubits have been deployed in some of the most sophisticated quantum processors, enabling demonstrations of quantum error correction [1–5], quantum many-body physics and entanglement dynamics [6–9], and quantum simulation [10]. Improvements in superconducting qubit coherence would help to enable large-scale quantum processors, potentially capable of executing useful tasks. Current superconducting qubits are limited by dielectric loss that is orders of magnitude higher than expected from bulk properties of the constituent materials [11–14]. This high dielectric loss indicates that qubit relaxation likely originates from uncontrolled surfaces, interfaces, and contaminants. Tantalum qubits have recently been demonstrated to exhibit record lifetimes and coherence times exceeding 0.3 ms [15], which has been reproduced with different fabrication methods [16] and substrates [17], indicating that major advances can be enabled by materials discovery. Tantalum qubits have also recently been deployed to achieve break-even quantum error correction [3], and further improvements in coherence could allow current processors and architectures to push beyond the threshold for fault tolerance [5,18]. The advantage of tantalum likely arises from its stoichiometric, kinetically limited oxide and its chemical robustness, allowing for extensive device cleaning [15,19]. However, little is known about the remaining sources of loss that limit state-of-the-art tantalum devices.

Prior work in other material systems has focused on the role of parasitic two-level systems (TLSs) in decoherence and dissipation [20]. TLSs were originally explored in the context of thermal transport in glasses [21,22] and are ubiquitous sources of loss and decoherence in myriad systems, including superconducting devices [23–30], microwave

---

[*] npdeleon@princeton.edu
[†] These authors contributed equally to this work.

kinetic inductance detectors [20,31,32], optomechanical cavities [33], and acoustic resonators [34,35]. However, the magnitude of the TLS contribution to device loss is difficult to quantitatively disentangle from other sources of loss, such as radiative losses [36], packaging [37], nonequilibrium quasiparticles [38], and nonsaturable absorption [39,40]. This identification is complicated by the likelihood that there are multiple TLS sources in a given device, whose relative contribution may depend on device geometry, fabrication and cleaning procedures, and subtle material choices.

Here, we quantitatively separate different contributions to microwave loss arising from TLSs, quasiparticles, and other channels by varying temperature and microwave photon number. We observe internal quality factor (Q_{int}) up to 2×10^8 at high power, giving a large dynamic range that allows us to measure subtle sources of loss. We observe a nonmonotonic temperature dependence in Q_{int}, with the low-temperature behavior well described by TLS loss. Our ability to capture the complex, nonmonotonic power and temperature dependence of the loss gives us confidence that we have accounted for the dominant sources of loss. Furthermore, we can quantify surface and bulk TLS contributions by varying device geometry. We find that smaller devices are dominated by surface TLSs, while contributions from TLSs residing in the bulk become evident in larger devices. By treating the devices with a postfabrication buffered oxide etch (BOE), we can decrease the surface TLS bath, and by comparing different surface treatments, we can quantitatively estimate the contribution of different material interfaces. Finally, we characterize the different components of TLS loss at single microwave photon powers as a proxy for qubit performance.

II. RESONATOR FABRICATION AND MEASUREMENTS

We deposit 200 nm of Ta epitaxially on 300- or 500-μm-thick sapphire substrates using dc magnetron sputtering at elevated temperatures to stabilize the bcc α phase [41] (Appendix Sec. A 1). The 200 nm film thickness matches our prior work [15] and is kept constant throughout device simulations. However, in contrast to our prior work, all films are $\langle 111 \rangle$ oriented, single crystal, with some films having a minority component of the $\langle 110 \rangle$ orientation (Appendix Sec. I). We pattern resonators using photolithography followed by metal etching, either using a selective wet chemical etch or dry etching in an inductively coupled plasma reactive ion etching system. We then strip the photoresist and clean the devices using a piranha solution composed of 1:2 hydrogen peroxide in sulfuric acid ("native" surface). Finally, in order to chemically alter the tantalum surface, some devices are treated with either a 10:1 buffered oxide etch for 20 min ("BOE" surface), a 10:1 buffered oxide etch for 120 min ("long BOE" surface), or a refluxing mixture of 1:1:1 concentrated sulfuric, nitric, and perchloric acids ("triacid" surface).

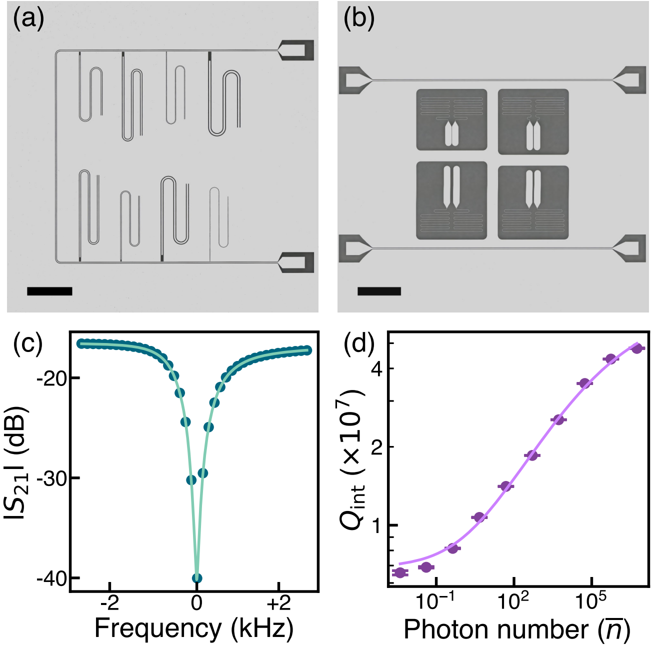

FIG. 1. Measuring superconducting resonators. (a) Optical microscope image of a CPW resonator chip, consisting of eight resonators with varying pitch capacitively coupled to a single rf feedline. The scale bar represents 1 mm. (b) Optical microscope image of an LE resonator chip, consisting of four resonators with varying capacitor spacing inductively coupled to a single rf feedline. The scale bar represents 1 mm. Scanning electron microscope images of representative devices are shown in Fig. 5. (c) Transmission spectrum of a single resonator, centered about the resonance frequency $f_0 = 4.484501$ GHz. The solid line indicates a fit to the data from which Q_{int}, f_0, and the coupling losses are extracted. (d) Q_{int} as a function of the applied microwave power, expressed in terms of the average intracavity photon number. The solid line indicates a fit to the data based on saturation of TLSs and an additional power-independent loss.

The fabricated devices consist of either coplanar waveguide (CPW) quarter-wave resonators or lumped element (LE) resonators. We vary their sizes to achieve different surface participation ratios (SPRs) [Figs. 1(a) and 1(b)], the fraction of the electric field energy residing in surface layers of the device [28,42]. The CPW resonators are shorted transmission lines with characteristic impedances of 50 Ω, and the LE resonators are LC oscillators with characteristic impedance of 300–400 Ω. Multiple resonators are coupled to a single feedline and are designed to have different resonant frequencies between 4 and 8 GHz to allow for spectrally selective interrogation.

We characterize the losses in each resonator by measuring transmission through the feedline in a dilution refrigerator with a base temperature around 17 mK and scanning the frequency of the probe tone around the resonant frequency (Appendix Sec. A 3). At the resonant frequency, the line shape of the transmission dip reflects

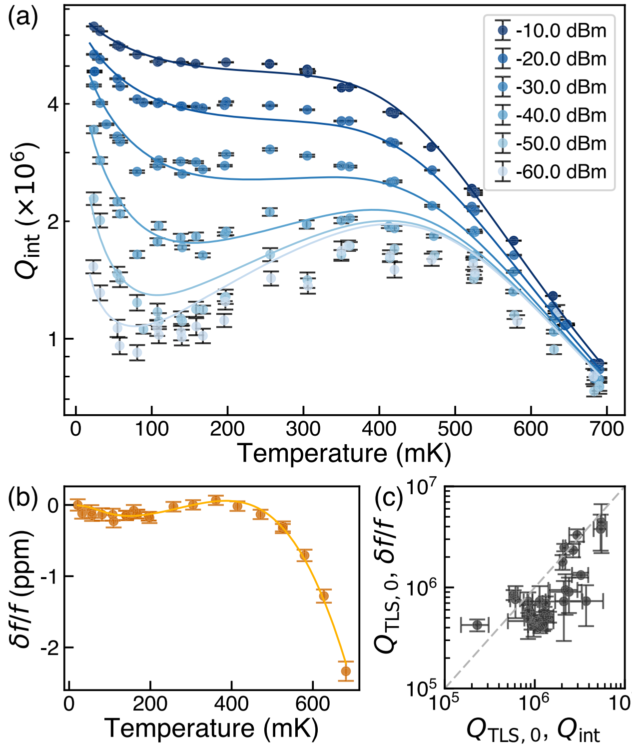

FIG. 2. Parametrizing losses. (a) Internal quality factor Q_{int} as a function of applied microwave power and temperature for a characteristic resonator. The traces are well separated at low temperatures and then collapse together and fall exponentially at high temperatures. The characteristic shape of the curves is fit to a model incorporating TLS loss and equilibrium quasiparticles. Solid lines show the best fit to the dataset. (b) Shift in the resonant frequency with temperature relative to the base temperature center frequency for a representative device. The solid line represents a fit to the data. (c) Comparison between estimates of $Q_{\mathrm{TLS,0}}$ extracted from two independent measurements: the power and temperature dependence of Q_{int} and the temperature dependence of the frequency. The dashed line is a guide to the eye showing the case where $Q_{\mathrm{TLS,0}}$ is equal for both measurements. Only 26 devices are shown in this plot, because we optimize our measurements to measure Q_{int} across temperature and power, with a sparser temperature sampling than is required to extract $Q_{\mathrm{TLS,0}}$ from the fractional frequency shift with high confidence.

both the internal losses and the coupling to the feedline. We fit the line shape to extract the internal quality factor Q_{int} [Fig. 1(c)] [43] (Appendix Sec. A 5).

A common observation in superconducting circuits is that losses decrease with increasing microwave power, indicating that the losses are from saturable TLSs [12]. We observe similar power-dependent loss in our devices [Fig. 1(d)], with low power Q_{int} ranging from 1×10^5 to 1×10^7 and high power Q_{int} ranging from 1×10^7 to 2×10^8 across different devices.

Different sources of loss can be distinguished by their power and temperature dependence. In order to further disentangle different physical mechanisms for loss, we characterize resonator losses over a wide range of temperatures and microwave powers [Fig. 2(a)]. The full power and temperature dependence is well described by a model that incorporates three sources of loss: TLSs (Q_{TLS}), equilibrium quasiparticles (Q_{QP}), and a separate power- and temperature-independent loss channel that limits Q_{int} at the highest microwave powers (Q_{other}). We fit the full dataset using the following model:

$$\frac{1}{Q_{\mathrm{int}}} = \frac{1}{Q_{\mathrm{TLS}}(\bar{n}, T)} + \frac{1}{Q_{\mathrm{QP}}(T)} + \frac{1}{Q_{\mathrm{other}}}. \tag{1}$$

The TLS and quasiparticle losses are parametrized by [20]

$$Q_{\mathrm{TLS}}(\bar{n}, T) = Q_{\mathrm{TLS},0}\frac{\sqrt{1+\left(\frac{\bar{n}^{\beta_2}}{DT^{\beta_1}}\right)\tanh\left(\frac{\hbar\omega}{2k_BT}\right)}}{\tanh\left(\frac{\hbar\omega}{2k_BT}\right)} \tag{2}$$

and

$$Q_{\mathrm{QP}}(T) = A_{\mathrm{QP}}\frac{e^{\Delta_0/k_BT}}{\sinh\left(\frac{\hbar\omega}{2k_BT}\right)K_0\left(\frac{\hbar\omega}{2k_BT}\right)}, \tag{3}$$

where ω is the center angular frequency of the resonator; T is the temperature; $\bar{n}$ is the intracavity photon number; $Q_{\mathrm{TLS,0}}$ is the inverse linear absorption from TLSs; D, β_1 [44], and β_2 [45] are parameters characterizing TLS saturation; A_{QP} is an overall amplitude proportional to the kinetic inductance ratio; Δ_0 is the superconducting gap ($\Delta_0 = 1.764k_BT_c$); T_c is the superconducting critical temperature of the film; K_0 is the zeroth-order modified Bessel function of the second kind; k_B is the Boltzmann constant; and $\hbar$ is the reduced Planck constant. There are seven free fit parameters: $Q_{\mathrm{TLS,0}}$, D, β_1, β_2, A_{QP}, T_c, and Q_{other}.

This model gives rise to three separate regimes in the data. At high temperatures above 500 mK, Q_{int} exhibits a weak power dependence and decreases exponentially with temperature. This behavior is consistent with equilibrium quasiparticle loss, as the temperature becomes an appreciable fraction of the superconducting critical temperature for α-Ta [20]. Although prior studies in other materials have observed power-dependent quasiparticle loss [46], we do not observe such behavior in tantalum. We note that our measurements are optimized to capture the low-power, low-temperature losses and, therefore, may not disambiguate the T_c and A_{QP} fitting parameters accurately. We also note that we observe a large range in T_c that is not observable by direct film characterization, indicating that the resonators are much more sensitive probes for minority phases. At intermediate temperatures 100–500 mK, Q_{int} increases with temperature by around a factor of 2 for the lowest microwave powers, with little variation with temperature at the highest microwave powers, consistent with thermal saturation of TLSs, where the characteristic temperature is given by the resonator frequency [20]. At the lowest temperatures, there is an apparent $1/T$ dependence of Q_{int}. This behavior is consistent with a decreasing TLS coherence time with increasing temperature and subsequent increase of TLS saturation power [20,47]. In our devices, the saturation power at base temperature for the TLS bath can be as low as 0.01 photons.

At base temperature and at the powers relevant for transmon operation where the average intracavity photon number is $\bar{n} = 1$, the dominant source of loss is $1/Q_{\mathrm{TLS}}$. We focus on $Q_{\mathrm{TLS},0}$ as a parameter that captures linear absorption due to TLSs and, therefore, reveals differences in the materials under different fabrication conditions. We can check that $Q_{\mathrm{TLS},0}$ is a robust parameter by independently measuring the temperature-dependent shift in the frequency of the resonator [Fig. 2(b)]. The resonance frequency shifts because of the change in the real part of the dielectric constant arising from losses associated with the entire spectral distribution of the TLS bath [48], as well as losses induced by quasiparticles. The frequency shift is given by [20]

$$\frac{\delta f(T)}{f_0} = \left(\frac{\delta f(T)}{f_0}\right)_{\mathrm{TLS}} + \left(\frac{\delta f(T)}{f_0}\right)_{\mathrm{QP}}, \tag{4}$$

where f_0 is the center frequency of the resonator at zero temperature and δf is the difference in the center frequency of the resonator at nonzero temperature.

The TLS and quasiparticle contributions to the frequency shift are given by [20]

$$\left(\frac{\delta f(T)}{f_0}\right)_{\mathrm{TLS}} = \frac{1}{\pi Q_{\mathrm{TLS},0}} \mathrm{Re}\left[\Psi\left(\frac{1}{2} + i\frac{\hbar\omega}{2\pi k_B T}\right) - \ln\left(\frac{\hbar\omega}{2\pi k_B T}\right)\right] \tag{5}$$

and

$$\left(\frac{\delta f(T)}{f_0}\right)_{\mathrm{QP}} = -\frac{\alpha}{2}\left(\frac{|\sigma(0,\omega)|}{|\sigma(T,\omega)|}\sin\left(\phi(T,\omega)\right) - 1\right), \tag{6}$$

where Ψ is the complex digamma function; $|\sigma|$ is the amplitude of the complex conductivity; ϕ is the phase of the complex conductivity; and α is the kinetic inductance fraction (Appendix Sec. A 7). The three free fit parameters are $Q_{\mathrm{TLS},0}$, T_c, and α.

We compare the extracted $Q_{\mathrm{TLS},0}$ from the two measurements [Fig. 2(c)] and find that they agree on average to within a factor of 2.6σ. We note that our measurements are optimized for measuring Q_{int} rather than $\delta f/f_0$. As $\delta f/f_0$ is not sensitive to the applied microwave power, we have a factor of 5–10 fewer data points with which to fit $\delta f/f_0$ than Q_{int} while the fit function contains only four fewer free parameters, and, thus, the fit to the Q_{int} data corresponds to a narrower χ^2 distribution than the fit to the $\delta f/f_0$ data. We conclude that, while the difference between the two measurements is statistically significant, $Q_{\mathrm{TLS},0}$ is a robust parameter that forms a quantitative basis of comparison across devices when fitted from Q_{int} data.

We note that several samples show an earlier onset of thermal quasiparticles with temperature, which we hypothesize is due to microscopic patches of β-Ta (Appendix Sec. I), and in these samples the TLS saturation with temperature can be masked by the earlier onset of thermal quasiparticle loss. As $\delta f/f_0$ is only a function of temperature, this effect makes it difficult to extract $Q_{\mathrm{TLS},0}$, while the power dependence of the Q_{int} data allows a robust extraction of $Q_{\mathrm{TLS},0}$ even with significant thermal quasiparticle loss.

Using both the power and temperature dependence of Q_{int} gives us confidence that we have accounted for the major sources of loss and that the main contribution to loss at low temperature and microwave power is TLS loss. With this understanding in hand, we note that, in principle, a measurement scheme optimized to measure the frequency shift versus temperature may be as robust at extracting $Q_{\mathrm{TLS},0}$ as the fits to $Q_{\mathrm{int}}(\bar{n}, T)$ that we consider here.

III. PARAMETRIZING SOURCES OF LOSS

TLSs that cause loss can occur in many different materials in the same device: surface oxides, surface contamination, the exposed sapphire surface, the tantalum-sapphire interface, the bulk of the sapphire, and other elements related to packaging. In order to identify the location and origin of TLS loss, we fabricate 26 chips containing a total of 113 devices with varying geometry and surface conditions and perform temperature- and power-dependent loss measurements to extract $Q_{\mathrm{TLS},0}$. Varying the device geometry changes the SPR [Fig. 3(a)]. By modeling interfaces as dielectrics with an assumed standard thickness (3 nm) and permittivity ($\epsilon = 10$), we can compute the fraction of electric field energy that overlaps with the interfaces of a device for a given electromagnetic mode, the SPR [42] (Appendix Sec. C).

For the CPW resonators, we tune the SPR by tuning the pitch of the shorted CPW transmission line. Fixing the impedance to be 50 Ω constrains the ratio between the center pin width and the gap width [49], so that the center pin also increases in width as the SPR is reduced. For the LE resonators, we tune SPR by changing the spacing and size of the capacitor pads, where larger spacings and pads correspond to lower SPR. Across both types of devices, we vary the SPR by a factor of 30 (Appendix Sec. C).

The extracted $Q_{\mathrm{TLS},0}$ increases with decreasing SPR [Fig. 3(b)]. The trend is approximately linear and then plateaus for low SPR, below 3×10^{-4}. We model this SPR dependence as arising from two different components, a surface-related TLS loss that scales with SPR and a bulk TLS loss that is SPR independent. The losses can be

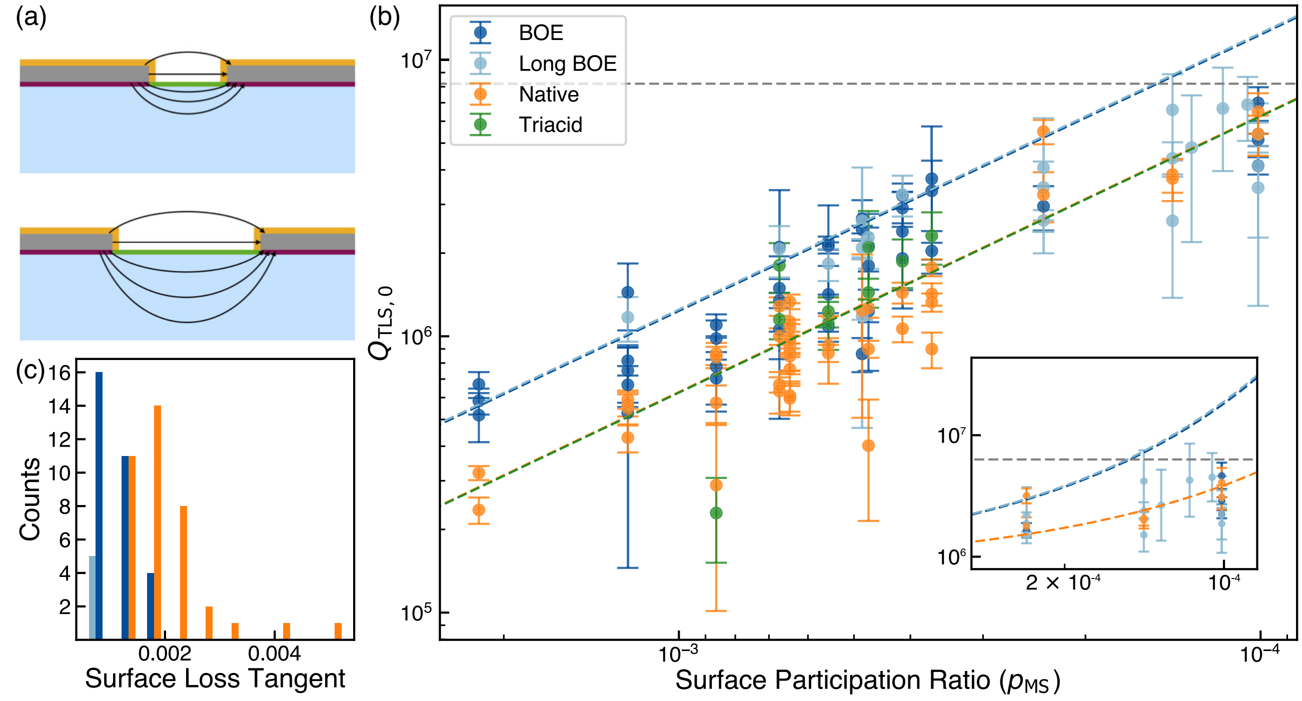

FIG. 3. Dependence of loss on SPR. (a) Cartoon cross section illustrating the dependence of SPR on device geometry. As the distance between capacitor pads (gray) increases, the fraction of the electric field (black arrows) energy overlapping with a thin layer at the three interfaces [metal-air (yellow), metal-substrate (purple), and substrate-air (green)] decreases. The fraction of the electric field energy in the sapphire substrate (blue) does not strongly depend on the distance between capacitor pads. (b) Dependence of the extracted $Q_{\mathrm{TLS},0}$ from Q_{int} measurements on SPR. For the highest SPR devices $Q_{\mathrm{TLS},0}$ exhibits linear scaling with SPR, but for lower SPR the $Q_{\mathrm{TLS},0}$ saturates, indicating that there are both surface and bulk TLS baths. Comparing the data for four different surface conditions allows for the estimate of surface loss tangents for each surface as well as the bulk substrate loss tangent. An inset is included in the bottom right showing the lowest SPR data plotted on a log-linear scale, which reveals the scale of the observed plateau for the lowest SPR devices. (c) Histogram of surface loss tangents for native, BOE, and long BOE CPW devices, showing that BOE and long BOE treatments result in surface loss tangents that are a factor of 1.96 times and 2.02 times, respectively, lower than the native surface on average.

parametrized as a loss tangent, which is the ratio of the imaginary and real components of the dielectric constant, $\tan\delta = \mathrm{Im}(\epsilon)/\mathrm{Re}(\epsilon)$. The apparent surface loss tangent varies across the four surface treatments. Fitting these two components to the full dataset across all 113 devices yields the surface and bulk loss tangents shown in the middle column in Table I. By contrast, for the largest devices, all devices have similar $Q_{\mathrm{TLS},0}$ and do not vary by SPR or by surface treatment, indicating that they are limited by a bulk loss component [Fig. 3(b), inset]. The fitted bulk loss tangent shown in Table I is an order of magnitude higher than recent bulk measurements on the same substrates [11]. This indicates that the bulk loss we observe is dominated by a surface damage layer within 50 μm of the surface rather than a uniform loss tangent throughout the bulk (Appendix Sec. D). We note that these hypothesized extended defects, while not evenly distributed throughout the depth of the substrate, are spread several orders of magnitude deeper than any disorder in the tantalum-sapphire interfacial region.

