

# Decoherence of superconducting qubits caused by quasiparticle tunneling

G. Catelani, S. E. Nigg, S. M. Girvin, R. J. Schoelkopf, and L. I. Glazman  
*Departments of Physics and Applied Physics, Yale University, New Haven, CT 06520, USA*  
(Dated: August 1, 2012)

In superconducting qubits, the interaction of the qubit degree of freedom with quasiparticles defines a fundamental limitation for the qubit coherence. We develop a theory of the pure dephasing rate  $\Gamma_\phi$  caused by quasiparticles tunneling through a Josephson junction and of the inhomogeneous broadening due to changes in the occupations of Andreev states in the junction. To estimate  $\Gamma_\phi$ , we derive a master equation for the qubit dynamics. The tunneling rate of free quasiparticles is enhanced by their large density of states at energies close to the superconducting gap. Nevertheless, we find that  $\Gamma_\phi$  is small compared to the rates determined by extrinsic factors in most of the current qubit designs (phase and flux qubits, transmon, fluxonium). The split transmon, in which a single junction is replaced by a SQUID loop, represents an exception that could make possible the measurement of  $\Gamma_\phi$ . Fluctuations of the qubit frequency leading to inhomogeneous broadening may be caused by the fluctuations in the occupation numbers of the Andreev states associated with a phase-biased Josephson junction. This mechanism may be revealed in qubits with small-area junctions, since the smallest relative change in frequency it causes is of the order of the inverse number of transmission channels in the junction.

PACS numbers: 74.50.+r, 85.25.Cp

## I. INTRODUCTION

Over the past several years significant efforts have been directed toward designing and implementing superconducting circuits with improved coherence properties. For quantum computation purposes, the coherence time  $T_2$  of a qubit must be sufficiently long as to allow for error correction.<sup>1</sup> The unavoidable couplings of the qubit with various sources of noise are responsible for decoherence, and different types of qubits have different sensitivities to a given noise source. For example, the phase and flux qubits coherence times are limited by flux noise,<sup>2,3</sup> while the transmon parameters are chosen to decrease the effect of charge noise in comparison with the Cooper pair box.<sup>4</sup> Flux and charge noise originate from the environment surrounding the qubits; in this paper, by contrast, we study an *intrinsic* mechanism of decoherence due to the coupling between the qubit and the quasiparticle excitations in the superconductor the qubit is made of. In general one can distinguish two contributions to the time  $T_2$ : first, the qubit can lose energy and the corresponding relaxation time  $T_1$  imposes an upper bound to the coherence time,  $T_2 \leq 2T_1$ . Second, additional pure dephasing mechanisms, characterized by the rate  $\Gamma_\phi$ , can shorten  $T_2$  below this upper limit. Recent theoretical<sup>5,6</sup> and experimental<sup>7-9</sup> works have highlighted the contribution of quasiparticle tunneling to the relaxation rate. Here we focus on the pure dephasing effect of quasiparticle tunneling.

The decoherence rates discussed above are related to the power spectral density  $S(\omega)$  of the noise source: the relaxation rate is proportional to the value of the spectral density at the qubit frequency  $\omega_{10}$ ,  $1/T_1 \propto S(\omega_{10})$ , while the pure dephasing rate is determined by the low-frequency part of the spectral density,  $\Gamma_\phi \sim S(0)$  – see, e.g., Ref. 10. Clearly the latter relationship cannot hold if

the power spectral density diverges as  $\omega \rightarrow 0$ . Because of its experimental relevance, a well-studied example of diverging spectral density is that of  $1/f$  noise; in the case of  $1/f$  flux noise, for instance, the decay of the qubit coherence is not exponential in time, but Gaussian-like<sup>10,11</sup> (up to a logarithmic factor that depends on the measurement protocol). In studying how quasiparticle tunneling affects dephasing we find another such example, since the quasiparticle current spectral density is logarithmically divergent at low frequencies when the gaps on the two sides of the junction have the same magnitudes (see Sec. III). We show that despite this divergence, a finite dephasing rate  $\Gamma_\phi$  can be determined. We then estimate the dephasing rate for a few different single- and multi-junction qubits and find that in most cases  $\Gamma_\phi$  is small compared to the the quasiparticle induced relaxation rate. An exception is the split transmon, in which the two rates can be of the same order of magnitude (see Sec. V A). Since it is known that quasiparticles limit the relaxation rate in this system at sufficiently high temperatures,<sup>9</sup> it may be possible to measure the quasiparticle dephasing rate if other sources of dephasing can be minimized.

