


Efficient quasiparticle traps with low dissipation through gap engineering

R.-P. Riwar and G. Catelani

JARA Institute for Quantum Information (PGI-11), Forschungszentrum Jülich, 52425 Jülich, Germany (Received 15 July 2019; revised manuscript received 10 September 2019; published 23 October 2019)

Quasiparticles represent an intrinsic source of perturbation for superconducting qubits, leading to both dissipation of the qubit energy and dephasing. Recently, it has been shown that normal-metal traps may efficiently reduce the quasiparticle population and improve the qubit lifetime, provided the trap surpasses a certain characteristic size. Moreover, while the trap itself introduces new relaxation mechanisms, they are not expected to harm state-of-the-art transmon qubits under the condition that the traps are not placed too close to extremal positions where electric fields are high. Here we study a different type of trap, realized through gap engineering. We find that gap-engineered traps relax the remaining constraints imposed on normal metal traps. First, the characteristic trap size, above which the trap is efficient, is reduced with respect to normal metal traps, such that here, strong traps are possible in smaller devices. Second, the losses caused by the trap are now greatly reduced, providing more flexibility in trap placement. The latter point is of particular importance, since for efficient protection from quasiparticles, the traps ideally should be placed close to the active parts of the qubit device, where electric fields are typically high.

DOI: [10.1103/PhysRevB.100.144514](https://doi.org/10.1103/PhysRevB.100.144514)**I. INTRODUCTION**

Superconducting circuits are prime candidates for a successful physical implementation of qubits, the fundamental building blocks of quantum computers [1–3]. In order to ensure a stable and sufficiently error-free computation, qubits need to be protected from possible perturbations. In particular, quasiparticle excitations pose a serious problem as an intrinsic source of errors [4–10]. Their harmful effect comes from the coupling to the qubit degrees of freedom, when they tunnel across a Josephson junction, as was shown in Ref [11]. This mechanism can be a particularly severe issue, because experiments have shown that quasiparticles occur at much higher densities than those expected in thermal equilibrium [8,10].

In order to mitigate the negative effects of quasiparticles, one needs to find ways to evacuate them from the active parts of the device. So far two classes of strategies have been proposed and tested. The first class, the one we focus on in this paper, includes introducing local regions containing subgap states into which quasiparticles can relax. Such quasiparticle traps have been realized through introducing vortices [12–15], tunnel coupling the device to normal metals [16–21], or through gap engineering [22,23]. A second strategy involves a time-dependent control of the device, in order to pump quasiparticles away through pulses [24].

In the present paper, we theoretically study traps implemented via gap engineering, that is, by coupling the superconductor S used to fabricate the active device part (usually, one or more Josephson junctions) to a different superconductor S' with a lower gap. The effect of gap engineering was extensively studied for Cooper-pair transistors [7,16,22,25,26]. In essence, the circuit can be efficiently protected from quasiparticles by making transistors where the central islands have a higher gap than the leads. For circuits with only single junctions, such as transmons, the gap engineering must

performed in a more sophisticated manner, e.g., by proximitizing parts of the circuit. This was attempted in Ref. [23], where the gap was locally increased at the junction; however, there was no unequivocal proof of a beneficial effect, possibly due to a nonoptimal choice of the design. Here we propose different designs of gap-engineered traps located further away from the junction, building on our previous insights from normal metal traps, which were shown to be efficient [19]. In particular, we predict an improvement compared to normal metal traps, in terms of both increased trapping efficiency and reduction of unwanted adverse effects.

In order to appreciate these improvements, we summarize our past works on normal metal traps [19,21,27]. The most important conclusions are the following. There is a minimal trap size above which the trap is “strong” [19], in the sense that the quasiparticle evacuation rate is limited by diffusion through the device. Importantly, this minimal trap size depends on the quality of the superconductor-normal metal (SN) interface and can become large due to a quasiparticle backflow stemming from the peak in the superconducting density of states just above the gap. This limits the effectiveness of normal-metal traps in small devices. For sufficiently large traps, the quasiparticle evacuation can be significantly accelerated by splitting the trap into many pieces and distributing them evenly over the device [21]. However, likewise the positive effect of trap splitting is limited, as soon as the individual trap pieces approach the minimal size. Finally, we have shown that placing traps close to active parts of the circuit is favorable for the qubit lifetime, as it suppresses the quasiparticle density and its fluctuations at the place where it does the most harm.

On the other hand, normal metal traps themselves introduce new processes that could limit the quality factor of the qubit [27,28]. The two most important processes are the dissipation in the normal metal bulk due to the ac charge redistribution in the circuit and the losses due to photo-assisted

tunneling at the SN interface. While the latter is expected to be sufficiently small, the former process can potentially pose problems if traps are placed too close to active parts of the circuit where electric fields are high. This issue is thus in competition with the optimization principle mentioned above, where traps close to the active parts increases the protection from the nonequilibrium quasiparticles. Normal metals can also induce, via the inverse proximity effect, subgap states that increase the qubit decay rate, but this effect is exponentially suppressed with distance from the active parts and can be neglected.

We here propose to replace the weakly coupled normal metal pieces, with strongly coupled superconducting parts with a significantly lower gap. For sufficiently good interfaces, the proximity effect will introduce a spatially dependent gap in the circuit (see, e.g., Refs. [28,29] and references therein) and provide a region where the quasiparticles can be trapped. Studying the quasiparticle diffusion dynamics, we predict that the minimal trap size to reach the strong trapping limit is reduced with respect to normal metal traps, making efficient traps possible in devices of smaller sizes. In particular, the unfavorable backflow of quasiparticles from normal metals is absent here. Second, we investigate the dominant dissipation processes induced by the gap-engineered traps. We find that charge redistribution currents no longer pose a threat, since the presence of the lower gap reduces the bulk resistivity by several orders of magnitude. Photo-assisted tunneling processes on the other hand can now become more important since the improved interfaces increase the tunneling rate. Overall, we show however that in practice gap-engineered traps do not cause any harmful side effect.

This paper is structured as follows. In Sec. II we present the diffusion model for qubits with gap-engineered traps and provide our results for the quasiparticle diffusion dynamics. In Sec. III we estimate the dominant dissipation processes induced by these traps. Our conclusions are presented in Sec. IV. In Appendix A we provide a microscopic justification for our diffusion model. In Appendix B we detail our estimate for the dissipation due to photo-assisted tunneling. Finally, in Appendix C we compute the geometric capacitances for coplanar circuits.

II. DIFFUSION AND TRAPPING OF QUASIPARTICLES IN AN SS' TRAP

A. The diffusion model

We study the following model. The starting point consists of a conventional transmon circuit, made from a certain superconducting material S with gap Δ . Part of this circuit is in contact with a second superconductor S' , with a lower gap $\Delta_{S'}$; see Fig. 1(a). We are interested in the limit, where there is a good contact between the two superconductors, such that S gets proximitized, making its gap position-dependent, $\Delta(y)$ [see Fig. 1(b)]. The proximity effect thus provides a gap-engineered quasiparticle trap in the transmon [30]. Quasiparticles at energy $\approx \Delta$ or above will then, once they diffuse into the lower gap region, relax, and consequently become confined to the proximitized region. The essential trapping principle is therefore the same as in the previously studied

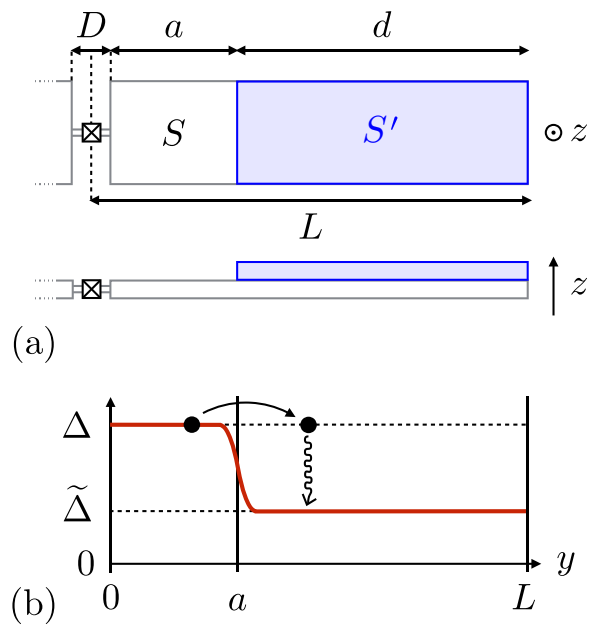


FIG. 1. (a) Transmon qubit made out of superconductor S with a gap Δ . The trap is realized through gap engineering with a second superconductor S' which has a lower gap $\Delta_{S'}$. The middle dashed line marks the origin of the y axis (horizontal direction). One transmon side has the total length L . The second superconductor S' is attached to the transmon with a distance a from edge to edge and has length L' . The two transmon plates are separated by a small distance $D \ll a, L$. For most of the discussion (unless specified otherwise) we set $D \rightarrow 0$. (b) The resulting position dependence of the gap inside the transmon, as a function of y . The change of the gap parameter is occurring on the length scale of the coherence length. Once the quasiparticles diffuse into the region with the lower gap, they may relax to energies below Δ and are thus trapped.

normal-metal traps, except that with proximitized traps we expect several improvements in terms of trapping performance, as well as a reduction of unwanted side effects, as we detail in the course of this paper.

As indicated in Fig. 1(b), deep within the proximitized region, the gap approaches a certain value $\tilde{\Delta}$, with $\Delta_{S'} \leq \tilde{\Delta} \leq \Delta$. The smooth change from Δ to $\tilde{\Delta}$ in the proximitized part of the superconductor occurs on the scale of the coherence length [31].

We will show below that the coherence length (~ 100 nm in thin aluminum films [32,33]) is small in comparison to the diffusion length scales, such that on the level of the diffusion equation, we can assume an immediate, discontinuous drop of the gap parameter $\Delta(y) = \Delta\theta(a-y) + \tilde{\Delta}\theta(y-a)$, for $0 \leq y \leq L$, where L is the total device size, and $a \leq y \leq L$ is the proximitized part.

Finally, in the following, we are interested in the quasiparticles located in the active parts of the circuit, which can actually disrupt the qubit state. This concerns the quasiparticles at energies above Δ , for which we define the dimensionless density (normalized by the density of Cooper pairs)

$$x_{\text{qp}}(y) = \frac{2}{\Delta} \int_{\Delta}^{\infty} d\epsilon \frac{\epsilon}{\sqrt{\epsilon^2 - \Delta^2(y)}} f_{\text{qp}}(\epsilon, y). \quad (1)$$

For this quantity, we can formulate the phenomenological diffusion equation

$$\dot{x}_{\text{qp}}(y) = \partial_y [D_{\text{qp}}(y) \partial_y x_{\text{qp}}(y)] - \tau_r^{-1} \theta(y-a) x_{\text{qp}}(y), \quad (2)$$

where D_{qp} is the (effective) quasiparticle diffusion constant. For the microscopic justification of this equation from a more general energy-dependent diffusion term, see Appendix A.

