Normal-metal quasiparticle traps for superconducting qubits

R.-P. Riwar,[1,2] A. Hosseinkhani,[1,3] L. D. Burkhart,[2] Y. Y. Gao,[2] R. J. Schoelkopf,[2] L. I. Glazman,[2] and G. Catelani[1]

[1]*Peter Grünberg Institut (PGI-2) and JARA Institute for Quantum Information, Forschungszentrum Jülich, 52425 Jülich, Germany*

[2]*Departments of Physics and Applied Physics, Yale University, New Haven, Connecticut 06520, USA*

[3]*JARA-Institute for Quantum Information, RWTH Aachen University, D-52056 Aachen, Germany*

The presence of quasiparticles in superconducting qubits emerges as an intrinsic constraint on their coherence. While it is difficult to prevent the generation of quasiparticles, keeping them away from active elements of the qubit provides a viable way of improving the device performance. Here we develop theoretically and validate experimentally a model for the effect of a single small trap on the dynamics of the excess quasiparticles injected in a transmon-type qubit. The model allows one to evaluate the time it takes to evacuate the injected quasiparticles from the transmon as a function of trap parameters. With the increase of the trap size, this time decreases monotonically, saturating at the level determined by the quasiparticles diffusion constant and the qubit geometry. We determine the characteristic trap size needed for the relaxation time to approach that saturation value.

I. INTRODUCTION

Ideal superconducting devices rely on dissipationless tunneling of Cooper pairs across a Josephson junction. For example, in a Cooper pair pump [1], the controlled transport of Cooper pairs across two or more junctions can in principle make it possible to relate frequency and current and hence enable metrological applications of such a device [2]. For quantum information purposes, the nonlinear relation between the supercurrent and the phase difference across a junction makes the junction an ideal nonlinear element to build a qubit [3]. However, in addition to the pairs tunneling, single-particle excitations known as quasiparticles can also tunnel. In the pumps this leads to "counting errors," limiting the accuracy of the current-frequency relation [1,2]. In qubits, quasiparticles interact with the phase degree of freedom, providing an unwanted channel for the qubit energy relaxation [4,5]. While in many cases it is impossible to prevent the creation of quasiparticles, one may keep them away from the Josephson junctions by trapping. Evacuation of the quasiparticles from the vicinity of the junction provides a way to extend the energy relaxation time (T_1) in the steady state, and to restore the steady state after a perturbation, whether caused by qubit operation or some uncontrolled environmental effect.

Quasiparticle trapping has been explored for a long time, and various proposals exist on how to implement such a trapping. For example, gap engineering takes advantage of the fact that quasiparticles accumulate in regions of the lower gap to steer them into or away from certain parts of the device. Gap engineering was used successfully to limit quasiparticle "poisoning" in a Cooper pair transistor [6], while it proved ineffective in a transmon qubit [7]. A vortex in a superconducting film can also act as a well-localized trap, since the gap is completely suppressed at the vortex position. Trapping by vortices has been demonstrated [8–11], but vortex motion may induce an unwanted dissipation. An island of a normal metal in contact with the superconductor may also serve as a quasiparticles trap [12,13]. In the limit of weak electron tunneling across the contact, the proximity effect is negligible. The quasiparticles tunneled into the normal metal are trapped there upon losing their energy by phonon emission or inelastic electron-electron scattering.

The majority of previous works concentrated on the control of a steady-state quasiparticle population [2,12,13]. In contrast, we are interested in the effect of a normal-metal trap on the dynamics of the quasiparticle density. Traps accelerate the evacuation of the excess quasiparticles injected in a qubit in the process of its operation. Our main goal is to determine how the characteristic time of the evacuation depends on the parameters of a small normal-metal island in contact with the superconducting qubit. The characteristic time shortens with the increase of the trap size, saturating at a value dependent on the qubit geometry and the quasiparticle diffusion coefficient. The size at which a trap becomes effective depends on the contact resistance, the energy relaxation rate in the normal-metal island, and the effective temperature of the quasiparticles. We develop a simple model allowing us to evaluate the time evolution of the quasiparticle density and find the characteristic evacuation time as a function of the trap parameters. The model is validated by measurements of the qubit T_1 relaxation time performed on a series of transmons with normal-metal traps of various sizes.

The paper is organized as follows: In Sec. II we develop a phenomenological quasiparticle diffusion and trapping model which includes the effect of a normal-metal trap. In Sec. III we study the dynamics of the density during injection and trapping in a simple configuration, and in Sec. IV we provide experimental data supporting our approach. We summarize the present work in Sec. V.

II. THE DIFFUSION AND TRAPPING MODEL

Let us consider a quasiparticle trap made of a normal (N) metal covering part of a superconducting (S) qubit. The contact between the two superconductors and the normal trap is provided by an insulating (I) layer characterized by a small electron transmission coefficient. In order to relate the quasiparticle tunneling rate to the conductance of the contact, we use the tunneling Hamiltonian formalism applied to a model N-I-S system, see Fig. 1,
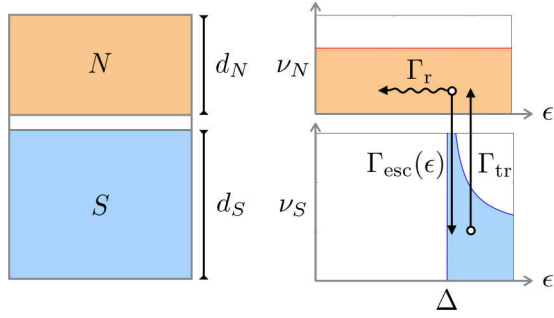
$$H = H_{\mathrm{qp}} + H_N + H_T, \tag{1}$$

FIG. 1. Left: A small superconductor S of thickness d_S separated from a normal metal N of thickness d_N by an insulating layer. Right: Depiction of the processes leading to trapping: tunneling from S to N with rate Γ_{tr} and from N to S with rate $\Gamma_{\text{esc}}(\epsilon)$, and relaxation in N with rate Γ_{r}.

$$H_{\text{qp}} = \sum_{n\sigma} \epsilon_n \gamma_{n\sigma}^\dagger \gamma_{n\sigma}, \tag{2}$$

$$H_N = \sum_{m\sigma} \xi_m c_{m\sigma}^\dagger c_{m\sigma}, \tag{3}$$

$$H_T = \frac{\widetilde{t}}{\sqrt{\Omega_N \Omega_S}} \sum_{m,n,\sigma} (c_{m\sigma}^\dagger d_{n\sigma} + d_{n\sigma}^\dagger c_{m\sigma}). \tag{4}$$

We denote with $\Omega_{N,S} = A \times d_{N,S}$ the volumes of the N and S layers, respectively (A is the area of interface, and $d_{N,S}$ are the layer thicknesses); $c_{m\sigma}^\dagger$ and $d_{n\sigma}^\dagger$ are the creation operators for electrons in the normal metal (energy ξ_m and spin σ) and superconductor. The electron operators in the superconductor are related by Bogoliubov's transformation to the quasiparticle annihilation (creation) operators $\gamma_{n\sigma}^{(\dagger)}$,

$$d_{n\uparrow} = u_n \gamma_{n\uparrow} + v_n \gamma_{n\downarrow}^\dagger, \tag{5}$$

$$d_{n\downarrow}^\dagger = -v_n \gamma_{n\uparrow} + u_n \gamma_{n\downarrow}^\dagger, \tag{6}$$

$$u_n^2 = 1 - v_n^2 = \frac{1}{2}\left(1 + \frac{\xi_n}{\epsilon_n}\right). \tag{7}$$

Here $\epsilon_n = \sqrt{\xi_n^2 + \Delta^2}$ is the energy of a quasiparticle, and ξ_n is the energy of an electron in the normal state of the superconductor. The tunneling constant $\widetilde{t}$ can be related, by Fermi's golden rule, to the resistance R_T of the contact,

$$\frac{R_q}{2\pi R_T} = 4\pi |\widetilde{t}|^2 \nu_{S0}\nu_{N0}, \quad R_q = \frac{2\pi\hbar}{e^2}, \tag{8}$$

where ν_{N0} and ν_{S0} are the densities of states in the normal metal and in the (normal state of the) superconductor, respectively. The tunnel conductance, $1/R_T$, is proportional to the area A of the junction; the intensive quantity characterizing the insulating layer is its conductance per unit area, $1/R_T A$.

