ARTICLE   OPEN

Saving superconducting quantum processors from decay and correlated errors generated by gamma and cosmic rays

John M. Martinis[1]✉

Error-corrected quantum computers can only work if errors are small and uncorrelated. Here, I show how cosmic rays or stray background radiation affects superconducting qubits by modeling the phonon to electron/quasiparticle down-conversion physics. For present designs, the model predicts about 57% of the radiation energy breaks Cooper pairs into quasiparticles, which then vigorously suppress the qubit energy relaxation time ($T_1 \sim 600$ ns) over a large area (cm) and for a long time (ms). Such large and correlated decay kills error correction. Using this quantitative model, I show how this energy can be channeled away from the qubit so that this error mechanism can be reduced by many orders of magnitude. I also comment on how this affects other solid-state qubits.

INTRODUCTION

Quantum computers are proposed to perform calculations that cannot be run by classical supercomputers, such as efficient prime factorization or solving how molecules bind using quantum chemistry[1,2]. Such difficult problems can only be solved by embedding the algorithm in a large quantum computer that is running quantum error correction.

Quantum computers have intrinsic errors, so algorithms can be natively run with typically only a few hundred to thousand logic operations[3,4]. In order to run the most powerful and useful algorithms, say with millions to billions of logic gates, errors must be reduced to a parts per million or billion range, or lower. Fortunately, this is possible using quantum error correction, where the qubit state is distributed to many physical qubits in a way similar to classical error correction, so that errors in the physical qubit states can be selectively measured, decoded, and corrected. For example, surface code error correction encodes a protected "logical" state with about 1000 physical qubits[5,6]. As long as physical errors are small, about 0.1%, and occur randomly and independently among these 1000 qubits, then the logical error can be less than 0.1 part per billion[7]. However, if errors are large or correlated, bunching together either in time or across the chip in space, then error decoding fails. With a logical error, the memory of the quantum computer is lost and the algorithm fails.

This paper explains how cosmic rays and background gamma ray radiation are pulsed energy sources that produce large and correlated errors in superconducting qubits. Although quasiparticle and radiation effects have been observed in previous experiments on qubits, their average effect is small and roughly the magnitude of other decoherence mechanisms ($T_1 \sim 100$ μs)[8,9]. On a per event basis, however, the effect is argued here to be large ($T_1 \sim 1$ μs) and correlated in area ($\sim$cm^2) and time ($\sim$ms), which would strongly kill quantum error correction. It is imperative to further understand this error mechanism, not just reducing the effects of radiation, but slashing the resultant quasiparticle density, time, and length scales by a factor of 100 or more each, summarized as an overall million-fold improvement in present designs. There is no guarantee that such a huge improvement can be extrapolated from present experimental knowledge, so this paper overviews and suggests the critical physics and design changes needed to solve this important bottleneck.

Cosmic rays naturally occur from high-energy particles impinging from space to the atmosphere, where they are converted into muon particles that deeply penetrate all matter on the surface of the earth. When the muons traverse the quantum chip, they deposit a large amount of energy in the substrate of the quantum processor, on average 460 keV[10], which then briefly "heats" the chip. Gamma rays from natural background sources have a somewhat larger rate and can deposit energy up to about 1 MeV[10]. Experiments on low-temperature detectors and qubits have observed such radiation and quasiparticle effects[9–18].

A model is presented here that quantitatively describes the phonon down-conversion process to quasiparticles, which then decays the qubit state. This model shows that greater than 90% of the radiation energy is converted into phonons[10,19]. For present designs, 57% of the phonon energy then breaks Cooper pairs, with significant consequences: these quasiparticles reduce the qubit energy decay time T_1 in the range 0.16–1.6 μs[20,21], have a large spatial extent of mm to cm range, a long duration of 100 μs–10 ms, and occur more than once per minute. Each of these parameters is troubling, but in combination they are large enough to kill a complex quantum computation by many orders of magnitude.

Using a quantitative model for the generation of quasiparticles and their decay, I show that one can reliably redesign the quantum processor by channeling the phonon energy away from the qubits. The most important change is using thick films of a normal metal or low-gap superconductor to channel energy away from qubits. This redesign should reduce the initial quasiparticle density by a factor of 100, usefully larger than for a previous detector experiment with thin films[14,22,23]. This work is also complementary to a recent paper that describes well the radiation physics and the effects of breaking electron–hole pairs in the silicon crystal as part of the down-conversion process[10]; such charge offsets should not be an issue with large transmon qubits[24].

It is tempting to work on various mitigation strategies to reduce the background radiation, such as using low-radiation materials or

[1]Quantala, Santa Barbara, CA, USA. ✉email: martinis@quantala.tech

running a quantum computer deep underground. However, as the error rate needs to be reduced to less than one per day, it is unlikely these strategies will be effective enough. Instead, the energy should be channeled away from the qubits at the chip level.

In this paper, I first review the basic conversion physics of radiation to phonons to quasiparticles, how quasiparticles affect the qubit, and how quasiparticles eventually relax by recombination. Detailed models are then presented that describe an event in five stages, predicting the approximate magnitude, time, and size scale of the effect. I then describe the critical stages and how redesign of the qubit can lessen the impact by a huge margin, the goal of this paper. Finally, the impact of such radiation is briefly discussed for other qubit systems.

RESULTS

Phonon down-conversion to quasiparticles

At low temperatures, the physics of thermalization often becomes slow, and sometimes the riskiest assumption is that a system can be simply described by a temperature. For the nonequilibrium physics described here, it is better to represent the system being mostly at a low background temperature T, but with a small number of excitations that can each be described by their energy E or their occupation $f(E)$. The basic idea is illustrated in Fig. 1, where after the radiation event (a) there are a few high-energy phonons, which then get down-converted over time (b) to a larger number of low-energy phonons and quasiparticles. The quasiparticles radiate phonons until they are within typically 50–100 mK of the superconducting gap Δ[20]. As shown in Fig. 1c, the excitations need to be channeled to the normal metal and away from the qubit superconductor.

Following Eqs. (3) and (4) of ref. [20], the density of superconducting quasiparticles n_{qp} is

$$\frac{n_{qp}}{n_{cp}} = \frac{2}{\Delta} \int_{\Delta}^{\infty} d\epsilon\, f(\epsilon)\, \rho(\epsilon), \tag{1}$$

$$\rho(\epsilon) = \epsilon/\sqrt{\epsilon^2 - \Delta^2}, \tag{2}$$

where $n_{cp} = 2.8 \cdot 10^6/\mu m^3$ is the density of Cooper pairs for aluminum and $\rho(\epsilon)$ is the normalized superconducting density of states. In the two-fluid model of a superconductor, the quasiparticles damp the qubit like a low density normal metal, with a quality factor for a phase or transmon qubit given by[20]

$$1/Q \simeq 1.2\, n_{qp}/n_{cp} \tag{3}$$

where the numerical factor 1.2 is computed for 5 GHz and uses the regular assumption that quasiparticle energies are near the gap. Here, what matters is the quasiparticle density at the junction, since the dissipation in the two-fluid model can be thought of as coming from the normal-state tunnel resistance $R_T \simeq 100\,\Omega$. The quasiparticles also provide damping from the superconducting leads of the qubit, but typically can be neglected because its normal resistance is small $\sim 10\,m\Omega \ll R_T$.