In the diffusion equation we have included a relaxation term—the last term on the right-hand side—responsible for trapping quasiparticles through relaxation to energies below Δ . We assume that the relaxation rate τ_r^{-1} is due to inelastic electron-phonon interactions and can therefore be computed according to Ref. [34]. For $\Delta \gg \tilde{\Delta}$ and temperatures small compared to $\tilde{\Delta}/k_B$, it results in

$$\tau_r^{-1} \approx \left\{ 1 - 6 \frac{\tilde{\Delta}^2}{\Delta^2} \left[\ln \left(2 \frac{\Delta}{\tilde{\Delta}} \right) - 1 \right] \right\} \tau_N^{-1} \quad (3)$$

and approaches the normal-metal relaxation rate $1/\tau_N$ [35]. For almost equivalent gaps, $\Delta \gtrsim \tilde{\Delta}$, we receive

$$\tau_r^{-1} \approx \frac{64\sqrt{2}}{35} \left(1 - \frac{\tilde{\Delta}}{\Delta} \right)^{\frac{7}{2}} \tau_N^{-1}, \quad (4)$$

that is, a relaxation rate much smaller than the normal-metal relaxation rate. Note that here we have assumed again sufficiently small temperature, $k_B T \ll \Delta - \tilde{\Delta}$, even though $\Delta - \tilde{\Delta} \ll \Delta$. Obviously, it is therefore advantageous to choose a sufficiently low gap and strong proximity effect to achieve fast relaxation. As a realistic example, we note that good contact can be achieved between aluminum and titanium films, and that the Al to Ti (bulk) critical temperature ratio is about 3. By varying the thicknesses of the two materials, and possibly adding a normal-metal layer (e.g., gold), the critical temperature and hence the gap can be tuned over a wide range [36,37].

Let us note at this point, that the relaxation term in Eq. (2) is analogous to the effective trapping rate Γ_{eff} used to model normal-metal traps [19], with a slight difference. Namely, for normal-metal traps, the trapping mechanism involves both quasiparticle tunneling (to the normal metal) and subsequent relaxation, hence Γ_{eff} is not simply equal to τ_r^{-1} (for more details, see Sec. II C). With gap-engineered traps on the other hand, we exclusively have to take into account the latter process, which is why the bare τ_r^{-1} appears in the quasiparticle dynamics. Therefore, we could in principle use interchangeably the nomenclature “relaxation rate” and “trapping rate” to refer to τ_r^{-1} . We will nonetheless stick to the term “trapping rate” in the remainder of this work to unify the language and to emphasize the formal analogy to the quasiparticle dynamics with normal metal traps.

As for the time τ_N , estimates for bulk metals [34] range from 10^{-10} to 10^{-6} s. However, in thin films the acoustic mismatch with the substrate leads to the so-called phonon-trapping effect [38] which, similarly to the effect on the quasiparticle recombination time, could increase the (effective) time τ_N above its bulk value. Here we note that experiments with normal-metal traps made of copper are compatible with a value $\tau_N \sim 4 \times 10^{-7}$ s [19]. One can also estimate τ_N from the quantity Σ relating the power P per unit volume dissipated by electron-phonon scattering to electron and phonon

temperatures via $P = \Sigma(T_e^5 - T_{ph}^5)$ [39]; the estimate takes the form $\tau_N \simeq 75\nu k_B^5 / \Sigma \Delta^3$, with ν the density of states in S . For the metals reported in Table I of Ref. [39], this gives τ_N ranging, e.g., from 8×10^{-8} s for copper to 8×10^{-7} s for aluminum.

B. The dynamics of quasiparticle trapping

In order to examine the quasiparticle diffusion dynamics, we need to find the eigenvalues of Eq. (2). They can be found when rewriting above diffusion equation separately for each side,

$$\dot{x}_{\text{qp}}(y) = D_{\text{qp}} \partial_y^2 x_{\text{qp}}(y), \quad y < a, \quad (5)$$

$$\dot{x}_{\text{qp}}(y) = \tilde{D}_{\text{qp}} \partial_y^2 x_{\text{qp}}(y) - \tau_r^{-1} x_{\text{qp}}(y), \quad y > a, \quad (6)$$

and at the interface $y = a$, the quasiparticle density has to be continuous and satisfy the boundary condition

$$\tilde{D}_{\text{qp}} \partial_y x_{\text{qp}}|_{y=a+0^+} = D_{\text{qp}} \partial_y x_{\text{qp}}|_{y=a-0^+}, \quad (7)$$

which expresses the conservation of current. In order to study the quasiparticle diffusion dynamics, we proceed similarly as in Ref. [19,21]. Namely, both Eqs. (5) and (6) have eigenmodes of the form $x_{\text{qp}}(y < a) \sim e^{-\lambda_k t} \cos(ky)$ and $x_{\text{qp}}(y > a) \sim e^{-\tilde{\lambda}_k t} \cos([L-y]\tilde{k})$, respectively, which decay with the constant rates $\lambda_k = D_{\text{qp}} k^2$ and $\tilde{\lambda}_k = \tilde{D}_{\text{qp}} \tilde{k}^2 + \tau_r^{-1}$. The respective cosine forms are chosen such as to take into account the hard wall boundary conditions at $y = 0$ and $y = L$. Requiring that these rates are the same for an eigenmode of the *total* system (i.e., for y from 0 to L), we find

$$\tilde{k} = \sqrt{\frac{D_{\text{qp}}}{\tilde{D}_{\text{qp}}} k^2 - \frac{1}{L_r^2}}, \quad (8)$$

where $L_r = \sqrt{\tilde{D}_{\text{qp}} \tau_r}$. Taking now into account the boundary condition at $y = a$, Eq. (7), we find that this imposes an additional condition on k ,

$$\frac{1}{L_r k} \tan(ka) = -\frac{L_r \tilde{k}}{1 + (L_r \tilde{k})^2} \tan([L-a]\tilde{k}). \quad (9)$$

As a result, this fixes the allowed values of k to discrete values and, consequently, also the eigenvalues of the diffusion equation are discrete, due to the finite system size. At sufficiently long times, the eigenmode with the lowest k will be the dominant one. In fact, one can show that for the lowest k , \tilde{k} is always imaginary, $\tilde{k} = i\kappa$, with $\kappa = \sqrt{1/L_r^2 - D_{\text{qp}} k^2 / \tilde{D}_{\text{qp}}}$, such that we can rewrite $-\tilde{k} \tan([L-a]\tilde{k}) = \kappa \tanh([L-a]\kappa)$.

We proceed next by assuming that $\tilde{D}_{\text{qp}} > D_{\text{qp}}$, meaning that the quasiparticles at energies $E \geq \Delta$ diffuse faster in the part with the lower gap. This is due to the quasiparticle diffusion constant decreasing close to the respective gap (see also Appendix A). We then arrive at closed expressions for the slowest density decay rate λ_0 in the limiting cases shown in Table I.

C. Comparison to normal-metal traps

With these results, we are now equipped to compare gap-engineered traps to normal-metal traps [19,21]. In the latter

TABLE I. The decay rate of the lowest mode for different regimes, as obtained from Eq. (9). Here large/small trap refers to the trap length $L - a$ in comparison to the length a of the nonproximitized region; see Fig. 1.

Trap size	Relaxation limit	Diffusion limit
$a \ll L - a$	$\lambda_0 \approx \tau_r^{-1}$	$\lambda_0 \approx \frac{\pi^2}{4} \frac{D_{\text{qp}}}{a^2}$
Large trap	$a \ll \sqrt{D_{\text{qp}} \tau_r}$	$a \gg \sqrt{D_{\text{qp}} \tau_r}$
$a \gg L - a$	$\lambda_0 \approx \frac{L-a}{L} \tau_r^{-1}$	$\lambda_0 \approx \frac{\pi^2}{4} \frac{D_{\text{qp}}}{L^2}$
Small trap	$L - a \ll \frac{D_{\text{qp}} \tau_r}{L}$	$L - a \gg \frac{D_{\text{qp}} \tau_r}{L}$

case, for typical values of the tunneling rate between superconductor and normal metal, the effective trapping rate of quasiparticles close to the trap is limited by the relaxation process in the normal metal, rather than by the tunneling process; the resulting trapping rate is $\tau_r^{-1} \sim \sqrt{k_B T / \Delta} \tau_N^{-1}$, with $T \ll \Delta / k_B$ the (effective) quasiparticle temperature. Note that this rate is smaller than τ_N^{-1} ; this suppression is due to a significant escape of quasiparticles back into the superconductor because of the strongly peaked density of states in the superconductor.

The suppression of the trapping rate results additionally in the need for longer normal-metal traps compared to gap-engineered ones. To quantify this statement, we point out that for both types of traps, there is a transition from weak to strong (efficient) trapping as a function of the trap size. When traps are small the decay rate of the quasiparticle density is roughly given by the trapping rate times the ratio of the trap length over the system size (cf. the central bottom entry in Table I). For sufficiently long traps, on the other hand, the density decay rate is limited by the diffusion time in the device (cf. the right bottom entry in Table I). By comparing the two density decay rates, one finds the minimal trap length for which the density decay rate is limited by diffusion, which for normal-metal traps is given by [19]

$$d_N \sim \sqrt{\frac{\Delta}{k_B T}} \frac{D_{\text{qp}} \tau_N}{L}, \quad (10)$$

where L is again the size of the superconductor. The above formula indicates (see also Ref. [21]) that when making the devices smaller and smaller, there comes a point where the diffusion-limited regime can no longer be reached, because it would require $d_N > L$.

Importantly, we can contrast this with the gap-engineered traps considered here. Due to the perfect interface between the Δ and $\tilde{\Delta}$ region, there is no limiting tunneling process. Moreover, there cannot occur any escape processes as for normal-metal traps: in the latter case, the trapping can take place only by relaxation in the normal-metal part of the device, while for gap-engineered traps both sides of the interface have the same low gap and therefore trapping due to relaxation can occur in both the S and S' parts of the trap. Therefore, the bottleneck process is necessarily the relaxation of quasiparticles, and its rate is no longer reduced due to escape processes by $\sqrt{k_B T / \Delta}$ as for normal-metal traps; indeed, the trapping rate is given directly by τ_r^{-1} , which is $\approx \tau_N^{-1}$ for $\Delta \gg \tilde{\Delta}$ [cf. Eq. (3)]. Consequently, we find that the saturation to the

diffusion-limited regime occurs at a different minimal trap size. As explained above, we can estimate the latter by equating the results for the density decay rate λ_0 in the relaxation and diffusion-limited cases; see Table I. We find

$$d_S \sim \frac{D_{\text{qp}} \tau_r}{L} \approx \frac{D_{\text{qp}} \tau_N}{L}. \quad (11)$$

Crucially, this means that with gap-engineered traps, one can achieve the diffusion-limited regime in smaller devices than would be possible with normal metal traps, since $d_N / d_S \sim \sqrt{\Delta / k_B T} \sim 10$ for typical experimental conditions [40].

We can generalize the above statement to an arbitrary ratio between d and L_r , where the minimal trap size becomes

$$d_S \approx L_r \tanh^{-1} \left(\frac{D_{\text{qp}} L_r}{\tilde{D}_{\text{qp}} L} \right). \quad (12)$$

The above result can be obtained from Eq. (9), by approximating it in the limit $\tilde{k} \approx 1 / L_r$ and $a \approx L$, such that

$$Lk \tan(Lk) \approx \frac{\tilde{D}_{\text{qp}} L}{D_{\text{qp}} L_r} \tanh \left(\frac{d}{L_r} \right). \quad (13)$$

The transition from relaxation to diffusion limit occurs when the right-hand side is ≈ 1 , which is obviously satisfied by d_S in Eq. (12).

Finally, we note that when making the trap region so large that $L - a > a$, then the length scale $L - a$ gets replaced by a , such that the transition from weak to strong (diffusion-limited) trapping is no longer dependent on $L - a$. Instead, the transition occurs at

$$a_S = \sqrt{D_{\text{qp}} \tau_r}; \quad (14)$$

that is, the diffusion-limited regime is entered when the region with higher gap Δ becomes longer than the diffusion length scale within that region. We provide the results for the density decay rate in the various regimes in Table I.