We may use Fermi's golden rule to evaluate also the rates of tunneling-induced change of the occupation factors of electrons, $f(\xi_m) = \sum_\sigma \langle c_{m\sigma}^\dagger c_{m\sigma}\rangle$, and quasiparticles, $f_{\text{qp}}(\epsilon_n) = \sum_\sigma \langle \gamma_{n\sigma}^\dagger \gamma_{n\sigma}\rangle$. We can distinguish two processes. Quasiparticles tunnel from the superconductor into the normal metal with rate $\Gamma_{\text{tr}} = 2\pi |\widetilde{t}|^2 \nu_{N0}/\Omega_S$. The transition rate is proportional to the density of the final states involved in the

transition; therefore the quasiparticle trapping rate does not have a pronounced energy dependence. The complementary process of a nonequilibrium electron escape into the superconductor, however, does display a strong energy dependence associated with the BCS singularity in the density of final states, $\Gamma_{\text{esc}}(\epsilon) = 2\pi |\widetilde{t}|^2 \nu_{S0}\nu_S(\epsilon)/\Omega_N$; here

$$\nu_S(\epsilon) = \frac{\epsilon}{\sqrt{\epsilon^2 - \Delta^2}} \tag{9}$$

is the normalized BCS density of states.

One can see from Eq. (8) that the rates Γ_{tr} and $\Gamma_{\text{esc}}(\epsilon)$ are independent of the area A at fixed conductance per unit area of the insulating layer. We may express the rates as

$$\Gamma_{\text{tr}} = \widetilde{\gamma}_{\text{tr}}/d_S, \quad \Gamma_{\text{esc}}(\epsilon) = \widetilde{\gamma}_{\text{esc}}(\epsilon)/d_N \tag{10}$$

in terms of quantities independent of geometry, $\widetilde{\gamma}_{\text{tr}}$ and $\widetilde{\gamma}_{\text{esc}}$,

$$\widetilde{\gamma}_{\text{tr}} = \frac{R_q}{4\pi(R_T A)\nu_{S0}}, \quad \widetilde{\gamma}_{\text{esc}} = \frac{R_q \nu_S(\epsilon)}{4\pi(R_T A)\nu_{N0}}, \tag{11}$$

with $(R_T A)$ being the contact resistance times the area of the contact. This product, with units of $\Omega\,\text{cm}^2$, is independent of A, being inversely proportional to the transmission coefficient through the insulating barrier.

The above formulas enable us to estimate the trapping and escape rates for an aluminum-copper interface for a typical experimental setup (cf. Sec. IV): aluminum has a density of states $\nu_{S0} = 0.73 \times 10^{47}/\text{J m}^3$ [14] and a direct measurement of the contact resistance yields $(R_T A) \sim 430\,\Omega\,\mu\text{m}^2$ (this corresponds to the transmission coefficient of order 10^{-5}). Taking $d_S \sim 80$ nm we find, using Eqs. (10) and (11), $\Gamma_{\text{tr}} \sim 8\times10^6\,\text{s}^{-1}$. The escape rate saturates at an energy-independent value, $\Gamma_{\text{esc}}(\epsilon) \to \Gamma_{\text{esc}}$ at energies $\epsilon \gg \Delta$. Since $d_S \approx d_N$ and $\nu_{S0} \approx \nu_{N0}$ in a typical experiment, one has $\Gamma_{\text{esc}} \approx \Gamma_{\text{tr}}$.

In writing the rate equations for the energy distribution functions of electrons and quasiparticles, we assume the continuum limit for energies ξ_m and ϵ_n. It is convenient to define the probability density to find an electron (quasiparticle) in the normal metal (superconductor) with energy $\epsilon \geqslant \Delta$ as

$$p_N(\epsilon) = \frac{\nu_{N0}\Omega_N}{\nu_{S0}\Omega_S} f(\epsilon), \tag{12}$$

$$p_S(\epsilon) = \nu_S(\epsilon) f_{\text{qp}}(\epsilon). \tag{13}$$

Without loss of generality, we normalize the probability with respect to $\nu_{S0}\Omega_S$. Note that eventually, the experimentally accessible quantity is the normalized quasiparticle density, which can be derived from p_S as

$$x_{\text{qp}} = \frac{2}{\Delta} \int_\Delta^\infty d\epsilon\, p_S(\epsilon). \tag{14}$$

In the absence of spatial dispersion of the distribution functions, the rate equations read (see Appendix A)

$$\dot{p}_N(\epsilon) = \Gamma_{\text{tr}} p_S(\epsilon) - \Gamma_{\text{esc}}(\epsilon) p_N(\epsilon) - \Gamma_{\text{r}} p_N(\epsilon), \tag{15}$$

$$\dot{p}_S(\epsilon) = \Gamma_{\text{esc}}(\epsilon) p_N(\epsilon) - \Gamma_{\text{tr}} p_S(\epsilon). \tag{16}$$

The terms proportional to Γ_{tr} describe trapping of quasiparticle excitations in the normal metal, and those proportional to $\Gamma_{\text{esc}}(\epsilon)$ the possible escape of electron excitations back to the

superconductor; these events take place with rates described by Eqs. (10) and (11).

Since the tunneling process is elastic, excitations appear in the normal metal at energies close to the gap Δ. At low temperature $T \ll \Delta$, there are many unoccupied states below Δ in the normal metal, into which the excitations can decay. These inelastic processes are mediated by electron-electron and electron-phonon interactions and lead to relaxation, which we capture in Eq. (15) with the phenomenological rate Γ_r. All the processes included in the rate equations (15) and (16) are represented in the right panel of Fig. 1.

If the relaxation is immediate, the quasiparticles get trapped in the normal metal with rate Γ_{tr}. However, the relaxation rate Γ_r due to electron-electron and electron-phonon interactions in the normal metal is of course finite. It has been estimated in the supplementary to Ref. [10] to be $\Gamma_r \sim 10^7$ s^{-1}; the measurements reported in Ref. [15] lead to a relaxation rate for electron-phonon interaction of the same order of magnitude, while an estimate based on Ref. [16] yields the faster relaxation rate $\Gamma_r \sim 10^8$ s^{-1}. In all cases, relaxation cannot be assumed immediate in comparison with the trapping and escape rates estimated above, especially taking into account that the escape rate quickly increases for energies approaching the gap due to the divergent BCS density of states in Eq. (9). In fact, for some energy interval close to the gap, the escape rate dominates the quasiparticle dynamics, such that the excitations do not have enough time to relax. Therefore, we cannot in general neglect the backflow of excitations from the normal trap to the superconductor.

The backflow may result in an effective rate which is slower than Γ_{tr}. Assuming a steady-state distribution of nonequilibrium electrons in the normal layer, we set $\dot{p}_N = 0$ in Eq. (15) and solve for p_N in terms of p_S (see also Appendix B). Substituting the solution into Eq. (16) and integrating over energy, we arrive at

$$\dot{x}_{qp} = -\Gamma_{eff} x_{qp}, \tag{17}$$

with the effective trapping rate defined by

$$\Gamma_{eff} = \frac{1}{\int_\Delta^\infty d\epsilon\, p_S(\epsilon)} \int_\Delta^\infty d\epsilon \frac{\Gamma_{tr}\Gamma_r}{\Gamma_{esc}(\epsilon) + \Gamma_r} p_S(\epsilon). \tag{18}$$

It is clear that Γ_{eff} is suppressed to a level below Γ_{tr}. The level of suppression depends on the typical width of the quasiparticle distribution function in energy space. Assuming $p_S(\epsilon)$ is characterized by an effective temperature, $T \ll \Delta$, we find that the trapping is not suppressed, $\Gamma_{eff} \approx \Gamma_{tr}$, only if the energy relaxation is fast enough ($\Gamma_r \gg (\Delta/T)^{1/2}\Gamma_{esc}$); in this case excitations in the normal metal quickly relax to energies below the gap and cannot return into the superconductor. In the opposite case ($\Gamma_r \lesssim (\Delta/T)^{1/2}\Gamma_{esc}$), the effective rate becomes T-dependent and suppressed below the nominal trapping rate, $\Gamma_{eff} \approx (2T/\pi\Delta)^{1/2}\Gamma_{tr}\Gamma_r/\Gamma_{esc}$. Note that in the slow relaxation regime the effective trapping rate Γ_{eff} is independent of the tunneling probability between superconductor and normal metal, the limiting value of Γ_{eff} being proportional to the relaxation rate.