The down-conversion of the cosmic or gamma ray radiation has been discussed previously for metals[25] and silicon[26]. The initial interaction of the radiation produces silicon electron–hole pairs and high-energy phonons up to the Debye energy 450 K[10]. These phonons down-convert to lower-energy phonons, but this process stops at an energy about 50 K (1 THz)[26] where the phonon dispersion relationship becomes linear. In silicon, most of the electron–hole pairs recombine rapidly (ns time scales), with less than 10% of the radiation energy contributing to their net occupation[10,19]. A sapphire substrate should be similar to silicon because both are "high-gap" insulators at cryogenic temperatures, but this assumption should be experimentally checked.

The most important part of the down-conversion process is for phonon energies below 50 K and the subsequent cascade of phonon and electron/quasiparticle scattering processes in metal. Here, a high-energy phonon is converted to electron/quasiparticle pairs, which then scatters again to produce another lower-energy phonon. The relevant scattering processes are listed in Table 1. These down-conversion processes are described by Kaplan et al. for superconductors[27], but can also be used for normal metals by setting the gap to zero. As these rates are written as integrals over the Fermi occupation factors, the formulas can thus be rewritten to express a single-particle scattering rate versus energy. Rates at $T = 0$ but for nonequilibrium quasiparticles are given below.

For q→q + p, the lifetime of a quasiparticle at energy ϵ to scatter to any energy ϵ' by emitting a phonon of energy $\epsilon - \epsilon'$ is given by Eq. (8) of ref. [27]

$$\Gamma_q^s(\epsilon) = \frac{1}{\tau_0} \int_{\Delta}^{\epsilon} d\epsilon' \frac{(\epsilon - \epsilon')^2}{(kT_c)^3} \rho(\epsilon') \left(1 - \frac{\Delta^2}{\epsilon \epsilon'}\right) \tag{4}$$

$$\simeq \frac{1.8}{\tau_0 \Delta^3} (\epsilon - \Delta)^3, \tag{5}$$

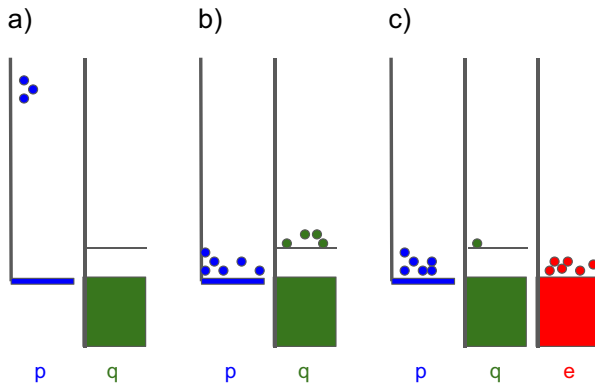**Fig. 1 Nonequilibrium phonon, quasiparticle, and electron excitations. a** Schematic representation of the phonon (p) and quasiparticle (q) excitations after the radiation event, with only a few high-energy phonons. The vertical axis represents the state energy. **b** After relaxation, the phonon energy has down-converted to quasiparticle excitations in the superconductor and low-energy (subgap) phonons. About 57% of the initial energy remains in the quasiparticles. **c** With a normal metal structure, most of the energy has been channeled away from the superconductor and towards the normal metal electrons (e), which is designed not to affect the qubit.

Table 1. Scattering processes.

Scattering	Kaplan (Al)	
p→q + q	(1/1.0 ns) (E_p/K)	
q→q + p	(1/1700 ns) $((E_q - \Delta)/K)^3$	
q + q→p	(22/440 ns) n_{qp}/n_{cp}	
Scattering	Power (n-Al)	Power (Cu)
p→e + e	(1/3.1 ns) (T_p/K)	(1/8.2 ns) (T_p/K)
e→e + p	(1/350 ns) (T_e/K)^3	(1/24 ns) (T_e/K)^3

Table of scattering processes for phonons (p), electrons (e), and quasiparticles (q), with initial energies E_p, E_e, and E_q, respectively, given in Kelvin. The "Kaplan" column represents the scattering rate given by ref. [27] for aluminum, using Eqs. (5), (7), and (10). The "Power" column is a rate from the power calculation given in Eqs. (12) and (13), for normal aluminum and copper. Both give estimates for scattering rates and their energy dependence.
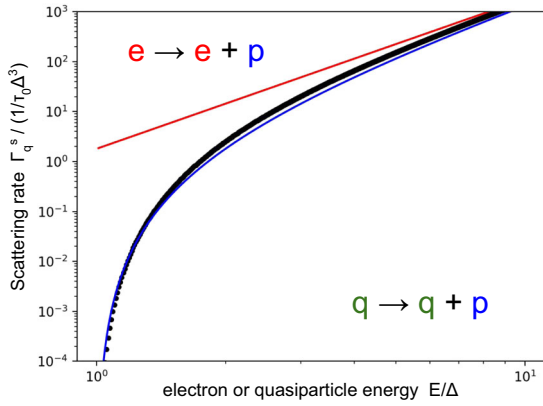
Fig. 2 Quasiparticle scattering rate. Plot showing the quasiparticle decay rate versus energy Eq. (4) (black dots), along with a fit formula Eq. (5) (blue line); the corresponding rate for a normal metal is the red line. The rates scale approximately as the final phonon density of states.

Fig. 3 Phonon decay rate. Plot showing the phonon decay rate versus phonon energy of Eq. (8) (black dots), compared to a normal metal (red line). No pair-breaking occurs below energy 2Δ, as expected. The rate above 2Δ is described well by the fit function of Eq. (9) (blue line).

where the second equation was fit to the results of numerical integration, plotted in Fig. 2. Here, the BCS relation is used $\Delta = 1.76\,kT_c$, and $\tau_0 = 440\,$ns is the inferred characteristic electron–phonon time for Al[27]. Note this rate scales as the available phonon density of states $(\epsilon - \Delta)^3$.