III. DISSIPATION IN A PROXIMITIZED TRAP

While quasiparticle traps improve the qubit relaxation time, at the same time they may provide new paths for the dissipation of the qubit energy as well as dephasing. For normal metal traps, it has been elaborated [27] that the main mechanisms are either due to ac charge redistribution within the normal metal or photo-assisted tunneling at the superconductor-trap interface. While in Ref. [27] we estimate that both mechanisms are sufficiently low such that they do not limit the quality factor of the best currently available qubits, they could impose limitations on future improved devices. We now show that we expect both mechanisms to be strongly reduced when using proximitized traps.

A. Dissipation due to ac resistance

The dissipation due to the ac resistance, when redistributing charges in the circuit, should be strongly reduced: the here considered gap-engineered traps have a nonzero gap everywhere, bringing the dc dissipation to zero. There will remain, however, a residual ac response. We briefly review the ac response of superconductors to estimate the degree of dissipation reduction. In the ac case [41,42], we find the real

part of the conductivity as

$$\frac{\text{Re}[\sigma_S(\omega)]}{\sigma_N} = \frac{1}{\omega} \int_{-\infty}^{\infty} d\epsilon v(\epsilon, \epsilon + \omega) [f(\epsilon) - f(\epsilon + \omega)], \quad (15)$$

with σ_N the normal-state conductivity and

$$v(\epsilon, \epsilon') = \frac{|\epsilon\epsilon' + \tilde{\Delta}^2|}{\sqrt{\epsilon^2 - \tilde{\Delta}^2} \sqrt{\epsilon'^2 - \tilde{\Delta}^2}}. \quad (16)$$

We compute the integral in the limit $T \ll \omega \ll \tilde{\Delta}$, in which we find

$$\frac{\text{Re}[\sigma_S(\omega)]}{\sigma_N} \approx \frac{1}{2} \left(\frac{2\tilde{\Delta}}{\omega} \right)^{\frac{3}{2}} \tilde{x}_{\text{qp}}, \quad (17)$$

where \tilde{x}_{qp} is the total normalized quasiparticle density in the proximitized region [obtained by the replacement $\Delta \rightarrow \tilde{\Delta}$ in Eq. (1)]. The left-hand side of Eq. (17) is small for small quasiparticle densities. Consequently, in order to find the resistivity (that is, the inverse of the conductivity), the imaginary part of the conductivity is needed. It accounts for the kinetic inductance and in the limit $\omega, T \ll \tilde{\Delta}$, is given by [43]

$$\frac{\text{Im}[\sigma_S]}{\sigma_N} = \frac{\pi \tilde{\Delta}}{\omega}. \quad (18)$$

Based on this, we can find the real part of the resistivity, which we need to estimate the dissipation due to ac currents,

$$\text{Re}[\rho(\omega)] = \frac{\text{Re}[\sigma_S(\omega)]}{|\sigma_S(\omega)|^2}. \quad (19)$$

Due to $\tilde{x}_{\text{qp}} \ll 1$, $\text{Re}[\sigma_S] \ll \text{Im}[\sigma_S]$, and therefore we find

$$\text{Re}[\rho(\omega)] \approx \frac{\sqrt{2}}{\pi^2} \sqrt{\frac{\omega}{\tilde{\Delta}}} \tilde{x}_{\text{qp}} \rho_N, \quad (20)$$

which is much smaller than the normal-state resistivity $\rho_N = 1/\sigma_N$: in the superconducting state, the resistivity is reduced by both the factor $\sqrt{\omega/\tilde{\Delta}} < 1$ and the factor $\tilde{x}_{\text{qp}} \ll 1$. In practice we expect $\tilde{x}_{\text{qp}} \lesssim 10^{-4}$, since either injection of quasiparticles [14] or relatively high temperature ($k_B T/\tilde{\Delta} > 0.1$) is needed to arrive at such densities, so we can safely deduce that the ac dissipation is reduced by at least four orders of magnitude, likely more.

This significant reduction allows for more flexibility regarding the trap placement. In particular, it enables the positioning of quasiparticle traps closer to the active parts of the circuits, where the evacuation of quasiparticles is crucial. For normal metal traps, this represented a serious limitation. Close to the active circuit parts, the electric fields are high and consequently so are the charge redistribution currents. In fact, for normal metal traps, we found [27] that when placing a trap close to these extreme points, the relaxation due to ac dissipation could give a contribution of the order 10% to the quality factor of present-day qubits. With the reduction for gap-engineered traps, we expect that this is no longer a concern, even if qubits relaxation time is further improved.

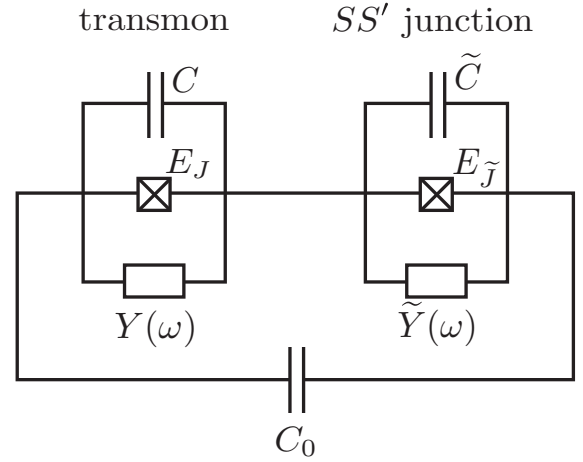


FIG. 2. Dissipative losses induced by the SS' trap. We model both the transmon, as well as the S - S' junction in terms of resistively shunted LC circuits. The Josephson energies of the respective junctions are E_J and $E_{\tilde{J}}$, and the capacitances are denoted as C and \tilde{C} . The dissipation due to quasiparticle tunneling (whereby energy ω is dissipated) is described through the real admittances $Y(\omega)$ and $\tilde{Y}(\omega)$, respectively. The circuits are coupled through the shunting capacitor C_0 .

B. Relaxation due to S - S' tunneling

We now consider the contribution to qubit relaxation from quasiparticle tunneling between superconductors S and S' . For this purpose, we treat the transmon as a nearly harmonic LC oscillator. Within this picture, the dissipation due to quasiparticle tunneling across the S - S Josephson junction can be modeled as a resistive shunt. The resulting qubit decay rate, in the absence of the S - S' junction, is proportional to the RC time of the junction, $\tau_{\text{qubit}}^{-1} \sim \text{Re}[Y(\omega_0)]/C$. Here $Y(\omega_0)$ is the admittance of the S - S junction at the qubit frequency ω_0 [11] and C is the corresponding capacitance.

Within the same picture, we can now think of the S - S' interface as a second LC resonator, coupled to the transmon with a finite capacitive coupling; see Fig. 2. Similarly, quasiparticles can tunnel across the S - S' contact and lead to a resistive shunt; this approach resembles the one used for tunneling at the N - S interface of normal-metal traps [27]. As a reminder, in order to effectively trap quasiparticles, we showed before that a good interface is advantageous, to create a trap with a sufficiently low gap. Crucially, the magnitude of the dissipation depends on the interface quality. In fact, there are several competing effects. Obviously, an improved interface increases Y due to the higher tunneling rate and thus increases the dissipation. Note in particular that the junction conductance can increase by several orders of magnitude, since in the tunneling regime the typical transparency is of order 10^{-5} , while it is of order one for a good contact. On the other hand, a good tunnel coupling can significantly increase the resonance frequency of the second oscillator, leading to a very small coupling between the two circuits, and thus to a decreasing dissipation.

Actually, a further competition occurs concerning the quasiparticle densities. While a good trap diminishes x_{qp} at the junction, and thus increases the RC time of the transmon, the excess quasiparticles will be located within the trap region,

thus decreasing the RC time of the S - S' junction. We therefore want to ensure that the trap does not become a victim of its own success, by reducing the quasiparticle density at the active device parts to the extent, where the additional dissipation caused by the trap could surpass the original dissipative process at the junction. As we will show in the following, this is not a concern for realistic parameters.

In order to estimate which of the competing processes dominates, we deploy for the circuit in Fig. 2 a simplified classical model. The equations of motion for the phase differences across the transmon junction (φ) and across the S - S' interface ($\tilde{\varphi}$) in frequency space are

$$(M_0 - M_{\text{diss}}) \begin{pmatrix} \varphi(\omega) \\ \tilde{\varphi}(\omega) \end{pmatrix} = 0 \quad (21)$$

with the matrix M_0 describing the dissipationless dynamics of the circuit

$$M_0 = \begin{pmatrix} \frac{C+C_0}{2e} \omega^2 - 2eE_J & -\frac{C_0}{2e} \omega^2 \\ -\frac{C_0}{2e} \omega^2 & \frac{\tilde{C}+C_0}{2e} \omega^2 - 2e\tilde{E}_J \end{pmatrix}, \quad (22)$$

where E_J and C are the Josephson energy and capacitance of the transmon junction, respectively. The tilde designates the same parameters for the S - S' junction. Finally, C_0 denotes the shunting capacitor coupling the two junctions. The matrix M_{diss} takes into account the dissipative part of the dynamics,

$$M_{\text{diss}} = \begin{pmatrix} i\frac{g}{e} Y(\omega) & 0 \\ 0 & i\frac{g}{e} \tilde{Y}(\omega) \end{pmatrix}. \quad (23)$$

Here Y and \tilde{Y} are the admittances of the corresponding Josephson junctions, which (as we see later) depend on the junction conductances g_T and \tilde{g}_T , respectively [44]. The Josephson energies can likewise be related to the respective conductances, $E_J = \frac{\Delta}{8} \frac{g_T}{g_0}$ and $\tilde{E}_J = \frac{\Delta}{8} \frac{\tilde{g}_T}{g_0}$, where $g_0 = e^2/2\pi$ is the conductance quantum [45]. For a good S - S' contact, \tilde{g}_T can be many orders of magnitudes higher than g_T , such that $\tilde{E}_J \gg E_J$. We will later discuss how this classical picture can be justified from a full quantum mechanical model (for details, see also Appendix B).

From Eq. (21) we find the condition for the resonance frequencies

$$\alpha^2 \omega^4 = \left[\omega^2 - 2i\omega \frac{Y(\omega)}{C+C_0} - \omega_0^2(1-\alpha^2) \right] \times \left[\omega^2 - 2i\omega \frac{\tilde{Y}(\omega)}{\tilde{C}+C_0} - \tilde{\omega}_0^2(1-\alpha^2) \right]. \quad (24)$$

Here the frequencies $\omega_0 = \sqrt{8E_C E_J}$ and $\tilde{\omega}_0 = \sqrt{8\tilde{E}_C \tilde{E}_J}$ reduce to the resonance frequencies of the two LC circuits in the absence of coupling and dissipation. The charging energies are

$$E_C = \frac{e^2}{2} \frac{\tilde{C} + C_0}{C\tilde{C} + [C + \tilde{C}]C_0}, \quad (25)$$

$$\tilde{E}_C = \frac{e^2}{2} \frac{C + C_0}{C\tilde{C} + [C + \tilde{C}]C_0}, \quad (26)$$

and the dimensionless coupling parameter is

$$\alpha = \frac{C_0}{\sqrt{(C+C_0)(\tilde{C}+C_0)}}. \quad (27)$$

We then make two main assumptions. We assume both a small coupling between the two resonators and a small dissipation (i.e., transmons with a high quality factor). In this limit, we expand ω around the transmon's resonance frequency, $\omega \approx \omega_0 + \delta\omega$, with a small correction $\delta\omega \ll \omega_0$.

This correction has both a real and an imaginary part. The real part corresponds to merely a small shift of the transmons resonance frequency, owing to the coupling. It amounts to

$$\text{Re } \delta\omega = -\frac{1}{2} \frac{\alpha^2 \tilde{\omega}_0^2}{\tilde{\omega}_0^2 - \omega_0^2} \omega_0. \quad (28)$$

Assuming $\tilde{\omega}_0 > \omega_0$, we find that the necessary condition to justify the perturbative expansion is a sufficiently large detuning $\tilde{\omega}_0^2 - \omega_0^2 \gg \alpha^2 \tilde{\omega}_0^2$, which is justified for realistic parameters (see below). Once this condition is satisfied, it is easy to see that $\delta\omega \ll \omega_0$. In fact, we will show that for our system $\tilde{\omega}_0 \gg \omega_0$, where the above condition is equivalent to $\alpha \ll 1$.