The quasistatic approximation ($\dot{p}_N = 0$) we used above becomes justified once we move from the model system of Fig. 1 to a more realistic geometry of a long superconducting strip in
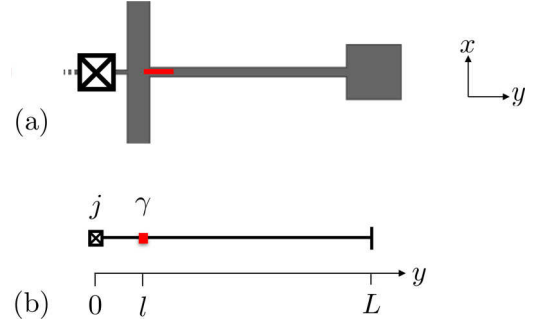

FIG. 2. (a) Figure of a realistic transmon qubit device close to the proportions of experiment. The Josephson junction is indicated with the crossed box, in gray is the superconductor, and in red the normal-metal trap. Shown is half the qubit (the dashed lines indicate that the superconducting structure including trap is mirrored on the left-hand side of the junction). (b) Simplified model of a 1D superconducting strip with small trap, described by Eq. (22).

contact with a metallic trap; see Fig. 2(a). In that geometry, the time variation of the quasiparticle distribution function p_S is controlled by the diffusion time in the strip, which is typically substantially longer than $1/\Gamma_r$. The generalization of the rate equations (15) and (16) to include diffusion is performed in Appendix B. In addition to diffusion, other processes such as quasiparticle recombination, generation, and trapping in the bulk must be generally taken into account. For sufficiently thin normal and superconducting layers, we find a generalized diffusion equation for the quasiparticle density x_{qp},

$$\dot{x}_{qp} = D_{qp}\nabla^2 x_{qp} - a(x,y)\Gamma_{eff} x_{qp} - r x_{qp}^2 - s_b x_{qp} + g, \tag{19}$$

where $x_{qp}(x,y)$ depends only on coordinates in the plane of the superconducting strip (and is assumed constant across its thickness) and the area function $a(x,y)$ equals 1 for x and y where the trap and the superconductor are in contact, and 0 elsewhere; see Fig. 2(a).

The diffusion constant D_{qp} in Eq. (19) is proportional to the normal-state diffusion constant for the electrons in the superconductor; the proportionality coefficient can in principle be calculated from the detailed information on the energy distribution of quasiparticles that we discard in using the phenomenological Eq. (19). The recombination term $r x_{qp}^2$ accounts for processes in which two quasiparticles recombine into a Cooper pair [17], again neglecting the details of the quasiparticle distribution. The relationship between recombination time, quasiparticle energy, and electron-phonon interaction strength can be found in Ref. [18]. Moreover, there is a background trapping term $s_b x_{qp}$ that describes any process that can localize a quasiparticle and hence remove its contribution to the bulk density x_{qp}. Trapping by vortices is an example of such a process, recently characterized in Ref. [10]. The generation rate g describes pair-breaking processes, both thermal and nonthermal; at low temperatures, nonequilibrium processes of unknown origin lead to a quasiparticle density orders of magnitude larger than the thermal equilibrium one [19–21].

In what follows we will neglect both background trapping and recombination: according to the measurements in Ref. [10] we expect $s_b < 0.2 \times 10^3$ s^{-1} as well as $r x_{qp} < 1.25 \times 10^3$ s^{-1} (having assumed $x_{qp} < 10^{-4}$). Both processes are orders of magnitudes slower than the effective trapping rate Γ_{eff}, even when the latter is highly reduced by backflow. Indeed, even for a low effective temperature $T = 10$ mK, using $\Gamma_r \sim 10^7$ s^{-1} and $\Delta/h = 44$ GHz for aluminum, we find $\Gamma_{eff} \sim 0.55 \times 10^6$ s^{-1}. Finally, we assume a long wire geometry, where the dimensions of the system in the x and z directions are sufficiently small such that the superconductor can be treated as (quasi-)one-dimensional, and we consider traps that are small (in a sense to be specified below), so that they are effectively zero-dimensional. In this case, from Eq. (19) we obtain

$$\dot{x}_{qp} = D_{qp} \partial_y^2 x_{qp} - \gamma \delta(y - l) x_{qp} + g, \qquad (20)$$

where the trap is at position $y = l$ and $\gamma = \Gamma_{eff} d$, with d the length of the trap in the y direction. To estimate when the trap is sufficiently small, we note that the trapping length

$$\lambda_{tr} \equiv \sqrt{D_{qp}/\Gamma_{eff}} \qquad (21)$$

gives the scale over which the density decays due to trapping, so the smallness condition is $d \ll \lambda_{tr}$. In the next section we study the dynamics of the quasiparticle density by solving Eq. (20) in various regimes.

III. QUASIPARTICLE DYNAMICS DURING INJECTION AND TRAPPING

In this section we compute the dynamics of the quasiparticle density in a simple geometry depicted in Fig. 2(b). It models a transmon qubit in Fig. 2(a) by neglecting for simplicity both the gap capacitor near the Josephson junction and the square pad at the opposite end of the long wire. Note that because of the spatial symmetry, it is sufficient to consider only half of the system, $0 \leqslant y \leqslant L$. After separating out the steady-state background density due to the finite generation rate g, the equation controlling the evolution of the excess density of quasiparticles takes the form

$$\frac{\partial x_{qp}(y,t)}{\partial t} = \left[D_{qp} \frac{\partial^2}{\partial y^2} - \gamma \delta(y - l) \right] x_{qp}(y,t) + j \delta(y - 0^+) \theta(-t) \theta(t + t_{inj}). \qquad (22)$$

This diffusion equation is supplemented by the boundary conditions $\partial_y x_{qp}(L,t) = 0$ and $\partial_y x_{qp}(0,t) = 0$. The former condition ensures that no quasiparticles leave the device (hard wall condition), while the latter reflects the spatial symmetry of the system.

In the experiments, quasiparticles are generated at the Josephson junction when injecting a high-power microwave pulse into the cavity hosting the qubit [10], resulting in a time-dependent source of quasiparticles localized at $y = 0$. In Eq. (22), this source is modeled by a term with a generation rate proportional to j active over the time interval $-t_{inj} < t < 0$. Clearly, there are two stages of time evolution: First, during the injection process, when the source term is switched on, the quasiparticle density will start to rise and distribute across the wire. Once the source term is switched off, the presence of the normal-metal trap ensures the decay of the excess density

back to zero. In the following, we provide analytical results for the time-dependent dynamics of the quasiparticle density, where we focus predominantly on the experimentally accessible [10] density at the junction, $y = 0$.

The time-dependent diffusion equation (22) can be solved via a decomposition in the modes $e^{\lambda_k t} n_k(y)$ of the homogeneous equation [i.e., Eq. (22) without the source term], with λ_k being the eigenvalue and n_k satisfying equation

$$\lambda_k n_k(y) = \left[D_{qp} \frac{\partial^2}{\partial y^2} - \gamma \delta(y - l) \right] n_k(y). \qquad (23)$$

For a strip of finite length L, the eigenvalues are discrete and the eigenmodes form an orthonormal basis,

$$\int_0^L \frac{dy}{L} n_k(y) n_{k'}(y) = \delta_{kk'}. \qquad (24)$$

In the presence of the trap at $y = l$, the eigenmodes are defined piecewise as

$$n_k(y) = \frac{1}{\sqrt{N_k}} \begin{cases} \cos(ky), & y < l, \\ a_k \cos(ky) + b_k \sin(ky), & y > l, \end{cases} \qquad (25)$$

with the normalization constant N_k (which will be provided explicitly later in some limiting cases) and the coefficients

$$a_k = 1 - \frac{\gamma}{D_{qp} k} \cos(kl) \sin(kl),$$

$$b_k = \frac{\gamma}{D_{qp} k} \cos^2(kl). \qquad (26)$$

The eigenvalue corresponding to eigenmode k is $\lambda_k = -D_{qp} k^2$. The boundary condition at $y = 0$ is satisfied by Eq. (25), while the one at $y = L$ gives the equation

$$\cot(kL) = \frac{1 - \frac{\gamma}{D_{qp} k} \cos(kl) \sin(kl)}{\frac{\gamma}{D_{qp} k} \cos^2(kl)}, \qquad (27)$$

which fixes the wave vector k to discrete values.