For q + q→p, the lifetime of a quasiparticle state with energy ϵ to recombine with another quasiparticle of any energy ϵ' is given by Eq. (8) of ref. [27]

$$\Gamma_q^r(\epsilon) = \frac{1}{\tau_0} \int_\Delta^\infty d\epsilon' \frac{(\epsilon + \epsilon')^2}{(kT_c)^3} f(\epsilon')\, \rho(\epsilon')\left(1 + \frac{\Delta^2}{\epsilon\epsilon'}\right) \quad (6)$$

$$\simeq \frac{22}{\tau_0} \frac{n_{qp}}{n_{cp}}, \quad (7)$$

where for the second equation the quasiparticles are assumed to be near the gap[20]. Note the similarity of this recombination rate to scattering, but here the energy of the emitted phonon $\epsilon + \epsilon'$ is greater than 2Δ. With the term $f(\epsilon')$, this rate is proportional to the density of quasiparticles, implying that the recombination rate slows down once the density of quasiparticles becomes small.

For p→q + q, the lifetime of a phonon with energy E_p to break two quasiparticles of energy ϵ and $E_p - \epsilon$ is given by Eqs. (27) and (30) of ref. [27]

$$\Gamma_p^b(E_p) = \frac{1}{\pi\tau_0^{ph}\Delta}$$
$$\times \int_\Delta^{E_p-\Delta} d\epsilon\, \rho(\epsilon)\rho(E_p - \epsilon)\left(1 + \frac{\Delta^2}{\epsilon(E_p-\epsilon)}\right) \quad (8)$$

$$\simeq \frac{1}{\pi\tau_0^{ph}\Delta}\left[E_p + \frac{3.8\,\Delta}{(E_p/\Delta + 2.3)^{0.8}}\right] \quad (9)$$

$$\sim \frac{1}{\pi\tau_0^{ph}\Delta}\, 1.4\, E_p\,, \quad (10)$$

where $E_p > 2\Delta$, and the second equation was fit to numerical integration. Equation (8) approaches $1/\tau_0^{ph}$ for $E_p = 2\Delta$. For Al, ref. [27] gives $\tau_0^{ph} = 0.24$ns. Figure 3 is a plot of the decay rate Eq. (8) (black dots) and the excellent approximation of Eq. (9) (blue line), and also compares to electron scattering without a gap (red line).

The phonon energy E_p must be greater than 2Δ to produce two quasiparticles; the excess energy is split randomly between them, on average $E_q = \Delta + (E_p - 2\Delta)/2$. Similarly for quasiparticle scattering, the excess quasiparticle energy is distributed between the quasiparticle and phonon, because the phonon density of states scales as the square of frequency, on average the resulting phonon takes most of the energy $(3/4)(E_q - \Delta)$. The last scattering
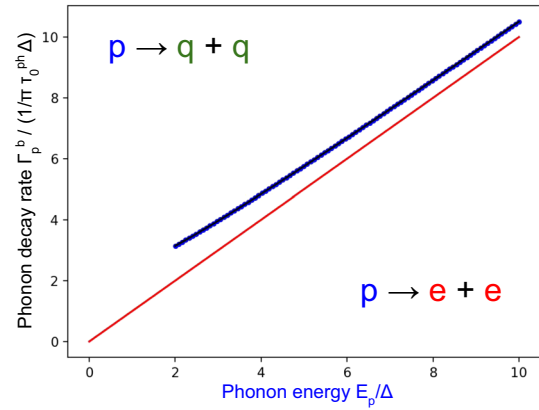
Table 2.	Energy parameters from heat capacities.		
Energy/Power		n-Al	Cu
U_p (10^{-9}nJ μm^{-3}K^{-4})		$2.5 \cdot (T_p^4/4)$	$6.6 \cdot (T_p^4/4)$
U_e (10^{-9}nJ μm^{-3}K^{-2})		$140 \cdot (T_e^2/2)$	$97 \cdot (T_e^2/2)$
Σ_{ep} (nW μm^{-3}K^{-5})		0.2	2

Table of energies for electrons U_e and phonons U_p taken from low-temperature heat capacities for normal aluminum and copper (https://nvlpubs.nist.gov/nistpubs/Legacy/MONO/nbsmonograph21.pdf). The electron–phonon coupling constant Σ_{ep} is also listed.

process in Table 1 is recombination, but since it scales as quasiparticle density the rate is small and can be ignored during the initial down-conversion cascade.

Electron–electron scattering is thought to be unimportant as its rate ($\sim E_e$) becomes relatively small compared to phonon scattering ($\sim E_e^3$) at energies above a few Kelvin. Also, retaining more energy in the electrons will not change significantly the results here.

Table 1 also lists scattering rates based on the power transfer between electrons and phonons, which has been directly measured in experiments. The power exchange between electrons and phonons within a volume V is given by[28]

$$P_{ep} = \Sigma_{ep}\, V\, (T_e^5 - T_p^5), \quad (11)$$

which looks like the Stefan–Boltzmann radiation law, except here it describes the volume radiation of phonons. The materials constant Σ has been measured for elements and alloys Cu[28], Au[29], CuAu[30], and NiCr and is found to be reasonably constant $\Sigma \simeq 2$ nW μm^{-3} K^{-5}. For aluminum the value is about ten times smaller[31]. The total energy of the electrons U_e and phonons U_p are calculated by their integrating heat capacities, with low-temperature values for normal aluminum and copper listed in Table 2. Along with P_{ep}, approximate scattering rates are

$$\Gamma_p = P_{ep}/U_p \sim T_p \quad (12)$$

$$\Gamma_e = P_{ep}/U_e \sim T_e^3\,. \quad (13)$$

These scattering rates are listed in Table 1. Note that the energy dependence for both the Kaplan and power formulas is the same. The Kaplan and power rates are similar, showing these are decent estimates for the time scale of electron–phonon physics.

Table 3. Scattering lengths.

Energy	20 K	4 K
Material	Cu/Al	Cu/Al
e velocity (mm/ns)	1.57/2.03	
p velocity (μ/ns)	4.8/6.4	
e/q rate (1/ns)	330/3.5	2.7/0.0052
e/q diffuse (μm) [for 3/0.1 μm]	3.8/7.6	42/200
p rate (1/ns)	2.4/20	0.49/4.0
p length (μm)	2.0/0.32	9.8/1.6

Table of electron, quasiparticle, and phonon scattering rates and lengths for an energy 20 K, well above the superconducting gap Δ of aluminum, and for 4 K $\simeq 2\Delta$ at phonon freeze out. The diffusion length scales are for a copper and aluminum thickness of 3 and 0.1 μm, respectively. These rates are only estimates and should be directly measured for proper predictions.

From these electron, quasiparticle, and phonon rates, a length scale for the cascade process can be estimated using the electron and phonon velocities. The results are summarized in Table 3 for a thin 0.1 μm aluminum film and a thicker 3 μm copper film, to be introduced later. The electrons are assumed to diffuse in the metal films; these length scales are typically small compared to phonon lengths.