The quasiparticle dephasing mechanism discussed above is due to tunneling of *free* quasiparticles across the junction. Another dephasing mechanism originates from quasiparticles weakly *bound* to a phase-biased junction that give rise to subgap Andreev states; the dephasing is caused by changes in the occupations of these states that make the Josephson coupling and hence qubit frequency  $\omega_q$  fluctuate. Because of this additional dephasing, the measured decoherence rate  $1/T_2^*$  acquires an inhomogeneous broadening contribution,  $1/T_2^* - 1/T_2$ , which can be suppressed using echo pulse sequences. When the average occupation  $x_{\text{qp}}^A$  of the Andreev states is small,  $x_{\text{qp}}^A \ll 1$ , the typical (i.e., root mean square) fluctuation of the occupations is given by the square root of  $x_{\text{qp}}^A$ .

Then for the phase qubit we show in Sec. VI that the typical frequency fluctuation is proportional to the typical fluctuation of the occupations divided by the square root of the (effective) number of transmission channels  $N_e$  in the junction,  $\langle(\Delta\omega_q)^2\rangle^{1/2}/\omega_q \propto \sqrt{x_{\text{qp}}^A/N_e}$ . For these fluctuations to measurably affect the decoherence rate  $1/T_2^*$ , the condition  $\langle\Delta\omega_q^2\rangle^{1/2}T_2 \gtrsim 1$  should be satisfied; using this condition we estimate that this mechanism is not a limiting factor to coherence in current experiments with phase qubits. On the contrary, it could contribute to decoherence in recent transmon experiments,<sup>7,12</sup> due to the small junction area (i.e., smaller  $N_e$  in comparison with phase qubits). However, this possibility will require a separate investigation, due to the lack of phase bias in the transmon.

The paper is organized as follows: in the next Section we introduce the effective description of a single-junction system. In Sec. III we present the master equation governing the qubit dynamics and we discuss the self-consistent regularization of the logarithmic divergence in the dephasing rate. Applications of our results to single- and multi-junctions qubits are in Secs. IV and V, respectively. The role of Andreev states is analyzed in Sec. VI. We summarize our work in Sec. VII. We use units  $\hbar = k_B = 1$  throughout the paper.

## II. EFFECTIVE MODEL

The effective Hamiltonian  $\hat{H}$  for a superconducting qubit can be split into two parts,

$$\hat{H} = \hat{H}_0 + \delta\hat{H}, \quad (1)$$

where the non-interacting Hamiltonian  $\hat{H}_0$  is the sum of qubit and quasiparticle terms,

$$\hat{H}_0 = \hat{H}_\varphi + \hat{H}_{\text{qp}}. \quad (2)$$

The Hamiltonian for the qubit degree of freedom accounts for the charging ( $E_C$ ), Josephson ( $E_J$ ), and inductive ( $E_L$ ) energies in a system comprising an inductive loop shunting a tunnel junction,

$$\hat{H}_\varphi = 4E_C \left( \hat{N} - n_g \right)^2 - E_J \cos \hat{\varphi} + \frac{1}{2} E_L (\hat{\varphi} - 2\pi\Phi_e/\Phi_0)^2, \quad (3)$$

with  $n_g$  the dimensionless gate voltage,  $\Phi_e$  the external magnetic flux threading the loop, and  $\Phi_0 = h/2e$  the flux quantum. The operator  $\hat{N} = -id/d\varphi$  counts the number of Cooper pairs passed through the junction. The quasiparticle Hamiltonian is given by

$$\hat{H}_{\text{qp}} = \sum_{j=L,R} \hat{H}_{\text{qp}}^j, \quad \hat{H}_{\text{qp}}^j = \sum_{l=1}^{N_{ch}} \sum_{n,\sigma} \epsilon_n^j \hat{\alpha}_{n\sigma l}^{j\dagger} \hat{\alpha}_{n\sigma l}^j, \quad (4)$$

where  $\hat{\alpha}_{n\sigma l}^j$  ( $\hat{\alpha}_{n\sigma l}^{j\dagger}$ ) are annihilation (creation) operators for quasiparticles with channel index  $l$  and spin  $\sigma = \uparrow, \downarrow$

in lead  $j = L, R$  to the left or right of the junction. We have assumed for simplicity the same number of channels  $N_{ch}$  and identical densities of states per spin direction  $\nu_0$  in both leads. Denoting with  $\Delta^j$  the superconducting gap, the quasiparticle energies are  $\epsilon_n^j = \sqrt{(\xi_n^j)^2 + (\Delta^j)^2}$ , with  $\xi_n^j$  single-particle energy level  $n$  in the normal state of lead  $j$ . The occupation probabilities of these levels are given by the distribution functions

$$f^j(\xi_n^j) = \langle\langle \hat{\alpha}_{n\uparrow l}^{j\dagger} \hat{\alpha}_{n\uparrow l}^j \rangle\rangle_{\text{qp}} = \langle\langle \hat{\alpha}_{n\downarrow l}^{j\dagger} \hat{\alpha}_{n\downarrow l}^j \rangle\rangle_{\text{qp}}, \quad j = L, R, \quad (5)$$

where double angular brackets  $\langle\langle \dots \rangle\rangle_{\text{qp}}$  denote averaging over quasiparticle states. We take the distribution functions to be independent of spin and equal in the two leads. We also assume that  $\delta E$ , the characteristic energy of the quasiparticles above the gap, is small compared to the gap, but the distribution function is otherwise generic, thus allowing for non-equilibrium conditions.