In terms of the eigenbasis introduced above, by solving Eq. (22) we find that the excess quasiparticle density immediately after the injection, at time $t = 0$, is given by

$$x_{qp}(y,0) = \sum_k c_k \frac{e^{\lambda_k t_{inj}} - 1}{\lambda_k} n_k(y) \qquad (28)$$

with

$$c_k = j \int_0^L \frac{dy}{L} n_k(y) \delta(y - 0^+) = \frac{j}{L} n_k(0), \qquad (29)$$

where we assumed that at times $t < -t_{inj}$, there were no excess quasiparticles in the system. Once the injection stage is finished, the subsequent trapping of the quasiparticles controls the evolution of their density,

$$x_{qp}(y,t) = \frac{j}{L} \sum_k n_k(0) e^{-D_{qp} k^2 t} \frac{1 - e^{-D_{qp} k^2 t_{inj}}}{D_{qp} k^2} n_k(y). \qquad (30)$$

The expressions for $x_{qp}(y,t)$ derived here are general and do not rely on any further simplifying assumption. Next, we consider in more detail several limiting cases.

A. The long-strip limit

If both the injection time t_{inj} and the time t after injection are short compared to the diffusion time scale $\sim L^2/D_{\text{qp}}$, the generated quasiparticles do not reach the far end of the strip, and we may take the limit $L \to \infty$. In this limit, all values of k are allowed and sums over k are replaced by an integral, $\frac{1}{L}\sum_k \to \int \frac{dk}{2\pi}$. Moreover, when letting $L \to \infty$ while keeping the distance l between trap and junction finite, the normalization constant N_k is dominated by the part of the mode with $y > l$, so that $N_k \simeq (a_k^2 + b_k^2)/2$. Clearly, a single trap suppresses the excess quasiparticle density at the junction best if the distance l is short. For simplicity, from now on we assume $l \to 0^+$. That leaves us with only one characteristic time scale, the saturation time

$$t_{\text{sat}} = D_{\text{qp}}/\gamma^2. \tag{31}$$

It gives the time scale over which the density near the junction approaches its steady-state value $x_0 = j/\gamma$, prescribed by the balance between generation and trapping, during the injection process. Indeed, after time τ from the start of the injection, quasiparticles have spread over a distance $\sim\sqrt{D_{\text{qp}}\tau}$ and the diffusive current at that time can be estimated as $D_{\text{qp}}\partial_y x_{\text{qp}}(0) \sim D_{\text{qp}} x_{\text{qp}}(0)/\sqrt{D_{\text{qp}}\tau}$. For $\tau = t_{\text{sat}}$ the diffusive current is therefore of the order of the trapping current $\gamma x_{\text{qp}}(0)$; as quasiparticles spread farther out, the diffusive current will decrease, indicating that indeed a steady state is (asymptotically) reached. It is important to note that the total number of quasiparticles in the device keeps growing for the entire duration of injection, despite the saturation of $x_{\text{qp}}(0)$ at $\tau \sim t_{\text{sat}}$.

The evolution in the relaxation stage, $t > 0$, depends on the ratio $t_{\text{sat}}/t_{\text{inj}}$. A straightforward use of Eq. (30) yields for the quasiparticle density close to the trap, $y \to 0$, in the long-time limit $t \gg t_{\text{sat}}$

$$x_{\text{qp}}(0,t) \approx \frac{x_0}{\sqrt{\pi}}\left(\sqrt{\frac{t_{\text{sat}}}{t}} - \sqrt{\frac{t_{\text{sat}}}{t + t_{\text{inj}}}}\right). \tag{32}$$

This asymptote is valid for any value of $t_{\text{sat}}/t_{\text{inj}}$. If $t_{\text{inj}} \gg t_{\text{sat}}$, one may distinguish between an intermediate asymptotic behavior, $x_{\text{qp}}(0,t) \propto 1/t^{-1/2}$, valid at times $t_{\text{sat}} \ll t \ll t_{\text{inj}}$, and a long-time asymptote, $x_{\text{qp}}(0,t) \propto t^{-3/2}$, at $t \gg t_{\text{inj}}$. Only the latter behavior is present for short injection times $t_{\text{inj}} \lesssim t_{\text{sat}}$.

B. The effect of finite diffusion time

We now turn to the case of a finite-length strip, so that the diffusion time across the whole device,

$$t_L = 4L^2/(\pi^2 D_{\text{qp}}), \tag{33}$$

provides yet another scale for the relaxation dynamics of x_{qp}. The comparison of the two time scales, t_L and t_{sat}, allows us to introduce the notion of a weak versus a strong trap. A weak trap corresponds to $t_{\text{sat}} \gg t_L$. The diffusion through the device occurs much faster than the local saturation at the trap, and consequently, the quasiparticle distribution is almost homogeneous throughout the device. A strong trap, $t_{\text{sat}} \ll t_L$, leads to a highly inhomogeneous spatial distribution of the quasiparticle density. Recalling that $\gamma = \Gamma_{\text{eff}}d$, this distinction can also be expressed in terms of a comparison of the trap

length d with the length scale

$$l_0 \equiv \frac{\pi}{2}\frac{D_{\text{qp}}}{L\Gamma_{\text{eff}}} = \frac{\pi}{2}\frac{\lambda_{\text{tr}}^2}{L}, \tag{34}$$

with λ_{tr} of Eq. (21); a weak (strong) trap is characterized by $d \ll l_0$ ($d \gg l_0$). Note that if $\lambda_{\text{tr}} \ll L$, l_0 is much smaller than λ_{tr}, so the crossover between the two limits occurs while the trap length remains short, $d \ll \lambda_{\text{tr}}$, and we can still use Eq. (22).

For a weak trap, $d \ll l_0$, we may neglect the y dependence of $x_{\text{qp}}(y,t)$ in Eq. (22), and integrating it over y we find

$$x_{\text{qp}}(y,t) \approx x_0(1 - e^{-t_{\text{inj}}/\tau_w})e^{-t/\tau_w}, \tag{35}$$

where

$$\frac{1}{\tau_w} = \frac{d}{L}\Gamma_{\text{eff}}. \tag{36}$$

As long as x_{qp} can be considered y-independent, the expression (36) for the density decay rate may be easily generalized: the ratio d/L on the right-hand side should be replaced by $A_{\text{tr}}/A_{\text{dev}}$, where A_{tr} is the total area of the trap and A_{dev} is the area of the entire device. Importantly, the decay rate here depends merely on the ratio of the total areas, whereas details of the geometry of the trap and device are unimportant.

In the opposite case of a strong trap, $d \gg l_0$, the approximation of a constant $x_{\text{qp}}(x,y)$ is no longer valid, and the decay rate will depend on the details of the trap geometry and placement. For simplicity, we concentrate again on the strip geometry. To obtain the eigenmodes, one may replace the right-hand side of Eq. (27) with zero. Therefore, k is simply given by $k = \frac{\pi}{2L}p$, where p is an odd integer [up to small corrections of order l_0/d; cf. Eq. (39)]. In contrast to the case of a weak trap, the relaxation is now limited by the diffusion time. From Eq. (30) we find the time-dependent quasiparticle density at the junction to be

$$x_{\text{qp}}(0,t) \approx \frac{4}{\pi}x_0\sqrt{\frac{t_{\text{sat}}}{t_L}}\sum_p\left[e^{-\frac{t}{t_L}p^2} - e^{-\frac{t+t_{\text{inj}}}{t_L}p^2}\right], \tag{37}$$

with $x_0 = j/\gamma$, and the sum over the odd integer p. For short times, $t \ll t_L$, the time evolution is insensitive to the boundary condition at $x = L$, and indeed we recover the results given in Sec. III A. (Note that of course, being able to observe the transition from a $t^{-1/2}$ to a $t^{-3/2}$ power law decay is contingent upon t_{inj} being much smaller than t_L.) For times exceeding the diffusion time, $t \gtrsim t_L$, the time evolution is dominated by the single exponential of the slowest mode, and we can write

$$x_{\text{qp}}(0,t) \approx \frac{4}{\pi}x_0\sqrt{\frac{t_{\text{sat}}}{t_L}}(1 - e^{-t_{\text{inj}}/\tau_w})e^{-t/\tau_w}, \tag{38}$$

where the decay time constant is now determined by the diffusion time (33), $\tau_w = t_L$ [22].

Concentrating on the long-time evolution, we can more generally relate the time constant τ_w to the wave number of the slowest mode. Thus, we are able to investigate the full crossover in τ_w from weak to strong trap as a function of d/l_0. Setting $l \to 0$ in Eq. (27), we may rewrite it as

$$\cot\left(\frac{\pi}{2}\tilde{k}\right) = \frac{l_0}{d}\tilde{k}, \quad \tilde{k} = \frac{2}{\pi}kL. \tag{39}$$

The time constant can be expressed in terms of the smallest positive solution $\tilde{k}_0$ of Eq. (39) as $\tau_w = t_L/\tilde{k}_0^2$. Therefore, the ratio t_L/τ_w is a function of a single variable, d/l_0. The full crossover function between the linear dependence at small d/l_0 and saturation at $d/l_0 \gg 1$ can be found by solving Eq. (39) numerically. In Fig. 5, we show t_L/τ_w as a function of d/l_0, together with experimental data that we discuss in the next section. The introduction of scaled variables t_L/τ_w and d/l_0 allows us to compare the trapping for a number of devices and for a set of different temperatures.