Under the same assumptions, we receive for the imaginary part of the correction

$$\text{Im } \delta\omega = \tau_{RC}^{-1} + \tilde{\tau}_{RC}^{-1}. \quad (29)$$

The first term in Eq. (29) corresponds to the RC time of the transmon itself,

$$\tau_{RC}^{-1} = \frac{\text{Re}[Y(\omega_0)]}{C+C_0}, \quad (30)$$

where the losses arise from quasiparticles tunneling across the junction [11]. The second term is the sought after correction,

$$\tilde{\tau}_{RC}^{-1} = \frac{\alpha^2 \omega_0^4}{(\tilde{\omega}_0^2 - \omega_0^2)^2} \frac{\text{Re}[\tilde{Y}(\omega_0)]}{\tilde{C} + C_0}, \quad (31)$$

due to quasiparticles tunneling across the S - S' interface.

In order to proceed, we first need an expression for the admittances which is valid for both low and high contact transparency, in order to account for the improved S - S' interface. We assume for both Y and \tilde{Y} a junction with the same gap parameter on either side. This means that for the S - S' junction, we need to assume a sufficiently strong proximity effect, such that the gap is equal to $\tilde{\Delta}$ on both sides. In this case, we can use the general formula provided by Ref. [46],

$$\text{Re}[Y(\omega)] = \sum_n \frac{1}{\omega} \frac{e^2 \tau_n}{\pi} \int_{\Delta}^{\infty} d\epsilon v(\epsilon) v(\omega + \epsilon) \times |z_n|^2 [f(\epsilon) - f(\omega + \epsilon)] \quad (32)$$

with

$$v(\epsilon) = \frac{\epsilon \sqrt{\epsilon^2 - \Delta^2}}{\epsilon^2 - \Delta^2 [1 - \tau_n \sin^2(\frac{\varphi}{2})]} \quad (33)$$

and

$$|z_n|^2 = \left| 1 + \frac{\Delta^2 [1 + \cos \varphi] - \Delta^2 [1 - \tau_n \sin^2(\frac{\varphi}{2})]}{\epsilon(\omega + \epsilon)} \right|^2. \quad (34)$$

The two distribution functions $f(\epsilon)$ and $f(\epsilon + \omega)$ represent respectively the absorption and emission of an energy quantum by the quasiparticle reservoir. In order to obtain \tilde{Y} we simply take above expressions and replace $\varphi \rightarrow \tilde{\varphi}$, $\Delta \rightarrow \tilde{\Delta}$, $\tau_n \rightarrow \tilde{\tau}_n$ (the latter standing for the channel transparencies of the S - S' junction) and $f \rightarrow \tilde{f}$ (likewise denoting the quasiparticle distribution function at the S - S' junction). Note that the authors in Ref. [46] use a different definition of the phase difference φ than we do. We account for this by dividing their original result by a factor 2.

In the limit $E_C \ll E_J$ and $\tilde{E}_C \ll \tilde{E}_J$, we have both $\varphi \approx 0$ and $\tilde{\varphi} \approx 0$, in accordance with the above approximation of a harmonic oscillator close to ground state, and we find

$$\text{Re}[\tilde{Y}(\omega)] \approx \sum_n \frac{1}{\omega} \frac{e^{2\tilde{\tau}_n}}{\pi} \int_{\tilde{\Delta}}^{\infty} d\epsilon \frac{\epsilon}{\sqrt{\epsilon^2 - \tilde{\Delta}^2}} \frac{\epsilon + \omega}{\sqrt{(\epsilon + \omega)^2 - \tilde{\Delta}^2}} \times \left| 1 + \frac{\tilde{\Delta}^2}{\epsilon(\omega + \epsilon)} \right|^2 [f(\epsilon) - f(\omega + \epsilon)]. \quad (35)$$

Let us now further assume that the quasiparticle distribution functions have a finite effective temperature T_{qp} , such that $f(\tilde{f})$ has only a finite support for ϵ between Δ ($\tilde{\Delta}$) and $\Delta + T_{\text{qp}}$ ($\tilde{\Delta} + T_{\text{qp}}$). In particular, the validity of this assumption for f has to be discussed with care. We expect it to be valid for two reasons. First, for the desired regime of a good contact and a large gap difference, the trapping rate of quasiparticles at energies $\sim \Delta$ approaches the normal metal limit [see Eq. (3)] and is therefore fast compared to the qubit relaxation time [19]. Second, even if there remains some occupation of quasiparticles at energies close to Δ , it is proportional to the steady-state occupation at $\epsilon \approx \tilde{\Delta}$, which will be smaller than the steady-state occupation at $\tilde{\Delta}$; the latter relaxes only via pair annihilation processes, which provide the bottleneck time scale at low temperatures [34]. Therefore, for the integral over all energies, such a residual occupation at high energies will provide but a small correction to \tilde{Y} . In any case, we find

$$\text{Re}[Y(\omega)] = \frac{2}{\omega} \Delta g_T \int_{\Delta}^{\infty} d\epsilon \frac{f(\epsilon) - f(\omega + \epsilon)}{\sqrt{\epsilon - \Delta} \sqrt{\epsilon + \omega - \Delta}}, \quad (36)$$

having in addition defined the normal metal conductance as $g_T = 2g_0 \sum_n \tau_n$. The expression for \tilde{Y} follows again in analogy, with the additional replacement $g_T \rightarrow \tilde{g}_T$, where $\tilde{g}_T = 2g_0 \sum_n \tilde{\tau}_n$. In fact, we have arrived at the same expression as in Ref. [11], where it was derived for tunneling SIS junctions. We thus verified that in the nearly harmonic regime we are considering, this expression is valid for channels with both high and low transparency.

To continue, we need to discuss the dependence of the admittance with respect to the quasiparticle temperature. We can consider both a regime of ‘‘cold’’ quasiparticles, $\omega \gg T_{\text{qp}}$, and ‘‘hot’’ quasiparticles $\omega \ll T_{\text{qp}}$. The former regime is simpler, which is why we discuss it first. Here the emission of energy quanta by the quasiparticle reservoir is suppressed, and we find

$$\text{Re}[Y(\omega)] \approx \frac{1}{2} \left(\frac{2\Delta}{\omega} \right)^{\frac{3}{2}} g_T x_{\text{qp}}, \quad (37)$$

which is in accordance with the expression found in [11]. In fact, with Y and \tilde{Y} in the limit of cold quasiparticles, it is

possible to show that the RC times in Eqs. (30) and (31) follow directly from the full quantum mechanical calculation, shown in Appendix B.

Recent measurements of transmon transition rates [47] have been interpreted in terms of hot nonequilibrium quasiparticles, although a more likely explanation is in terms of photon-assisted Cooper pair breaking [48]. In any case, if there are hot quasiparticles (in or out of equilibrium), the emission process can no longer be ignored. What is more, the classical picture breaks down, a fact we have already discussed in the different context of a model for normal-metal traps [27]. Hence, the circuit relaxation time can no longer be strictly considered an RC time. One can nonetheless restore an effectively classical model with a modified admittance, Y_{mod} , where the absorption and emission processes are *added* instead of subtracted, as in Eq. (36). One can justify this procedure by noting that in the actual quantum mechanical model there occur the rates of qubit relaxation $\Gamma_{1 \rightarrow 0}$ and excitation $\Gamma_{0 \rightarrow 1}$. The relaxation rate of the qubit is then given by the *sum*, $\tau^{-1} = \Gamma_{1 \rightarrow 0} + \Gamma_{0 \rightarrow 1}$. We provide more details on this argument in Appendix B. For the modified admittance, with $\omega \ll T_{\text{qp}}$, we find

$$\text{Re}[Y_{\text{mod}}(\omega)] \approx \sqrt{\frac{\omega}{4T_{\text{qp}}}} \ln \left(\frac{4T_{\text{qp}}}{\omega} \right) \left(\frac{2\Delta}{\omega} \right)^{\frac{3}{2}} g_T x_{\text{qp}}. \quad (38)$$

We see that this result differs from the cold quasiparticle result, Eq. (37), merely by a prefactor $\ln(r)/\sqrt{r}$, depending on the ratio $r = 4T_{\text{qp}}/\omega_0$, which contains no parameters specific to the S - S' junction; i.e., it is the same for both Y and \tilde{Y} . This will simplify the discussion at the end of this section.

Finally, we need to provide estimates for the capacitances. First, both C and C_0 are coplanar capacitors, with different distances between the two plates, D for C and $D + a$ for C_0 ; see also Fig. 1(a). Consequently, we can approximate their ratio by means of the electrostatic result for coplanar capacitors (see Appendix C). In order to determine C , we compute the relation between the accumulated charge across the transmon junction with respect to an applied voltage difference V , such that

$$C \approx \frac{\epsilon_0}{\pi} W \ln \left(\frac{4L}{D} \right), \quad (39)$$

valid for $D \ll L$. With W we denote the width of the superconducting plates and with ϵ_0 the vacuum permittivity. Note that due to the logarithmic dependence on L/D , this geometric factor contributes but a numerical prefactor, which we can discard for our estimate, $C \sim \epsilon W$. As for the capacitance C_0 , we may in principle take into account that the trap part can have a varying length $L - a$; it can be either close to L , such that $a \lesssim D$ and $D + a \ll L - a$, or we could also consider a small trap $D + a \gg L - a$. However, both scenarios provide again only a logarithm of geometric factors, such that likewise $C_0 \sim \epsilon W$. We observe, that while C is generally larger than C_0 , the two are realistically within one order of magnitude of each other (see Appendix C). We therefore use $C \sim C_0$ in the following.

Concerning the capacitance of the second oscillator \tilde{C} , we treat it likewise as a geometric capacitance. However, unlike the previous two capacitances, we here assume that the S'

layer is grown on top of the S layer, such that we have an ordinary parallel plate capacitor,

$$\tilde{C} = \frac{\epsilon_0 \tilde{A}}{d}, \quad (40)$$

with $\tilde{A} = W(L - a)$ being the contact area, and d the thickness of the barrier between the two layers. (see Fig. 1). Because of the good S - S' contact, d is very small, $L - a \gg d$, such that we find $\tilde{C} \gg C$, C_0 , irrespective of the geometric details. As a consequence, we have $\alpha \sim \sqrt{C/\tilde{C}} \ll 1$, because of the dominant \tilde{C} . This validates our initial hypothesis of the circuits being weakly coupled, in spite of $C \sim C_0$.