The most important length is for the phonon interaction, which shows that even at a high energy 20 K, well above the superconducting gap for aluminum, the interaction length (0.32 μm) is somewhat larger than the film thickness (0.1 μm). Thus, the phonon in the aluminum will down-convert only a fraction of the time. When impinging on the aluminum–silicon interface, the phonons will escape the aluminum film about 50% of the time, travel through the silicon wafer, and are then eventually reabsorbed after another impingement into the metal. The rate of scattering lowers as the energy approaches $2\Delta = 4$ K, so that the phonon on average will diffuse laterally by several thicknesses of the silicon substrate, of order a few mm's.

Present qubit design

Superconducting qubits are typically fabricated with superconducting metal fabricated on the surface of a silicon chip. To first calculate the efficiency of high-energy phonons being down-converted to quasiparticles, note that the scattering rates are not important. This is because high-energy phonons ($E_p > 2\Delta$) have to eventually scatter in the superconductor, and quasiparticles eventually have to scatter in the superconductor.

The down-conversion process is numerically simulated using a Monte-Carlo method, where the distribution of final state energies is taken from the integrands of Eqs. (4) and (8). The results are shown in Fig. 4. The quasiparticle down-conversion efficiency $n_{qp}/(E_p/\Delta)$, which is normalized to the maximum possible number E_p/Δ, is plotted versus the initial phonon energy E_p/Δ. As expected, at low phonon energies $E_p < 2\Delta$, there is no breaking of Cooper pairs, but from $2\Delta < E_p < 4\Delta$ one quasiparticle pair is created. The numerical simulation shows that at larger phonon energies the average quasiparticle fraction is 0.57, as found previously[25,32,33], which is reasonable since one expects the initial energy to be shared between both phonons and quasiparticles.

Assuming the qubits are embedded in a superconducting ground plane, typical for integrated circuits, these down-converted quasiparticles are expected to fill some area of the chip. The calculations for the down-conversion are shown in Table 4. An initial radiation energy is assumed to be 0.2 MeV, a typical value for muons and gamma rays. The next row is the number of quasiparticles using the aluminum gap and conversion efficiency 0.57. The quasiparticle density is next computed
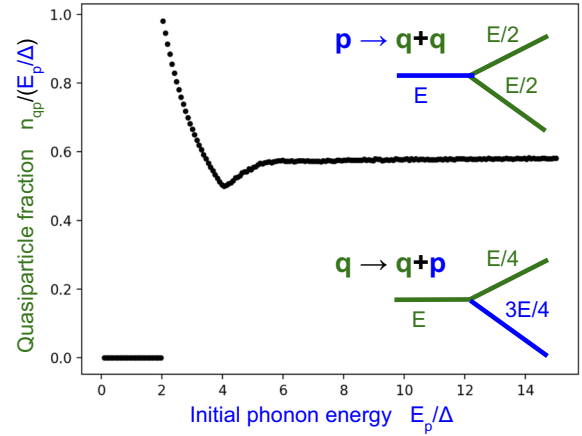

Fig. 4 Phonon to quasiparticle down-conversion efficiency. Plot of normalized number of quasiparticles versus initial phonon energy, solved via numerical simulation for the random energy branching of the scattering processes p→q + q and q→q + p. No quasiparticles are generated for $E_p < 2\Delta$, and 2 for $2\Delta < E_p < 4\Delta$, as expected. At high energies, 57% of the initial phonon energy is converted to quasiparticles on average. The insets depict the scattering process, listing the average kinetic energies.

Table 4. Quasiparticle down-conversion and qubit decay.

Initial radiation energy	0.2 MeV
Number of quasiparticles	$0.67 \cdot 10^9$
Density for 1 cm^2 × 0.1 μm	67/μm^2
Density n_{qp}/n_{cp}	$2.4 \cdot 10^{-5}$
Qubit Q	51 k
T_1 for area 1 cm^2	1.6 μs
T_1 for area 10 mm^2	160 ns

Table of parameters for down-converting radiation energy into quasiparticles, which then damps the qubits. The aluminum film is assumed to be 0.1 μm thick over the entire 1 cm^2 chip. Also listed is T_1 for a smaller area of the quasiparticle hotspot.

using an aluminum film thickness of 0.1 μm and a chip area of 1 cm^2. Next is displayed the fraction density to Cooper pairs, and the estimated transmon quality factor Q. The computed $T_1 = 1.6$ μs from Eq. (3) is short enough to significantly damp the qubit. As discussed previously, the initial hotspot of the down-conversion is several times the silicon thickness, so an area of 10 mm^2 may be more appropriate, in which case the decay time is severe, 160 ns.

With an understanding of the basic scattering physics, the various stages of the energy down-conversion process are described next. We assume a low background temperature (~20 mK) well below the gap energy. The approximate time and length scales are included to estimate spatial and time correlations to the qubit error, along with the qubit decay T_1 to understand its magnitude.

Stage 1, fireball (10 ns, 1 mm): The energy spectrum of gamma and cosmic rays are given in Fig. S2 of ref. [10]. Average cosmic and gamma energies are 0.46 and 0.1 MeV, but range to about 2.5 and 1 MeV, respectively. Cosmic rays deposit energy in a line through the chip, with a spatial extent roughly the thickness of the wafer 400 μm, whereas gamma ray energy is in a roughly spherical volume of radius 50–100 μm. Since these length scales are smaller than for the various down-conversion and diffusion processes, the two types of events can be treated as equivalent, especially since the purpose here is to minimize the effects of radiation, not measure it.

After a cosmic ray or gamma ray interacts with the substrate, high-energy phonons and silicon electron/holes are created. Most of the electron/holes recombine, placing greater than 90% of the initial energy into phonons. The phonon energy can down-convert by itself until about 50 K, when the phonon dispersion relation becomes linear. These phonons will spread through the substrate, moving away from the creation site.

Stage 2, freeze out (300 ns, 3 mm, $T_1 \simeq 160$ ns): The phonons move through the silicon substrate at a velocity of 9.6 µm/ns. When impinging the superconducting metal, about 50% of phonons will transmit into the metal[26]. In aluminum, the scattering length of 0.1–0.3 µm will cause phonons to down-convert with reasonably high probability ($\gtrsim$50%). Thus, the spatial extent of the freeze-out stage is several times the substrate thickness.

The quasiparticles and phonons continue to generate a cascade of scattering events, lowering the energy of the phonons until it drops below $2\Delta \simeq 4$ K. At this time, no additional quasiparticles are generated. About 57% of the initial energy is converted to quasiparticles. As shown in Table 3, the electron diffusion distance is small ~10 µm, so the lateral extent is set by the phonon diffusion as for the fireball stage. With ~10 mm² quasiparticle area, the corresponding T_1 is given in Table 4.