TABLE I. Loss tangents extracted from $Q_{\mathrm{TLS},0}$ (middle column) and $Q_{\mathrm{TLS}}(\bar{n}=1)$ data (right column).

Dielectric	tan δ	tan $\delta(\bar{n}=1)$
Native	$(15.9 \pm 0.7) \times 10^{-4}$	$(13.3 \pm 0.5) \times 10^{-4}$
BOE	$(8.1 \pm 0.6) \times 10^{-4}$	$(7.7 \pm 0.4) \times 10^{-4}$
Long BOE	$(8 \pm 1) \times 10^{-4}$	$(7 \pm 2) \times 10^{-4}$
Triacid	$(16 \pm 3) \times 10^{-4}$	$(13 \pm 1) \times 10^{-4}$
Bulk	$(1.3 \pm 0.2) \times 10^{-7}$	$(1.0 \pm 0.2) \times 10^{-7}$

While we etch our devices with three different processes (Appendix Sec. A 1), we observe that the surface loss tangent for a given surface treatment is robust when we stratify by etch type. This observation indicates that the dominant contribution to surface loss is the nature of the surface oxide or contamination at the surface (Appendix Sec. J 1) and not morphology of the tantalum edge.

We study the correlation between loss tangent and tantalum oxide thickness after the four different surface treatments to localize the source of surface-related TLS loss. The native oxide is an approximately 3-nm-thick, kinetically limited, stoichiometric oxide that is remarkably robust to chemical processing, as measured using x-ray photoelectron spectroscopy and transmission electron microscopy [15,19] (Appendix Sec. E). In Ref. [19], we observe a reduction in oxide thickness after BOE treatment to approximately 2.4 nm, while after triacid treatment, the oxide grows to nearly 6 nm. We correlate results from Ref. [19] with lab-based XPS results to estimate that the total oxide thickness after the long BOE treatment is 1.5 nm ± 0.3 nm (Appendix Sec. D). We note that, in Ref. [19], BOE treatments were found to not completely etch away the oxide layer, and, thus, the BOE and long BOE treatments do not affect the underlying tantalum film. The long BOE and BOE treated devices exhibit 2.02 and 1.96 times, respectively, higher $Q_{\mathrm{TLS},0}$ than the devices with a native surface [Figs. 3(b) and 3(c)]. Since the tantalum oxide layer is amorphous, and it is thinner after either BOE treatment, a likely hypothesis for the origin of TLS loss is the oxide layer. However, we observe that the triacid-treated samples have similar values of $Q_{\mathrm{TLS},0}$ to those of the native samples, despite their thicker oxide. Therefore, the TLS loss is not proportional to the volume of the oxide, possibly because another bath of TLSs decreases to compensate the additional oxide-related loss in the triacid-treated samples.

We hypothesize that the triacid and BOE treatments are highly effective at removing residual hydrocarbon contamination from fabrication, resulting in a reduction in the parasitic hydrocarbon TLS loss commensurate with the increased oxide loss from triacid treatment. Treatment in BOE removes some surface oxide, and, therefore, any contamination on that surface should be removed concurrently; similarly, the triacid treatment has previously been shown to be strongly oxidizing and effective at removing hydrocarbons [50]. As all devices are treated in 1:2 hydrogen peroxide to sulphuric acid, a self-heating mixture, and heated sulphuric acid etches sapphire [51], we assume that there is no significant fabrication residue remaining on the substrate-air interface. We, therefore, model the observed surface loss tangents for the three surface conditions as arising from three components: a hydrocarbon component at the metal-air interface, an oxide-related loss tangent that is proportional to the oxide thickness, and a component related to the metal-substrate and substrate-air interfaces (Appendix Sec. F):

$$\begin{aligned}\frac{1}{Q_{\mathrm{TLS},0}} &= p_{\mathrm{MA}} \tan \delta_{\mathrm{MA}} + p_{\mathrm{MS}} \tan \delta_{\mathrm{MS}} + p_{\mathrm{SA}} \tan \delta_{\mathrm{SA}} \\ &= p_{\mathrm{MA}} \left(\left(\frac{t_{\mathrm{TaOx},i}}{t_0} \right) \tan \delta_{\mathrm{TaOx}} + \gamma_i \tan \delta_{\mathrm{HC}} \right) \\ &\quad + p_{\mathrm{MS}} \tan \delta_{\mathrm{MS}} + p_{\mathrm{SA}} \tan \delta_{\mathrm{SA}}, \end{aligned} \tag{7}$$

where t_0 is the assumed thickness of the substrate-air and metal-substrate interfaces used to simulate participation ratios (3 nm), $t_{\mathrm{TaOx},i}$ is the oxide thickness of treatment i, $\gamma_i \in \{0, 1\}$ is a factor determining if hydrocarbon loss is considered for surface condition i, and the subscripts MA, SA, MS, and HC refer to metal-air, substrate-air, metal-substrate, and hydrocarbons, respectively.

To estimate the different components of the loss, we assume that the hydrocarbon loss is completely eliminated after any of the triacid or BOE treatments ($\gamma_i = 0$) and is present only for the native condition ($\gamma_{\mathrm{native}} = 1$). By quantitatively comparing the extracted loss tangents for our four conditions with Eq. (7), we calculate a putative rescaled hydrocarbon-related loss tangent for an assumed standard 3-nm-thick interface, $(p_{\mathrm{MA}}/p_{\mathrm{MS}}) \tan \delta_{\mathrm{HC}} = (5.9 \pm 0.6) \times 10^{-4}$, and an intrinsic loss tangent for the tantalum oxide, $\tan \delta_{\mathrm{TaOx}} = (5 \pm 1) \times 10^{-3}$. When rescaled to account for the difference in participation ratios of different surfaces, $(p_{\mathrm{MA}}/p_{\mathrm{MS}}) \tan \delta_{\mathrm{TaOx}} = (6 \pm 1) \times 10^{-4}$.

Assuming that the triacid and two BOE treatments do not affect the metal-substrate and substrate-air interfaces, our model also provides an estimate for the loss contributions of those two interfaces and finds that they give a combined value of $\tan \delta_{\mathrm{MS}} + (p_{\mathrm{SA}}/p_{\mathrm{MS}}) \tan \delta_{\mathrm{SA}} = (4 \pm 1) \times 10^{-4}$. The rescaled loss tangents are all comparable, indicating that all surfaces may play a critical role in determining overall loss. We note that the model we hypothesize can explain our data with the fewest number of assumptions about our system. Other models, which could include differing loss tangents from different tantalum oxide species or more complex spatial variations of hydrocarbons, involve making more assumptions than the model we propose. Our methodology illustrates that such models could be quantitatively tested by varying a larger number of surface processing parameters. We further note that some models for the effects of surface treatment can be ruled out from our data, as detailed in Appendix Sec. F.

Our data rule out a model for the TLS loss that is purely extensive in the oxide thickness. Here, we hypothesize that a second bath residing in fabrication-related contaminant hydrocarbons can account for the difference. However, there are other possible microscopic models, such as the possibility that a single chemical species or suboxide is responsible for all the TLS loss in the oxide, and the native and triacid samples have equal amounts of that species despite the large difference in total oxide thickness. Testing such hypotheses would require the measurement of many more surface conditions that independently vary each candidate TLS component.

While $Q_{\mathrm{TLS},0}$ parametrizes the linear absorption in the device, for transmon operation the steady state photon occupation is around $\bar{n} = 1$, and, thus, the relevant parameter is $Q_{\mathrm{TLS}}(\bar{n} = 1)$. Our parametrization allows us to calculate $Q_{\mathrm{TLS}}(\bar{n} = 1)$ and make quantitative statements about transmon operating conditions. For the largest devices at base temperature, $Q_{\mathrm{TLS}}(\bar{n} = 1)$ ranges from 5 to 15×10^6 (Fig. 4), in line with state-of-the-art qubits [15,16].

Examining the dependence of $Q_{\mathrm{TLS}}(\bar{n} = 1)$ on SPR reveals that BOE treatment also leads to an improvement.

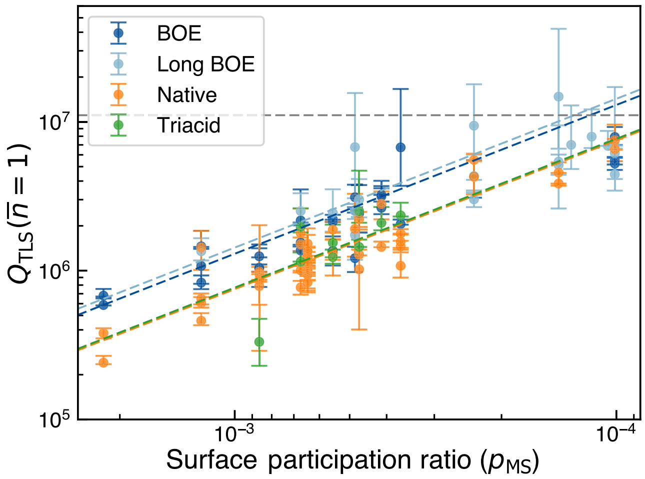

FIG. 4. Dependence of $Q_{\mathrm{TLS}}(\bar{n} = 1)$ on SPR. $Q_{\mathrm{TLS}}(\bar{n} = 1)$ is calculated at base temperature. Data from four different surface conditions are shown. While Q_{TLS} is a nonlinear function of $\bar{n}$, we nevertheless observe a roughly linear dependence between $Q_{\mathrm{TLS}}(\bar{n} = 1)$ and SPR. We fit an apparent loss tangent to each surface condition, as well as the bulk substrate loss tangent. Data are calculated from Eq. (7) with errors propagated from errors in fit parameters. The error bars are truncated at the lower end by $Q_{\mathrm{TLS},0}$.

The SPR dependence is roughly linear despite the nonlinear dependence of Q_{TLS} on the microwave photon number. By performing the same analysis on the SPR scaling of $Q_{\mathrm{TLS}}(\bar{n} = 1)$, we extract apparent loss tangents at single-photon powers for the four surface conditions, shown in the rightmost column in Table I. Again, for the largest devices, surface and bulk loss are comparable, indicating that improvements in both are required to improve overall device performance.

IV. CONCLUSION

We observe that state-of-the-art tantalum devices are limited by TLS loss. Using systematic measurements and parametrization of losses in superconducting circuits, we show that there are multiple sources of TLSs: a surface-related TLS bath associated with the tantalum oxide that can be reduced by around a factor of 2 with BOE treatment and a substrate-related TLS bath. Furthermore, the surface-related TLS loss is not extensive in the oxide volume, indicating that there may be at least one additional TLS bath, such as fabrication-related hydrocarbon contamination. Critically, each of these components is of similar magnitude for state-of-the-art devices, and future improvements in superconducting qubits will require material improvements that address all of these sources of loss. Two natural avenues to pursue based on our findings would be to passivate the Ta surface to avoid oxide formation entirely and to study subsurface damage from polishing and surface processing in sapphire substrates.

The observed temperature dependence also points to two paths for improving the performance of superconducting qubits: reducing the density of TLSs and improving the coherence time of the TLS bath. Our ongoing work includes studying the dynamics of the TLS bath using pump-probe spectroscopy [36] and other time-domain methods [52].

Correlations between the measurements presented here and direct materials spectroscopy may identify atomistic origins of TLS loss. For example, the losses in the tantalum oxide could arise from particular suboxides or interface states, and detailed chemical profiling using x-ray photoelectron spectroscopy could elucidate the particular chemical species responsible for TLS loss [19,53]. More broadly, the parametrization presented here isolates and identifies the material-related loss, thereby enabling quantitative comparisons among different material systems, such as new superconducting metals [54] and metal heterostructures, alternative substrates such as high-purity silicon [55], and different fabrication and postprocessing techniques.

ACKNOWLEDGMENTS

We thank Rob Schoelkopf, Suhas Ganjam, Michel Devoret, Steve Girvin, Will Oliver, Chen Wang, Mike Hatridge, and Jeff Thompson for helpful discussions. This work was primarily supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Co-design Center for Quantum Advantage (C2QA) under Contract No. DE-SC0012704. Film characterization and processing was partly supported by the National Science Foundation (RAISE DMR-1839199). The authors acknowledge the use of Princeton's Imaging and Analysis Center (IAC), which is partially supported by the Princeton Center for Complex Materials (PCCM), a National Science Foundation (NSF) Materials Research Science and Engineering Center (MRSEC; DMR-2011750), as well as the Princeton Micro/Nano Fabrication Laboratory. We also acknowledge MIT Lincoln Labs for supplying a traveling wave parametric amplifier.

APPENDIX A: SUPPLEMENTARY EXPERIMENTAL METHODS

1. Sample fabrication

3” diameter sapphire substrates are cleaned in a 2:1 $H_2SO_4:H_2O_2$ piranha solution for 20 min, then rinsed in 3 cups of deionized water followed by 1 cup of 2-propanol, and then blow dried in N_2. Then, the sapphire substrate is loaded into a dc magnetron sputtering system (AJA Orion 8). The substrate is heated *in situ* at 850°C before tantalum sputtering. The film deposition parameters are as follows: rf power of 250 W, Ar flow rate of 30 sccm, ambient pressure, temperature ramp rate 1°C per minute, and steady state temperature of 750° C, which results in a film growth rate of approximately 8 nm per minute. Postdeposition, the tantalum films are confirmed to be predominantly ⟨111⟩ orientation in the α phase using a Bruker D8 Advance x-ray diffractometer. The deposited tantalum film is dehydration baked at 110° C and then cooled for about a minute on a metal plate. Following this, AZ1518 is spun on at 4000 rpm for 45 s with a ramp rate rate of 1000 rpm/s for an approximate resist thickness of 3 μm and soft baked at 95° C for 1 min. The photoresist is patterned using a Heidelberg DL66+ laser writer with a 1.8 μm spot size with a 50% attenuator, intensity setting of 30%, and focus offset setting of 10%. The photoresist is developed in AZ300MIF solution for 90 s and rinsed in deionized water for 30 s. After development, the mask is hard baked at 110° C for 2 min and then cooled on a metal plate for 1 min.

Using the patterned photoresist as a mask, we etch each device with one of three different etch types. One type is a wet chemical etch, 1∶1∶1 ratio of HF:HNO_3:H_2O (Transene Tantalum Etchant 111), in which a sample is swirled for 21 s before being rinsed in 3 cups of deionized water and 1 cup of 2-propanol and then blow dried in N_2. The second etch type is a chlorine-based dry chemical etch in an inductively coupled plasma reactive ion etcher (PlasmaTherm Takachi). The etching parameters for the chlorine dry etch are as follows: ambient pressure of 5.4 mTorr, chlorine flow rate of 5 sccm, argon flow rate of 5 sccm, rf power of 500 W, and bias power of 50 W, which results in an etch rate of approximately 100 nm/ min. The third etch type is a fluorine-based dry etch, using the same reactive ion etcher as the chlorine etch, with parameters: ambient pressure 50 mTorr, CHF_3 flow rate 40 sccm, SF_6 flow rate 15 sccm, Ar_3 flow rate 10 sccm, rf power of 100 W, and bias power of 100 W. We note small variations in the sidewall angle (Fig. 5).

After etching, the photoresist mask is stripped in a Remover PG bath at 80° C for 1 h followed by rinsing in 2-propanol. The patterned Ta film is coated with hard-baked AZ1518 using the same parameters mentioned above to act as a protective layer for dicing. The wafer is diced (Advanced Dicing Technologies proVectus 7100 dicing saw) into 10- or 7-mm pieces, depending on the packaging used in the dilution refrigerator. Following dicing, the photoresist is stripped in a Remover PG bath at 80° C for 1 h, followed by 2 min each sonication in toluene, acetone, and 2-propanol. Some chips are sonicated in methanol for 2 min between the acetone and 2-propanol sonication to remove zinc contamination. The chips are blow dried in N_2 and then cleaned in a 2∶1 H_2SO_4∶H_2O_2 piranha solution for 20 min followed by rinsing in 3 cups of deionized water and 1 cup of 2-propanol and then blow dried in N_2. We note that fabrication residue is visible with scanning electron microscope images before this final piranha treatment but is removed afterward (Fig. 5).

After fabrication, the samples are treated in BOE or triacid as detailed in Sec. A 2. Then the chips are bonded to a printed circuit board (PCB) using an automatic wire bonder (Questar Q7800). We use two types of packages for our resonator chips. One comprises a Cu-plated PCB and a Cu puck and penny coated with 1 μm aluminum. The second comprises a commercial microwave package (QDevil QCage.24) with an associated Au-plated PCB. The mounting of the package to the dilution refrigerator is described in Sec. A 3.

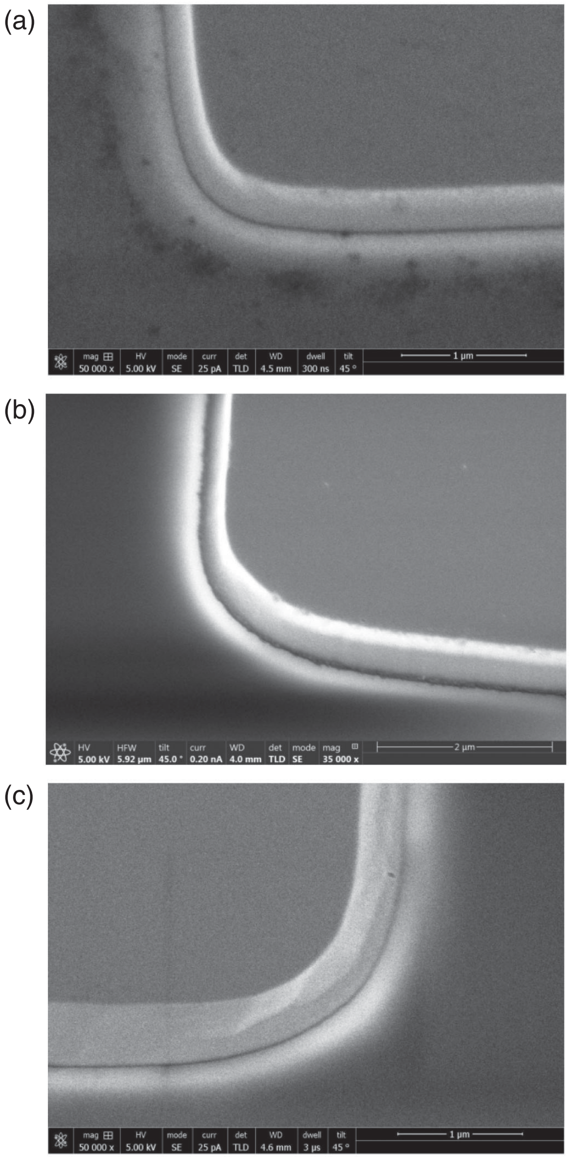

FIG. 5. Scanning electron microscope images of (a) a wet-etched device before piranha treatment, (b) a wet-etched device after piranha treatment, and (c) a chlorine-etched device after piranha treatment.

2. Surface processing

10:1 buffered oxide etch (BOE, Transene) is a mixture of ten parts 40% NH_4F solution to one part 49% HF solution by volume. BOE treated samples are placed in buffered oxide etch at room temperature and are not agitated.

After 20 min (BOE treatment) or 120 min (long BOE treatment), the samples are removed and triple rinsed in deionized water and 2-propanol before being blow dried in N_2.

The triacid treatment is 1∶1∶1 equal mix by volume of 95%–98% H_2SO_4, 70% HNO_3, and 70% $HClO_4$ solutions (all percentages by weight). We procure all solutions from Sigma Aldrich (catalog numbers H_2SO_4 - 258105, HNO_3 - 225711, and $HClO_4$ - 244252). The mixture is refluxed for 2 h and then allowed to cool for 1 h. After cooling, the sample is removed and triply rinsed in deionized water and 2-propanol before being blow dried in N_2.