Having expressions for the capacitances in our hands, we now proceed by comparing the resonance frequencies of the two oscillators. Importantly, both $\tilde{C} \gg C$ and $\tilde{E}_J \gg E_J$, such that it is not yet obvious whether $\tilde{\omega}_0$ is much larger than ω_0 or not. As for the first factor, we expect d to be at most on the nanometer scale, whereas $L - a$ can be as large as hundreds of micrometers to a millimeter, resulting in $\tilde{C}/C \lesssim 10^6$. As for the Josephson energies, we estimate their ratio as follows. We use the relation between the Josephson energies and their respective contact conductances, given above. We therefore need to estimate the ratio g_T/\tilde{g}_T . In order to do so, let us assume for simplicity that all channels have roughly the same transparency, such that $g_T/g_0 \approx N_{\text{ch}}\langle\tau\rangle$, where N_{ch} is the number of channels, and $\langle\tau\rangle$ is the average over the channel ensemble of the transmission probabilities (and similarly for \tilde{g}_T). For weakly coupled tunnel junctions, we take $\langle\tau\rangle \sim 10^{-5}$ [19]; this estimate applies to g_T . For very good contacts (i.e., for \tilde{g}_T) on the other hand, $\langle\tau\rangle$ should in principle be much larger, of the order 10^{-1} to 1; this is needed to ensure strong proximity effect, and it is confirmed by experiments [49,50]. In order to estimate the ratio of channel numbers, we note that the number of channels scales with the contact area, such that $N_{\text{ch}}/\tilde{N}_{\text{ch}} = A/\tilde{A}$, where A is the contact area of the transmon junction. This ratio depends strongly on the transmon and trap geometry. Typically, $A \sim (0.1 \mu\text{m})^2$. For a small trap, \tilde{A} could be as low as $10 \mu\text{m} \times 100 \mu\text{m}$. A large trap, covering a good portion of the transmon, can be at least one order of magnitude larger ($\sim 100 \mu\text{m} \times 100 \mu\text{m}$). Let us take the small trap and $\langle\tilde{\tau}\rangle \sim 10^{-1}$ as an upper bound for the ratio, such that

$$\frac{g_T}{\tilde{g}_T} < 10^{-9}. \quad (41)$$

The ratio E_J/\tilde{E}_J between the Josephson energies may have a slightly less stringent upper bound, due to the additional factor $\Delta/\tilde{\Delta} > 1$; however, the two energy gaps will differ in practice by at most one order of magnitude, since we require $\omega_0 < 2\tilde{\Delta}$ to avoid pair-breaking processes. Consequently, we find that

$$\frac{\tilde{\omega}_0}{\omega_0} > 10, \quad (42)$$

indicating that the increase of the contact conductance is the dominating effect. This means that the first excited state of the transmon remains the lowest energetically accessible state, such that we do not have to concern ourselves with any relaxation process of the qubit through exciting the SS' oscillator.

Finally, we have all elements at disposal to estimate the RC times. We first discuss the contribution of the S - S' junction, in

particular its dependence on the junction conductance. Due to $\tilde{\omega}_0 \gg \omega_0$ and $\tilde{C} \gg C \sim C_0$, we find

$$\tilde{\tau}_{RC}^{-1} \sim \frac{1}{C} \sqrt{\frac{\tilde{\Delta}}{\Delta}} \frac{g_T}{\tilde{g}_T} \left(\frac{\Delta}{\omega_0}\right)^{\frac{3}{2}} g_T \tilde{x}_{\text{qp}}. \quad (43)$$

We note that the above result is correct as such for cold quasiparticles, whereas for the hot regime we would have to include the prefactor depending on T_{qp}/ω_0 , as identified in Eq. (38). We discard it for lack of relevance: as mentioned right after Eq. (38), this prefactor does not depend on junction specific parameters (that is, it is independent of \tilde{g}_T). With the above result, we can answer the first question whether an increase in the interface quality increases or decreases the qubit relaxation rate. As we see, even though the admittance increases linearly with \tilde{g}_T , the prefactor $1/\tilde{\omega}_0^4 \sim 1/\tilde{g}_T^2$ decreases faster than that; hence the positive effect of the decoupling due to frequency mismatch dominates over the negative effect of the increased tunneling rate. Improving the S - S' interface is therefore beneficial.

We now compare the two RC times, in order to ensure that $\tilde{\tau}_{RC} > \tau_{RC}$. As mentioned above, otherwise we would encounter a counterproductive case where the trap dissipation due to quasiparticle tunneling would be larger than the original quasiparticle process the trap was designed to mitigate. The RC time of the transmon can likewise be expressed in terms of circuit parameters and quasiparticle density as (for cold quasiparticles)

$$\tau_{RC}^{-1} \sim \frac{1}{C} \left(\frac{\Delta}{\omega_0}\right)^{\frac{3}{2}} g_T x_{\text{qp}}. \quad (44)$$

Through this, we find

$$\frac{\tau_{RC}}{\tilde{\tau}_{RC}} \sim \sqrt{\frac{\tilde{\Delta}}{\Delta}} \frac{g_T}{\tilde{g}_T} \frac{\tilde{x}_{\text{qp}}}{x_{\text{qp}}}. \quad (45)$$

This result is valid for both cold and hot quasiparticles [as the prefactor identified in Eq. (38) drops out from the ratio]. Importantly, the ratio between the two RC times has both large and small factors. While the ratio of gaps satisfies $\sqrt{\tilde{\Delta}/\Delta} > 1$, it will not be a significant factor and can be simplified to ~ 1 . The improved contact conductance works in our favor and provides a factor of at least 10^{-9} ; see Eq. (41). As far as the quasiparticle density is concerned, it has been shown [21] that a good trap can reduce x_{qp} at the junction by about two orders of magnitude with respect to the stationary density $x_{\text{qp}}^{\text{st}}$ in the absence of the trap [51]; that is, for a typical value $x_{\text{qp}}^{\text{st}} \sim 10^{-6}$ we can expect ideally $x_{\text{qp}} \sim 10^{-8}$ at the junction. The quasiparticle density at the S - S' junction \tilde{x}_{qp} on the other hand can be significantly higher than $x_{\text{qp}}^{\text{st}}$. After all, the trap has a lower gap, $\tilde{\Delta} < \Delta$ and is designed to attract and collect nonequilibrium quasiparticles from all over the device, which recombine very slowly. However, in order to reach the regime $\tau_{RC}^{-1} \sim \tilde{\tau}_{RC}^{-1} \tilde{x}_{\text{qp}}$ would have to be larger than unity, implying the full suppression of superconductivity in the trap. We expect this to be very far from a realistic scenario for transmon qubits: as we have already noted, even in experiments with quasiparticle injection, the nonequilibrium quasiparticle density accumulated at the junction does not exceed 10^{-3} [14,19]. We can therefore safely conclude that

$\tilde{\tau}_{RC} \gg \tau_{RC}$. Consequently, the dissipation due to quasiparticle tunneling across the S - S' junction cannot severely impede the qubit quality factor, in spite of the increased contact conductance and the high quasiparticle density in the trap. The same conclusion is also reached, using Eq. (41) in Eq. (45), in the limiting case of an ineffective trap with $\tilde{\Delta} \rightarrow \Delta$, in which relaxation cannot take place and hence $\tilde{x}_{qp} = x_{qp}$. We finally note that even for traps as short as a fraction of a micron, our result still holds.

IV. SUMMARY

We have theoretically examined the quasiparticle dynamics and dissipative processes of gap-engineered traps in transmon qubits. In our model, the gap engineering is realized through the proximity effect, by coupling a second, lower gap superconductor to the transmon, with a good contact of high transparency.

We have identified several advantages with respect to normal-metal traps studied in earlier works. As far as the quasiparticle diffusion is concerned we have predicted that gap-engineered traps, unlike normal-metal traps, do not suffer from quasiparticle backflow (i.e., escape from the trap back to the qubit). As a consequence, the minimal trap size above which the traps are in a regime of strong and efficient quasiparticle evacuation is reduced. Therefore, strong gap-engineered traps are available for even smaller devices, so that the present proposal should be viable not only for 3D transmons as considered here, but also C-shunted flux qubits [52] and X-mons [53]. For qubits comprising only small islands, such as a “regular” (not shunted) flux qubit or a fluxonium with the superinductor implemented via a junction array [3], the design considered is likely difficult to implement directly, since the minimal trap size is of the order of the typical island size; still, our approach could be used in devices coupled to radiation via a larger antenna [54]. Similarly, in devices based on proximitized nanowires such as Andreev level qubits [55] (which can be “poisoned” by a single quasiparticle), the fabrication of traps on the nanowire itself could be challenging, but they could be useful in the superconductors contacting the nanowire. We also note that alternatives to quasiparticle traps are already under consideration, such as quasiparticle pumping [24] to expel quasiparticles from the small islands, phonon traps [56] which can reduce the generation rate of quasiparticles, and “quasiparticle filters” for nanowire devices [57].

Concerning dissipative processes, we have studied the ac resistivity of the bulk superconductor and found that the resistivity drops by many orders of magnitudes with respect to the normal metal case. Consequently, bulk dissipation arising due to charge redistribution, which was a critical issue for normal metal traps, is no longer of any concern. Finally, we have investigated losses due to quasiparticle tunneling at the proximity interface. We have found that even though the admittance of the interface increases significantly, due to the combination of a good contact and an increased quasiparticle density in the trap region, the overall influence of the trap is negligible. The main reason for this is that a good contact strongly detunes the transmon and trap resonance frequencies, and effectively decouples the two. Given the above advantages, we come to the conclusion that gap engineering

may be the preferable strategy to implement quasiparticle traps.

ACKNOWLEDGMENTS

We gratefully acknowledge useful discussions with A. Hosseinkhani and K. Serniak. This work was supported in part by a Max Planck award (R.-P.R.).

APPENDIX A: JUSTIFICATION OF DIFFUSION EQUATION FROM MICROSCOPIC EQUATIONS

In the main text, we provide heuristic boundary conditions for the quasiparticle density x_{qp} at the trap edge, where the order parameter drops. We here show how to justify these boundary conditions from a microscopic vantage point.

For a proximitized piece of a dirty superconductor with an inhomogeneous gap parameter $\Delta(x)$, the diffusion equation providing the dynamics of the quasiparticle occupation number as a function of energy $f_{qp}(\epsilon)$ can be given as [29]

$$\dot{f}_{qp}(\epsilon, y) = \partial_y [D_{qp}(\epsilon, y) \partial_y f_{qp}(\epsilon, y)], \quad (A1)$$

where we assume that the charge imbalance contribution is negligible. In the above equation, the inhomogeneous gap $\Delta(y)$ enters through the y -dependent diffusion coefficient $D_{qp}(\epsilon, y) = D_0 / v_{BCS}(\epsilon)$, where D_0 is the normal state diffusion coefficient, and $v_{BCS}(\epsilon) = \epsilon / \sqrt{\epsilon^2 - \Delta^2(y)}$ is the normalized BCS density of states.

We know that the gap parameter for a proximitized slab of superconductor varies on the length scale given by the superconducting coherence length ξ (see, e.g., Ref. [58]), which is much smaller than the length scales relevant for the diffusion of quasiparticles. We can therefore separate above equation into two parts with a constant gap, a first part for $0 < y < a$ with gap $\Delta(y) = \Delta$ and a second part for $a < y < L$ with $\Delta(y) = \tilde{\Delta}$ (we assume $a, L - a \gg \xi$). The region where the gap abruptly jumps can now be incorporated through the boundary conditions [which are straightforwardly derived from Eq. (A1)],

$$f_{qp}^-(\epsilon) = f_{qp}^+(\epsilon), \quad (A2)$$

$$D_{qp}^-(\epsilon) \partial_y f_{qp}^-(\epsilon) = D_{qp}^+(\epsilon) \partial_y f_{qp}^+(\epsilon), \quad (A3)$$

with the notation $q^- = q|_{y=a-0^+}$ and $q^+ = q|_{y=a+0^+}$. We stress that in Eq. (A3), the diffusion constant does not cancel, because it is not the same for $y = a - 0^+$ and $y = a + 0^+$.

We note that the derivation of an effective boundary for x_{qp} based on above Eqs. (A2) and (A3) is not trivial, due to the density of states being position dependent. We can however use a certain set of assumptions on the function f_{qp} through which we can still find an effective boundary condition for x_{qp} . The procedure goes as follows.