IV. EXPERIMENTAL DATA

In this section we compare the model developed in the previous sections with experiments measuring the dynamics of injected quasiparticles in 3D transmon qubits [5]. The qubit, similar to the device sketched in Fig. 2, consists of a single Al/AlO$_x$/Al Josephson junction shunted by a coplanar gap capacitor, with long ($\sim$1 mm), narrow antenna leads which connect to a pair of small ($80 \times 80~\mu$m^2) pads; see Fig. 3. One or two chips containing qubits are mounted in a superconducting aluminum rectangular waveguide cavity. All measurements are performed in an Oxford cryogen-free dilution refrigerator, with magnetic field shielding, infrared shielding, and filtering described in Ref. [23].

After fabrication of the qubits, normal-metal traps are patterned via optical lithography, which gives control of trap location and size to better than 1 μ m. The heavily oxidized aluminum surface of the qubit is treated with an ion etch, and 100 nm of copper is deposited in a liftoff process thereafter. Through independent dc measurements, we find the Al-Cu interface resistance to be between 200 and 430 $\Omega~\mu$m^2. As shown in Fig. 3(c), one edge of the trap is located a short, fixed distance ($\sim$35 μ m) away from the junction. The trap has a width of 8 μm, and it is placed symmetrically on the 12 μm wide lead. For this study, we focus on devices in which the trap length, d, along the lead is varied from 20 to 400 μm. The qubits' T_1 times measured at 13 mK vary (nonmonotonically) between 10 and 22 μs for d between 20 and 80 μm, while the two devices with longer traps ($d = 200$

and 400 μm) have shorter relaxation times (5 and 7 μs, respectively). Comparisons with a control device without traps ($T_1 = 19~\mu$s) and with earlier experiments [10] indicate that that short traps do not negatively affect the qubit coherence, while longer traps might be somewhat detrimental. Here we focus on the effect of traps on quasiparticle (QP) dynamics and do not give further consideration to the possible trap-induced loss mechanism.

The QP dynamics of these devices is studied using the contactless, *in situ* method described in Ref. [10], where QPs are introduced into the qubit by applying a large microwave tone at the bare cavity resonance. This injection pulse creates a voltage across the Josephson junction greater than 2Δ, generating many ($\gtrsim 10^5$ per μs) quasiparticles near the junction. The subsequent decay of x_{qp} is probed by monitoring the recovery of the qubit relaxation time T_1 measured as a function of time after the injection, in light of the simple relation

$$\Gamma(t) = 1/T_1(t) = Cx_{qp}(0,t) + \Gamma_0, \tag{40}$$

where Γ_0 is the steady-state relaxation rate of the qubit, which includes the effects of residual quasiparticle population and other relaxation mechanisms such as dielectric losses, and C is a known proportionality constant [24] (whose value we do not need here). In other words, we exploit the fact that the time-dependent part of the qubit decay rate Γ is directly proportional to the excess quasiparticle density at the junction, $y = 0$.

Figure 4 shows a typical measurement of the qubit decay rate in a device with a small normal-metal trap. The decay time constant τ_w is estimated by fitting the data with a single exponential of the form

$$\Gamma(t) = Ae^{-t/\tau_w} + \Gamma_0. \tag{41}$$

As discussed in Sec. III B, we are considering only the slowest decay mode of x_{qp}, so we fit the data to the above expression at long times $t \gtrsim t_L$ [with t_L of Eq. (33)], where we find
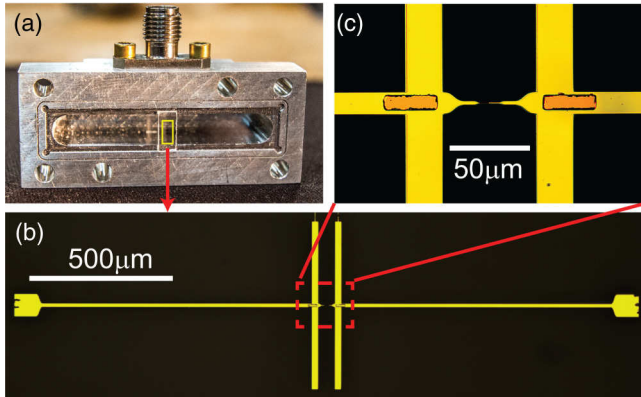

FIG. 3. (a) Photograph of a 3D aluminium cavity loaded with a transmon qubit. (b) Optical image of an example of the devices used for this study. (c) Zoomed-in image of the Cu-trap deposited near the junction
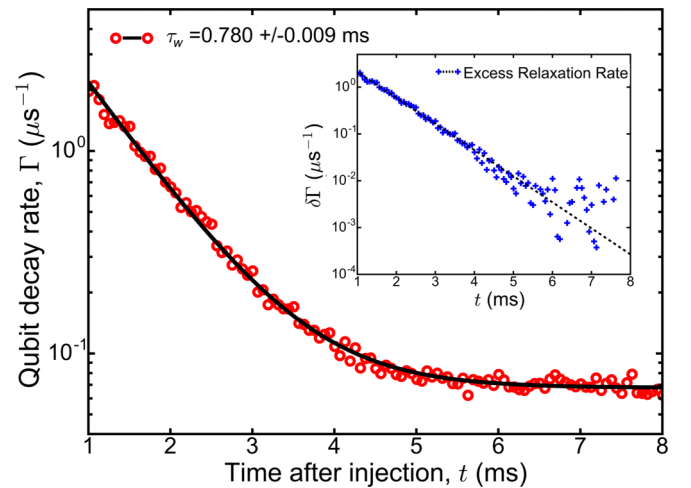

FIG. 4. Qubit energy relaxation rate Γ after quasiparticle injection. The solid line is a fit to the data by a single exponential with time constant τ_w, see Eq. (41). The inset shows the relaxation rate after subtracting a constant background in logarithmic scale, displaying good agreement with the predicted functional form.
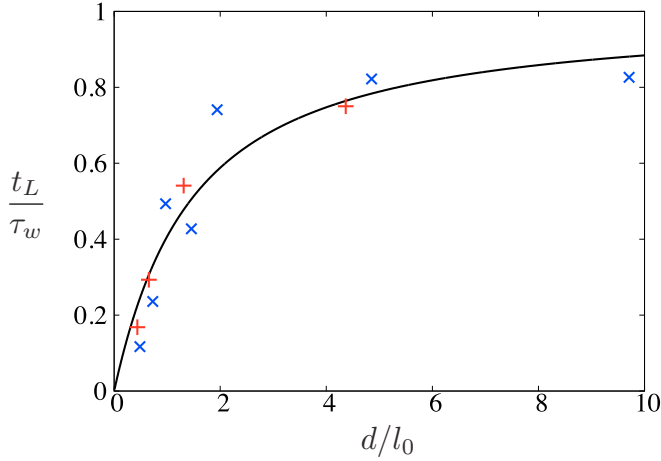
FIG. 5. Dimensionless density decay rate $1/\tau_w$ normalized by the diffusion time t_L, cf. Eq. (33), as a function of trap length d measured in units of l_0, see Eq. (34) for the definition. The solid line is calculated by solving Eq. (39) numerically. The experimental data are taken at two different fridge temperatures: the blue symbol "x" is used for $T_{\text{fr}} = 13$ mK and the red symbol "+" for $T_{\text{fr}} = 50$ mK. Note the transition from a linear dependence to the saturated diffusive limit at $d \sim l_0$.

good agreement between the data and the predicted single-exponential decay.