3. Measurement apparatus

All devices are measured in a BlueFors XLD dilution refrigerator with a base mixing chamber temperature of approximately 17 mK. There are four independent input lines and four corresponding output lines. A fridge diagram showing the layout for all four input and output lines is given in Fig. 6.

Each input line has between 60 and 85 dB of attenuation from discrete cryogenic XMA attenuators (above mixing chamber, PN: 2082-604X-dB-CRYO) and cryogenic attenuators from Quantum Microwave (at mixing chamber, PNs: QMC-CRYOATTF-06 and QMC-CRYOATTF-03), as well as attenuation from stainless-steel coaxial transmission line cables, subminiature version A (SMA) connections, and insertion losses from filters. The total input line attenuation varies across the lines from 86.7 to 108.7 dB at resonator frequencies. Two types of low pass filter are used at the mixing chamber, a commercial filter from K&L Microwave (PN: 6L250-00089) outside of the magnetic shield, and an eccosorb filter placed inside of the magnetic shield. Two types of magnetic shield are used across our experiments. One type is a custom-fabricated can made of mu-metal, with which we used a custom-made eccosorb filter with upper cutoff frequency approximately 8 GHz, and the other is a prototype product (QCan) from QDevil with which we use an eccosorb filter supplied by QDevil with a similar passband.

Each output line contains filters, isolators, and amplifiers. At the mixing chamber, we use an eccosorb filter, with part number matching that on the input line and the same K&L filter as the input line. Two isolators are placed in series, both from QuinStar Technology (QCI-075900XM00). At the 4 K stage, a high-electron mobility transistor (HEMT) amplifier is used (Low Noise Factory LNF-LNC4_8F). Superconducting NbTi wire is used between the isolators and the HEMT to reduce signal attenuation. Additional filters are sometimes placed in the output line, with passbands which contain all resonators that are being measured.

Several devices use a traveling wave parametric amplifier (TWPA) sourced from the MIT Lincoln Laboratory. The TWPA is placed in a separate magnetic shield at the mixing chamber and placed in the signal path immediately after the second isolator. The TWPA is pumped using a separate input line.

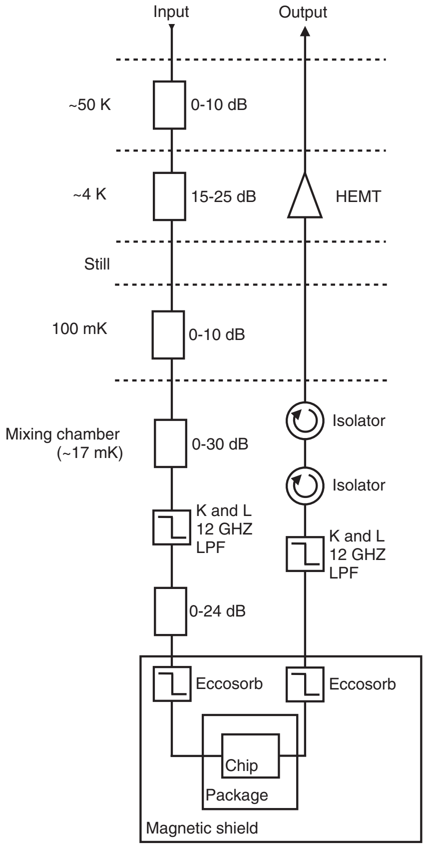

FIG. 6. Wiring diagram for each of our measurement lines. Ranges of attenuation are given where the attenuation varies from line to line. The magnetic shields in our experiments vary between a QCan supplied by QDevil and a custom-made mu-metal can. A traveling wave parametric amplifier (TWPA) is sometimes used on the output line in our experiments and is placed in a separate magnetic shield and wired after the second isolator.

Outside of the fridge, we use additional amplifiers on the output line. We use amplifiers from Mini-Circuits (PNs: CMA-83LN+ and ZVE-3W-183+), RF-Lambda (PN: RLNA02G08G30), and Miteq (PN: AFS4-00101200-18-10P-4). The configuration of these amplifiers varies between experiments, but approximately 50 dB of gain is used in all cases.

All measurements are conducted with a vector network analyzer from Keysight (PNA-X Network Analyzer N5241A). For most experiments, the measurement parameters are 201 points across the frequency axis, IF bandwidth of 30 Hz, and an integration time per resonator varying from 1 min (high power) to 3 h (low power). Integration times are adjusted for each resonator chip, and measurement parameters differ slightly for early experiments. The frequency span of the sweeps is adjusted at each temperature to be 5 times the linewidth of a trace measured with a power near the middle of the power range. As the

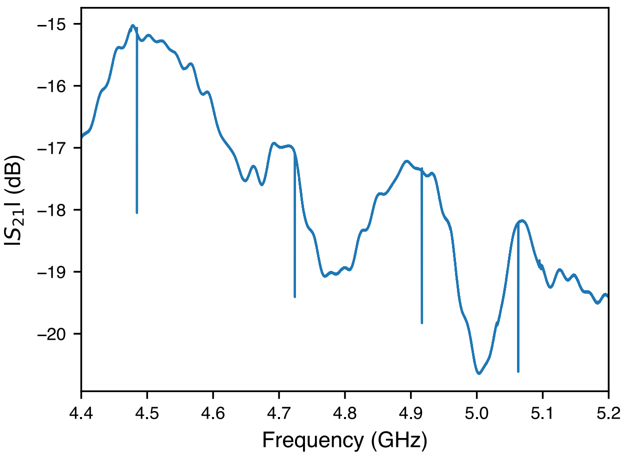

FIG. 7. Wide frequency transmission sweep of a chip with four resonators coupled to a single feedline. The four sharp dips correspond to the location of the four resonators.

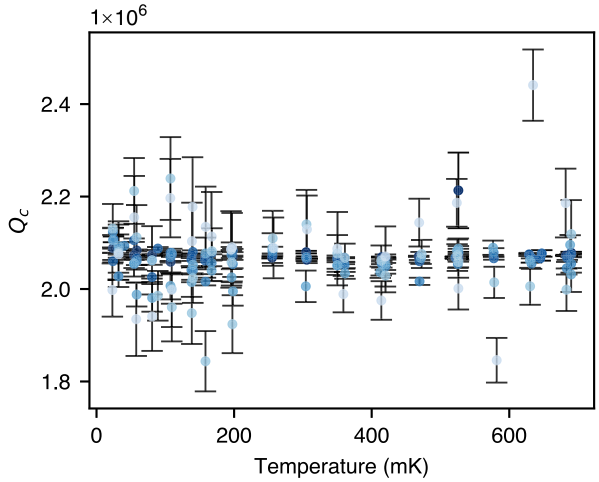

FIG. 8. Fitted Q_c parameters for a power and temperature sweep. This dataset corresponds to the same power and temperature sweep shown in Fig. 2(a) in the main text. Colors correspond to power applied at the input, with the highest power being the darkest shade and the lowest power being the lightest shade. All powers are spaced 10 dB apart.

quality factor decreases with decreasing power, some traces at the lowest power are excluded due to insufficiently broad frequency axes.

4. Resonator spectroscopy

Resonators are easily located in frequency space due to their high quality factor relative to all other features. Figure 7 shows a wide frequency sweep of a chip with four resonators coupled to a feedline. The wide frequency ripples may be caused by standing waves or reflections from connections on our measurement setup; however, given that the width of these ripples is on the order of 10 MHz and the width of the resonators is on the order of 1 kHz, we ignore these ripples and assume that a flat background exists when measuring each resonator.

5. Measuring Q_{int}

We used the following model to fit each resonator trace, such as the one shown in Fig. 1(c) [43]:

$$|S_{21}(\omega_{\text{probe}})| = \left|1 - \frac{\frac{Q_{\text{tot}}}{Q_c} - \frac{2iQ_{\text{tot}}\alpha}{\omega_0}}{1 + 2iQ_{\text{tot}}\frac{\omega_{\text{probe}} - \omega_0}{\omega_0}}\right| + S_{21,\text{baseline}}, \tag{A1}$$

where $|S_{21}|$ is the magnitude of the transmission through the feedline, Q_c is the coupling quality factor, Q_{tot} is the total quality factor ($Q_{\text{tot}}^{-1} = Q_{\text{int}}^{-1} + Q_c^{-1}$), α is the asymmetry of the resonator, ω_0 is the center angular frequency of the resonator, ω_{probe} is the angular frequency of the probe tone, and $|S_{21,\text{baseline}}|$ is the transmission through the feedline when no resonator is present. We assume that $|S_{21,\text{baseline}}|$ is a constant, which is approximately correct for resonators with a small linewidth. The derivation of this model is given in the appendix in Ref. [43] with a minimal assumption set.

The coupling quality factor Q_c parametrizes the loss from the resonator to the feedline. In order to characterize our material losses, we must be able to separately determine Q_c and Q_{int} across an entire temperature and power sweep. The value of Q_c is determined by the capacitive or inductive coupling between each resonator and the feedline, and, therefore, we expect Q_c to be independent of power and temperature.

In our analysis, each resonator $|S_{21}|$ trace is fit independently. To show that we can separately extract Q_c and Q_{int} from the same $|S_{21}|$ trace, we examine the fitted values of Q_c for each temperature sweep. We find that our fitted values of Q_c are constant across power and temperature, and so we conclude that we have extracted an accurate value of Q_c and, therefore, of Q_{int}. An example plot of Q_c versus power and temperature, corresponding to the same sweep shown in Fig. 2(a), is shown in Fig. 8.

To balance signal-to-noise ratio with sensitivity to Q_{int}, we design our resonators to match Q_C approximately equal to Q_{int}. As Q_{int} varies by orders of magnitude across the power and temperature sweeps, we compare our fitted values of $Q_{\text{TLS},0}$ (the approximate lower bound on measured Q_{int}) and Q_{other} (the upper bound on measured Q_{int}) in Fig. 9. As Q_C lies approximately halfway between $Q_{\text{TLS},0}$ and Q_{other} for all resonators, we conclude that our coupling magnitude is well designed for our experiment.

6. Nonlinear behavior at high microwave power

When measuring the highest powers, we occasionally are unable to fit a resonator trace (Fig. 10), which we attribute to nonlinear behavior of the resonator. Potential sources of this nonlinearity are the saturation of an amplifier or an effect of the superconducting state such as the nonlinear kinetic inductance of Cooper pairs [56].

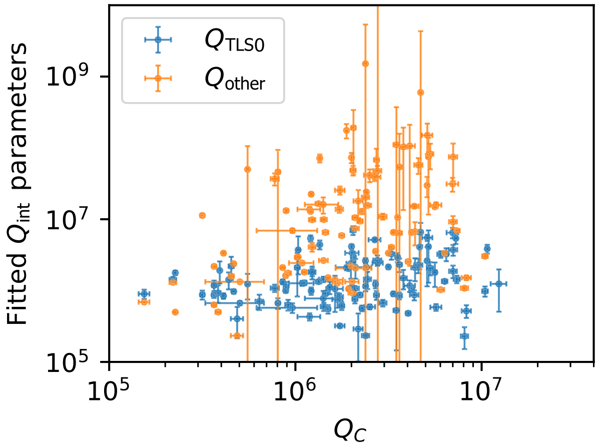

FIG. 9. $Q_{\rm TLS,0}$ and $Q_{\rm other}$ versus Q_C. The value and uncertainty for Q_C plotted is the mean and standard deviation of fitted Q_C values for all $|S_{21}|$ traces of a given resonator.

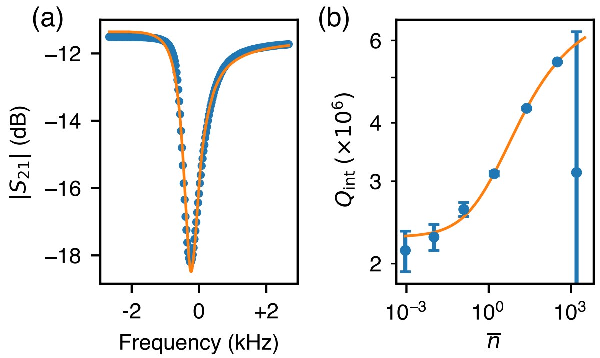

FIG. 10. (a) An example of high-power resonator transmission trace with observable nonlinearity. The best fit line to the data using Eq. (A1) is shown in orange. This trace is excluded from our analysis. (b) Power sweep showing $Q_{\rm int}$ for the nonlinear trace in (a) at the maximum $\bar n$. The orange line is a fit to the data with the nonlinear data point excluded.

As we are most concerned with the behavior of our devices at low power, we exclude traces showing the nonlinear behavior from our analysis.

7. Model for $Q_{\rm TLS}$

It can be shown that the loss induced by an ensemble of TLSs coupled to an electromagnetic mode takes the form [20]

$$\frac{1}{Q_{\rm TLS}(\bar n, T)} = \frac{1}{Q_{\rm TLS,0}}\left(\frac{\tanh(\frac{\hbar\omega}{2k_BT})}{\sqrt{1+\frac{\bar n}{n_c}}}\right), \qquad \text{(A2)}$$

where $\bar n$ is the average photon number in the mode, T is the temperature of the mode-ensemble system, ω is the frequency of the mode, and n_c is the critical photon number of the ensemble. In order to obtain the model we use to fit the TLS component of our $Q_{\rm int}$ data, we make a few substitutions. First, it can be shown that $n_c \propto (T_1T_2)^{-1}$, where T_1 and T_2 are the average relaxation and decoherence times, respectively, of the ensemble. Second, the average T_1 of the ensemble can be shown to follow a thermal distribution [20]:

$$T_1 \propto \tanh\left(\frac{\hbar\omega}{2k_BT}\right). \qquad \text{(A3)}$$

Third, TLS-TLS interactions can be modeled as state changes in one TLS causing dephasing in neighboring TLSs. As the temperature is reduced, thermal fluctuations in the states of the TLSs in the ensemble reduce as more and more members of the ensemble occupy the ground state. We, therefore, expect an inverse relationship between the TLS coherence time T_2 and temperature, which we model as [44]

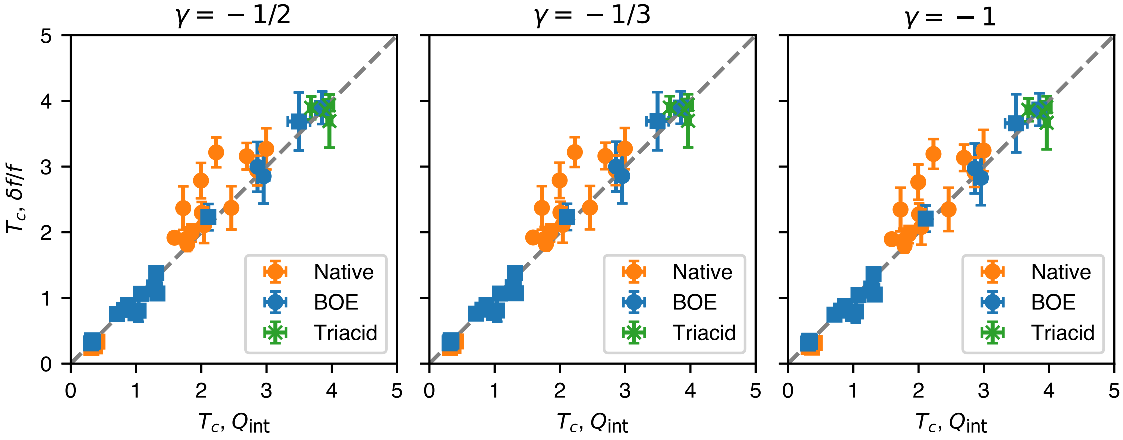

FIG. 11. Plots of T_c estimates provided by the frequency shift and $Q_{\rm int}$ fitting methods for the three different quasiparticle frequency shift regimes.

$$\frac{1}{T_2} \propto k_B T^{\beta_1}, \tag{A4}$$

where β_1 is an empirical parameter. Finally, different mode shapes overlap with and saturate the ensemble differently as they are populated with increasing numbers of photons, and we account for this by introducing another empirical fit parameter, β_2 [45]:

$$\bar{n} \to \bar{n}^{\beta_2}. \tag{A5}$$

Putting all of these substitutions together gives our TLS loss model:

$$\frac{1}{Q_{\mathrm{TLS}}(\bar{n},T)} = \frac{1}{Q_{\mathrm{TLS},0}} \left(\frac{\tanh(\frac{\hbar\omega}{2k_BT})}{\sqrt{1+(\frac{n^{\beta_2}}{DT^{\beta_1}})\tanh(\frac{\hbar\omega}{2k_BT})}} \right). \tag{A6}$$

The model we use to fit quasiparticle losses has been discussed in other works [57].

8. Model for frequency shift

The model we use to fit the frequency shift is

$$\frac{\delta f(T)}{f_0} = \left(\frac{\delta f(T)}{f_0}\right)_{\mathrm{TLS}} + \left(\frac{\delta f(T)}{f_0}\right)_{\mathrm{QP}}, \tag{A7}$$

where

$$\left(\frac{\delta f(T)}{f_0}\right)_{\mathrm{TLS}} = \frac{1}{\pi Q_{\mathrm{TLS},0}} \mathrm{Re}\left[\Psi\left(\frac{1}{2} + i\frac{\hbar\omega}{2\pi k_B T}\right) - \ln\left(\frac{\hbar\omega}{2\pi k_B T}\right)\right] \tag{A8}$$

is the TLS contribution to the frequency shift and

$$\left(\frac{\delta f(T)}{f_0}\right)_{\mathrm{QP}} = -\frac{\alpha}{2}\left(\left|\frac{\sigma(0,\omega)}{\sigma(T,\omega)}\right| \sin\left(\phi(T,\omega)\right) - 1\right) \tag{A9}$$

is the quasiparticle contribution to the frequency shift. In Eqs. (A8) and (A9), Ψ is the complex digamma function; σ_1 and σ_2 are the real and imaginary parts of the complex conductivity; ϕ is the phase between the real and imaginary parts of the complex conductivity; and α is the kinetic inductance fraction. The derivation of the TLS contribution can be found in, for example, Ref. [20]. The expression for the quasiparticle contribution is based on Ref. [20] but is not explicitly stated, so we derive it in detail below.

The frequency shift from quasiparticles is defined as

$$\frac{\delta f(T)}{f_0} = -\frac{\alpha}{2}\left(\frac{X_S(T) - X_S(0)}{X_S(0)}\right), \tag{A10}$$

where X_S is the imaginary part of the surface impedance of the superconductor, otherwise known as the reactance. In general, the surface impedance has a cumbersome form, but in three superconducting material limits it takes the simpler form

$$Z_s(T) = A\sigma(T)^\gamma, \tag{A11}$$

where A is a constant prefactor, $\sigma(T)$ is the superconducting complex conductivity, and γ is a parameter that takes a different value depending on which of the three limits the superconductor is in. As the quasiparticles of interest in our system have a thermal distribution, the complex conductivity takes the form

$$\sigma(T) = \sigma_1(T) + i\sigma_2(T), \tag{A12}$$

where

$$\frac{\sigma_1(T)}{\sigma_n} = \frac{4\Delta_0}{\hbar\omega} e^{-\Delta_0/k_BT} \sinh\left(\frac{\hbar\omega}{2k_BT}\right) K_0\left(\frac{\hbar\omega}{2k_BT}\right) \tag{A13}$$

and

$$\frac{\sigma_2(T)}{\sigma_n} = \frac{\pi\Delta_0}{\hbar\omega}\left[1 - \sqrt{\frac{2\pi k_BT}{\Delta_0}} e^{-\Delta_0/k_BT} - 2e^{-\Delta_0/k_BT} e^{-\hbar\omega/2k_BT} I_0\left(\frac{\hbar\omega}{2k_BT}\right)\right]. \tag{A14}$$

In the above equations, $\hbar$ is the reduced Planck constant, k_B is the Boltzmann constant, ω is the resonator center angular frequency, T is the temperature, $\Delta_0 = 1.764 k_B T_c$ is the superconducting gap, T_c is the superconducting critical temperature, I_0 is the zeroth-order modified Bessel function of the first kind, K_0 is the zeroth-order modified Bessel function of the second kind, and σ_n is the normal-state conductivity of the superconductor just above T_c. The surface impedance can be rewritten in a more convenient form using these quantities:

$$\begin{aligned} Z_s(T) &= A(\sigma_1(T) + i\sigma_2(T))^\gamma \\ &= A(|\sigma(T)|e^{i\phi(T)})^\gamma \\ &= A(|\sigma(T)|^\gamma e^{i\gamma\phi(T)}) \\ &= A|\sigma(T)|^\gamma(\cos(\gamma\phi(T)) + i\,\sin(\gamma\phi(T))), \end{aligned} \tag{A15}$$

where

$$|\sigma(T)| = \sqrt{\sigma_1(T)^2 + \sigma_2(T)^2} \tag{A16}$$

and

$$\phi(T) = \arctan\left(\frac{\sigma_2(T)}{\sigma_1(T)}\right). \tag{A17}$$

The reactance is then

$$X_S(T) = \mathrm{Im}[Z_S(T)] \\ = A|\sigma(T)|^{\gamma}(\sin(\gamma\phi(T))). \quad \text{(A18)}$$

The frequency shift can then be written in terms of the complex conductivities:

$$\frac{\delta f(T)}{f_0} = -\frac{\alpha}{2}\left(\frac{X_S(T) - X_S(0)}{X_S(0)}\right) \\ = -\frac{\alpha}{2}\left(\frac{\sin(\gamma\phi(T))}{\sin(\gamma\pi/2)}\left|\frac{\sigma(T)}{\sigma(0)}\right|^{\gamma} - 1\right). \quad \text{(A19)}$$

There are three possible values of γ depending on the electron mean free path (ℓ), coherence length (ξ_0), film thickness (d), and London penetration depth (λ_{LO}) of the superconductor [20]:

$$\gamma = -1/3, \text{thick film, extreme anomalous limit} \\ (\xi_0 \gg \lambda_{\mathrm{LO}} \text{ and } \ell \gg \lambda_{\mathrm{LO}}), \\ \gamma = -1/2, \text{thick film, dirty limit} \\ (\ell \ll \xi_0 \text{ and } \ell \ll \lambda_{\mathrm{LO}}), \\ \gamma = -1, \text{thin film, dirty limit} \\ (d \sim \ell \ll \xi_0 \text{ and } d \sim \ell \ll \lambda_{\mathrm{LO}}). \quad \text{(A20)}$$

However, we do not directly measure the relevant parameters to determine whether or not we are in any of these three regimes, so instead we fit our data to all three to see if the assumed regime makes a difference to the outcome of the fit. The results of this analysis are shown in Fig. 11, where the consistency between the superconducting critical temperature T_c estimated by our Q_{int} fits and our frequency shift fits is plotted for all three values of γ.