First, we assume that the quasiparticle distribution function can be written as a product, $f_{qp}(\epsilon, y) = g(\Delta\epsilon) * h(y)$, where $\Delta\epsilon = \epsilon - \Delta(y)$. We furthermore assume that $g(\Delta\epsilon)$ has a finite support within the window $0 < \Delta\epsilon < k_B T_{qp}$, which we refer to as the quasiparticle temperature T_{qp} . We focus on the regime of cold quasiparticles, $k_B T_{qp} \ll \Delta(y)$ (we note that this set of approximations has already been proven useful when describing experiments; see, e.g., [19]). For the free

and the proximitized parts we get independently the diffusion equations on either side of $y = a$ as

$$\dot{x}_{\text{qp}}(y) = D_{\text{qp}}^{\mp} \partial_y^2 x_{\text{qp}}(y), \quad (\text{A4})$$

where the phenomenological diffusion coefficient is given as

$$D_{\text{qp}}^{\mp} = \frac{\int_{\Delta^{\mp}}^{\infty} d\epsilon \nu_{\text{BCS}}(\epsilon) g(\epsilon - \Delta^{\mp}) D_{\text{qp}}(\epsilon)}{\int_{\Delta^{\mp}}^{\infty} d\epsilon \nu_{\text{BCS}}(\epsilon) g(\epsilon - \Delta^{\mp})}, \quad (\text{A5})$$

with $\Delta^- = \Delta$ and $\Delta^+ = \tilde{\Delta}$, for $y < a$ and $y > a$, respectively. Under the above assumptions, we indeed find $D_{\text{qp}}^- \ll D_{\text{qp}}^+$. We can show this explicitly, by making the crude simplification, $g(\Delta\epsilon) \approx \theta(\Delta\epsilon)\theta(k_{\text{B}}T_{\text{qp}} - \Delta\epsilon)$, such that $D_{\text{qp}}^{\mp} \sim \sqrt{k_{\text{B}}T_{\text{qp}}/\Delta^{\mp}} D_0$. We stress, however, that the above result holds qualitatively (up to numerical prefactors) also for more realistic assumptions for g , as long as g does not have a too pronounced energy dependence (apart from the cutoff).

We now derive explicitly the boundary conditions. We choose to multiply Eqs. (A2) and (A3) with the BCS density of states on the nonproximitized side, where $\Delta(y) = \Delta$, because x_{qp} at the junction is of most interest to us. Under the same assumptions as above, integrating the resulting equations provides us with the following expressions:

$$x_{\text{qp}}^- = \sqrt{2} \sqrt{\frac{\Delta^2 - \tilde{\Delta}^2}{\Delta T_{\text{qp}}}} x_{\text{qp}}^+, \quad (\text{A6})$$

$$D_{\text{qp}}^- \partial_y x_{\text{qp}}^- = \sqrt{2} \sqrt{\frac{\Delta^2 - \tilde{\Delta}^2}{\Delta T_{\text{qp}}}} D_{\text{qp}}^+ \partial_y x_{\text{qp}}^+. \quad (\text{A7})$$

To arrive at above equations, we have furthermore assumed that $T_{\text{qp}} \ll \Delta - \tilde{\Delta}$. Importantly, both Eqs. (A6) and (A7) contain the same correction prefactor. We can thus conclude that the heuristic boundary condition given in the main text, Eq. (7), can be justified from the microscopic starting point here, Eq. (A1), when renormalizing x_{qp} for $y > a$ by this prefactor.

APPENDIX B: DISSIPATION DUE TO PHOTO-ASSISTED TUNNELING IN THE QUANTUM CIRCUIT APPROACH

In the main text, we deploy a classical model for the dynamics of the transmon with an SS' trap, as coupled LC circuits. Here, we indicate how one can justify the classical model by means of a full quantum mechanical model.

For this purpose, we first compute the nondissipative part of the dynamics, disregarding the effects of the quasiparticles. Our starting point is the Lagrangian for the phase differences φ and $\tilde{\varphi}$ across the transmon and S - S' junction, respectively,

$$\begin{aligned} L = & \frac{1}{2} \frac{C}{(2e)^2} \dot{\varphi}^2 + E_J \cos(\varphi) \\ & + \frac{1}{2} \frac{\tilde{C}}{(2e)^2} \dot{\tilde{\varphi}}^2 + \sum_n \tilde{\Delta} \sqrt{1 - \tau_n \sin^2\left(\frac{\tilde{\varphi}}{2}\right)} \\ & + \frac{1}{2} \frac{C_0}{(2e)^2} (\dot{\varphi} - \dot{\tilde{\varphi}})^2, \end{aligned} \quad (\text{B1})$$

with the same capacitances as in the model given in Fig. 2. We perform the Legendre transformation to arrive at the

Hamiltonian. The variables conjugate to the phase difference operators are

$$n = \frac{\partial L}{\partial \dot{\varphi}} = \frac{C + C_0}{(2e)^2} \dot{\varphi} - \frac{C_0}{(2e)^2} \dot{\tilde{\varphi}}, \quad (\text{B2})$$

$$\tilde{n} = \frac{\partial L}{\partial \dot{\tilde{\varphi}}} = \frac{\tilde{C} + C_0}{(2e)^2} \dot{\tilde{\varphi}} - \frac{C_0}{(2e)^2} \dot{\varphi}, \quad (\text{B3})$$

with the quantization conditions $[n, \varphi] = i$ and $[\tilde{n}, \tilde{\varphi}] = i$, respectively. Furthermore, we assume small phase differences $\varphi \approx 0$ and $\tilde{\varphi} \approx 0$. Thus we arrive at a system of two coupled harmonic oscillators,

$$H = \frac{1}{2} E_C n^2 + \frac{1}{2} E_J \varphi^2 + \frac{1}{2} \tilde{E}_C \tilde{n}^2 + \frac{1}{2} \tilde{E}_J \tilde{\varphi}^2 + E_{C_0} n \tilde{n}, \quad (\text{B4})$$

with

$$E_C = (2e)^2 \frac{\tilde{C} + C_0}{\tilde{C}\tilde{C} + [C + \tilde{C}]C_0}, \quad (\text{B5})$$

$$\tilde{E}_C = (2e)^2 \frac{C + C_0}{\tilde{C}\tilde{C} + [C + \tilde{C}]C_0}, \quad (\text{B6})$$

$$E_{C_0} = (2e)^2 \frac{C_0}{\tilde{C}\tilde{C} + [C + \tilde{C}]C_0}. \quad (\text{B7})$$

Here to simplify the comparison with textbook formulas for the harmonic oscillator, we use a different definition of charging energies compared to the standard one adopted in the main text. With these definitions, we see that the harmonic oscillator approximation is valid for $E_C \ll E_J$ as well as $\tilde{E}_C \ll \tilde{E}_J$ (we remind readers that $\tilde{E}_J = \tilde{\Delta}/4 \sum_n \tau_n$). We continue by introducing ladder operators for the harmonic oscillators,

$$n = \frac{i}{\sqrt{2}} \left(\frac{E_J}{E_C} \right)^{1/4} (a^\dagger - a), \quad (\text{B8})$$

$$\varphi = \frac{1}{\sqrt{2}} \left(\frac{E_C}{E_J} \right)^{1/4} (a^\dagger + a). \quad (\text{B9})$$

The operators \tilde{n} and $\tilde{\varphi}$ are obtained in analogy, by replacing $E_C \rightarrow \tilde{E}_C$ and $E_J \rightarrow \tilde{E}_J$. We arrive at

$$\begin{aligned} H = & \omega_0 \left(a^\dagger a + \frac{1}{2} \right) + \tilde{\omega}_0 \left(\tilde{a}^\dagger \tilde{a} + \frac{1}{2} \right) \\ & - \frac{1}{2} \alpha \sqrt{\omega_0 \tilde{\omega}_0} (a^\dagger - a) (\tilde{a}^\dagger - \tilde{a}) \end{aligned} \quad (\text{B10})$$

with

$$\alpha = \frac{E_{C_0}}{\sqrt{E_C \tilde{E}_C}} = \frac{C_0}{\sqrt{(C + C_0)(\tilde{C} + C_0)}}, \quad (\text{B11})$$

just as in the classical equation in the main text, Eq. (27). The eigenstates for zero coupling can be denoted as $|m, \tilde{m}\rangle = (a^\dagger)^m (\tilde{a}^\dagger)^{\tilde{m}} |0, 0\rangle$, where $|0, 0\rangle$ is the ground state without excitations.

We now perform a perturbation theory for a small coupling parameter, $\alpha \ll 1$. In lowest order, we receive the corrections for the general circuit states

$$|m, \tilde{m}\rangle \rightarrow |m, \tilde{m}\rangle + |\delta_{m, \tilde{m}}\rangle. \quad (\text{B12})$$

The corrections have the following nonzero elements:

$$\langle m+1, \tilde{m}+1 | \delta_{m, \tilde{m}} \rangle = \frac{\alpha}{2} \sqrt{\omega_0 \tilde{\omega}_0} \frac{\sqrt{(m+1)(\tilde{m}+1)}}{\tilde{\omega}_0 + \omega_0}, \quad (\text{B13})$$

$$\langle m+1, \tilde{m}-1 | \delta_{m, \tilde{m}} \rangle = \frac{\alpha}{2} \sqrt{\omega_0 \tilde{\omega}_0} \frac{\sqrt{(m+1)\tilde{m}}}{\tilde{\omega}_0 - \omega_0}, \quad (\text{B14})$$

$$\langle m-1, \tilde{m}+1 | \delta_{m, \tilde{m}} \rangle = -\frac{\alpha}{2} \sqrt{\omega_0 \tilde{\omega}_0} \frac{\sqrt{m(\tilde{m}+1)}}{\tilde{\omega}_0 - \omega_0}, \quad (\text{B15})$$

$$\langle m-1, \tilde{m}-1 | \delta_{m, \tilde{m}} \rangle = -\frac{\alpha}{2} \sqrt{\omega_0 \tilde{\omega}_0} \frac{\sqrt{m\tilde{m}}}{\tilde{\omega}_0 + \omega_0}. \quad (\text{B16})$$

We now have all ingredients at hand to compute the dissipation due to quasiparticles. We are in particular interested in transitions between the qubit states,

$$|00\rangle, \quad |10\rangle = a^\dagger |00\rangle, \quad (\text{B17})$$

assuming that the SS' resonator remains in its ground state. This is a realistic assumption, since $\omega_0 \ll \tilde{\omega}_0$. As shown, e.g., in Ref. [11], the quasiparticle tunneling process couples to the operators $\sin(\varphi/2)$ (for tunneling across the transmon junction) or $\sin(\tilde{\varphi}/2)$ (across the SS' junction). In the limit of harmonic dynamics, it suffices to consider the approximation $\sin(\varphi/2) \approx \varphi/2$ (and likewise for $\tilde{\varphi}$). We then receive

$$|\langle 10 | \varphi | 00 \rangle|^2 \approx \frac{1}{2} \frac{E_C}{\omega_0}, \quad (\text{B18})$$

$$|\langle 10 | \tilde{\varphi} | 00 \rangle|^2 \approx \frac{1}{2} \frac{\tilde{E}_C}{\omega_0} \frac{\alpha^2 \omega_0^4}{(\tilde{\omega}_0^2 - \omega_0^2)^2}, \quad (\text{B19})$$

up to leading order in the respective transition matrix elements. We note that within the approximation $\alpha \ll 1$, the respective charging energies reduce to $E_C \approx (2e)^2/(C + C_0)$ and $\tilde{E}_C \approx (2e)^2/(\tilde{C} + C_0)$. Already on this level we see the appearance of the same prefactor structure as in the classical equation in the main text, Eq. (31).