Stage 3, quasiparticle diffusion (100 µs, 6 mm, $T_1 \simeq 0.6$ µs): The quasiparticles at first have energy somewhat larger than Δ and continue to shed their extra energy by emitting phonons. These phonons cannot be reabsorbed since they have energy below 2Δ, and thus bounce around the chip. For an initial quasiparticle energy 1 K above the gap, Table 1 gives a decay time of 1.7 µs, with a velocity slightly below the Fermi value v_e. For a 0.1 µm film, the diffusion distance D_q grows as the square root of time t

$$D_q \simeq \sqrt{v_e t \cdot 0.1 \ \mu m} \qquad (14)$$

$$= \sqrt{t/\mu s} \cdot 0.4 \ mm. \qquad (15)$$

Quasiparticle relaxation times scale as the cube of the energy E above the gap. At 100 µs, the quasiparticle energy is ~0.25 K above Δ. With the quasiparticle velocity dropping as $\sqrt{1 - (\Delta/E)^2}$[34], they will diffuse a distance ~ $0.68 D_q$.

This expanding quasiparticle area decreases their density and increases the qubit T_1 approximately with t. With velocity lowering with relaxation, the diffusion distance is between 2 and 4 mm, giving a resulting quasiparticle radius of about 6 mm. Note that quasiparticles in the superconducting island of a differential transmon will not lower its density by diffusion, so the qubit decay rate $1/T_1$ may not lower over time except by recombination.

Stage 4, quasiparticle recombination and rebreaking (1 ms, chip, $T_1 \gtrsim 1.6$ µs): This stage is concurrent with the last stage of quasiparticle diffusion, but contains additional physics of quasiparticles recombining into Cooper pairs. Being a two-particle process, the rate is proportional to the quasiparticle density as given in Table 1, which produces a density that decreases inversely with time as[21]

$$\frac{n_{qp}}{n_{cp}} = \frac{400 \ ns/43.6}{t + t_0}. \qquad (16)$$

The time is $t + t_0 \sim 100$ µs for the density in Table 4 with area 10 mm², giving the time and size scales for the end of the quasiparticle diffusion stage. Note that density does not decrease exponentially, so recombination is slow and T_1 scales as $1/t$. Assuming constant quasiparticle area, T_1 increases to 20 µs only after 6 ms. Such a long time is roughly consistent with 18 ms reported in ref. [35].

In addition, it is possible that after recombination the resulting phonon has no loss mechanism other that rebreaking another Cooper pair. Thus, the total number of quasiparticles can be roughly constant in this stage. Here, the phonons diffuse throughout the chip because of their long 3.2 µm interaction length in aluminum. The recreated quasiparticles will have lower density, and the qubit $T_1 = 1.6$ µs is now computed using the area 1 cm² of the substrate ground plane.

Because of this effect, one must be careful in discussing the concept of "quasiparticle recombination time," since rebreaking physics might artificially increase this time. It is imperative to describe the design of the chip mount and how phonons escape, as done in ref. [36].

Quasiparticle recombination is more complicated because of trapping effects[34]. The quasiparticle density at the junction dominates qubit loss since the normal resistance of the junction (~100 Ω) is so much higher than wiring (~10 mΩ). Quasiparticles will flow to and trap at the aluminum junction if its gap is less than the wiring and ground plane. Materials defects in the superconducting films can produce localized traps. Aluminum tends not to trap since scattering raises its transition temperature[37], whereas defects in niobium tend to lower its gap. Magnetic vortices also provide sites for lowering the gap for trapping[35].

Stage 5, phonon escape (4 ms, chip): The mechanism for the 2Δ phonons to escape is by connection to the thermal ground of the chip mount, which is typically copper. Our chips are floating to reduce electromagnetic coupling, and only thermalized via wire bonds[38]. The physics for predicting the phonon escape rate is illustrated in Fig. 5, where the phonons are moving throughout the chip of volume V and can only escape through the N_w wire bonds each of area A. As they must diffuse down the wire bond of length L, the exit probability is ℓ/L, where ℓ is the mean free path, here given by the diameter of the wire. The net phonon escape rate is

$$\Gamma_p = (N_w A/V) v_p (\ell/L). \qquad (17)$$

This rate is $\Gamma_p = 1/4$ ms for parameters $N_w = 300$, $A = \pi(12 \ \mu m)^2$, $V = (10 \ mm)^2$ (0.4 mm), $\ell = 25 \ \mu m$, and $L = 2$ mm. At time scales much greater than $1/\Gamma_p$ all the phonons, both 2Δ and below, can escape and return the chip to thermal equilibrium.

It has been assumed here that the superconducting gap of the AlSi wire bonds is slightly higher than for the aluminum in the device, so that 2Δ phonons do not break Cooper pairs in the wire bonds. If they do, the quasiparticles will diffuse in the wire bond with an escape probability that is roughly equal to ℓ/L, making the above estimate still reasonable.

Improved qubit design

The first priority when redesigning the qubit chip is to reduce the density of quasiparticles produced in the freeze-out stage. This can be readily done by adding a normal metal to the substrate so that the phonon energy will be channeled to the normal metal, away from the superconductor. This normal metal also helps the stage of quasiparticle recombination and rebreaking, channeling the 2Δ phonons to the normal metal instead of back into the superconductor.

The effect of the normal metal can be numerically simulated with a similar method as used for Fig. 4, except a fraction of the
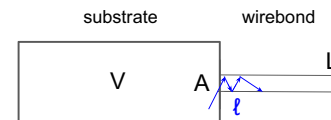

Fig. 5 Device schematic for phonon escape. Phonons are uniformly moving in a substrate of volume V, and can escape through N_w wire bonds each with area A. The escape rate is proportional to the phonon velocity v_p. The phonons must diffuse down the wire bond, eventually escaping with a probability ℓ/L, where L is the wire bond length and ℓ is the mean free path taken as the wire bond diameter. The escape rate is given by Eq. (17).

phonon down-conversion is for electron excitations, not just quasiparticles. Here, the fraction is estimated from the relative thicknesses for the superconducting and normal films because the electron–phonon interaction is roughly equal for many metals. Although precise ratios of phonon scattering rates are needed for quantitative designs, this should give a good starting point.