As can be seen, the assumed regime makes no difference to the fit outcome, so we choose to work in the thin film local limit ($\gamma = -1$) for our quasiparticle frequency shift fits:

$$\left(\frac{\delta f(T)}{f_0}\right)_{\mathrm{QP}} = -\frac{\alpha}{2}\left(\left|\frac{\sigma(0,\omega)}{\sigma(T,\omega)}\right| \sin(\phi(T,\omega)) - 1\right). \quad \text{(A21)}$$

9. Comparison to base temperature power sweeps

We compare the extracted $Q_{\mathrm{TLS},0}$ using our method to the value obtained from merely measuring the power dependence of Q_{int} at base temperature. In order to extract $Q_{\mathrm{TLS},0}$ at a single temperature point, we fit our base temperature data to

$$\frac{1}{Q_{\mathrm{int}}} = \frac{1}{Q_{\mathrm{TLS}}(\bar{n})} + \frac{1}{Q_{\mathrm{other}}}, \quad \text{(A22)}$$

where

$$Q_{\mathrm{TLS}}(\bar{n}) = Q_{\mathrm{TLS},0}\left(\frac{\sqrt{1 + (\frac{\bar{n}}{n_c})^{\beta}}}{\tanh(\frac{\hbar\omega}{2k_B T})}\right). \quad \text{(A23)}$$

In the above equation, $Q_{\mathrm{TLS},0}$ is the linear absorption from TLSs, $\bar{n}$ is the mean photon number in the resonator, n_c is the critical photon number of the TLS ensemble, $\hbar$ is the reduced Planck constant, ω is the angular frequency of the photons in the resonator, and T is the temperature of the mode-ensemble system.

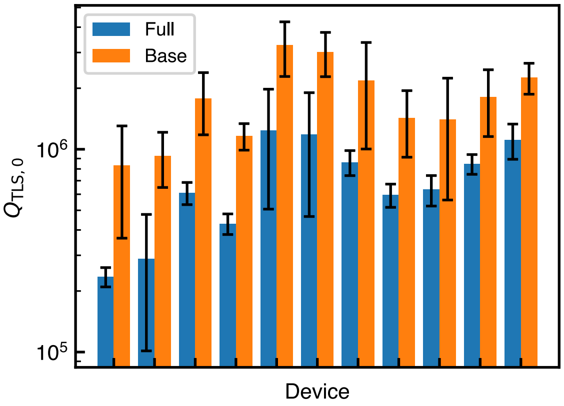

FIG. 12. $Q_{\mathrm{TLS},0}$ extracted from a full temperature and power sweep (blue) and a base temperature power sweep (orange) for a random selection of devices. The base temperature power sweep systematically overestimates $Q_{\mathrm{TLS},0}$.

Figure 12 shows $Q_{\mathrm{TLS},0}$ extracted by both methods. We find that, while the model used to fit the base temperature power sweeps contains fewer free fit parameters, the fractional size of the error bars on the associated $Q_{\mathrm{TLS},0}$ values are 24% larger on average than those associated with the full power and temperature sweeps. Furthermore, the base temperature extraction provides a systematically larger estimate of $Q_{\mathrm{TLS},0}$ than the full power and temperature extraction. The origin of this discrepancy is the linear decrease in Q_{int} with increasing temperature near base temperature, which arises from decreasing TLS coherence and correspondingly increasing TLS saturation power [for example, as can be seen in Fig. 2(a)].

In order to understand the effect of the tendency of base temperature power sweeps to overestimate $Q_{\mathrm{TLS},0}$ on our extracted quantities, we perform the same analysis we report in the main text to extract surface and bulk loss tangents, restricting the data to only using the base temperature datasets. The results of this analysis are shown in Table II alongside the results from the full temperature sweep dataset. As can be seen, the reduced dataset typically underestimates the surface loss tangent and, more importantly, does not capture the relative values among surface conditions. In particular, the native and BOE treated

TABLE II. Loss tangents extracted from $Q_{\mathrm{TLS},0}$ fitted using the full temperature sweep data set (middle column) and $Q_{\mathrm{TLS},0}$ fitted using the base temperature sweep dataset (right column).

Dielectric	$\tan\delta$ (full)	$\tan\delta$ (base)
Native	$(15.9 \pm 0.7) \times 10^{-4}$	$(9.1 \pm 0.8) \times 10^{-4}$
BOE	$(8.1 \pm 0.6) \times 10^{-4}$	$(9.1 \pm 0.4) \times 10^{-4}$
Long BOE	$(8 \pm 1) \times 10^{-4}$	$(4.2 \pm 0.3) \times 10^{-4}$
Triacid	$(16 \pm 3) \times 10^{-4}$	$(6 \pm 2) \times 10^{-4}$
Bulk	$(1.3 \pm 0.2) \times 10^{-7}$	$(1.5 \pm 0.1) \times 10^{-7}$

samples have equal surface losses using the restricted dataset, because using the base temperature sweeps alone effectively convolves the linear absorption and the saturation parameter, making it difficult to make quantitative conclusions about underlying material differences.

APPENDIX B: RESONATOR DESIGN

1. CPW resonators

Our CPW resonators are quarter-wave resonators constructed by shorting one end of a transmission line. Our design sets the characteristic impedance (Z_0) of the resonators to 50 Ω, which dictates a relationship between the center pin width and the gap width [49]. This means that if the distance between the center pin and ground plane (pitch) of the resonator is specified, the center pin width is fully constrained. The resonators are designed to have resonance frequencies between 6 and 8 GHz, where the resonance frequency is dictated by

$$f_0 = \frac{v}{4l\sqrt{\epsilon_{\mathrm{eff}}}}, \tag{B1}$$

where ϵ_{eff} is the effective dielectric constant defined in Ref. [49], l is the length of the resonator, and v is the speed of electric field propagation down the transmission line. We generally assume $v = c$, where c is the speed of light in vacuum.

When designing both the LE and CPW resonators, we aim to have the coupling rate of the resonator to the feedline ($1/Q_c$, sometimes written as $1/Q_{\mathrm{ext}}$) be equal to the expected internal loss rate of the resonator (Q_{int}). If Q_c is too small, the measurement is not sensitive to changes in Q_{int}; and if Q_c is too large, the photon lifetime in the resonator is short and signal-to-noise (SNR) decreases. Our CPW resonators are capacitively coupled to the feedline, and we compute this coupling using an equation from Ref. [58]:

$$C_c = \sqrt{\frac{\pi}{4Q_c}}\frac{1}{2\pi f_0 Z_0}, \tag{B2}$$

where C_c is the capacitance between the center pin and the feedline. We compute this capacitance using finite element analysis (Ansys Maxwell 3D) and generally find good agreement between predicted and measured resonance frequencies and external loss rates.

2. LE resonators

The LE resonators consist of a meander inductor in series with a dipole capacitor, with a resonance frequency given by

$$f_0 = \frac{1}{\sqrt{LC}} \tag{B3}$$

and an impedance given by

$$Z_0 = \sqrt{\frac{L}{C}}, \tag{B4}$$

where L and C are the inductance of the inductor and the capacitance of the capacitor, respectively. We account for stray capacitance across the inductor by modeling a stray capacitor in parallel with the lumped inductor. Therefore, the total capacitance in the resonator is the sum of the lumped capacitance C_L and the stray capacitance C_S, $C_{\mathrm{tot}} = C_L + C_S$. The resonance frequency and impedance are then

$$f_0 = \frac{1}{\sqrt{(C_L + C_S)L}} \tag{B5}$$

and

$$Z_0 = \sqrt{\frac{L}{C_L + C_S}}. \tag{B6}$$

For a given resonator design, we compute these three unknowns using three separate simulations. The first is a capacitance simulation in Ansys Maxwell 3D of only the dipole capacitor pads. We take the modeled capacitance to be equal to the lumped capacitance C_L. The second and third simulations are Ansys high frequency structure simulator (HFSS) eigenmode simulations of the meander and the full resonator. The resonance frequency of the meander can be written as

$$f_{0,\mathrm{meander}} = \frac{1}{\sqrt{LC_S}}, \tag{B7}$$

and the resonance frequency of the full resonator is given by Eq. (B5). With C_L, $f_{0,\mathrm{meander}}$, and $f_{0,\mathrm{resonator}}$ calculated from the three simulations, the remaining unknowns (C_S and L) and the fundamental resonator parameters (f_0 and Z_0) can be computed.

The external coupling rate of the LE resonators is determined empirically by cooling down an initial design and then adjusting the distance from the feedline to better match the external coupling rate to the internal loss rate. The distance from the feedline is adjusted by assuming the coupling would fall off proportional to $1/r^3$, as the coupling is inductive. After this initial cooldown, finer adjustments are made for subsequent designs, but, in general, the external coupling rate of the LE resonators matches the internal loss rate, which is the aforementioned condition for optimizing both SNR and sensitivity to changes in Q_{int}.

APPENDIX C: SPR CALCULATIONS

For both the CPW and LE resonators, the SPRs reported in the main text are computed by simulating the electric field energy stored in 3-nm-thick dielectric interface layers with dielectric constants of $\epsilon = 10$. For the CPW resonators, the simulation is done using a single cross section of the center pin and ground plane, and for the LE resonators the simulation is done using a single cross section of the dipole capacitor pads. We use dc finite element simulations (Ansys Maxwell) for both kinds of single cross section simulations.

The single cross section approximation is appropriate for the CPW resonators, because their geometry is a single cross section extruded along a path. However, the degree to which the single cross section simulation is a good approximation for the LE resonators is not as obvious, as the LE resonators have nontrivial structure in the direction normal to the cross section plane. To check that the single cross section simulations accurately estimate the SPRs of the LE resonators, we separately compute the SPRs for a handful of LE resonators using the method outlined in Ref. [42], which involves an eigenmode simulation supplemented with a dc cross section simulation of the metal edges. We find that the single cross section and 2D sheet methods agree to within 15%, indicating that the single dipole capacitor cross section simulation is a suitable approximation for the full LE resonator SPR. This also implies that the meander inductor does not contribute significantly to the total SPR of the LE resonators. We note that, as the meander inductor and the capacitor pads are fabricated simultaneously and both features are large enough to be efficiently accessed by our wet chemical treatments (Fig. 5), the surface loss tangent should not vary between the meander inductor and capacitor pads, and, thus, we can ignore the contribution of the meander inductor to our device loss.

Two potential limitations of our use of surface participation in this work are (i) the arbitrary choice of the MS interface as the interface used to extract quantities and (ii) the assumption that the sidewalls of our devices are right angles. In order to quantify the effect these two assumptions have on our conclusions, we analyze our data using all three interfaces (MS, SA, and MA) at sidewall angles of 0 and 45°, and the values of our extracted quantities are shown for all possible combinations of interface and sidewall angle in Fig. 13. We find that, while sidewall angle and interface choice affect the absolute values of our extracted surface loss tangents, their relative magnitudes are preserved. Furthermore, we find that our extracted substrate loss tangent, and correspondingly the motivation for a two-component loss model, only increases in magnitude when other interfaces and angles are considered.

APPENDIX D: BULK LOSSES

In the main text, we describe how we extract loss independent of SPR by fitting $Q_{\mathrm{TLS},0}$ versus SPR. To further justify this model, we perform an explicit model comparison between one that incorporates a bulk loss component and another that includes only surface loss (Fig. 14). The surface-only model is a poor fit to the data at high surface participation ratio, particularly for BOE treated devices. The reduced chi-squared of the fit shown

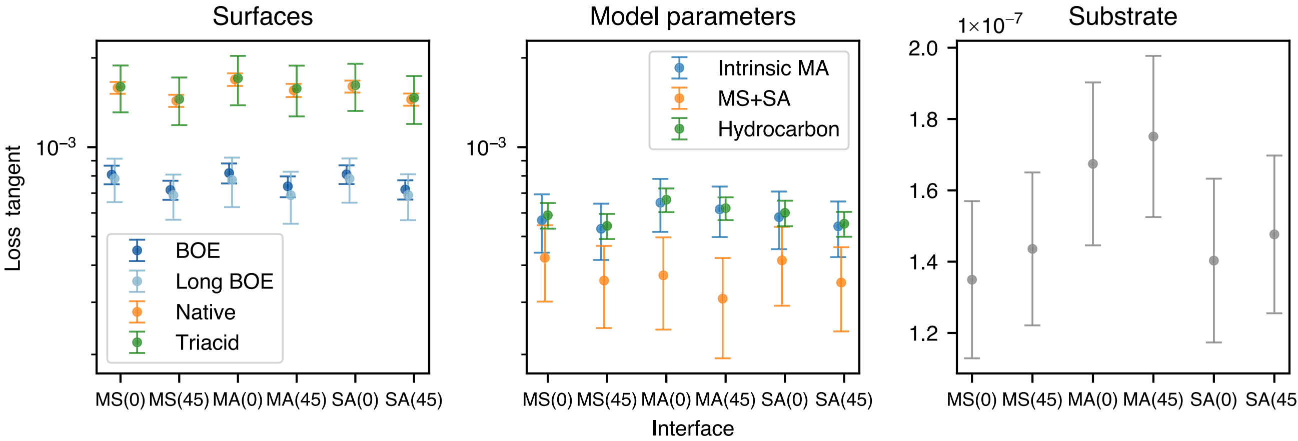

FIG. 13. Plot of extracted quantities versus chosen interface and assumed sidewall angle. Small horizontal perturbations are added to points that share the same horizontal coordinate in order to increase readability of the figure.

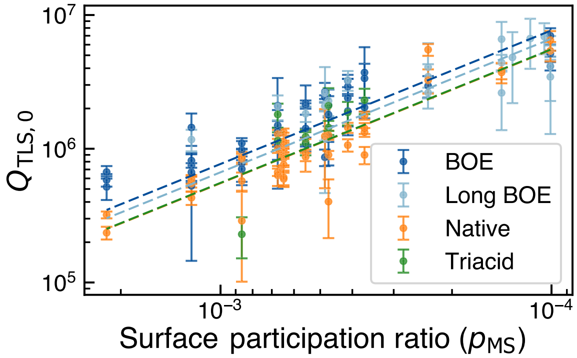

FIG. 14. Fit to data using a model for loss that includes only a contribution from surfaces.

in the main text is 5.7, whereas the reduced chi-squared of the fit to the surface-only model is 9.5.

We note that the extracted low-power bulk loss tangent is an order of magnitude larger than that measured in Ref. [11]. One hypothesis for this difference is that the "bulk" loss to which our measurements are sensitive is not the same as the volumetric average bulk loss measured by Ref. [11]. In our experiments, the device with the lowest SPR is an LE device with 65 μm spacing between capacitor pads. Our experiment, therefore, cannot distinguish between "bulk" and depths below the surface comparable to this spacing. In Ref. [11], by contrast, the experiment probes the loss tangent averaged over the bulk of a 440-μm-thick sapphire sample. We hypothesize that a near surface layer hosts a higher concentration of defects that give rise to TLS behavior. Since our measured bulk loss tangent is an order of magnitude higher than that measured in Ref. [11], in order to reconcile the two measurements this highly damaged layer would need to be around 10 times thinner than the bulk substrate measured in Ref. [11], around 50 μm.

We note that any defects at the tantalum-sapphire interface would be localized to the top few atomic layers of the sapphire surface. Our observations are consistent with the bulk loss occurring over much longer length scales, several orders of magnitude deeper than this interface, and, thus, would not be caused by a poor lattice match between the tantalum and sapphire. We further note that our films show highly coherent epitaxial growth of tantalum on sapphire, as reported in Ref. [15].

These hypothesized extended defects could be caused by polishing, damage from etching, or other fabrication-induced damage. Direct materials characterization of the polished sapphire could elucidate potential microscopic sources of TLS associated with this damage.

APPENDIX E: SURFACE CHARACTERIZATION AFTER CHEMICAL PROCESSING

We use x-ray photoelectron spectroscopy (XPS) to characterize the surface of our tantalum films before and after surface processing. We start with samples that have hard-baked photoresist applied and stripped off in solvent following the procedures outlined in Sec. A 1. We scan a sample before any further chemical processing, after a piranha treatment ("native" surface), and after both a piranha and a 20-min BOE treatment (BOE surface), as well as a separate sample after triacid treatment (triacid surface). To reduce the amount of adventitious carbon accumulated on the samples after chemical cleaning, we attempt to keep the length of time between chemical treatments and XPS measurements low. We take measurements within 30 min of the piranha treatment and BOE treatment and take measurements 4 h after triacid treatment. All XPS measurements are taken on a ThermoFisher K-Alpha XPS spectrometer with an aluminum Kα x-ray source.

We take a broad survey scan on each sample and observe Ta, O, and C peaks on all samples and a Na1s peak on the untreated sample only. We take fine scans of the Ta4f, O1s, and C1s peaks for all samples with a binding energy step size of 0.1 eV and a dwell time of 50 ms. We subtract a Shirley background from the Ta4f and O1s peaks [59] and a linear background from the C1s peaks. To account for potentially different x-ray flux between different measurements, we normalize all intensity data to the total intensity of the Ta4f spectrum for each sample. In addition, we calibrate the binding energy scale by setting the lowest binding energy Ta4f peak to 21.2 eV.

In the Ta4f spectrum, we can resolve two pairs of two peaks. We attribute the symmetric pair of peaks between 26 and 30 eV to the dominant Ta^{5+} oxidation state and the asymmetric pair of peaks between 21 and 24 eV to the tantalum metal [60]. Each state generates two peaks due to the strong spin-orbit coupling in tantalum [61]. The relative intensity of the Ta^{5+} peaks is smallest for the untreated sample, increases slightly after a piranha treatment, decreases slightly after a BOE treatment, and is largest after a triacid treatment [Fig. 15(a), qualitatively matching what is described in Ref. [19]].

For the C1s peak, the intensity is maximized for the untreated sample and is significantly reduced by each acid treatment. Performing a BOE treatment after piranha treatment reduces the C1s intensity over that of just piranha. The measurement on our triacid-treated sample shows the strongest C1s signal out of the three acid-treated measurements [Fig. 15(c)].

The relative intensity of the Ta^{5+} doublet and the intensity of the O1s peak both indicate that the oxide thickness grows slightly after piranha treatment, is etched slightly after the BOE treatment, and is grown significantly after triacid treatment. In Ref. [19], we measure that the BOE treatment reduces the oxide thickness by 20% and the triacid treatment grows the oxide thickness by over a factor of 2.

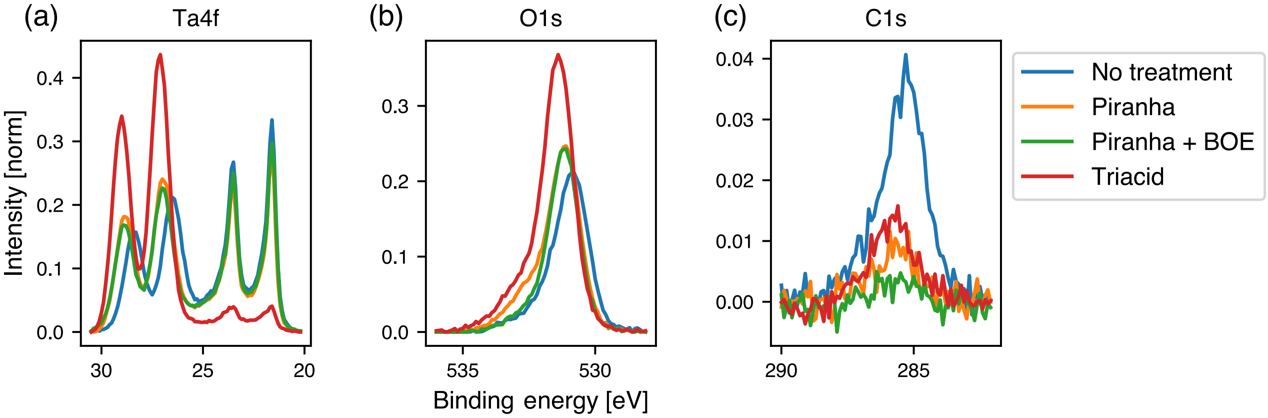

FIG. 15. XPS data of the surface of tantalum films after various surface treatments. Scans are of the Ta4f (a), O1s (b), and C1s (c) spectra. The "no treatment" data are taken after photoresist is stripped from a sample. "Piranha," "Piranha + BOE," and "Triacid" correspond to the "Native," "BOE," and "Triacid" surface conditions, respectively. Ta4f and O1s data have a Shirley background subtracted [59], and all C1s data have a linear background subtracted. All data are normalized to the total Ta4f intensity measured on the sample.

The sources of carbon in our system are adventitious carbon and photoresist residue. Therefore, the intensity change of the C1s peak is related to removal of fabrication residue but can be complicated by the duration of air exposure, which leads to adventitious carbon accumulation. Our data show that piranha treatment is effective at removing carbon from the surface, but performing BOE in addition to piranha can remove more carbon than piranha alone. We attribute this further reduction to carbon being removed from the surface of the tantalum oxide as it is etched away. We expect the triacid treatment to be extremely effective at cleaning the surface [50]; however, the measurement of the triacid-treated sample does not show as much reduction in the C1s signal. We attribute the larger triacid signal to the increased length of time between cleaning and measurement, which would allow more adventitious carbon to deposit on the surface.