In order to arrive at the full RC time including the quasiparticle density, we follow Ref. [11], where it is shown how the above computed transition matrix element enters in the total transition rate. Namely, the transition rates are

$$\Gamma_{1 \rightarrow 0} = \frac{1}{4} |\langle 10 | \varphi | 00 \rangle|^2 S_{\text{qp}}(\omega_0), \quad (\text{B20})$$

$$\tilde{\Gamma}_{1 \rightarrow 0} = \frac{1}{4} |\langle 10 | \tilde{\varphi} | 00 \rangle|^2 \tilde{S}_{\text{qp}}(\omega_0), \quad (\text{B21})$$

with

$$S_{\text{qp}}(\omega) \approx \frac{16E_J}{\pi} \int_{\Delta}^{\infty} d\epsilon \frac{f_{\text{qp}}(\epsilon)}{\sqrt{\epsilon - \Delta} \sqrt{\epsilon + \omega - \Delta}}, \quad (\text{B22})$$

$$\tilde{S}_{\text{qp}}(\omega) \approx \frac{16\tilde{E}_J}{\pi} \int_{\tilde{\Delta}}^{\infty} d\epsilon \frac{\tilde{f}_{\text{qp}}(\epsilon)}{\sqrt{\epsilon - \tilde{\Delta}} \sqrt{\epsilon + \omega - \tilde{\Delta}}}. \quad (\text{B23})$$

In the reversed rates $\Gamma_{0 \rightarrow 1}$ and $\tilde{\Gamma}_{0 \rightarrow 1}$, there enter the time-reversed correlation functions,

$$S_{\text{qp}}^{\text{rev}}(\omega) \approx \frac{16E_J}{\pi} \int_{\Delta}^{\infty} d\epsilon \frac{f_{\text{qp}}(\epsilon + \omega)}{\sqrt{\epsilon - \Delta} \sqrt{\epsilon + \omega - \Delta}}, \quad (\text{B24})$$

$$\tilde{S}_{\text{qp}}^{\text{rev}}(\omega) \approx \frac{16\tilde{E}_J}{\pi} \int_{\tilde{\Delta}}^{\infty} d\epsilon \frac{\tilde{f}_{\text{qp}}(\epsilon + \omega)}{\sqrt{\epsilon - \tilde{\Delta}} \sqrt{\epsilon + \omega - \tilde{\Delta}}}. \quad (\text{B25})$$

Note that we chose $\omega > 0$ by default. The remaining task is to relate the quasiparticle correlation functions to the admittance.

They are generally related to each other as

$$S_{\text{qp}}(\omega) - S_{\text{qp}}^{\text{rev}}(\omega) = \frac{2\omega \text{Re}Y(\omega)}{\pi g_0}, \quad (\text{B26})$$

and similarly for \tilde{Y} . Let us now again assume that the quasiparticles occupy a finite energy window, given by the effective quasiparticle temperature T_{qp} . If the quasiparticles are cold with respect to the qubit energy scale, $\omega_0 \gg T_{\text{qp}}$, then we can neglect excitation processes, and the above relation simplifies to the so-called ‘‘high-frequency’’ case,

$$S_{\text{qp}}(\omega) \approx \frac{2\omega \text{Re}Y(\omega)}{\pi g_0}. \quad (\text{B27})$$

We finally get

$$\Gamma_{1 \rightarrow 0} = \frac{1}{4} \frac{E_C}{\pi} \frac{\text{Re}Y(\omega_0)}{g_0}, \quad (\text{B28})$$

$$\tilde{\Gamma}_{1 \rightarrow 0} = \frac{1}{4} \frac{\tilde{E}_C}{\pi} \frac{\alpha^2 \omega_0^4 \text{Re}\tilde{Y}(\omega_0)}{(\tilde{\omega}_0^2 - \omega_0^2)^2 g_0}, \quad (\text{B29})$$

which can easily be identified as the respective RC times in the main text; see Eqs. (30) and (31).

As already pointed out in the main text, the quasiclassical approximation fails with certainty, if the quasiparticle distribution is hot, $\omega < T_{\text{qp}}$. Here we justify the appearance of the modified admittance introduced in the main text, by means of the here described transition rates. We note that we have already commented on a similar scenario in a previous paper [27]. In essence, for hot quasiparticles, both the qubit relaxation and excitation are equally probable, $S_{\text{qp}} \approx S_{\text{qp}}^{\text{rev}}$. The decay rate of the qubit system can now be captured by the ‘‘new’’ RC time, $\tau_{RC}^{-1} = \Gamma_{1 \rightarrow 0} + \Gamma_{0 \rightarrow 1}$. This justifies the definition of a modified admittance

$$\text{Re}[Y^{\text{mod}}(\omega)] = \frac{2\Delta}{\omega} g_T \int_{\Delta}^{\infty} d\epsilon \frac{f(\epsilon) + f(\omega + \epsilon)}{\sqrt{\epsilon - \Delta} \sqrt{\epsilon + \omega - \Delta}} \quad (\text{B30})$$

(and likewise for $\tilde{\tau}_{RC}$), where the two fermi functions at different energies are added rather than subtracted. This equation gives rise to Eq. (38).

APPENDIX C: DERIVATION OF COPLANAR CAPACITANCES

We here derive the capacitances of finite size coplanar capacitors, as used in the main text. This calculation is a generalization of the result obtained in Ref. [27], where the coplanar capacitors were semi-infinite. We start from infinitesimally thin plates in the z direction (positioned at $z = 0$), infinitely large in x and with finite size in the y direction. The distance of the two plates is given by D and their length as L (see also Fig. 1). We choose to set the origin of y in the middle of the two plates, such that there is a left plate for $-L - D/2 < y < -D/2$ and a right plate for $D/2 < y < D/2 + L$. This problem can be mapped to a parallel plate capacitor through the conformal map,

$$\zeta = K \int^{\tau} d\omega \frac{1}{\sqrt{\omega^2 - 1} \sqrt{(l+1)^2 - \omega^2}}, \quad (\text{C1})$$

where $\tau = 2(x + iz)/D$ is the target space of the coplanar capacitors, and ζ is the map to the space of the parallel plate

capacitor, with the prefactor K to be determined below. We defined furthermore $l = 2L/D$. We identify the integral as the incomplete elliptic integral of the first kind, defined as

$$F\left(\frac{\tau}{l+1}; l+1\right) = \int_0^{\frac{\tau}{l+1}} dt \frac{1}{\sqrt{1-t^2}\sqrt{1-(l+1)^2t^2}}, \quad (\text{C2})$$

such that

$$\zeta(\tau) = iKF\left(\frac{\tau}{l+1}; l+1\right). \quad (\text{C3})$$

The elliptic integral has the following asymptotic behaviors. For $l \gg 1$,

$$F\left(\frac{\tau}{l+1}; l+1\right) \rightarrow i\frac{\pi}{2} \frac{1}{(l+1)}, \quad (\text{C4})$$

whereas for $l \ll 1$,

$$F\left(\frac{\tau}{l+1}; l+1\right) \approx \frac{1}{2} \ln\left(\frac{1}{L}\right). \quad (\text{C5})$$

We now need to fix the prefactor K . We do so by demanding that the point $y + iz = D/2$, that is, $\tau = 1$, on the coplanar capacitor corresponds to the point $\zeta = i/2$ in the parallel capacitor space

$$\zeta(\tau = 1) \stackrel{!}{=} \frac{i}{2}, \quad (\text{C6})$$

resulting in

$$K = \frac{1}{2F\left(\frac{1}{l+1}; l+1\right)}. \quad (\text{C7})$$

Eventually, in order to compute the capacitance, we need to know the surface charge on the capacitor plates. We find them as follows. We first solve the parallel plate capacitor in ζ space for a given voltage difference. In this space, the solution of the potential is simply

$$\varphi'(\zeta) = -V\text{Im}[\zeta], \quad (\text{C8})$$

assuming that the parallel plates are situated at $\text{Im}\zeta = -1/2$ and $\text{Im}\zeta = 1/2$. This is translated to the coplanar capacitor in τ space as $\varphi(\tau) = \varphi'[\zeta(\tau)]$. The surface charge σ is then given in the standard way as the normal component of the electric field,

$$\sigma(y) = \epsilon_0 E_n(y) = -\epsilon_0 \lim_{\delta \rightarrow 0} \frac{\varphi(\tau + i\delta) - \varphi(\tau)}{\delta} \Big|_{\tau=2y/D} \quad (\text{C9})$$

with ϵ_0 the vacuum permittivity. We receive

$$\sigma(y) = \frac{\epsilon_0 V}{2F\left(\frac{1}{\frac{l}{2}+1}; \frac{2L}{D} + 1\right)} \frac{1}{\sqrt{y^2 - \left(\frac{D}{2}\right)^2} \sqrt{\left(\frac{D}{2} + L\right)^2 - y^2}}. \quad (\text{C10})$$

In the limit of infinitely large capacitor plates, $L \gg D/2$, we recover our result from Ref. [27]

$$\sigma(y) \approx \frac{\epsilon_0 V}{\pi} \frac{1}{\sqrt{y^2 - \left(\frac{D}{2}\right)^2}}. \quad (\text{C11})$$

This is the result we need to estimate the transmon capacitance C and the cross capacitance C_0 for large traps. The opposite limit of $L \ll D/2$ can be used to estimate the cross capacitance C_0 of small traps, where

$$\sigma(y) \approx \frac{\epsilon_0 V}{2 \ln\left(\frac{D}{2L}\right)} \frac{1}{\sqrt{y - \frac{D}{2}} \sqrt{\frac{D}{2} + L - y}}. \quad (\text{C12})$$

A coplanar capacitance C_{co} can be found in the standard way, by taking the ratio between the total charge stored in the capacitor and the voltage difference. In order to regularize the total charge, we introduce the finite (but large) plate width W in the x direction, such that

$$C_{\text{co}} = \frac{Q_{\text{co}}}{V} = \frac{\int_0^W dt \int_{\frac{D}{2}}^{L+\frac{D}{2}} dy \sigma(y)}{V}. \quad (\text{C13})$$

For large plates we thus receive

$$C_{\text{co}} \approx \frac{\epsilon_0}{\pi} W \ln\left(\frac{4L}{D}\right). \quad (\text{C14})$$

For small plates, the capacitance is

$$C_{\text{co}} \approx \frac{\pi}{2} \epsilon_0 \frac{W}{\ln\left(\frac{D}{2L}\right)}. \quad (\text{C15})$$

Note that this result does not precisely correspond to the cross conductance C_0 for small traps. Namely, in this Appendix we consider symmetric coplanar capacitors. The actual setup treated in the main text is in contrast an asymmetric coplanar capacitor, where charges are separated between a large transmon plate and a small trap plate. However, we note that the capacitance of both the small and large (symmetric) coplanar capacitors provides the same scaling, differing only up to logarithmic factors. We can likewise see when doing the spatial integral exactly, that even when crossing over from one regime to the other there appear no additional scaling factors other than W . We have therefore no reason to expect that the scaling will be different for the asymmetric case, both on physical grounds as well as from the above scaling argument.

Finally, let us comment on the assumption $C \sim C_0$ used in Sec. III. Equation (C14) applies to both C and C_0 if, for the latter, $D + a \ll L - a$ and the substitutions $L \rightarrow L - a$ and $D \rightarrow D + a$ are made; then we immediately see that $C/C_0 \gtrsim 1$. In the opposite regime $D + a \gg L - a$, we must use Eq. (C15) for C_0 (again with the above substitutions); realistically, even for long plates at a typical distance ($L = 500 \mu\text{m}$, $D = 30 \mu\text{m}$) and a very small trap $L - a = 1 \mu\text{m}$), we estimate $C/C_0 \simeq 4.7$, therefore justifying the assumption $C \sim C_0$.

[1] M. H. Devoret and J. M. Martinis, in *Experimental Aspects of Quantum Computing*, edited by H. O. Everitt (Springer US, Boston, 2005), pp. 163–203.