Repeating the measurement for several trap lengths d, we find that the experimental decay rate $1/\tau_w$ varies with the length of the trap in qualitative agreement with the rate calculated by solving Eq. (39); see Fig. 5. Indeed, for short traps we approximately find the linear dependence of $1/\tau_w$ on the trap length predicted by Eq. (36), while for longer traps the rate saturates to the the diffusion limit, $1/\tau_w \approx 1/t_L$. To scale the experimental data so that they can be compared to the theoretical expectation, we use l_0 and t_L as fitting parameters, and allow them to be different for data taken at different fridge temperatures T_{fr}, thus assuming that both D_{qp} as well as Γ_{eff} depend on T_{fr}. The fitting parameters are $l_0 = 41.2 \pm 17.1$ μm and $t_L = 184 \pm 29$ μs for $T_{\text{fr}} = 13$ mK and $l_0 = 45.8 \pm 16.7$ μm and $t_L = 125 \pm 20$ μs for $T_{\text{fr}} = 50$ mK [25]. Note that the relative change in l_0 is smaller than that in t_L and that this is in qualitative agreement with theoretical expectations: since l_0 is proportional to $D_{\text{qp}}/\Gamma_{\text{eff}}$, the expected increases of both D_{qp} and Γ_{eff} with effective temperature can partially compensate each other, while no such compensation is possible for $t_L \propto 1/D_{\text{qp}}$.

As discussed after Eq. (36), in the linear regime we can take into account the actual geometry of the transmon by modifying that expression for the decay rate, which becomes $1/\tau_w = \Gamma_{\text{eff}} A_{\text{tr}}/A_{\text{dev}}$. We use this formula to estimate Γ_{eff} using the short-trap data and find $\Gamma_{\text{eff}} \approx 2.42 \times 10^5$ s^{-1} for $T_{\text{fr}} = 13$ mK (corresponding to the blue data points in Fig. 5) and $\Gamma_{\text{eff}} \approx 3.74 \times 10^5$ s^{-1} for $T_{\text{fr}} = 50$ mK (red points). These numbers are close to the order-of-magnitude estimate for Γ_{eff} given at the end of Sec. II, where we assumed that the backflow of quasiparticles must be taken into account and strongly suppresses the effective trapping rate. In that section we have also shown that $\Gamma_{\text{eff}} \sim \Gamma_r (T/\Delta)^{1/2}$, indicating that

Γ_{eff} should grow with temperature. While we observe an increase in the Γ_{eff} extracted from the data with increasing fridge temperature, this increase is smaller than the factor of 2 expected from theory. This discrepancy is not surprising, since it is known that at low temperatures the quasiparticles are not in thermal equilibrium at the fridge temperature [19]. Moreover, the injection pulse can cause additional heating in the qubit [26], further weakening the relationship between fridge temperature and quasiparticle effective temperature.

V. SUMMARY

In this work we develop a basic model enabling us to predict the effect of a normal-metal trap on the dynamics of the nonequilibrium quasiparticles population in a superconducting qubit. The model accounts for the tunneling between the superconductor and the trap, as well as for the electron energy relaxation in the trap; see Eq. (18). The surprising finding is that the effective trapping rate Γ_{eff} is sensitive to the energy of the quasiparticles and is constrained by their backflow from the normal-metal trap on time scales shorter than the electron energy relaxation rate. Furthermore, we find the dependence of the time needed to evacuate the injected quasiparticles on the trap size. The evacuation time saturates at the lowest, diffusion-limited value upon extending the trap above a certain characteristic length l_0; the dependence of l_0 on the parameters of the trap and qubit is given in Eq. (34).

The experimental findings reported in Sec. IV validate the theoretical model. The relaxation rate $1/T_1$ of a transmon qubit is proportional to the quasiparticle density in the vicinity of the Josephson junction, making it possible to measure the dynamics of the quasiparticle population. We find that the population decay rate increases with the length of the normal-metal traps, in agreement with the predicted crossover from weak to strong trapping; see Fig. 5. For small traps we can estimate the effective trapping rate Γ_{eff}: both its order of magnitude and its increase with temperature indicate indeed a limitation due to the backflow of quasiparticles.

Utilizing traps is a viable strategy of mitigating the detrimental effect of quasiparticles on the qubit's T_1 time. Further improvement of normal-metal traps may benefit from finding ways to shorten the electron energy relaxation time in them. Based on the experiments of Refs. [15,16], using a different pure metal (e.g., silver or gold) for the trap is unlikely to result in substantially shorter relaxation time; metals hosting magnetic impurities might be helpful in this regard, but such impurities could harm the qubit by opening other relaxation channels.

APPENDIX A: TUNNELING RATE EQUATIONS

In this Appendix, we derive the rate equations for quasiparticles and electrons accounting for tunneling between a superconductor and a normal metal. Here we assume that both the superconductor and the normal metal are of sufficiently small volumes (Ω_S and Ω_N, respectively) that the diffusion of excitations occurs on a fast time scale and the occupation probabilities are hence uniform in space. Within these volumes we define the probabilities

$$f(\xi_m) = \sum_\sigma \langle c_{m\sigma}^\dagger c_{m\sigma} \rangle, \tag{A1}$$

$$f_{qp}(\epsilon_n) = \sum_\sigma \langle \gamma_{n\sigma}^\dagger \gamma_{n\sigma} \rangle \tag{A2}$$

of finding an electron excitation of energy ξ_m in the normal metal and a quasiparticle excitation of energy ϵ_n in the superconductor, respectively. The tunnel coupling between the two, see Eq. (4), gives rise to a change in both occupation probabilities for energies above the gap, via processes whose rates can be computed using Fermi's golden rule:

$$\dot{f}(\xi_m) = \sum_{n\sigma} [W_{n\sigma \to m\sigma} - W_{m\sigma \to n\sigma}$$
$$+ W_{0 \to m\sigma, n-\sigma} - W_{m\sigma, n-\sigma \to 0}], \tag{A3}$$

$$\dot{f}_{qp}(\epsilon_n) = \sum_{m\sigma} [-W_{n\sigma \to m\sigma} + W_{m\sigma \to n\sigma}$$
$$+ W_{0 \to m\sigma, n-\sigma} - W_{m\sigma, n-\sigma \to 0}], \tag{A4}$$

with

$$W_{n\sigma \to m\sigma} = \frac{2\pi}{\hbar} \frac{|\tilde{t}|^2}{\Omega_S \Omega_N} u_n^2 f_{qp}(\epsilon_n)[1 - f(\xi_m)] \delta(\epsilon_n - \xi_m), \tag{A5}$$

$$W_{0 \to m\sigma, n-\sigma} = \frac{2\pi}{\hbar} \frac{|\tilde{t}|^2}{\Omega_S \Omega_N} v_n^2 [1 - f_{qp}(\epsilon_n)]$$
$$\times [1 - f(\xi_m)] \delta(\epsilon_n + \xi_m). \tag{A6}$$

The reverse processes are found by replacing $f_{(qp)} \to 1 - f_{(qp)}$. Assuming particle-hole symmetry, $f(-\xi) = 1 - f(\xi)$, we summarize the rate equations as

$$\dot{f}(\xi_m) = \frac{2\pi}{\hbar} \frac{|\tilde{t}|^2}{\Omega_S \Omega_N} \sum_n [f_{qp}(\epsilon_n) - f(\xi_m)] \delta(\epsilon_n - \xi_m), \tag{A7}$$

$$\dot{f}_{qp}(\epsilon_n) = \frac{2\pi}{\hbar} \frac{|\tilde{t}|^2}{\Omega_S \Omega_N} \sum_m [f(\xi_m) - f_{qp}(\epsilon_n)] \delta(\epsilon_n - \xi_m). \tag{A8}$$

The tunneling processes considered above are elastic. In the normal metal, for temperatures $T \ll \Delta$ there is a large interval of unoccupied states below the gap. Inelastic processes, such as electron-phonon and electron-electron interactions, can relax the excitations in the normal metal to energies below the gap, so that they cannot return to the superconductor. We phenomenologically account for this relaxation by adding the term $-\Gamma_r f(\xi_m)$ to the right-hand side of Eq. (A7). The relaxation rate Γ_r is assumed energy-independent, which is justified if the interval of nonzero excitations above the gap is within a narrow energy strip of width $\ll \Delta$.