We note that, as only oxygen, tantalum, and carbon are present in the survey scans for the piranha, piranha + BOE, and triacid-treated samples, changes observed in the O1s spectra may originate from changes in the tantalum oxide or the hydrocarbons but may also be complicated by adsorbed water. We note that both the tantalum oxide and the hydrocarbons are considered in our analysis of the loss tangents in our system. Further experiments and discussion of the interpretation of XPS spectra of the tantalum-air interface can be found in Ref. [19].

APPENDIX F: MODEL FOR SURFACE LOSSES

In the limit that surface losses dominate, dielectric loss in superconducting resonators can be expressed as

$$\frac{1}{Q_{\mathrm{TLS},0}} = p_{\mathrm{MS}} \tan\delta_{\mathrm{MS}} + p_{\mathrm{MA}} \tan\delta_{\mathrm{MA}} + p_{\mathrm{SA}} \tan\delta_{\mathrm{SA}}, \quad \text{(F1)}$$

where $\tan\delta_i$ is the loss tangent of interface i and MS, MA, and SA are the metal-substrate, metal-air, and substrate-air interfaces, respectively. The above expression can be rearranged as follows:

$$\begin{aligned}\frac{1}{Q_{\mathrm{TLS},0}} &= p_{\mathrm{MS}} \tan\delta_{\mathrm{MS}} + p_{\mathrm{MA}} \tan\delta_{\mathrm{MA}} + p_{\mathrm{SA}} \tan\delta_{\mathrm{SA}} \\ &= p_{\mathrm{MS}}(\tan\delta_{\mathrm{MS}} + \beta_{\mathrm{MA}} \tan\delta_{\mathrm{MA}} + \beta_{\mathrm{SA}} \tan\delta_{\mathrm{SA}}) \\ &= p_{\mathrm{MS}} \tan\delta, \quad \text{(F2)}\end{aligned}$$

where $\beta_i = p_i/p_{\mathrm{MS}}$ and $\tan\delta$ is the parameter we fit for. We can recast the above in terms of p_{MA}:

$$\begin{aligned}\frac{1}{Q_{\mathrm{TLS},0}} &= p_{\mathrm{MS}} \tan\delta \\ &= p_{\mathrm{MA}}\left(\frac{p_{\mathrm{MS}}}{p_{\mathrm{MA}}}\right) \tan\delta \\ &= p_{\mathrm{MA}}\alpha_{\mathrm{MS}} \tan\delta, \quad \text{(F3)}\end{aligned}$$

where $\alpha_i = p_i/p_{\mathrm{MA}}$.

We now consider a model in which the BOE, long BOE, and triacid samples have a source of loss on the MA interface that scales linearly with the oxide thickness, and the native samples suffer from both this oxide-thickness-dependent loss and an additional source of loss on the MA interface which we hypothesize is due to fabrication-related hydrocarbons (Fig. 16). We can recast losses in terms of the true oxide thickness and the hydrocarbon-related loss by

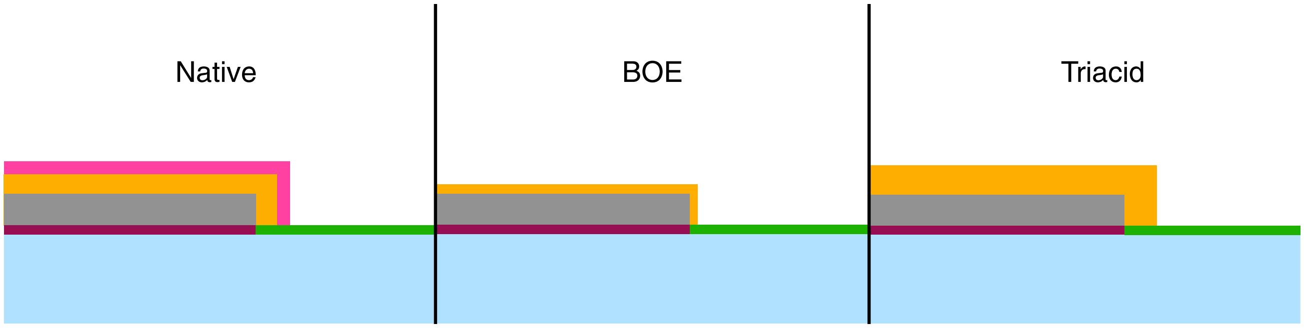

FIG. 16. A model for hydrocarbon losses in which BOE and triacid treatments remove residual hydrocarbons left over from photoresist. The native samples suffer from losses from both the native oxide and the hydrocarbons on the MA interface. In the above cartoon, the pink layer is hydrocarbons, and the orange, green, and purple layers are the MA, SA, and MS interfaces, respectively. The BOE diagram corresponds to both the BOE and long BOE treatments, with the difference being the thickness of the oxide layer. This model assumes that piranha cleaning is effective at removing hydrocarbons on the sapphire but not on the oxide surface.

$$\begin{aligned}\left(\frac{1}{Q_{\mathrm{TLS},0}}\right)_i &= p_{\mathrm{MA}}\alpha_{\mathrm{MS}}\tan\delta_i \\ &= p_{\mathrm{MA}}\left(\frac{t_i}{t_0}\right)\left(\frac{t_0}{t_i}\right)\alpha_{\mathrm{MS}}\tan\delta_i \\ &= p_{\mathrm{MA},i}\left(\frac{t_0}{t_i}\right)\alpha_{\mathrm{MS}}\tan\delta_i \\ &= p_{\mathrm{MA},i}\left(\frac{t_0}{t_i}\right)(\tan\delta_{\mathrm{MA}} + \alpha_{\mathrm{SA}}\tan\delta_{\mathrm{SA}} \\ &\quad + \alpha_{\mathrm{MS}}\tan\delta_{\mathrm{MS}} + \gamma_i\tan\delta_{\mathrm{HC}}) \\ &= p_{\mathrm{MA},i}\left(\tan\delta_{\mathrm{MA},0} + \left(\frac{t_0}{t_i}\right)(\alpha_{\mathrm{SA}}\tan\delta_{\mathrm{SA}}\right. \\ &\quad \left. + \alpha_{\mathrm{MS}}\tan\delta_{\mathrm{MS}} + \gamma_i\tan\delta_{\mathrm{HC}})\right), \end{aligned} \tag{F4}$$

where $t_0 = 3$ nm is the standard assumed oxide thickness, t_i is the measured oxide thickness for the ith surface processing technique, $p_{\mathrm{MA},i}$ is the true MA surface participation of the ith surface processing technique (up to a factor of the assumed oxide dielectric constant, $\epsilon = 10$), and $\gamma_i \in \{0, 1\}$ determines if hydrocarbon loss is considered for the ith surface processing technique ($\gamma_{\mathrm{native}} = 1$ and 0 otherwise). Equating the third and fifth lines of the above gives

$$\begin{aligned}\alpha_{\mathrm{MS}}\tan\delta_i = \left(\frac{t_i}{t_0}\right)\tan\delta_{\mathrm{MA},0} + \alpha_{\mathrm{SA}}\tan\delta_{\mathrm{SA}} \\ + \alpha_{\mathrm{MS}}\tan\delta_{\mathrm{MS}} + \gamma_i\tan\delta_{\mathrm{HC}}.\end{aligned} \tag{F5}$$

By considering the native surface and any two of the BOE, long BOE, and triacid surface, we can solve the above equations for $\tan\delta_{\mathrm{MA},0}$, $\alpha_{\mathrm{SA}}\tan\delta_{\mathrm{SA}} + \alpha_{\mathrm{MS}}\tan\delta_{\mathrm{MS}}$, and $\tan\delta_{\mathrm{HC}}$.

Consider a set of surface treatments $\{\mathrm{Native}, a, b\}$, where a and b are any pair of BOE, long BOE, and triacid. The system of equations described by Eq. (F5) for this set of treatments is solved for $\tan\delta_{\mathrm{MA},0}$ and $\alpha_{\mathrm{SA}}\tan\delta_{\mathrm{SA}} + \alpha_{\mathrm{MS}}\tan\delta_{\mathrm{MS}}$ by

$$\tan\delta_{\mathrm{MA},0} = \left(\frac{t_0}{t_a - t_b}\right)\alpha_{\mathrm{MS}}(\tan\delta_a - \tan\delta_b)$$

and

$$\alpha_{\mathrm{SA}}\tan\delta_{\mathrm{SA}} + \alpha_{\mathrm{MS}}\tan\delta_{\mathrm{MS}} = \alpha_{\mathrm{MS}}\left(\frac{t_a\tan\delta_b - t_b\tan\delta_a}{t_a - t_b}\right). \tag{F6}$$

Note that these solutions involve only the surface treatments a and b. Solving for $\tan\delta_{\mathrm{HC}}$,

$$\begin{aligned}\tan\delta_{\mathrm{HC}} = \alpha_{\mathrm{MS}}\tan\delta_{\mathrm{Native}} - \left(\frac{t_{\mathrm{Native}}}{t_0}\right)\tan\delta_{\mathrm{MA},0} \\ - (\alpha_{\mathrm{SA}}\tan\delta_{\mathrm{SA}} + \alpha_{\mathrm{MS}}\tan\delta_{\mathrm{MS}}).\end{aligned} \tag{F7}$$

In order to have a better basis of comparison to our extracted quantities, we rescale the quantities computed above to p_{MS}, which is the conventional metric by which surface-dependent losses are compared. For the hydrocarbon loss, we have

$$\begin{aligned}\frac{1}{Q_{\mathrm{HC}}} &= p_{\mathrm{MA}}\tan\delta_{\mathrm{HC}} \\ &= p_{\mathrm{MS}}\beta_{\mathrm{MA}}\tan\delta_{\mathrm{HC}},\end{aligned} \tag{F8}$$

where Q_{HC} is the inverse loss associated with the hydrocarbons and $\beta_{\mathrm{MA}} = p_{\mathrm{MA}}/p_{\mathrm{MS}}$.

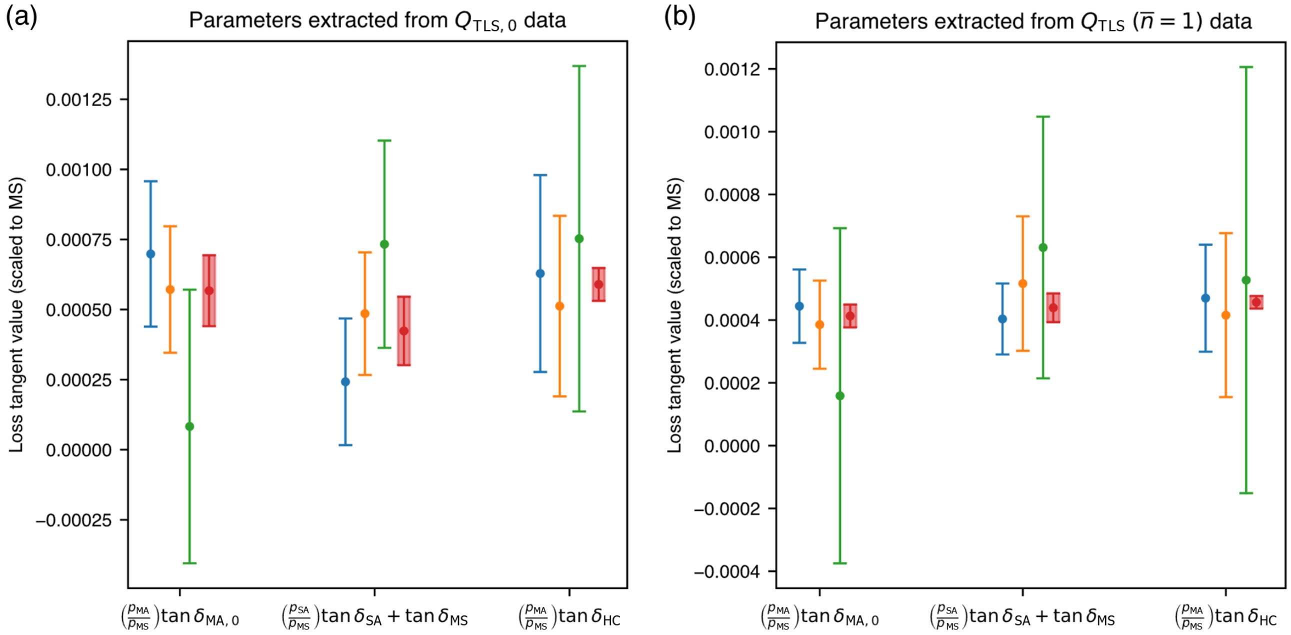

FIG. 17. Model estimates provided by the three possible combinations of hydrocarbon-free species. Blue is BOE and triacid, orange is long BOE and triacid, and green is BOE and long BOE. The best fit value is shown in red, with a shaded box to distinguish it from the different extracted estimates. (a) Estimates from $Q_{\mathrm{TLS},0}$ data. (a) Estimates from $Q_{\mathrm{TLS}}(\bar{n}=1)$ data.

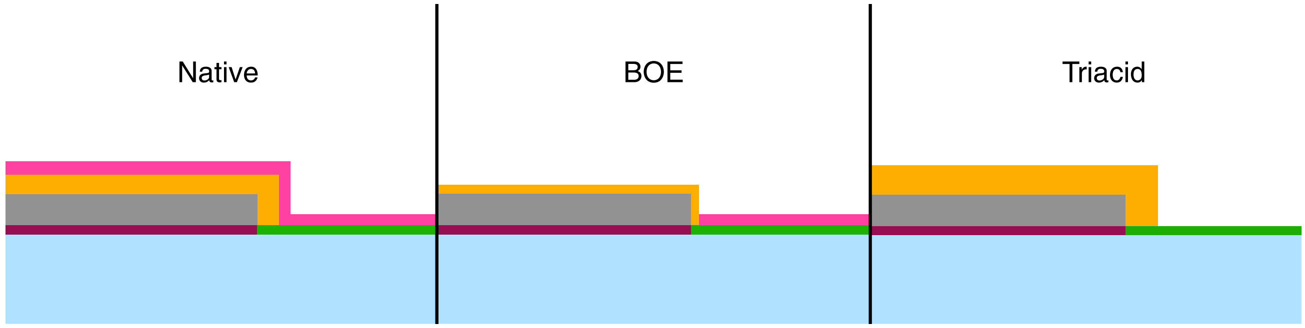

FIG. 18. An example of a model excluded by our data. Since BOE does not etch sapphire and does not etch hydrocarbons, one possibility is that hydrocarbons reside on the MA and SA interfaces of the native samples and the SA interface of the BOE and long BOE. This model could be achieved if piranha cleaning is ineffective at removing hydrocarbons from the sapphire surface, while the triacid treatment is highly effective. We continue to assume that hydrocarbons are lifted off from tantalum with BOE etching of the oxide. The BOE diagram corresponds to both the BOE and long BOE treatments, with the difference being only the thickness of the oxide layer. The parameter values we extract from this model given our data are unphysical.

For the SA and MS loss,

$$\frac{1}{Q_{\mathrm{MS}}}+\frac{1}{Q_{\mathrm{MA}}}=p_{\mathrm{MA}}(\alpha_{\mathrm{SA}}\tan\delta_{\mathrm{SA}}+\alpha_{\mathrm{MS}}\tan\delta_{\mathrm{MS}})$$
$$=p_{\mathrm{MS}}\beta_{\mathrm{MA}}(\alpha_{\mathrm{SA}}\tan\delta_{\mathrm{SA}}+\alpha_{\mathrm{MS}}\tan\delta_{\mathrm{MS}}), \tag{F9}$$

where Q_{SA} and Q_{MS} are the inverse loss associated with the SA and MS interfaces, respectively.

Oxide thicknesses for the native, BOE, and triacid treatments are determined in Ref. [19], and we estimate the oxide thickness for the long BOE in Sec. G. In all cases, we consider the total oxide thickness to be a sum of the Ta^{5+}, Ta^{3+}, and Ta^{1+} species. We compare the solutions for $\tan\delta_{\mathrm{MA},0}$, $\alpha_{\mathrm{SA}}\tan\delta_{\mathrm{SA}}+\alpha_{\mathrm{MS}}\tan\delta_{\mathrm{MS}}$, and $\tan\delta_{\mathrm{HC}}$ for different choices of three surface treatments in Fig. 17 and find that the solutions agree to within uncertainties. For each parameter, we fit the best single value to the three values reported by the three possible sets of surface treatments and report these fitted values in the main text.

Other assumptions about the configuration of hydrocarbons after the three surface treatments can be made, but we find that our data exclude certain configurations of hydrocarbons. For example, if we assume that piranha solution is ineffective at removing hydrocarbons from the sapphire surface, we assume the distribution of hydrocarbons depicted in Fig. 18. We note that hydrofluoric acid solutions, such as BOE, do not appreciably etch crystalline sapphire and, therefore, would not lift off hydrocarbons from the sapphire surface [62]. With this assumption set, we recover unphysical (negative) values for certain loss tangents, which implies that the model is incorrect. This result suggests that piranha cleaning is effective at removing fabrication-related hydrocarbons from the sapphire surface.

APPENDIX G: OXIDE THICKNESS AFTER LONG BOE

In Ref. [19], we measure the oxide thickness of tantalum films under three surface conditions: native, BOE treated for 20 min, and BOE treated for 40 min. The technique used, variable energy XPS (VEXPS), requires a synchrotron light source and so could not be replicated in our lab to measure the thickness of the 120 min BOE treated surface. Instead, to estimate the total oxide thickness, we correlate XPS measurements done on our laboratory system to oxide thickness measurements from Ref. [19].

We measure the Ta4f spectrum for four samples with four surface treatments: native and treated in BOE for 20, 40, and 120 min [Fig. 19(a)]. We subtract a Shirley background from all spectra [59] and normalize all data so that the metallic $\mathrm{Ta}_{7/2}$ peak height is unity. Similar to our observations in Sec. E, we see a decrease in the photoelectron fraction from the Ta^{5+} species with BOE treatment. We fit all Ta4f spectra with doublets associated with the Ta^0, $\mathrm{Ta}^0_{\mathrm{int}}$, Ta^{1+}, Ta^{3+}, and Ta^{5+} states [19,60]. The Ta^0 and $\mathrm{Ta}^0_{\mathrm{int}}$ peaks are all fit with asymmetric Voigt peaks, while the Ta^{1+}, Ta^{3+}, and Ta^{5+} peaks are all fit with symmetric Gaussians.

We consider the total oxide thickness to be the sum of the Ta^{1+}, Ta^{3+}, and Ta^{5+} species. We expect the fraction of the photoelectron intensity corresponding to these peaks to be proportional to their thickness with some unknown rate parameter. To find this rate parameter, we take the photoelectron intensity fraction of all oxide species for each sample and compare them to the oxide thicknesses measured in VEXPS [Fig. 19(c)]. We find an approximate linear relationship between the photoelectron intensity fraction of the oxide and the measured oxide thickness in VEXPS. We extrapolate this line to the photoelectron intensity fraction of the 120 min BOE treated device and

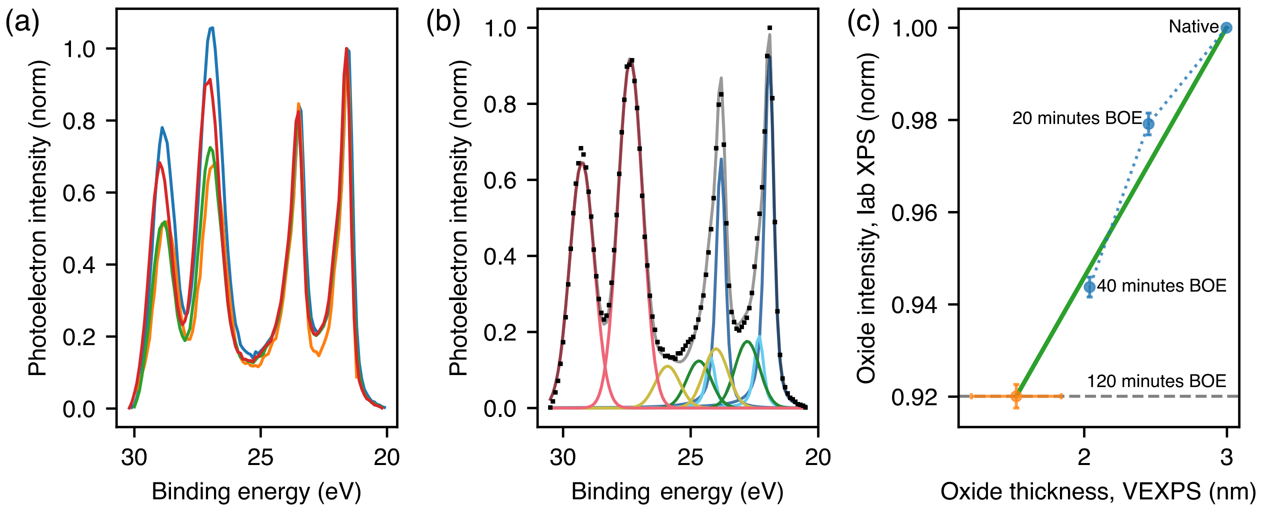

FIG. 19. (a) Ta4f spectra for native (blue), 20 min BOE treated (red), 40 min BOE treated (green), and 120 min BOE treated (orange) samples. Data are Shirley background corrected [59] and normalized so the peak height of the $\mathrm{Ta}_{7/2}$ metallic peak is unity. (b) Fit to the XPS peaks for the 20 min BOE treated sample. The peaks used to fit the spectrum are doublets of Ta^0 (dark blue), $\mathrm{Ta}^0_{\mathrm{int}}$ (cyan), Ta^{1+} (green), Ta^{3+} (yellow), and Ta^{5+} (pink). Ta^0 and $\mathrm{Ta}^0_{\mathrm{int}}$ peaks are fit with asymmetric Voigt profiles; others are fit with symmetric Gaussians. The lower binding energy peak in each doublet corresponds to the $\mathrm{Ta}_{7/2}$ spin state and the higher to the $\mathrm{Ta}_{7/2}$ spin state [61]. (c) Correlation of oxide photoelectron intensity fit to our lab-based XPS data to the oxide thickness measured in VEXPS [19]. Data (blue) are available from VEXPS for native, 20 min BOE treated, and 40 min BOE treated samples. Green is the best fit line to these data points, extrapolated to the oxide photoelectron intensity fraction of the 120 min BOE treated sample (gray dashed line) to give an estimate of the oxide thickness after a 120 min BOE treatment (orange). Photoelectron intensity fractions are normalized so that the native intensity fraction is unity.