Figure 6 shows the simulation data for the quasiparticle fraction of down-conversion, plotted as the relative thicknesses of the superconducting (t_s) and normal (t_n) metal. The participation ratio formula $t_s/(t_s + 1.65\, t_n)$ is chosen with the constant 1.65 so that the data lies on a line. These data show that as the fraction of superconducting metal decreases, so does its fraction of quasiparticles from phonon down-conversion, as expected. Prior experiments with thin films (~0.1 μm) have shown modest reductions (2–5) of quasiparticles[14,22,23]. Here, T_1 must increase by about a factor of 100 based on estimates from the last section. Choosing the thickness of the normal metal will ultimately be optimized from experimental measurements of T_1, but the above estimates implies a thickness of about 6 μm. As this is significantly thicker than standard thin-film deposition techniques, electro-deposited copper, silver, or gold film is likely the best solution. These metals also have the advantage of low stress.
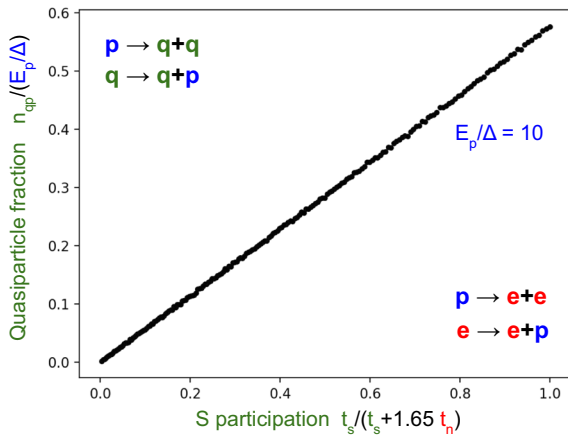

Fig. 6 Reduction of quasiparticles from channeling energy to normal metal. Plot of normalized number of quasiparticles versus participation of the superconductor, showing the reduction of quasiparticles using normal metal. The participation formula is expressed with the thickness of the superconductor t_s and normal t_n metal, fitting the constant 1.65 to give a linear dependence. The initial phonon energy is $E_p/\Delta = 10$. The insets show the four scattering processes for the superconductor and the normal metal.

The natural placement of this normal metal film is the backside of the qubit substrate. However, such a design has problems with electromagnetic radiation. Figure 7a illustrates the idea with a simplified schematic of the design. A differential transmon qubit typically is embedded in a ground plane, so that the normal metal on the backside of the substrate forms a transmission line with low impedance, estimated to be ~6 Ω. The pads of transmon capacitor have an area of about 0.1 mm^2 which form parallel plate capacitors that drive the transmission line mode. Since these capacitors are about 20% of the capacitance of the transmon, the qubit and transmission line are strongly coupled, so one expects a large effect on the transmon. This can be estimated by considering the equivalent circuit of Fig. 7b, where the short transmission line between the capacitors is equivalently ~0.3 nH inductor, with an impedance of 10 Ω at 5 GHz. This has a small effect on the equivalent resistance (6 Ω) representing the transmission line radiation. The effect of the coupling and this resistor is to significantly damp the qubit with $Q \sim 1$ k. Note that with strong coupling, this design will not work with any reasonable change to the transmission line impedance.

A straightforward solution is to break up the normal metal into isolated islands so that the small capacitance between the islands stops the transmission line radiation. As this only modestly decreases the volume of the normal metal, the thermalization physics will not change much. For this design the equivalent circuit is given by Fig. 7b but with the resistance given by that across the copper island ~0.01 Ω. For this low resistance the qubit will be lightly damped with $Q \sim 3$ M. If a higher Q is needed, an improvement would be to deposit a thin superconducting film between the substrate and the normal metal, for example titanium.

An alternative design would be to use a continuous super-conducting ground plane on the backside, with many via connections to the qubit ground plane to short the transmission line mode. Then, a continuous normal metal film could be used on the backside. The chip could even be placed on a copper mount with vacuum grease to better connect the substrate to thermal ground.

There is a remaining serious problem: quasiparticles that have been created still have a recombination time that is much longer than the error-correction cycle (~1 μs), creating errors correlated in time. This time can be estimated based on Eq. (16), which for a qubit $Q = 1$ M would require a time to decay $t + t_0 \sim 10$ ms. This is clearly too long, even if there is some other mechanism that decays the quasiparticles 10–100 times faster.

The solution is to build quasiparticle traps into the qubit layer[39–41], so that any quasiparticles near the Josephson junction can diffuse into a lower-gap superconductor, relax its energy by the emission of
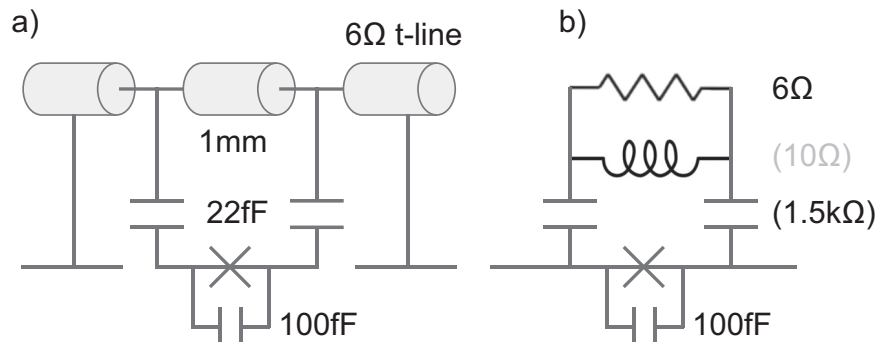

Fig. 7 Equivalent circuit with backside metal. a Circuit showing differential transmon embedded in a ground plane on the qubit layer. The continuous normal metal on the backside then acts as a low impedance transmission line of ~6 Ω impedance, which couples to the transmon via parallel plate capacitors coming from the qubit pads that have an area about 0.1 mm^2. **b** Simplified circuit, replacing the transmission line with an equivalent resistor, and the short transmission line between capacitors with an inductor, with impedances at 5 GHz shown in parenthesis. This heavily damps the qubit with a $Q \sim 1$ k.

a phonon, and thus be trapped away from the junction. For a transition temperature 0.5 K, the quasiparticle scattering time is about 1.7 μs (see Table 1), giving a diffusion length of about 0.52 μm[41]. With such a length, a significant fraction of quasiparticles might leave the trap before down-conversion, so the design of the trap should have a significant volume fraction. In order to keep these quasiparticles away from the edges that are most susceptible to dissipation, a simple solution is to make the islands of the lower-gap superconductor, as illustrated in Fig. 8.

Reference[41] has a good discussion on the constraints on the resistance of the quasiparticle traps and superconductor interface. As stated previously for the aluminum leads, qubit loss scales as component resistance: since both the trap and interface resistances can be quite small, about $10^3 - 10^4$ times less than the junction, their dissipation should be negligible. For a transition temperature of 0.5 K, the gap would still be high enough to ensure low dissipation from thermally generated quasiparticles. Other systems constraints might prevail for the wiring and ground, like choosing superconductors with small surface loss.

For this improved design, the stages of energy decay should be the following:

Stage 1, fireball (10 ns, 1 mm): The initial cascade will have less diffusion laterally because the electron–phonon cascade can down-convert efficiently in the thick normal metal layer.