In the next step, we are interested in the probabilities to find excitations in the states within a small energy interval $\delta\epsilon$. We define the probability densities

$$p_N(\epsilon) = \frac{1}{N_S} \sum_{\epsilon < \xi_m < \epsilon + \delta\epsilon} f(\xi_m), \tag{A9}$$

$$p_S(\epsilon) = \frac{1}{N_S} \sum_{\epsilon < \epsilon_n < \epsilon + \delta\epsilon} f_{qp}(\epsilon_n), \tag{A10}$$

which are normalized with respect to the normal-state number of states in the superconductor $N_S = \nu_{S0} \Omega_S \delta\epsilon$. In the continuum limit $\delta\epsilon \to 0$ these definitions lead to Eqs. (12) and (13), respectively. From Eqs. (A7) and (A8) plus the phenomenological relaxation term discussed above, we obtain Eqs. (15) and (16) with the rates

$$\Gamma_{esc}(\epsilon) = \frac{2\pi}{\hbar} |\tilde{t}|^2 \frac{\nu_{S0}}{\Omega_N} \frac{\epsilon}{\sqrt{\epsilon^2 - \Delta^2}}, \tag{A11}$$

$$\Gamma_{tr} = \frac{2\pi}{\hbar} |\tilde{t}|^2 \frac{\nu_{N0}}{\Omega_S}. \tag{A12}$$

APPENDIX B: GENERALIZED DIFFUSION EQUATION AND EFFECTIVE TRAPPING RATE

In this Appendix we discuss the generalization of the rate equations (15) and (16) to include diffusion. In disordered metals, the effect of elastic impurity scattering on the distribution function is accounted for by a diffusion term; for quasiparticles in superconductors, the diffusion constant in the so-called "hydrodynamical approach" [27] is energy-dependent:

$$\dot{p}_N(\epsilon, \vec{r}, t) = D_N \vec{\nabla}^2 p_N(\epsilon, \vec{r}, t) + a(x, y) \delta(z - d_S)$$
$$\times [\tilde{\gamma}_{tr} p_S(\epsilon, \vec{r}, t) - \tilde{\gamma}_{esc}(\epsilon) p_N(\epsilon, \vec{r}, t)]$$
$$- \Gamma_r p_N(\epsilon, \vec{r}, t), \tag{B1}$$

$$\dot{p}_S(\epsilon, \vec{r}, t) = D_S(\epsilon) \vec{\nabla}^2 p_S(\epsilon, \vec{r}, t) - a(x, y) \delta(z - d_S)$$
$$\times [\tilde{\gamma}_{tr} p_S(\epsilon, \vec{r}, t) - \tilde{\gamma}_{esc}(\epsilon) p_N(\epsilon, \vec{r}, t)], \tag{B2}$$

where D_N is the diffusion constant in the normal-metal trap, and

$$D_S(\epsilon) = D_S / \nu_S(\epsilon) \tag{B3}$$

with D_S being the normal-state diffusion constant in the superconductor; $\tilde{\gamma}_{tr}$, $\tilde{\gamma}_{esc}$, and Γ_r are defined in Sec. II. The function $a(x, y)$ is 1 if coordinates x, y belong to the normal-superconductor contact, 0 otherwise. For the z coordinate, we assume that the superconductor occupies the interval $0 < z < d_S$ and the interval $d_S < z < d_S + d_N$ corresponds to the normal metal.

The above is a set of coupled linear differential equations that, for each energy ϵ, can be in principle solved in terms of eigenmodes, as done in Sec. III for Eq. (22). Here, to justify that equation we consider the conditions under which Eqs. (B1) and (B2) can be simplified, starting with the assumption that the normal metal is a thin layer.

1. Thin normal metal

In a sufficiently thin trap, the electron density within the normal metal should not change significantly in the z direction. This is the case if the length scale on which p_N varies in the z direction is much larger than the thickness of the trap d_N. Then we can expand p_N as a function of the distance $z - d_N - d_S$ from the upper surface:

$$p_N(\epsilon,\vec{r},t) = \frac{1}{d_N}\widetilde{p}_N(\epsilon,x,y,t)$$
$$+ \frac{(z - d_N - d_S)^2}{2d_N^3}p_N^{(2)}(\epsilon,x,y,t) + \cdots. \quad \text{(B4)}$$

The linear term is absent as to satisfy the boundary condition $\partial_z p_N = 0$ at $z = d_S + d_N$. For the expansion to be applicable, we require $p_N^{(2)} \ll \widetilde{p}_N$; this condition will lead to a limit on the thickness d_N, as we show in what follows.

The diffusion equation (B1) for the normal metal may be alternatively expressed as

$$\dot{p}_N(\epsilon,\vec{r},t) = D_N\vec{\nabla}^2 p_N(\epsilon,\vec{r},t) - \Gamma_r p_N(\epsilon,\vec{r},t) \quad \text{(B5)}$$

for $z > d_S$, with the boundary condition at $z = d_S$

$$D_N\partial_z p_N(\epsilon,x,y,d_S,t) + \widetilde{\gamma}_{tr} p_S(\epsilon,x,y,d_S,t)$$
$$- \widetilde{\gamma}_{esc}(\epsilon)p_N(\epsilon,x,y,d_S,t) = 0. \quad \text{(B6)}$$

Using the ansatz (B4), from the two equations above we find in the leading order

$$\dot{\widetilde{p}}_N = D_N\vec{\nabla}^2\widetilde{p}_N + \frac{D_N}{d_N^2}p_N^{(2)} - \Gamma_r\widetilde{p}_N \quad \text{(B7)}$$

and

$$-\frac{D_N}{d_N}p_N^{(2)} + d_N\widetilde{\gamma}_{tr}p_S - \widetilde{\gamma}_{esc}(\epsilon)\widetilde{p}_N = 0. \quad \text{(B8)}$$

Solving Eq. (B8) for $p_N^{(2)}$ and substituting it into Eq. (B7) we arrive at

$$\dot{\widetilde{p}}_N = D_N\vec{\nabla}^2\widetilde{p}_N + \widetilde{\gamma}_{tr}p_S - \Gamma_{esc}(\epsilon)\widetilde{p}_N - \Gamma_r\widetilde{p}_N; \quad \text{(B9)}$$

here the energy-dependent escape rate $\Gamma_{esc}(\epsilon)$ is defined by Eqs. (9) and (10). We can formally solve this equation for $\widetilde{p}_N$ by introducing an appropriate set of eigenmodes, which are discrete due to the finite size of the normal metal. Choosing for simplicity a rectangular trap with $0 < x < d_x$ and $0 < y < d_y$, we have eigenmodes of the form

$$n(k_x,k_y) = \sqrt{N_xN_y}\cos(k_xx)\cos(k_yy), \quad \text{(B10)}$$

with $k_{x,y} = \pi n_{x,y}/d_{x,y}$, $n_{x,y} \in \mathbb{N}$, the normalization constant $N_{x,y} = 1 + \delta_{0n_{x,y}}$, and the corresponding eigenvalues

$$\lambda(k_x,k_y) = -D_Nk_x^2 - D_Nk_y^2 - \Gamma_{esc}(\epsilon) - \Gamma_r. \quad \text{(B11)}$$

We can now write the solution to Eq. (B9) as

$$\widetilde{p}_N(x,y,t) = \Gamma_{tr}\int\frac{d\omega}{2\pi}\sum_{k_x,k_y}\frac{e^{-i\omega t}}{-i\omega - \lambda(k_x,k_y)}$$
$$\times n(k_x,k_y)\tilde{p}_S(k_x,k_y,\omega), \quad \text{(B12)}$$

where we introduced the Fourier transform of p_S at the interface,

$$\tilde{p}_S(k_x,k_y,\omega) = \int_0^{d_x}\frac{dx}{d_x}\int_0^{d_y}\frac{dy}{d_y}n(k_x,k_y)$$
$$\times \int dt\, e^{i\omega t}p_S(x,y,d_S,t). \quad \text{(B13)}$$

In the solution given in Eq. (B12) we discarded any transient terms, which are exponentially suppressed for times $t \gg [\Gamma_{esc}(\epsilon) + \Gamma_r]^{-1}$.

Let us assume that the length L of the (largest) part of the superconductor not covered by the trap is sufficiently long (see end of Sec. B 2); then at the long times relevant to experiments, the time scale for the evolution of p_S is determined by diffusion in the uncovered part. Similarly, diffusion in the region under the trap makes the long-time part of p_S a smooth function of x and y. This means that in Eq. (B12) we can neglect all but the lowest mode and set $\omega = 0$ and $k_x = k_y = 0$ in the denominator. We thus arrive at

$$\widetilde{p}_N \approx \frac{\widetilde{\gamma}_{tr}p_S}{\Gamma_{esc}(\epsilon) + \Gamma_r}. \quad \text{(B14)}$$

Using this estimate and the solution to Eq. (B8), the condition $p_N^{(2)} \ll \widetilde{p}_N$ can be written as a condition on the normal-metal thickness,

$$d_N \ll \sqrt{\frac{D_N}{\Gamma_r}}. \quad \text{(B15)}$$

Even for diffusion as slow as that of quasiparticles ($D_{qp} > 10\,\text{cm}^2/\text{s}$ [10]), using $\Gamma_r \sim 10^7\,\text{s}^{-1}$ (see Sec. II), the right-hand side is of order 10 μm, much thicker than the film thickness in the experiments.