[2] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, *Phys. Rev. A* **76**, 042319 (2007).

- [3] V. E. Manucharyan, J. Koch, L. I. Glazman, and M. H. Devoret, *Science* **326**, 113 (2009).
- [4] R. Lutchyn, L. Glazman, and A. Larkin, *Phys. Rev. B* **72**, 014517 (2005).
- [5] G. Catelani, J. Koch, L. Frunzio, R. J. Schoelkopf, M. H. Devoret, and L. I. Glazman, *Phys. Rev. Lett.* **106**, 077002 (2011).
- [6] J. Leppäkangas and M. Marthaler, *Phys. Rev. B* **85**, 144503 (2012).
- [7] M. D. Shaw, R. M. Lutchyn, P. Delsing, and P. M. Echternach, *Phys. Rev. B* **78**, 024503 (2008).
- [8] J. M. Martinis, M. Ansmann, and J. Aumentado, *Phys. Rev. Lett.* **103**, 097002 (2009).
- [9] H. Paik, D. I. Schuster, L. S. Bishop, G. Kirchmair, G. Catelani, A. P. Sears, B. R. Johnson, M. J. Reagor, L. Frunzio, L. I. Glazman, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf, *Phys. Rev. Lett.* **107**, 240501 (2011).
- [10] D. Ristè, C. C. Bultink, M. J. Tiggelman, R. N. Schouten, K. W. Lehnert, and L. DiCarlo, *Nat. Commun.* **4**, 1913 (2013).
- [11] G. Catelani, R. J. Schoelkopf, M. H. Devoret, and L. I. Glazman, *Phys. Rev. B* **84**, 064517 (2011).
- [12] J. N. Ullom, P. A. Fisher, and M. Nahum, *Appl. Phys. Lett.* **73**, 2494 (1998).
- [13] I. Nsanzeza and B. L. T. Plourde, *Phys. Rev. Lett.* **113**, 117002 (2014).
- [14] C. Wang, Y. Y. Gao, I. M. Pop, U. Vool, C. Axline, T. Brecht, R. W. Heeres, L. Frunzio, M. H. Devoret, G. Catelani, L. I. Glazman, and R. J. Schoelkopf, *Nat. Commun.* **5**, 5836 (2014).
- [15] M. Taupin, I. M. Khaymovich, M. Meschke, A. S. Mel'nikov, and J. P. Pekola, *Nat. Commun.* **7**, 10977 (2016).
- [16] N. A. Court, A. J. Ferguson, R. Lutchyn, and R. G. Clark, *Phys. Rev. B* **77**, 100501(R) (2008).
- [17] S. Rajauria, H. Courtois, and B. Pannetier, *Phys. Rev. B* **80**, 214521 (2009).
- [18] S. Rajauria, L. M. A. Pascal, P. Gandit, F. W. J. Hekking, B. Pannetier, and H. Courtois, *Phys. Rev. B* **85**, 020505(R) (2012).
- [19] R.-P. Riwar, A. Hosseinkhani, L. D. Burkhardt, Y. Y. Gao, R. J. Schoelkopf, L. I. Glazman, and G. Catelani, *Phys. Rev. B* **94**, 104516 (2016).
- [20] U. Patel, I. V. Pechenezhskiy, B. L. T. Plourde, M. G. Vavilov, and R. McDermott, *Phys. Rev. B* **96**, 220501(R) (2017).
- [21] A. Hosseinkhani, R.-P. Riwar, R. J. Schoelkopf, L. I. Glazman, and G. Catelani, *Phys. Rev. Appl.* **8**, 064028 (2017).
- [22] J. Aumentado, M. W. Keller, J. M. Martinis, and M. H. Devoret, *Phys. Rev. Lett.* **92**, 066802 (2004).
- [23] L. Sun, L. DiCarlo, M. D. Reed, G. Catelani, L. S. Bishop, D. I. Schuster, B. R. Johnson, G. A. Yang, L. Frunzio, L. Glazman, M. H. Devoret, and R. J. Schoelkopf, *Phys. Rev. Lett.* **108**, 230509 (2012).
- [24] S. Gustavsson, F. Yan, G. Catelani, J. Bylander, A. Kamal, J. Birenbaum, D. Hover, D. Rosenberg, G. Samach, A. P. Sears, S. J. Weber, J. L. Yoder, J. Clarke, A. J. Kerman, F. Yoshihara, Y. Nakamura, T. P. Orlando, and W. D. Oliver, *Science* **354**, 1573 (2016).
- [25] A. J. Ferguson, N. A. Court, F. E. Hudson, and R. G. Clark, *Phys. Rev. Lett.* **97**, 106603 (2006).
- [26] O. Naaman and J. Aumentado, *Phys. Rev. B* **73**, 172504 (2006).
- [27] R.-P. Riwar, L. I. Glazman, and G. Catelani, *Phys. Rev. B* **98**, 024502 (2018).
- [28] A. Hosseinkhani and G. Catelani, *Phys. Rev. B* **97**, 054513 (2018).
- [29] W. Belzig, F. K. Wilhelm, C. Bruder, G. Schön, and A. D. Zaikin, *Superlattices Microstruct.* **25**, 1251 (1999).
- [30] In fact, the same principle could be implemented with a good contact to a sufficiently thin normal metal; see, e.g., Refs. [28,29] and references therein.
- [31] Since the coherence length is much larger than the Fermi wavelength (in low critical-temperature superconductors), we can neglect quasiparticle scattering events across this interface.
- [32] K. Steinberg, M. Scheffler, and M. Dressel, *Phys. Rev. B* **77**, 214517 (2008).
- [33] G. Catelani, X. S. Wu, and P. W. Adams, *Phys. Rev. B* **78**, 104515 (2008).
- [34] S. B. Kaplan, C. C. Chi, D. N. Langenberg, J. J. Chang, S. Jafarey, and D. J. Scalapino, *Phys. Rev. B* **14**, 4854 (1976).
- [35] By assumption, the rate τ_N^{-1} depends on the electron-phonon coupling strength; it is related to the rate τ_0^{-1} of Ref. [34] by $\tau_N^{-1} = \tau_0^{-1} (\Delta/k_B T_c)^3 / 3 \approx 1.83 \tau_0^{-1}$.
- [36] P. M. Echternach, H. G. LeDuc, A. Skalare, and W. R. McGrath, in *Proceedings of the Tenth International Symposium on Space Terahertz Technology*, edited by T. Crowe and R. Weikle (University of Virginia, Charlottesville, Virginia, USA, 1999), pp. 261–268.
- [37] S. Zhao, D. J. Goldie, C. N. Thomas, and S. Withington, *Supercond. Sci. Technol.* **31**, 105004 (2018).
- [38] S. B. Kaplan, *J. Low Temp. Phys.* **37**, 343 (1979).
- [39] F. Giazotto, T. T. Heikkilä, A. Luukanen, A. M. Savin, and J. P. Pekola, *Rev. Mod. Phys.* **78**, 217 (2006).
- [40] As an example, using $D_{qp} \simeq 18 + \text{cm}^2/\text{s}$ [14], $\tau_N \simeq 8 \times 10^{-7}$ s for aluminum (which we take as the superconductor S) as discussed at the end of Sec. II A, and $L \simeq 2$ mm as the (effective) device length [21], we find $d_S \simeq 0.7 \mu\text{m}$, about a factor of 8 smaller than the estimate for $d_N \simeq 6 \mu\text{m}$ given at the end of Sec. III of Ref. [21] (where the notation l_0 is used).
- [41] D. C. Mattis and J. Bardeen, *Phys. Rev.* **111**, 412 (1958).
- [42] G. Catelani, L. I. Glazman, and K. E. Nagaev, *Phys. Rev. B* **82**, 134502 (2010).
- [43] M. Tinkham, *Introduction to Superconductivity* (Courier Corporation, Indianapolis, 2004).
- [44] These admittances do not include the parts related to Josephson energy and capacitance, already accounted for by M_0 , but rather only the real parts of the quasiparticle corrections to those main terms; see Refs. [11] and [46].
- [45] We remind that we assume that the phase difference across the S - S' junction is sufficiently small that it can be treated within linear response; then the current-phase relation for short junctions is the same for any transparency [see, e.g., Ref. [59]; cf. also the discussion after Eq. (36)].
- [46] F. Kos, S. E. Nigg, and L. I. Glazman, *Phys. Rev. B* **87**, 174521 (2013).
- [47] K. Serniak, M. Hays, G. de Lange, S. Diamond, S. Shankar, L. D. Burkhardt, L. Frunzio, M. Houzet, and M. H. Devoret, *Phys. Rev. Lett.* **121**, 157701 (2018).
- [48] M. Houzet, K. Serniak, G. Catelani, M. H. Devoret, and L. I. Glazman, *Phys. Rev. Lett.* **123**, 107704 (2019).
- [49] J. Aarts, J. M. E. Geers, E. Brück, A. A. Golubov, and R. Coehoorn, *Phys. Rev. B* **56**, 2779 (1997).
- [50] K. Serniak (private communication, 2019).

- [51] The relevant trap size for effective suppression of the steady-state quasiparticle density is given not by d_S of Eq. (11), which pertains to the efficient evacuation of excess quasiparticles, but by a_S of Eq. (14), which plays the same role of λ_{tr} of Ref. [21]; we refer the reader to Sec. IV of that reference for further details. Here we note that using parameters for Al, we estimate a_S in the few tens of microns, which justifies the distinctions between small and large traps made in the text above Eq. (41).
- [52] F. Yan, S. Gustavsson, A. Kamal, J. Birenbaum, A. P. Sears, D. Hover, T. J. Gudmundsen, D. Rosenberg, G. Samach, S. Weber, J. L. Yoder, T. P. Orlando, J. Clarke, A. J. Kerman, and W. D. Oliver, *Nat. Commun.* **7**, 12964 (2016).
- [53] R. Barends, J. Kelly, A. Megrant, D. Sank, E. Jeffrey, Y. Chen, Y. Yin, B. Chiaro, J. Mutus, C. Neill, P. O'Malley, P. Roushan, J. Wenner, T. C. White, A. N. Cleland, and J. M. Martinis, *Phys. Rev. Lett.* **111**, 080502 (2013).
- [54] I. M. Pop, K. Geerlings, G. Catelani, R. J. Schoelkopf, L. I. Glazman, and M. H. Devoret, *Nature (London)* **508**, 369 (2014).
- [55] M. Hays, G. de Lange, K. Serniak, D. J. van Woerkom, D. Bouman, P. Krogstrup, J. Nygård, A. Geresdi, and M. H. Devoret, *Phys. Rev. Lett.* **121**, 047001 (2018).
- [56] F. Henriques, F. Valenti, T. Charpentier, M. Lagoin, C. Gouriou, M. Martínez, L. Cardani, M. Vignati, L. Grünhaupt, D. Gusenkova, J. Ferrero, S. T. Skacel, W. Wernsdorfer, A. V. Ustinov, G. Catelani, O. Sander, and I. M. Pop, [arXiv:1908.04257](https://arxiv.org/abs/1908.04257).
- [57] G. C. Ménard, F. K. Malinowski, D. Puglia, D. I. Pikulin, T. Karzig, B. Bauer, P. Krogstrup, and C. M. Marcus, [arXiv:1902.02689](https://arxiv.org/abs/1902.02689).
- [58] V. Cherkez, J. C. Cuevas, C. Brun, T. Cren, G. Mênard, F. Debontridder, V. S. Stolyarov, and D. Roditchev, *Phys. Rev. X* **4**, 011033 (2014).
- [59] C. W. J. Beenakker, Transport phenomena in mesoscopic systems, in *Proceedings of the 14th Taniguchi Symposium*, edited by H. Fukuyama and T. Ando (Springer, Berlin, 1992), pp. 235–253.