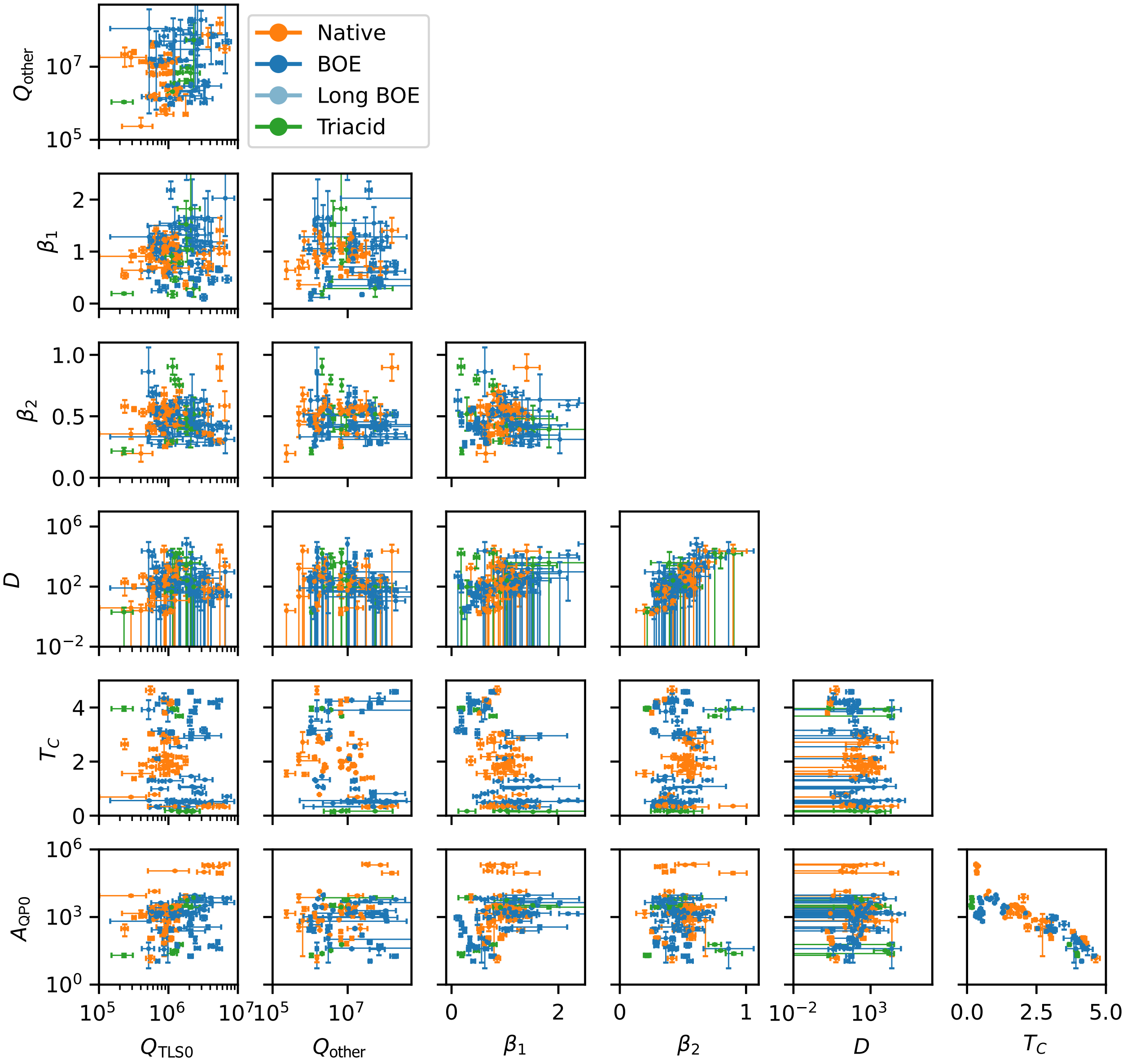

FIG. 20. Correlations between all seven fitted parameters used in Eqs. (1)–(3) in the main text. Lower uncertainty bounds in Q_{other} are truncated to $Q_{\mathrm{TLS},0}$.

find an estimated oxide thickness of 1.5 ± 3 nm. The error in this estimate is dominated by the linear fit shown in Fig. 19(c).

APPENDIX H: CORRELATIONS AMONG OTHER PARAMETERS

The model described in Eqs. (1)–(3) in the main text contains seven free fit parameters. To check that we can independently extract all seven parameters from our dataset, we plot each pair of fitted parameters in Fig. 20. We see no correlations among parameters except between T_c and A_{QP} and between D and β_2.

The correlation between T_c and A_{QP} is likely due to the limited amount of high-temperature data that we record (Sec. K). We conclude that, with our current dataset, we cannot quantitatively separate T_c and A_{QP}. For a few devices, we take detailed measurements over a larger temperature range in order to disambiguate T_c and A_{QP}, and we observe that Q_{int} decreases exponentially with temperature across three orders of magnitude, indicating that the T_c is indeed anomalously low (Fig. 21).

The correlation between D and β_2 is an artifact of the parametrization of Eq. (2). The correlation is a straight line on a log-linear plot, and, therefore, if we consider $D^{1/\beta 2}$ as our fit parameter instead of D, we see no correlation.

We note that no correlation is apparent between $Q_{\mathrm{TLS},0}$ and any of the other six parameters. Therefore, we are able

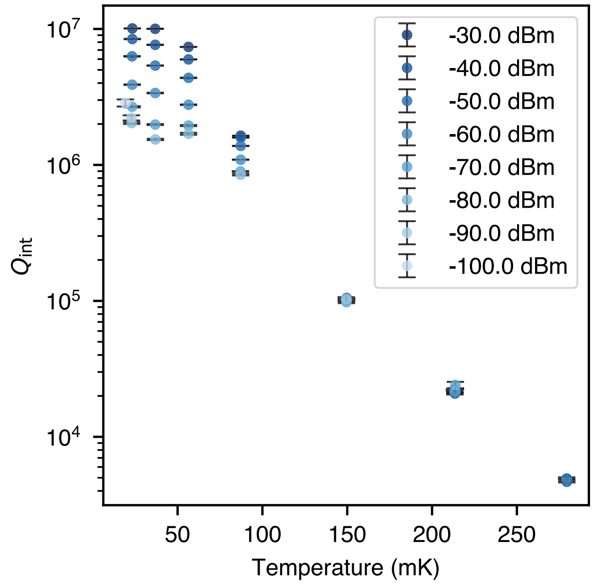

FIG. 21. Data from a power-temperature sweep that exhibits an exponential decrease in $Q_{\rm int}$ with temperature across three orders of magnitude, implying the origin of excess thermal quasiparticles is a low T_c.

to meaningfully differentiate the linear absorption of our TLS bath from its saturation behavior and from the effects of thermal quasiparticles and the constant loss parametrized by $Q_{\rm other}$. In particular, we note that the fitted value of T_c can range from 0.14 to above 4 K, which can affect the available temperature range of the data (Sec. K), but the fitted $Q_{\rm TLS,0}$ value does not correlate with T_c. Therefore, even though there is a significant variation in T_c across our devices, our method can disentangle this loss mechanism from TLS losses, and, thus, we include all devices in our analysis, regardless of T_c.

In Fig. 3(b) in the main text, we show that $Q_{\rm TLS,0}$ correlates with surface participation ratio. In Fig. 22, we plot the other six fitted parameters to Eqs. (1)–(3) in the main text. We first note that there is no stratification of any parameter with surface treatment. Furthermore, there is no correlation between any fitted parameter and surface participation ratio. This lack of correlation is expected for $A_{\rm QP}$ and T_C, as these parameters relate to the bulk properties of the tantalum film and not to interface properties, and for $Q_{\rm other}$, as $Q_{\rm other}$ is likely related to package losses (Sec. J 2).

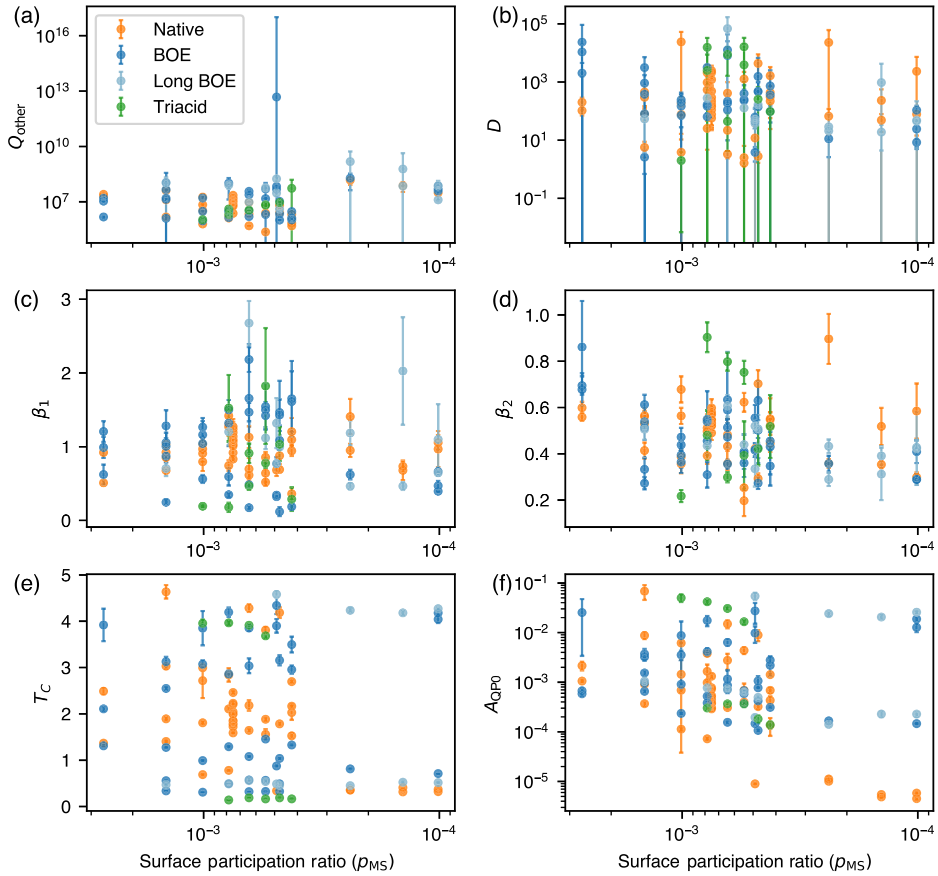

FIG. 22. Fitted parameters from Eqs. (1)–(3) in the main text versus surface participation ratio. $Q_{\rm TLS,0}$ is not shown here, as it is shown in Fig. 3(b) in the main text.

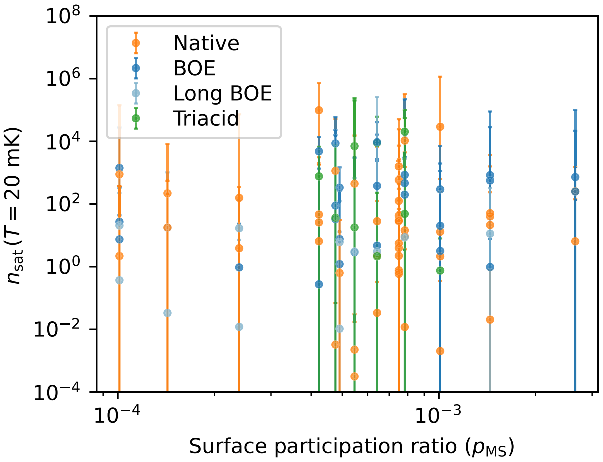

FIG. 23. Critical photon number (n_{sat}) at $T = 20$ mK versus surface participation ratio. n_{sat} is defined in Eq. (H2).

The lack of correlation between D, β_1, or β_2 and the surface participation ratio indicates that the saturation properties of the bath are not affected by the amount the bath participates in the mode. However, we note that, due to the interrelation of these three parameters, any individual parameter does not indicate the saturation photon number of the bath. Instead, we can rewrite Eq. (3) in the main text as

$$Q_{\mathrm{TLS}}(\bar{n}, T) = Q_{\mathrm{TLS},0} \frac{\sqrt{1 + \left(\frac{\bar{n}}{n_{\mathrm{sat}}}\right)^{\beta_2} \tanh\left(\frac{\hbar\omega}{2k_B T}\right)}}{\tanh\left(\frac{\hbar\omega}{2k_B T}\right)}, \tag{H1}$$

defining the saturation photon number n_{sat} as

$$n_{\mathrm{sat}} = (D T^{\beta_1})^{1/\beta_2}. \tag{H2}$$

We plot n_{sat} versus surface participation ratio in Fig. 23. Error bars in n_{sat} are propagated from the parameter uncertainties and covariances between D, β_1, and β_2. Because of the difficulty of separately fitting each of these three parameters, the uncertainties in n_{sat} are quite large. These large error bars in Fig. 23 do not allow us to know whether the saturation photon number correlates with surface participation ratio.

We further plot n_{sat} versus the fit parameters from Eqs. (1)–(3) in Fig. 24. We exclude plotting against D, β_1, and β_2, as these three parameters are used in Eq. (H2) to define n_{sat}, and so correlations are expected. Again, due to the large error bars, we cannot say whether n_{sat} correlates with other fitted parameters.

Large uncertainties in D, β_1, and β_2, which propagate to large uncertainties in n_{sat}, are likely due to the limited amount of saturation data available at low powers. As seen in Fig. 31, usually Q_{int} varies only by one order of magnitude before a power-independent source of loss begins to dominate, which makes disambiguating D, β_1, and β_2 challenging.

APPENDIX I: CRYSTALLOGRAPHIC PHASE AND ORIENTATION

We measure our tantalum film phase and orientation using x-ray diffraction (XRD). Our tantalum films are primarily α phase $\langle 111 \rangle$ oriented, but we also sometimes observe minority α phase $\langle 110 \rangle$ tantalum and β phase $\langle 002 \rangle$ species with XRD. We can observe the amount of these minority species vary across a single wafer, even to the point of not observing them near the center of a wafer and observing a significant fraction near the edge of the wafer. An example of this phenomenon is shown in Fig. 25(a). We also note that, even when no β phase tantalum is observed in XRD in the specific portion of the wafer on which we fabricate devices, the superconducting critical temperature fit to a power and temperature sweep can be significantly suppressed due to the lower critical temperature of β phase tantalum [Fig. 25(b)] [63].

We attribute the variation in T_C across our devices to the presence of a minority fraction of β phase tantalum. The combination of our XRD measurements with the fitted T_C suggests that our resonator devices are a more sensitive detector of the presence of β phase Ta, and the variability in XRD data across a wafer implies that a resonator made on a portion of the wafer does not indicate that devices made from other portions of the wafer will see a high T_C. These findings motivate the development of a more sensitive screening tool for tantalum crystallographic phase.

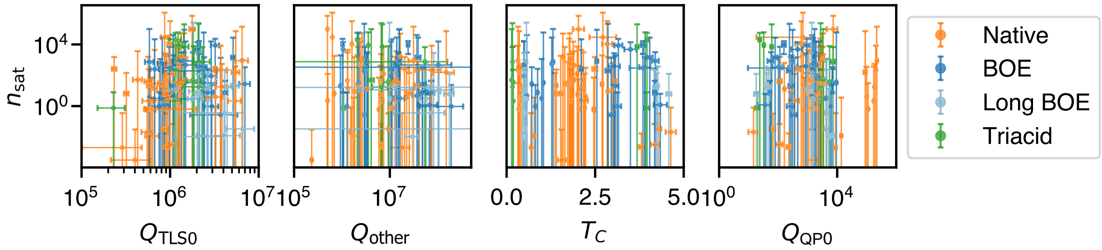

FIG. 24. Critical photon number (n_{sat}) at $T = 20$ mK versus fit parameters from Eqs. (1)–(3). n_{sat} is defined in Eq. (H2).

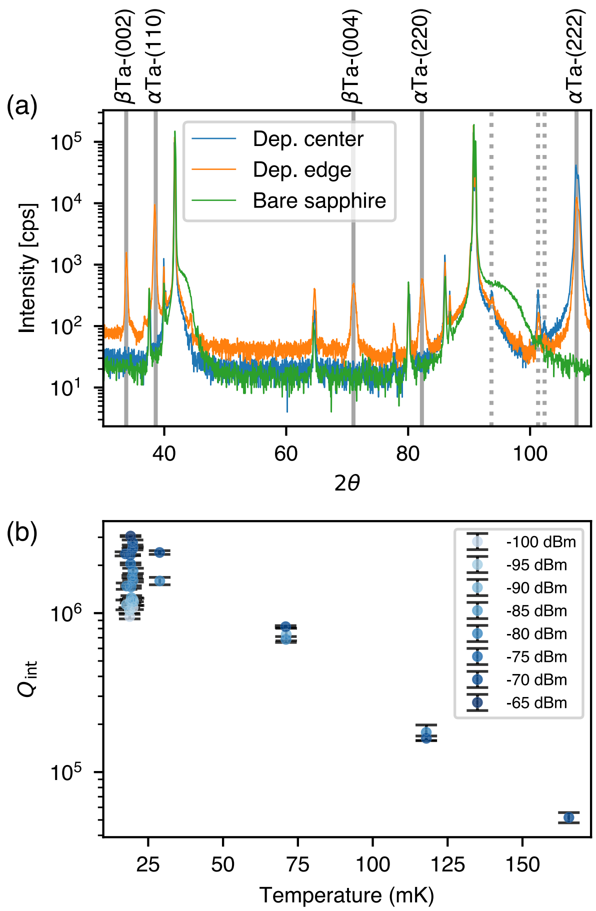

FIG. 25. (a) XRD data from the center and edge of a single tantalum deposition. XRD data taken on a bare sapphire wafer included for comparison. All data taken on a Bruker D8 Discover diffractometer. Solid gray vertical lines are crystallographic planes for tantalum which match our data for the primary copper K-α XRD wavelength [64]. Dotted lines are XRD peaks corresponding to the αTa-(222) plane and the copper K-β, tungsten L-$\alpha 1$, and tungsten L-$\alpha 2$ wavelengths. These wavelengths are also emitted from our x-ray source with less intensity and cause visible secondary peaks offset from primary primary peaks. (b) Power and temperature sweep of a resonator fabricated from the center of the wafer measured in (a). A suppressed T_C is observed despite no β phase tantalum being visible in the XRD.

APPENDIX J: ADDITIONAL VARIABLES

1. Effects of etch type on loss

In addition to the surface treatments we discuss in the main text, we also vary the method of etching the tantalum, as described in Sec. A 1. The plot of $Q_{\mathrm{TLS},0}$ versus p_{MS} in Fig. 3 in the main text is reproduced here but with the data points further stratified by etch type (Fig. 26). It is possible that dry etching with either the Cl- or F-based recipe leads to additional surface damage and TLS loss or that different etch types create different edge qualities; however, our data do not have enough statistical power to conclusively identify another source of loss arising from etch type.

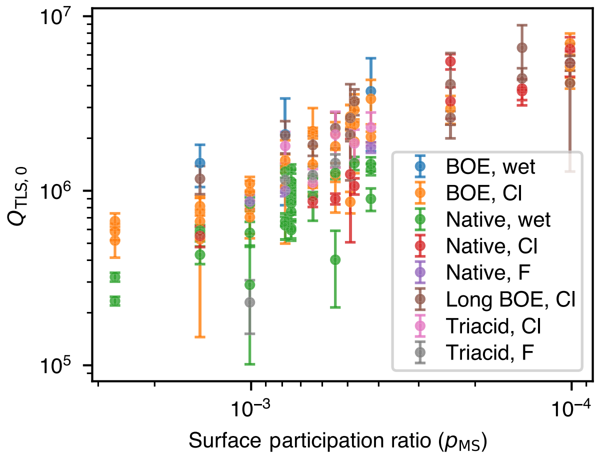

FIG. 26. Dependence of $Q_{\mathrm{TLS},0}$ on SPR separated by etch type.

We note that, as shown in Fig. 5, the wet-etched devices have steeper sidewalls than our dry-etched devices. We attribute the slope of the dry-etched sidewalls to resist edge erosion by the physical component of the etch. Further optimization in our dry etch could lead to different sidewall morphology.

2. Packaging

We use three types of package in our experiments. The first is a copper "puck and penny" assembly, the second is the commercially available QCage.24 from QDevil, and the third is a modified version of the QCage.24 with an aluminum-flashed coating on the surfaces of the package which face the device.

We compare the measured values of Q_{other} achieved for devices in each of the three packages in Fig. 27. We find that higher values of Q_{other} can be achieved for the QCage.24 package, with the highest values achieved with the aluminum flashing. Nine total devices packaged in the QCage.24

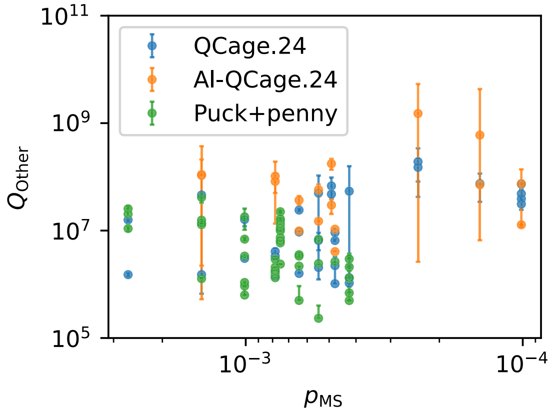

FIG. 27. Dependence of extracted Q_{other} on SPR, separated into devices packaged into the puck and penny assembly, the QCage.24 with bare copper surfaces, and the QCage.24 with aluminum-flashed surfaces. Lower error bars are truncated to the value of $Q_{\mathrm{TLS},0}$.