2. Effective trapping rate

We are interested in finding the equation governing the dynamics of p_S. To this end, we substitute Eqs. (B4) and (B14) into Eq. (B2) to get at leading order the following equation:

$$\dot{p}_S(\epsilon,\vec{r},t) = D_S(\epsilon)\vec{\nabla}^2 p_S(\epsilon,\vec{r},t)$$
$$- a(x,y)\delta(z - d_S)\widetilde{\gamma}_{eff}(\epsilon)p_S(\epsilon,\vec{r},t) \quad \text{(B16)}$$

with

$$\widetilde{\gamma}_{eff}(\epsilon) = \widetilde{\gamma}_{tr}\frac{\Gamma_r}{\Gamma_{esc}(\epsilon) + \Gamma_r}. \quad \text{(B17)}$$

Let us concentrate on the experimentally relevant case of a thin superconductor in which p_S varies slowly in the z direction with respect to the thickness d_S. In analogy to the normal-metal case of the previous section, the z dependence may be neglected if $d_S \ll \sqrt{D_S(\epsilon)/\Gamma_{eff}(\epsilon)}$ with $\Gamma_{eff} = \widetilde{\gamma}_{eff}/d_S$. Using the inequality $D_S(\epsilon)/\Gamma_{eff}(\epsilon) > D_S\Gamma_{esc}/\Gamma_{tr}\Gamma_r$ (we recall that parameters D_S and Γ_{eff} without energy arguments correspond to the high-energy limiting values at $\epsilon \gg \Delta$), we find the sufficient condition

$$d_S \ll \sqrt{\frac{D_S\Gamma_{esc}}{\Gamma_{tr}\Gamma_r}}. \quad \text{(B18)}$$

For $d_S \approx d_N$ and $\nu_{S0} \approx \nu_{N0}$, this condition reduces to $d_S \ll \lambda_S \equiv \sqrt{D_S/\Gamma_r}$ which, as discussed after Eq. (B15),

is generically satisfied in the experiments. Then for a thin superconductor Eq. (B16) simplifies to

$$\dot{p}_S(\epsilon,x,y,t) = D_S(\epsilon)\vec{\nabla}^2 p_S(\epsilon,x,y,t)$$
$$- a(x,y)\Gamma_{\text{eff}}(\epsilon)p_S(\epsilon,x,y,t). \quad (B19)$$

In closing this section, we note that the length scale λ_S also determines the validity of the assumption made above that the uncovered part of the device is sufficiently long, which reads $L > \lambda_S$.

3. Dynamics of the quasiparticle density

The quantity that can be measured is the (normalized) quasiparticle density x_{qp}, which is related to p_S by integration over energy, $x_{\text{qp}} = \frac{2}{\Delta}\int_\Delta^\infty d\epsilon\, p_S(\epsilon)$. Focusing on the case of thin films, using Eq. (B19) as reference we write

$$\dot{x}_{\text{qp}} = D_{\text{qp}}\vec{\nabla}^2 x_{\text{qp}} - a(x,y)\Gamma_{\text{eff}}x_{\text{qp}}, \quad (B20)$$

where x_{qp} depends only on x and y. Comparing this equation to the integral over energy of Eq. (B19) we identify the coefficient

Γ_{eff} with

$$\Gamma_{\text{eff}} = \frac{\int_\Delta^\infty d\epsilon\, \Gamma_{\text{eff}}(\epsilon)p_S(\epsilon)}{\int_\Delta^\infty d\epsilon\, p_S(\epsilon)} \quad (B21)$$

[cf. Eq. (18)], and an analogous relation holds between the quasiparticle diffusion coefficients D_{qp} and $D_S(\epsilon)$. These relations are exact if the energy dependence in p_S can be factorized from its temporal and spatial ones. Although the structure of Eqs. (B16) and (B19) does not favor such factorization, the latter may be enforced by the process of quasiparticle thermalization due to their interaction with phonons. The temperature dependence of Γ_{eff} quoted in the text after Eq. (18) assumes the quasiparticles are thermalized at an effective temperature T. Phenomenological equations such as Eqs. (B20) and (B21) are widely used in the literature [8–10,12,13,28] as they successfully describe experiments as in Sec. IV.

[1] L. J. Geerligs *et al.*, Z. Phys. B **85**, 349 (1991).

[2] J. P. Pekola *et al.*, Rev. Mod. Phys. **85**, 1421 (2013).

[3] M. H. Devoret and J. M. Martinis, Quantum Inf. Process. **3**, 163 (2004).

[4] G. Catelani, J. Koch, L. Frunzio, R. J. Schoelkopf, M. H. Devoret, and L. I. Glazman, Phys. Rev. Lett. **106**, 077002 (2011).

[5] H. Paik *et al.*, Phys. Rev. Lett. **107**, 240501 (2011).

[6] J. Aumentado, M. W. Keller, J. M. Martinis, and M. H. Devoret, Phys. Rev. Lett. **92**, 066802 (2004).

[7] L. Sun *et al.*, Phys. Rev. Lett. **108**, 230509 (2012).

[8] J. N. Ullom, P. A. Fisher, and M. Nahum, Appl. Phys. Lett. **73**, 2494 (1998).

[9] I. Nsanzineza and B. L. T. Plourde, Phys. Rev. Lett. **113**, 117002 (2014).

[10] C. Wang *et al.*, Nat. Commun. **5**, 5836 (2014).

[11] M. Taupin, I. M. Khaymovich, M. Meschke, A. S. Melnikov, and J. P. Pekola, Nat. Commun. **7**, 10977 (2016).

[12] S. Rajauria, H. Courtois, and B. Pannetier, Phys. Rev. B **80**, 214521 (2009).

[13] S. Rajauria, L. M. A. Pascal, Ph. Gandit, F. W. J. Hekking, B. Pannetier, and H. Courtois, Phys. Rev. B **85**, 020505(R) (2012).

[14] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Harcourt, Orlando, FL, 1976).

[15] A. Anthore, Ph.D. thesis, Université de Paris 6, 2003 (available at http://iramis.cea.fr/spec/Pres/Quantro/static/publications/phd-theses/index.html).

[16] B. Huard, A. Anthore, N. O. Birge, H. Pothier, and D. Esteve, Phys. Rev. Lett. **95**, 036802 (2005); B. Huard, Ph.D. thesis,

Université de Paris 6, 2006 (available at http://iramis.cea.fr/spec/Pres/Quantro/static/publications/phd-theses/index.html).

[17] A. Rothwarf and B. N. Taylor, Phys. Rev. Lett. **19**, 27 (1967).

[18] S. B. Kaplan, C. C. Chi, D. N. Langenberg, J. J. Chang, S. Jafarey, and D. J. Scalapino, Phys. Rev. B **14**, 4854 (1976).

[19] D. Ristè *et al.*, Nat. Commun. **4**, 1913 (2013).

[20] P. J. de Visser, J. J. A. Baselmans, P. Diener, S. J. C. Yates, A. Endo, and T. M. Klapwijk, Phys. Rev. Lett. **106**, 167004 (2011).

[21] M. Bal, M. H. Ansari, J.-L. Orgiazzi, R. M. Lutchyn, and A. Lupascu, Phys. Rev. B **91**, 195434 (2015).

[22] We note that while this result for τ_w was derived assuming a short trap, $d \ll \lambda_{\text{tr}}$, it is more generally valid even for longer trap with $d > \lambda_{\text{tr}}$ as long as we identify L with the longest section of the wire not covered by the normal metal, and the condition $L \gg \lambda_{\text{tr}}$ is satisfied. The effects of a not-small d/λ_{tr} ratio and of the device geometry will be considered in more detail elsewhere.

[23] A. P. Sears *et al.*, Phys. Rev. B **86**, 180504 (2012).

[24] G. Catelani, R. J. Schoelkopf, M. H. Devoret, and L. I. Glazman, Phys. Rev. B **84**, 064517 (2011).

[25] While these fitted values for t_L are in reasonable agreement with the values estimated using Eq. (33), we note that in general they can be affected by details of the device geometry and decay measurement process. We will return to these aspects elsewhere.

[26] U. Vool *et al.*, Phys. Rev. Lett. **113**, 247001 (2014).

[27] A. G. Aronov, Yu. M. Galperin, V. L. Gurevich, and V. I. Kozub, Adv. Phys. **30**, 539 (1981).

[28] G. C. O'Neil, P. J. Lowell, J. M. Underwood, and J. N. Ullom, Phys. Rev. B **85**, 134504 (2012).