Stage 2, freeze out (300 ns, 3 mm, $T_1 \simeq 16$ μs): About 99% of the phonon energy will be down-converted in the normal metal. Most phonons will not escape this layer even as their energy drops below 1 K. Some phonons escaping into the silicon and superconductor may diffuse laterally, but less than for the original design. Before the quasiparticle energy in the superconductor relaxes to the gap, the quasiparticle density at the junctions is estimated from the energy partitioning.
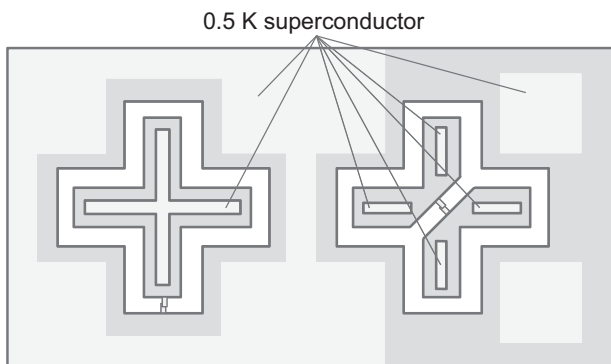
Stage 3, quasiparticle diffusion and down-conversion (1.7 μs 3 mm, $T_1 > 16$ μs): The quasiparticles at the junction should diffuse rapidly into the superconductor wiring. With the distance to the low-gap superconductor islands less than 100 μm, the diffusion time to the traps is short <100 ns. The down-conversion time once in the 0.5 K superconductor is about 1.7 μs. The T_1 of the qubit should reset to its background rate during this time.

Stage 4, quasiparticle recombination (1 ms, chip): During this time, all of the quasiparticles are in the low-gap superconductor, and will recombine over a long time scale as discussed previously. They should have little effect on the qubit because they are neither in the junction nor in the critical edges of the superconducting wiring or ground plane. Phonons created by recombination will be absorbed by the normal metal.

Stage 5, phonon escape (4 ms, chip): Phonons continue down-conversion in the normal metal and eventually escape to the copper chip mount.

A summary of the stages and a comparison between the present and future design is shown in Table 5.

The qubit lifetime is reduced for a time less than one surface code cycle. Qubit errors will be small and weakly correlated in space, maybe over one adjacent qubit. Error correction thus should work properly.

As a side benefit, note that surface loss is presently a dominant damping mechanism, coming from two-level states in surface oxides. They are likely producing the 0.1–1% thermal population observed in the qubit excited state, as the two-level states can be excited by nonequilibrium phonons at a frequency ~ 5 GHz. Strong down-conversion of phonons in the chip should reduce the effective temperature of these or other[8,42] two-level-states and therefore the qubit. Stray quasiparticles are still a concern[8,36,43].

DISCUSSION

Cosmic rays and background gamma radiation is clearly important physics to understand in the systems design of a quantum computer. Although it has been discussed here in detail for superconducting transmon devices, this physics might be important for other low-temperature qubits. Here, are some comments and questions:

For semiconductor qubits, the charge offsets produced by silicon electron–hole pairs[10] is as important as the phonon heating discussed here. This prior work shows that significant charge offsets would be seen over large areas (100 μm) for charge-sensitive transmon devices. Will this introduce correlated qubit errors?

Photonic quantum computers are being designed based on superconducting transition detectors[44]. Might the phonon pulse described here be enough to trip these detectors in a way correlated in time and space, and thus detrimental to error correction?

0.5 K superconductor

Fig. 8 Design of quasiparticle traps. Illustration of quasiparticle traps for a Xmon (left) and differential (right) transmon, where additional islands in the wiring and ground planes are made from a lower-gap superconductor. These islands can be directly connected to the wiring or through a thin tunnel barrier.

Table 5.	Summary of heating event.					
stage	Present			Future		
	Time (μs)	Size (mm)	T_1 (μs)	Time (μs)	Size (mm)	T_1 (μs)
1. Fireball	0.01	1		0.01	1	
2. Freeze out	0.3	3	0.16	0.3	3	16
3. Qp diffusion	100	6	0.6	1.7	3	>16
4. Qp recom. & rebrk.	1000	chip	>1.6	1000	chip	bl
5. Phonon escape	4000	chip		4000	chip	bl

For the five stages, the table summarizes the time scale, size scale, and qubit decay time. Shown is a comparison of present and improved design. The entry "bl" indicates T_1 decay approaches the baseline value.

Majorana qubits are made from a special quasiparticle state of a super- and semiconductor. It would seem that one stray quasiparticle could disrupt this protected state[45,46]. Is it possible to reduce the quasiparticle number from billions to much less than one so that these states might stay protected?

More generally, will errors from radiation events disrupt any error-correction scheme that does not use many qubits to protect the state? Because radiation events produce nonequilibrium energy pulses of size about 100 μm, are large qubits necessary for errors not to be spatially correlated?

Research into thermal particle detectors has enabled low-temperature physicists to understand how radiation injects energy into quantum devices. The energy pulses from cosmic and gamma rays are clearly an important issue for superconducting qubits, since qubit errors and correlations in time and space will kill error correction by many orders of magnitude. A model is presented here for this physics, and a solution based on channeling this energy away from the qubit and into benign structures like thick normal metals and quasiparticle traps. There are many interesting experiments to do soon to demonstrate that an effective solution is possible.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

CODE AVAILABILITY

The Python code that supports the findings of this study is available from the corresponding author upon reasonable request.