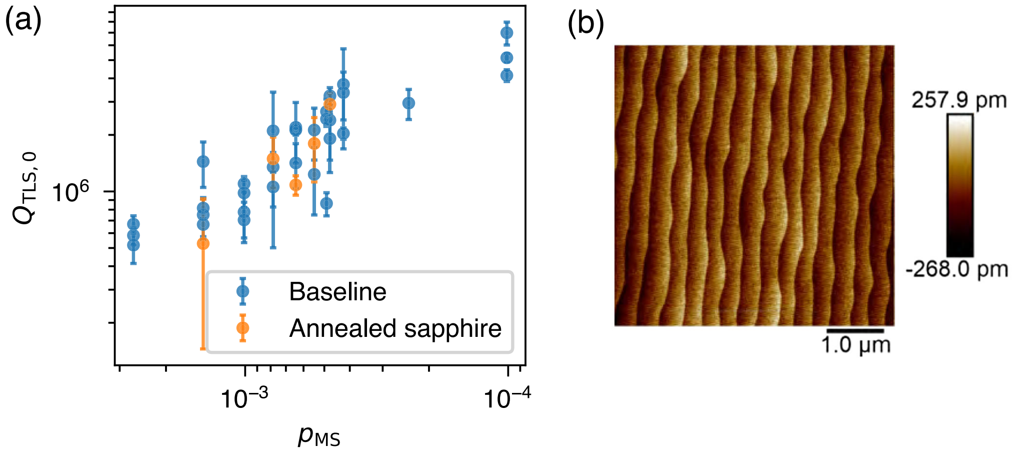

FIG. 28. (a) Dependence of extracted $Q_{\rm TLS,0}$ on SPR, separated into devices fabricated on annealed substrates and those fabricated on unannealed substrates. All devices are treated in BOE for 20 min. No significant difference in performance is seen. (b) AFM image of annealed sapphire surface. Scanned in 512 lines with a 1 Hz scan rate and a 7 mm tip.

and aluminum-flashed QCage.24 are excluded from Fig. 27, as $Q_{\rm other}$ is too large relative to the measured values of $Q_{\rm int}$ to be able to be fit confidently.

We conclude that, in some cases, $Q_{\rm other}$ is limited by packaging loss which is present on the puck and penny assembly but not present on the QCage.24. Based on the fact that the highest values of $Q_{\rm other}$ are found with aluminum flashing on the inside of the QCage.24, we further conclude that, in the QCage.24, the electric field of the modes of our devices have a non-negligible overlap with the packaging material. The improvement in $Q_{\rm other}$ is achieved by having the nearest surface of the package be a superconducting metal.

3. Annealing sapphire

One possible source of TLS loss is the disordered sapphire surface. We explore sapphire annealing to

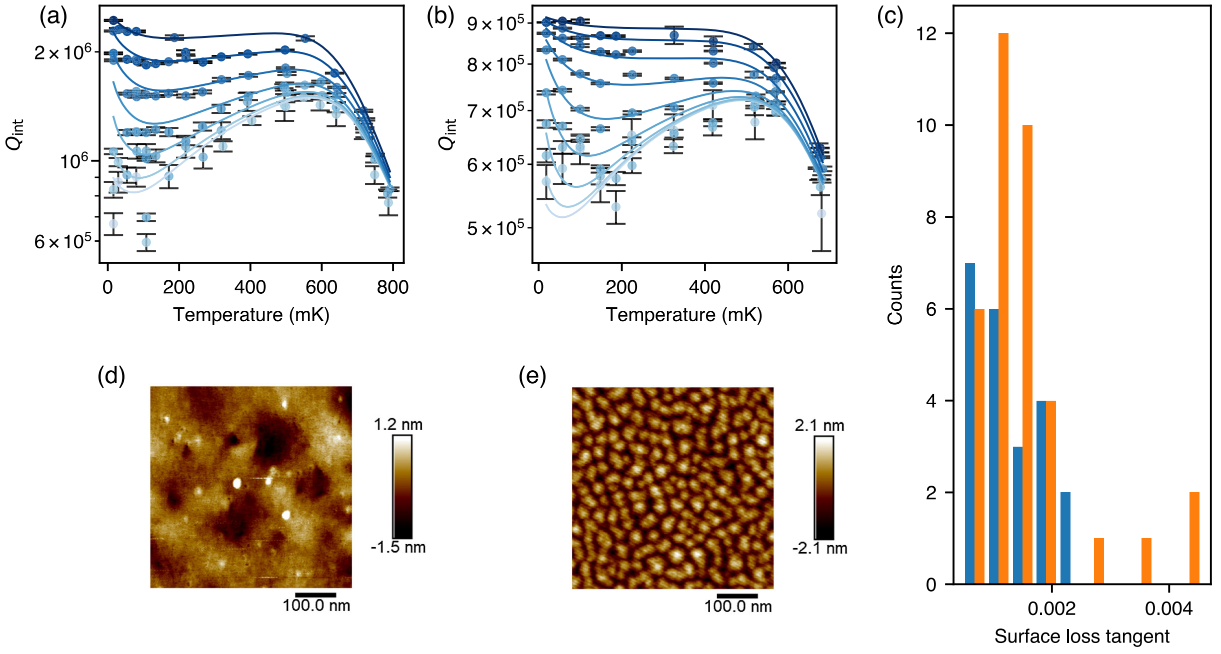

FIG. 29. Effect of surface morphology on device performance. Tantalum films used in our experiment are deposited by our group or by Star Cryoelectronics, and films from the two different sources show qualitatively different surface morphologies. (a),(b) Example temperature sweeps from devices fabricated with tantalum deposited by our group (a) or Star Cryoelectronics (b). Both devices are BOE treated and have surface participation ratios of approximately 10^{-3}. The color represents input power, with the darkest shade being the highest power. The spacing between powers is 10 dB. (c) Histogram of surface loss tangents from devices fabricated on films deposited by our group (blue) and Star Cryoelectronics (orange). Only devices with a BOE treatment are included. (d),(e) Example atomic force microscopy images showing surface morphology on a film deposited by our group (d) and by Star Cryoelectronics (e). The color scale represents depth.

interrogate the contribution of this surface. Prior to tantalum deposition, we process some of our sapphire wafers to achieve a near atomically flat surface with observable step edges to probe the impact on the metal-substrate and substrate-air losses. Atomic terraces have been previously observed on sapphire after high-temperature annealing [51].

After cleaning the sapphire wafers in a 2∶1 piranha bath and rinsing dry, we then treat the sapphire in 146° C sulphuric acid (Sigma Aldrich catalog number 258105) for 20 min, then triply rinse the wafer in deionized water, rinse once in 2-propanol, and then blow it dry in N_2. Finally, we anneal the sapphire in an air furnace with a temperature ramp of 4.1° C per minute to 1100° C and then held at 1100° C for one hour.

We confirm that we are able to achieve a flat surface by performing atomic force microscopy (AFM) on a sapphire sample prior to deposition [Fig. 28(b)]. The AFM used is a Bruker ICON3 with a 7 nm AFM tip. We observe discrete steps in height, each approximately 250 pm. This step size is on the order of a single lattice constant, and so we conclude we are observing atomic terraces.

We measure several resonators fabricated from films deposited on this annealed sapphire. All of these resonators are treated in BOE for 20 min. The dependence of $Q_{\mathrm{TLS},0}$ on SPR for these resonators is not distinguishable from the dependence seen for other BOE treated resonators [Fig. 28(a)]. As there is no measurable effect, we conclude that the sapphire anneal does not affect TLSs that are limiting our devices. We note that all resonators that are measured on annealed sapphire have high SPRs, in the linear region in Fig. 28(a), and, thus, are not sensitive to changes in bulk loss. An interesting avenue for future exploration would be to see if high-temperature annealing can change the bulk loss in sapphire.

4. Surface morphology

Our tantalum films are deposited both by our group and by Star Cryoelectronics. Both sources show a body-centered cubic α-Ta phase with majority $\langle 111 \rangle$ orientation when measured with an x-ray diffractometer; however, the surface morphology as measured with AFM is qualitatively different. Figures 29(d) and 29(e) show AFM images (Bruker ICON3) taken on the tantalum surface of films deposited by our group and Star Cryoelectronics, respectively.

We see no qualitative difference in the temperature sweep data between the two types of films [Figs. 29(a) and 25(b)]. We compare the fitted values of the surface loss tangent from devices with the same surface treatment fabricated on films from the two sources and see no significant difference (Fig. 29). We conclude that any losses associated with this variation in observed surface morphology difference do not limit device performance.

5. Rapid thermal annealing

With XPS, we can observe a shoulder peak at approximately 0.4 eV higher binding energy than the metallic tantalum peaks. Peaks in this location have been observed in Ref. [60], in which they are attributed to the closest layer of tantalum metal atoms to the oxide and have a differing

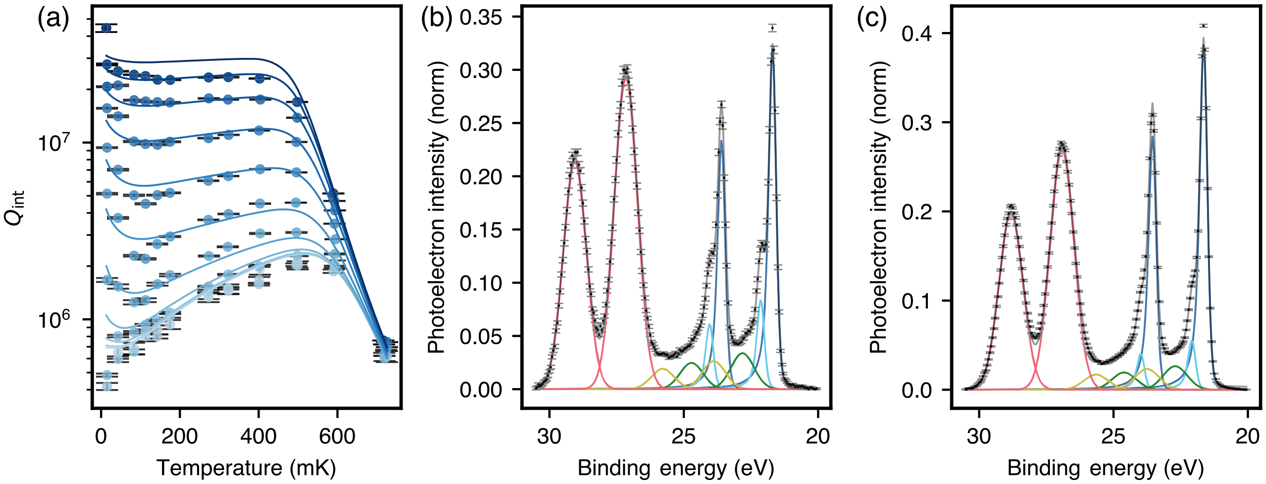

FIG. 30. (a) Results from temperature sweep fitted to a device treated with a rapid thermal anneal followed by a BOE treatment. The calculated surface participation ratio is approximately 2.6×10^{-3}. The fitted value of $Q_{\mathrm{TLS},0}$ is $(6.97 \pm 0.36) \times 10^5$. (b),(c) XPS data and fits for the Ta4f peaks performed on native films without (b) and with (c) the RTA process. All data are Shirley background corrected [59] and normalized so the total intensity for the spectrum is unity. The peaks used to fit the spectrum are doublets of Ta^0 (dark blue), $\mathrm{Ta}^0_{\mathrm{int}}$ (cyan), Ta^{1+} (green), Ta^{3+} (yellow), and Ta^{5+} (pink). Ta^0 and $\mathrm{Ta}^0_{\mathrm{int}}$ peaks are fit with asymmetric Voigt profiles; others are fit with symmetric Gaussians. The lower binding energy peak in each doublet corresponds to the $\mathrm{Ta}_{7/2}$ spin state and the higher to the $\mathrm{Ta}_{7/2}$ spin state [61]. There is an approximately 20% decrease in the fitted intensity of the $\mathrm{Ta}^0_{\mathrm{int}}$ peaks. Data are taken at the Spectroscopy Soft and Tender 2 (SST-2) end station at the National Synchrotron Light Source II at Brookhaven National Lab with X-ray energy 2000 eV. Data are collected with the same methodology described in Ref. [19].

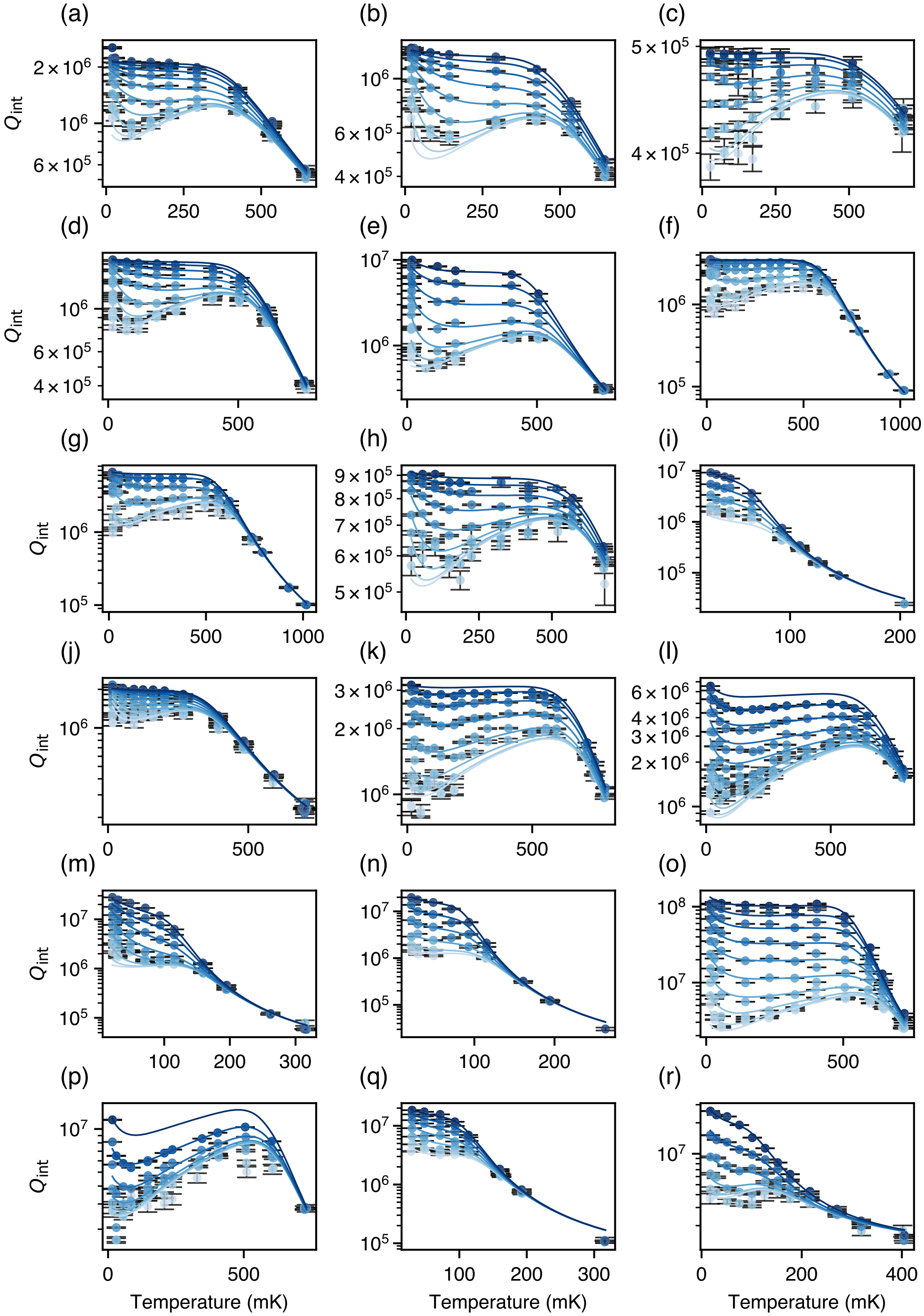

FIG. 31. Examples of representative fits to temperature sweep data. Data are taken from a variety of LE and CPW devices, as well as from native, BOE treated, long BOE treated, and triacid-treated surfaces. Colors indicate circulating power in the feedline, with the darkest shade representing the highest power and the lightest shade the lowest power. All traces are spaced 10 dB apart.

coordination number to those in the bulk. A plausible hypothesis for a location of TLSs is in this interfacial tantalum layer.

Rapid thermal annealing (RTA) is used in semiconductor processing to increase the ordering of interfacial layers in thin films [65] and is shown to have an effect on tantalum oxide thin films [66]. We use RTA to change the metal-oxide interface. Our process consists of a ramp to 800° C in 30 s and holding at 800° C for a further 30 s. The process is completed in a Thermo Scientific Lindberg Blue M furnace (PN: STF55433C-1).

We perform XPS on native samples with and without the RTA process and observe a decrease of approximately 20% in the fitted intensity of the interfacial tantalum shoulder peak [Figs. 30(b) and 26(c)]. We perform this RTA process on a resonator chip and measure a temperature sweep [Fig. 30(a)]. We observe no qualitative difference between the temperature sweep data on this device and data from temperature sweeps on devices without the RTA process. For the device fabricated on the film with the RTA process, the fitted $Q_{\mathrm{TLS},0}$ is $(6.97 \pm 0.36) \times 10^5$ with a SPR of 2.6×10^{-3}.

We measure three other devices with an SPR of 2.6×10^{-3} that are BOE treated, and we extract a mean $\mathrm{Q}_{\mathrm{TLS},0}$ of $(6.0 \pm 0.4) \times 10^5$. The device which undergoes RTA has a $Q_{\mathrm{TLS},0}$ over 2σ higher than the mean $Q_{\mathrm{TLS},0}$ for the control devices; we conclude that we may have seen a significant performance difference due to RTA. However, given that we measure only one RTA device, we cannot rule out the possibility that there would be a change in the extracted loss tangent of a family of RTA processed devices.

APPENDIX K: EXAMPLES OF TEMPERATURE FITS

Additional examples of fits to temperature sweep data are shown in Fig. 31. The SPR, surface treatment, resonator type, and packaging corresponding to each resonator are given in Table III.

We note that some devices [Figs. 31(g), 27(i), 27(j), 27(m), 27(n), 27(q), and 27(r)] begin to be dominated by equilibrium quasiparticles at a lower temperature, which we attribute to a minority phase of β-Ta that is below the detectable limit for our x-ray diffractometer setup. We also note that some resonators show a small range of Q_{int} [Figs. 31(f), 27(g), 27(j), and 27(q)], indicating that the TLS loss and nonsaturable loss mechanisms, parametrized by Q_{other}, are becoming comparable. Lastly, we note that some resonators [Figs. 31(p) and 27(r)] do not show evidence of a power- and temperature-independent loss mechanism in the ranges of Q_{int} shown. We are unable to fit Q_{other} to these devices.

TABLE III. Device and measurement parameters corresponding to data shown in Fig. 31. Surfaces are native (N), triacid (T), BOE (BOE), or long BOE (LB). Packaging is either puck and penny (P), QCage.24 (Q), or QCage.24 with aluminum flashing (QAl). Etch type is wet (W), dry chlorine based (Cl), or dry fluorine based (F).

Subplot	SPR ($\times 10^{-4}$)	Surface	Device type	Packaging	Etch
(a)	6.4	N	CPW	P	W
(b)	4.2	N	CPW	P	W
(c)	4.2	N	CPW	P	F
(d)	7.8	N	CPW	P	W
(e)	14.4	N	CPW	P	W
(f)	6.4	T	CPW	P	F
(g)	5.4	T	CPW	P	Cl
(h)	10.1	BOE	CPW	P	Cl
(i)	10.1	BOE	CPW	Q	Cl
(j)	5.5	BOE	CPW	Q	Cl
(k)	7.8	BOE	CPW	Q	Cl
(l)	6.4	N	CPW	Q	Cl
(m)	6.4	BOE	CPW	QAl	Cl
(n)	13.3	LB	CPW	QAl	Cl
(o)	4.9	LB	LE	QAl	Cl
(p)	4.9	BOE	LE	Q	Cl
(q)	4.9	LB	LE	QAl	Cl
(r)	1.4	N	LE	Q	Cl

APPENDIX L: DEPHASING

The frequency domain measurement setup used for all measurements in the main text and described in Sec. A 3 is sensitive to both energy decay and dephasing. To determine whether the measured quality factors are dominated by energy decay or dephasing, we perform experiments on four low surface participation resonators.

First, we perform ring down experiments on each resonator. We fully ring up the resonators with a high input power to approximately $\langle n \rangle = 10^3$ and record the power as the resonators relax. We average the incoming signals both coherently (including the phase information) and incoherently (averaging the amplitude only), as shown in Fig. 32(a). The expected form of the decay for coherently averaged signals is [11]

$$|\langle a(t) \rangle|^2 = e^{-\omega t/Q_{\mathrm{tot}}} |\langle e^{i\theta(t)} \rangle|^2, \tag{L1}$$

where $a(t)$ is the measured field, ω is the resonator angular frequency, Q_{tot} is the total quality factor, and $\theta(t)$ is the phase of the field. For incoherently averaged signals, the expected form of the decay is

$$\langle |a(t)|^2 \rangle = e^{-\omega t/Q_{\mathrm{tot}}} + b, \tag{L2}$$

where b is a background corresponding to the squared noise.

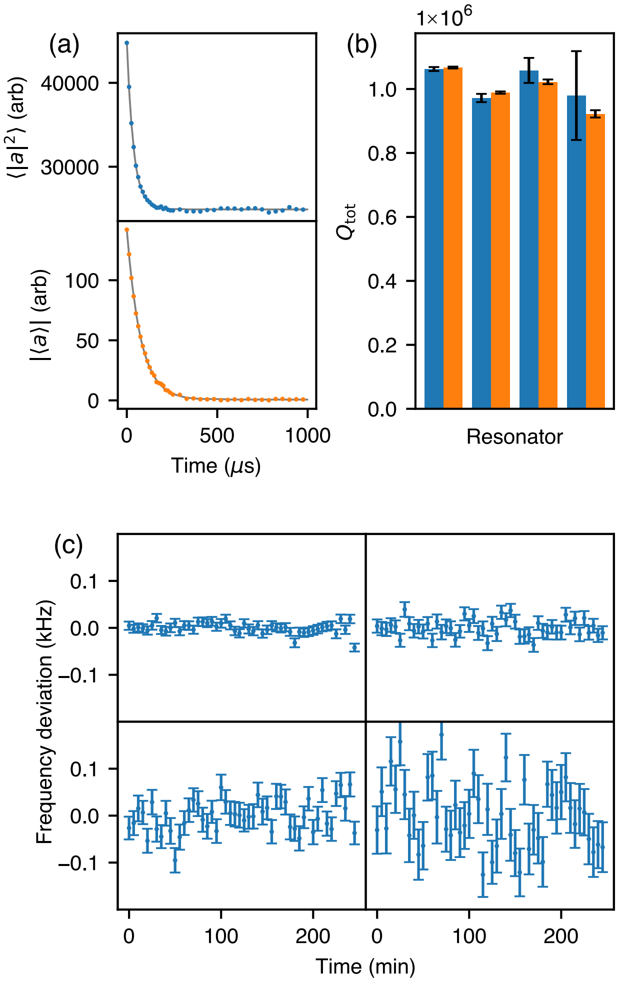

FIG. 32. Dephasing experiments on low SPR resonators. (a) Example incoherently (upper) and coherently (lower) averaged signals from a ring down experiment. Solid lines are fits to exponential decays. The fit to incoherently averaged data includes a constant offset to capture the squared noise. (b) Comparison of Q_{tot} fit to the incoherently and coherently averaged signals for four different resonators. (c) Fitted resonator center frequencies for measurements taken across different times. The vertical axis is the deviation of each data point from the mean. Each panel represents a different resonator.

If we assume no dephasing [$\theta(t)$ constant], then both the coherently and incoherently averaged signal have the same decay rate when fit to a simple exponential. Fitted total quality factors are shown in Fig. 32(b). For all four resonators, the two fitted values of Q_{tot} agree within 1σ. Using the smallest uncertainty fit, we find a lower limit on the internal dephasing quality factor $Q_\phi = (Q^{-1}_{\mathrm{tot,\,incoherent}} + Q^{-1}_{\mathrm{tot,\,coherent}})^{-1}$ to be over 1×10^8 for the first and second resonators. The third and fourth resonators give a lower limit on Q_ϕ of 5×10^7; however, these error bars are dominated by measurement noise, because the frequencies of the two resonators are close to the stop band of the measurement setup.

In addition, we measure all resonators every 5 min for 4 h and extracted the fitted center frequency [Fig. 32(c)]. The standard deviation of the center frequencies is slightly larger than the average uncertainty in the frequencies. The excess variance can be attributed to spectral diffusion; the largest excess standard deviation is 50 Hz, which would imply a dephasing quality factor of over 1×10^8 at our resonator frequencies.

As our fitted internal quality factors are only rarely above 1×10^8, we conclude that our measurements of Q_{int} are dominated by energy decay in our resonators.

[1] S. Krinner, N. Lacroix, A. Remm, A. Di Paolo, E. Genois, C. Leroux, C. Hellings, S. Lazar, F. Swiadek, J. Herrmann *et al.*, *Realizing Repeated Quantum Error Correction in a Distance-Three Surface Code*, Nature (London) **605**, 669 (2022).

[2] M. Gong, X. Yuan, S. Wang, Y. Wu, Y. Zhao, C. Zha, S. Li, Z. Zhang, Q. Zhao, Y. Liu *et al.*, *Experimental Exploration of Five-Qubit Quantum Error-Correcting Code with Superconducting Qubits*, Natl. Sci. Rev. **9**, nwab011 (2022).

[3] V. V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. L. Brock, A. Z. Ding, L. Frunzio *et al.*, *Real-Time Quantum Error Correction beyond Break-Even*, Nature (London) **616**, 50 (2023).

[4] Y. Zhao, Y. Ye, H.-L. Huang, Y. Zhang, D. Wu, H. Guan, Q. Zhu, Z. Wei, T. He, S. Cao *et al.*, *Realization of an Error-Correcting Surface Code with Superconducting Qubits*, Phys. Rev. Lett. **129**, 030501 (2022).

[5] R. Acharya, I. Aleiner, R. Allen, T. I. Andersen, M. Ansmann, F. Arute, K. Arya, A. Asfaw, J. Atalaya, R. Babbush *et al.*, *Suppressing Quantum Errors by Scaling a Surface Code Logical Qubit*, arXiv:2207.06431.

[6] X. Mi, M. Sonner, M. Y. Niu, K. W. Lee, B. Foxen, R. Acharya, I. Aleiner, T. I. Andersen, F. Arute, K. Arya *et al.*, *Noise-Resilient Edge Modes on a Chain of Superconducting Qubits*, Science **378**, 785 (2022).

[7] X. Mi, M. Ippoliti, C. Quintana, A. Greene, Z. Chen, J. Gross, F. Arute, K. Arya, J. Atalaya, R. Babbush *et al.*, *Time-Crystalline Eigenstate Order on a Quantum Processor*, Nature (London) **601**, 531 (2022).

[8] T. I. Andersen, Y. D. Lensky, K. Kechedzhi, I. Drozdov, A. Bengtsson, S. Hong, A. Morvan, X. Mi, A. Opremcak, R. Acharya *et al.*, *Observation of Non-Abelian Exchange Statistics on a Superconducting Processor*, arXiv:2210.10255.

[9] K. J. Satzinger, Y.-J. Liu, A. Smith, C. Knapp, M. Newman, C. Jones, Z. Chen, C. Quintana, X. Mi, A. Dunsworth *et al.*, *Realizing Topologically Ordered States on a Quantum Processor*, Science **374**, 1237 (2021).

[10] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, *Hardware-Efficient Variational Quantum Eigensolver for Small Molecules and Quantum Magnets*, Nature (London) **549**, 242 (2017).

[11] A. P. Read, B. J. Chapman, C. U. Lei, J. C. Curtis, S. Ganjam, L. Krayzman, L. Frunzio, and R. J. Schoelkopf, *Precision Measurement of the Microwave Dielectric Loss of*

Sapphire in the Quantum Regime with Parts-per-Billion Sensitivity, Phys. Rev. Appl. **19**, 034064 (2023).

[12] N. P. de Leon, K. M. Itoh, D. Kim, K. K. Mehta, T. E. Northup, H. Paik, B. S. Palmer, N. Samarth, S. Sangtawesin, and D. W. Steuerman, *Materials Challenges and Opportunities for Quantum Computing Hardware*, Science **372**, eabb2823 (2021).

[13] J. Krupka, K. Derzakowski, M. Tobar, J. Hartnett, and R. G. Geyer, *Complex Permittivity of Some Ultralow Loss Dielectric Crystals at Cryogenic Temperatures*, Meas. Sci. Technol. **10**, 387 (1999).

[14] J. Krupka, K. Derzakowski, A. Abramowicz, M. Tobar, and R. Geyer, *Use of Whispering-Gallery Modes for Complex Permittivity Determinations of Ultra-low-Loss Dielectric Materials*, IEEE Trans. Microwave Theory Tech. **47**, 752 (1999).

[15] A. P. M. Place, L. V. H. Rodgers, P. Mundada, B. M. Smitham, M. Fitzpatrick, Z. Leng, A. Premkumar, J. Bryon, A. Vrajitoarea, S. Sussman *et al.*, *New Material Platform for Superconducting Transmon Qubits with Coherence Times Exceeding 0.3 Milliseconds*, Nat. Commun. **12**, 1779 (2021).

[16] C. Wang, X. Li, H. Xu, Z. Li, J. Wang, Z. Yang, Z. Mi, X. Liang, T. Su, C. Yang *et al.*, *Towards Practical Quantum Computers: Transmon Qubit with a Lifetime Approaching 0.5 Milliseconds*, npj Quantum Inf. **8**, 3 (2022).

[17] D. P. Lozano, M. Mongillo, X. Piao, S. Couet, D. Wan, Y. Canvel, A. M. Vadiraj, T. Ivanov, J. Verjauw, R. Acharya *et al.*, *Manufacturing High-Q Superconducting α-Tantalum Resonators on Silicon Wafers*, arXiv:2211.16437.

[18] J. D. Teoh, P. Winkel, H. K. Babla, B. J. Chapman, J. Claes, S. J. de Graaf, J. W. O. Garmon, W. D. Kalfus, Y. Lu, A. Maiti *et al.*, *Dual-Rail Encoding with Superconducting Cavities*, arXiv:2212.12077.

[19] R. A. McLellan, A. Dutta, C. Zhou, Y. Jia, C. Weiland, X. Gui, A. P. M. Place, K. D. Crowley, X. H. Le, T. Madhavan *et al.*, *Chemical Profiles of the Oxides on Tantalum in State of the Art Superconducting Circuits*, Adv. Sci. **10**, 2300921 (2023).

[20] J. Gao, *The Physics of Superconducting Microwave Resonators*, Ph.D. thesis, California Institute of Technology, 2008.

[21] W. A. Phillips, *Two-Level States in Glasses*, Rep. Prog. Phys. **50**, 1657 (1987).

[22] C. Müller, J. H. Cole, and J. Lisenfeld, *Towards Understanding Two-Level-Systems in Amorphous Solids: Insights from Quantum Circuits*, Rep. Prog. Phys. **82**, 124501 (2019).

[23] J. M. Martinis, K. B. Cooper, R. McDermott, M. Steffen, M. Ansmann, K. D. Osborn, K. Cicak, S. Oh, D. P. Pappas, R. W. Simmonds *et al.*, *Decoherence in Josephson Qubits from Dielectric Loss*, Phys. Rev. Lett. **95**, 210503 (2005).

[24] E. Paladino, Y. M. Galperin, G. Falci, and B. L. Altshuler, *$1/f$ Noise: Implications for Solid-State Quantum Information*, Rev. Mod. Phys. **86**, 361 (2014).

[25] S. E. de Graaf, A. A. Adamyan, T. Lindström, D. Erts, S. E. Kubatkin, A. Y. Tzalenchuk, and A. V. Danilov, *Direct Identification of Dilute Surface Spins on* Al_2O_3*: Origin of Flux Noise in Quantum Circuits*, Phys. Rev. Lett. **118**, 057703 (2017).

[26] P. V. Klimov, J. Kelly, Z. Chen, M. Neeley, A. Megrant, B. Burkett, R. Barends, K. Arya, B. Chiaro, Y. Chen *et al.*, *Fluctuations of Energy-Relaxation Times in Superconducting Qubits*, Phys. Rev. Lett. **121**, 090502 (2018).

[27] J. Lisenfeld, A. Bilmes, A. Megrant, R. Barends, J. Kelly, P. Klimov, G. Weiss, J. M. Martinis, and A. V. Ustinov, *Electric Field Spectroscopy of Material Defects in Transmon Qubits*, npj Quantum Inf. **5**, 105 (2019).

[28] C. R. H. McRae, H. Wang, J. Gao, M. R. Vissers, T. Brecht, A. Dunsworth, D. P. Pappas, and J. Mutus, *Materials Loss Measurements Using Superconducting Microwave Resonators*, Rev. Sci. Instrum. **91**, 091101 (2020).

[29] W. Woods, G. Calusine, A. Melville, A. Sevi, E. Golden, D. K. Kim, D. Rosenberg, J. L. Yoder, and W. D. Oliver, *Determining Interface Dielectric Losses in Superconducting Coplanar-Waveguide Resonators*, Phys. Rev. Appl. **12**, 014012 (2019).

[30] B. G. Christensen, C. D. Wilen, A. Opremcak, J. Nelson, F. Schlenker, C. H. Zimonick, L. Faoro, L. B. Ioffe, Y. J. Rosen, J. L. DuBois, B. L. T. Plourde, and R. McDermott, *Anomalous Charge Noise in Superconducting Qubits*, Phys. Rev. B **100**, 140503(R) (2019).

[31] J. Gao, J. Zmuidzinas, B. A. Mazin, H. G. LeDuc, and P. K. Day, *Noise Properties of Superconducting Coplanar Waveguide Microwave Resonators*, Appl. Phys. Lett. **90**, 102507 (2007).

[32] S. Kumar, J. Gao, J. Zmuidzinas, B. A. Mazin, H. G. LeDuc, and P. K. Day, *Temperature Dependence of the Frequency and Noise of Superconducting Coplanar Waveguide Resonators*, Appl. Phys. Lett. **92**, 123503 (2008).

[33] R. Rivière, S. Deléglise, S. Weis, E. Gavartin, O. Arcizet, A. Schliesser, and T. J. Kippenberg, *Optomechanical Sideband Cooling of a Micromechanical Oscillator close to the Quantum Ground State*, Phys. Rev. A **83**, 063835 (2011).

[34] E. A. Wollack, A. Y. Cleland, P. Arrangoiz-Arriola, T. P. McKenna, R. G. Gruenke, R. N. Patel, W. Jiang, C. J. Sarabalis, and A. H. Safavi-Naeini, *Loss Channels Affecting Lithium Niobate Phononic Crystal Resonators at Cryogenic Temperature*, Appl. Phys. Lett. **118**, 123501 (2021).

[35] G. S. MacCabe, H. Ren, J. Luo, J. D. Cohen, H. Zhou, A. Sipahigil, M. Mirhosseini, and O. Painter, *Nano-acoustic Resonator with Ultralong Phonon Lifetime*, Science **370**, 840 (2020).

[36] J. M. Sage, V. Bolkhovsky, W. D. Oliver, B. Turek, and P. B. Welander, *Study of Loss in Superconducting Coplanar Waveguide Resonators*, J. Appl. Phys. **109**, 063915 (2011).

[37] S. Huang, B. Lienhard, G. Calusine, A. Vepsäläinen, J. Braumüller, D. K. Kim, A. J. Melville, B. M. Niedzielski, J. L. Yoder, B. Kannan *et al.*, *Microwave Package Design for Superconducting Quantum Processors*, PRX Quantum **2**, 020306 (2021).

[38] K. Serniak, M. Hays, G. de Lange, S. Diamond, S. Shankar, L. D. Burkhart, L. Frunzio, M. Houzet, and M. H. Devoret, *Hot Nonequilibrium Quasiparticles in Transmon Qubits*, Phys. Rev. Lett. **121**, 157701 (2018).

[39] V. B. Braginsky, V. S. Ilchenko, and K. S. Bagdassarov, *Experimental Observation of Fundamental Microwave Absorption in High-Quality Dielectric Crystals*, Phys. Lett. A **120**, 300 (1987).

[40] V. Gurevich and A. Tagantsev, *Intrinsic Dielectric Loss in Crystals*, Adv. Phys. **40**, 719 (1991).

[41] D. W. Face and D. E. Prober, *Nucleation of Body-Centered-Cubic Tantalum Films with a Thin Niobium Underlayer*, J. Vac. Sci. Technol. A **5**, 3408 (1987).

[42] C. Wang, C. Axline, Y. Y. Gao, T. Brecht, Y. Chu, L. Frunzio, M. H. Devoret, and R. J. Schoelkopf, *Surface Participation and Dielectric Loss in Superconducting Qubits*, Appl. Phys. Lett. **107**, 162601 (2015).

[43] K. L. Geerlings, *Improving Coherence of Superconducting Qubits and Resonators*, Ph.D. thesis, Yale University, 2013.

[44] J. Burnett, L. Faoro, I. Wisby, V. L. Gurtovoi, A. V. Chernykh, G. M. Mikhailov, V. A. Tulin, R. Shaikhaidarov, V. Antonov, P. J. Meeson *et al.*, *Evidence for Interacting Two-Level Systems from the 1/f Noise of a Superconducting Resonator*, Nat. Commun. **5**, 4119 (2014).

[45] H. Wang, M. Hofheinz, J. Wenner, M. Ansmann, R. C. Bialczak, M. Lenander, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank *et al.*, *Improving the Coherence Time of Superconducting Coplanar Resonators*, Appl. Phys. Lett. **95**, 233508 (2009).

[46] L. Grünhaupt, N. Maleeva, S. T. Skacel, M. Calvo, F. Levy-Bertrand, A. V. Ustinov, H. Rotzinger, A. Monfardini, G. Catelani, and I. M. Pop, *Loss Mechanisms and Quasiparticle Dynamics in Superconducting Microwave Resonators Made of Thin-Film Granular Aluminum*, Phys. Rev. Lett. **121**, 117001 (2018).

[47] C. M. Quintana, A. Megrant, Z. Chen, A. Dunsworth, B. Chiaro, R. Barends, B. Campbell, Y. Chen, I.-C. Hoi, E. Jeffrey *et al.*, *Characterization and Reduction of Microfabrication-Induced Decoherence in Superconducting Quantum Circuits*, Appl. Phys. Lett. **105**, 062601 (2014).

[48] D. P. Pappas, M. R. Vissers, D. S. Wisbey, J. S. Kline, and J. Gao, *Two Level System Loss in Superconducting Microwave Resonators*, IEEE Trans. Appl. Supercond. **21**, 871 (2011).

[49] R. Simons, *Coplanar Waveguide Circuits, Components, and Systems*, 1st ed. (Wiley-IEEE, New York, 2008).

[50] S. Sangtawesin, B. L. Dwyer, S. Srinivasan, J. J. Allred, L. V. H. Rodgers, K. De Greve, A. Stacey, N. Dontschuk, K. M. O'Donnell, D. Hu *et al.*, *Origins of Diamond Surface Noise Probed by Correlating Single-Spin Measurements with Surface Spectroscopy*, Phys. Rev. X **9**, 031052 (2019).

[51] F. Dwikusuma, D. Saulys, and T. F. Kuech, *Study on Sapphire Surface Preparation for III-Nitride Heteroepitaxial Growth by Chemical Treatments*, J. Electrochem. Soc. **149**, G603 (2002).

[52] P. Heidler, C. M. F. Schneider, K. Kustura, C. Gonzalez-Ballestero, O. Romero-Isart, and G. Kirchmair, *Non-Markovian Effects of Two-Level Systems in a Niobium Coaxial Resonator with a Single-Photon Lifetime of 10 Milliseconds*, Phys. Rev. Appl. **16**, 034024 (2021).

[53] A. Premkumar, C. Weiland, S. Hwang, B. Jäck, A. P. M. Place, I. Waluyo, A. Hunt, V. Bisogni, J. Pelliciari, A. Barbour *et al.*, *Microscopic Relaxation Channels in Materials for Superconducting Qubits*, Commun. Mater. **2**, 1 (2021).

[54] J. B. Chang, M. R. Vissers, A. D. Córcoles, M. Sandberg, J. Gao, D. W. Abraham, J. M. Chow, J. M. Gambetta, M. B. Rothwell, G. A. Keefe, M. Steffen, and D. P. Pappas, *Improved Superconducting Qubit Coherence Using Titanium Nitride*, Appl. Phys. Lett. **103**, 012602 (2013).

[55] D. L. Creedon, Y. Reshitnyk, W. Farr, J. M. Martinis, T. L. Duty, and M. E. Tobar, *High Q-Factor Sapphire Whispering Gallery Mode Microwave Resonator at Single Photon Energies and Millikelvin Temperatures*, Appl. Phys. Lett. **98**, 222903 (2011).

[56] J. Gittleman, B. Rosenblum, T. E. Seidel, and A. W. Wicklund, *Nonlinear Reactance of Superconducting Films*, Phys. Rev. **137**, A527 (1965).

[57] J. Zmuidzinas, *Superconducting Microresonators: Physics and Applications*, Annu. Rev. Condens. Matter Phys. **3**, 169 (2012).

[58] A. Palacios-Laloy, *Superconducting Qubit in a Resonator: Test of the Legget-Garg Inequality and Single-Shot Readout*, Ph.D. thesis, CEA Saclay, 2010.

[59] M. H. Engelhard, D. R. Baer, A. Herrera-Gomez, and P. M. A. Sherwood, *Introductory Guide to Backgrounds in XPS Spectra and Their Impact on Determining Peak Intensities*, J. Vac. Sci. Technol. A **38**, 063203 (2020).

[60] F. J. Himpsel, J. F. Morar, F. R. McFeely, R. A. Pollak, and G. Hollinger, *Core-Level Shifts and Oxidation States of Ta and W: Electron Spectroscopy for Chemical Analysis Applied to Surfaces*, Phys. Rev. B **30**, 7236 (1984).

[61] J. F. Moulder, *Handbook of X-Ray Photoelectron Spectroscopy: A Reference Book of Standard Spectra for Identification and Interpretation of XPS Data* (Physical Electronics Division, Perkin-Elmer Corporation, Waltham, MA, 1992).

[62] S. Juodkazis, K. Nishimura, H. Misawa, T. Ebisui, R. Waki, S. Matsuo, and T. Okada, *Control over the Crystalline State of Sapphire*, Adv. Mater. **18**, 1361 (2006).

[63] M. H. Read and C. Altman, *A New Structure in Tantalum Thin Films*, Appl. Phys. Lett. **7**, 51 (1965).

[64] R. D. Burbank, *An X-Ray Study of β-Tantalum*, J. Appl. Crystallogr. **6**, 217 (1973).

[65] R. Singh, *Rapid Isothermal Processing*, J. Appl. Phys. **63**, R59 (1988).

[66] S. Ezhilvalavan and T.-Y. Tseng, *Short-Duration Rapid-Thermal-Annealing Processing of Tantalum Oxide Thin Films*, J. Am. Ceram. Soc. **82**, 600 (1999).