REFERENCES

1. Shor, P. W. Algorithmic Number Theory: First International Symposium, ANTS-I (eds Adleman, L. M. & Huang, M.-D. A.) 289 (Springer, 1994). Lecture Notes in Computer Science, 877.
2. Berry, D. W., Gidney, C., Motta, M., McClean, J. R. & Babbush, R. Qubitization of arbitrary basis quantum chemistry leveraging sparsity and low rank factorization. Quantum 3, 208 (2019).
3. Preskill, J. Quantum computing in the NISQ era and beyond. Quantum 2, 79 (2018).
4. Arute, F. Quantum supremacy using a programmable superconducting processor. Nature 574, 505 (2019).
5. Bravyi, S. B. & Kitaev, A. Y. Quantum codes on a lattice with boundary. Preprint at https://arxiv.org/abs/quant-ph/9811052 (1998).
6. Fowler, A. G., Mariantoni, M., Martinis, J. M. & Cleland, A. N. Surface codes: towards practical large-scale quantum computation. Phys. Rev. A 86, 032324 (2012).
7. Martinis, J. M. Qubit metrology for building a fault-tolerant quantum computer. NPJ Quantum Inf. 1, 15005 (2015).
8. Serniak, K. Hot nonequilibrium quasiparticles in transmon qubits. Phys. Rev. Lett. 121, 157701 (2018).
9. Vepsäläinen, A. P. Impact of Ionizing radiation on superconductinng qubit coherence. Nature 584, 551 (2020).
10. Wilen, C. D. et al. Correlated charge noise and relaxation errors in superconducting qubits. Preprint at https://arxiv.org/abs/2012.06029 (2012).
11. Irwin, K. D., Hilton, G. C., Wollman, D. A. & Martinis, J. M. X-ray detection using a superconducting transition-edge sensor microcalorimeter with electrothermal feedback. Appl. Phys. Lett. 69, 1945 (1996).
12. Grunhaupt, L. Loss mechanisms and quasiparticle dynamics in superconducting microwave resonators made of thin-film granular aluminum. Phys. Rev. Lett. 121, 117001 (2018).
13. Cardani, L. et al. Reducing the impact of radioactivity on quantum circuits in a deep-underground facility. Preprint at https://arxiv.org/abs/2005.02286 (2005).
14. Karatsu, K. Mitigation of cosmic ray effect on microwave kinetic inductance detector arrays. Appl. Phys. Lett. 114, 032601 (2019).
15. McEwen, M. et al. Resolving catastrophic error bursts from cosmic rays in large arrays of superconducting qubits. Preprint at https://arxiv.org/abs/2104.05219 (2021).
16. Sergeev, A. & Mitin, V. Phonon traps reduce the quasiparticle density in superconducting circuits. Appl. Phys. Lett. 80, 817 (2002).
17. De Visser, P. J. Generation-recombination noise: the fundamental sensitivity limit for kinetic inductance detectors. J. Low. Temp. Phys. 167, 335 (2012).
18. Catelani, G., Schoelkopf, R. J., Devoret, M. H. & Glazman, L. I. Relaxation and frequency shifts induced by quasiparticles in superconducting qubits. Phys. Rev. B 84, 064517 (2011).
19. Shutt, T. Simultaneous high resolution measurement of phonons and ionization created by particle interactions in a 60 g germanium crystal at 25 mK. Phys. Rev. Lett. 69, 3531 (1992).
20. Martinis, J. M., Ansmann, M. & Aumentado, J. Energy decay in superconducting josephson-junction qubits from nonequilibrium quasiparticle excitations. Phys. Rev. Lett. 103, 097002 (2009).
21. Lenander, M. Measurement of energy decay in superconducting qubits from nonequilibrium quasiparticles. Phys. Rev. B 84, 024501 (2011).
22. Valenti, F. Interplay between kinetic inductance, nonlinearity, and quasiparticle dynamics in granular aluminum microwave kinetic inductance detectors. Phys. Rev. Appl. 11, 054087 (2019).
23. Henriques, F. Phonon traps reduce the quasiparticle density in superconducting circuits. Appl. Phys. Lett. 115, 212601 (2019).
24. Koch, J. Charge-insensitive qubit design derived from the Cooper pair box. Phys. Rev. A 76, 042319 (2007).
25. Kozorezov, A. G. et al. Quasiparticle-phonon downconversion in nonequilibrium superconductors. Phys. Rev. B 61, 11807 (2000).
26. Martinez, M. et al. Measurements and simulations of athermal phonon transmission from silicon absorbers to aluminum sensors. Phys. Rev. Appl. 11, 06402 (2019).
27. Kaplan, S. B. Quasiparticle and phonon lifetimes in superconductors. Phys. Rev. B 14, 4854 (1976).
28. Roukes, M. L., Freeman, M. R., Germain, R. S., Richardson, R. C. & Ketchen, M. B. Hot electrons and energy transport in metals at millikelvin temperatures. Phys. Rev. Lett. 55, 422 (1985).
29. Schmidt, D. R., Yung, C. S. & Cleland, A. N. Temporal measurement of hot-electron relaxation in a phonon-cooled metal island. Phys. Rev. B 69, 140301 (2004).
30. Wellstood, F. C., Urbina, C. & Clarke, J. Hot-electron effects in metals. Phys. Rev. B 49, 5942 (1994).
31. Kautz, R. L., Zimmerli, G. & Martinis, J. M. Self-heating in the coulomb-blockade electrometer. J. Appl. Phys. 73, 2386 (1993).
32. Kurakado, M. Possibility of high resolution detectors using superconducting tunnel junctions. Nucl. Instrum. Method 196, 275 (1982).
33. Rando, N. The properties of niobium superconducting tunneling junctions as X-ray detectors. Nucl. Instrum. Method 313, 173 (1992).
34. Aumentado, J., Keller, M. W., Martinis, J. M. & Devoret, M. H. Nonequilibrium quasiparticles and 2e periodicity in single-cooper-pair transistors. Phys. Rev. Lett. 92, 066802 (2004).
35. Wang, C. Measurement and control of quasiparticle dynamics in a superconducting qubit. Nat. Commun. 5, 5836 (2014).
36. Vool, U. Non-Poissonian quantum jumps of a fluxonium qubit due to quasiparticle excitations. Phys. Rev. Lett. 113, 247001 (2014).
37. Cohen, R. W. & Abeles, B. Superconductivity in granular aluminum films. Phys. Rev. 168, 144 (1968).
38. Wenner, J. Wirebond crosstalk and cavity modes in large chip mounts for superconducting qubits. Supercond. Sci. Technol. 24, 065001 (2011).
39. Lang, K. M., Nam, S., Aumentado, J., Urbina, C. & Martinis, J. M. Banishing quasiparticles from Josephson-junction qubits: why and how to do it. IEEE Trans. Appl. Supercond. 13, 989 (2003).
40. Riwar, R.-P. Normal-metal quasiparticle traps for superconducting qubits. Phys. Rev. B 94, 104516 (2016).
41. Riwar, R.-P. & Catelani, G. Efficient quasiparticle traps with low dissipation through gap engineering. Phys. Rev. B 100, 144514 (2019).
42. de Graaf, S. E. Two-level systems in superconducting quantum devices due to trapped quasiparticles. Sci. Adv. 6, eabc5055 (2020).
43. Houzet, M., Serniak, K., Catelani, G., Devoret, M. H. & Glazman, L. I. Photon-assisted charge-parity jumps in a superconducting qubit. Phys. Rev. Lett. 123, 107704 (2019).
44. Verevkin, A., Zhang, J. & Sobolewski, R. Detection efficiency of large-active-area NbN single-photon superconducting detectors in the ultraviolet to near-infrared range. Appl. Phys. Lett. 80, 4687 (2002).
45. Rainis, D. & Loss, D. Majorana qubit decoherence by quasiparticle poisoning. Phys. Rev. B 85, 174533 (2012).
46. Karzig, T., Cole, W. S. & Pikulin, D. I. Quasiparticle poisoning of Majorana qubits. Preprint at https://arxiv.org/abs/2004.01264 (2020).

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to J.M.M.

Reprints and permission information is available at http://www.nature.com/reprints