The interaction term  $\delta\hat{H}$  in Eq. (1) accounts for tunneling and, as discussed in Appendix A of Ref. 6, is the sum of three parts: quasiparticle tunneling  $\hat{H}_T$ , pair tunneling  $\hat{H}_T^p$ , and the Josephson energy counterterm  $\hat{H}_{E_J}$ . When the superconducting gaps are larger than all other energy scales, the only effect of the last two terms is to contribute to the renormalization of the qubit frequency<sup>6</sup> [see also the discussion after Eq. (13)]; therefore, we neglect those terms and consider only the quasiparticle tunneling Hamiltonian,  $\delta\hat{H} = \hat{H}_T$  with

$$\hat{H}_T = \sum_{l,k=1}^{N_{ch}} \tilde{t}_{lk} \sum_{n,m,\sigma} \left( e^{i\hat{\varphi}/2} u_n^L u_m^R - e^{-i\hat{\varphi}/2} v_m^R v_n^L \right) \hat{\alpha}_{n\sigma l}^{L\dagger} \hat{\alpha}_{m\sigma k}^R + \text{H.c.} \quad (6)$$

Here the Bogoliubov amplitudes  $u_n^j, v_n^j$  are real quantities, since their dependence on the phases of the order parameters appears explicitly through the gauge-invariant phase difference  $\varphi$ . The elements  $\tilde{t}_{lk} \ll 1$  of the electron tunneling matrix  $\tilde{t}$  are related to the junction conductance by  $g_T = 2g_K \sum_{p=1}^{N_{ch}} T_p$ , where  $g_K = e^2/h$  is the conductance quantum and the transmission probabilities  $T_p$  ( $p = 1, \dots, N_{ch}$ ) are the eigenvalues of the matrix  $(2\pi\nu_0)^2 \tilde{t}\tilde{t}^\dagger$ .

Since we are interested in the dynamics of the qubit only, rather than that of a multi-level system, we project the Hamiltonian  $\hat{H}$  onto the qubit states  $|0\rangle$  and  $|1\rangle$ , which we represent by the vectors  $(0, 1)^T$  and  $(1, 0)^T$  for the ground and excited states, respectively; the two-level approximation is justified under the conditions that permit the operability of the system as a qubit<sup>13</sup> (i.e., anharmonicity large compared to linewidth). Then in terms of the Pauli matrices we can write

$$\hat{H}_\varphi = \frac{\omega_{10}}{2} \hat{\sigma}^z, \quad (7)$$

where the qubit frequency in general depends on all the parameters present in Eq. (3), and, dropping for nota-

tional simplicity the channel indices,<sup>14</sup>

$$\hat{H}_T = \tilde{t} \sum_{n,m,\sigma} \left[ A_{nm}^d \hat{\sigma}^z + A_{nm}^r (\hat{\sigma}^+ + \hat{\sigma}^-) + A_{nm}^f \hat{I} \right] \hat{\alpha}_{n\sigma}^{L\dagger} \hat{\alpha}_{m\sigma}^R + \text{H.c.}, \quad (8)$$

where the coefficients  $A_{nm}^k$ ,  $k = d, r, f$ , have the structure

$$A_{nm}^k = A_c^k (u_n^L u_m^R - v_n^L v_m^R) + i A_s^k (u_n^L u_m^R + v_n^L v_m^R). \quad (9)$$

Here  $A_{c,s}^k$  denote combinations of matrix elements for the operators  $e^{\pm i\hat{\varphi}/2}$  associated with the transfer of a single charge across the junction,

$$s_{ij} = \langle i | \sin \frac{\hat{\varphi}}{2} | j \rangle \quad (10)$$

$$A_s^d = \frac{1}{2} (s_{11} - s_{00}) \quad (11)$$

$$A_s^r = s_{10} \quad (12)$$

$$A_s^f = \frac{1}{2} (s_{11} + s_{00}) \quad (13)$$

and the  $A_c^k$  are obtained by replacing sine with cosine in the above definitions. As it will become evident in the next section, only the terms with  $k = d$  and  $k = r$  contribute to pure dephasing and relaxation of the qubit, respectively.

The term with  $k = f$  (in combination with the  $k = r$  one) contributes to the average frequency shift. More precisely, the average frequency shift  $\delta\omega = \delta\omega_{E_J} + \delta\omega_{\text{qp}}$  has two parts,<sup>6</sup> originating from the quasiparticle renormalization of the Josephson energy and virtual transitions between qubit states mediated by quasiparticles, respectively. The latter part ( $\delta\omega_{\text{qp}}$ ) is discussed further in Appendix A. Here we note that in the leading ( $\propto \tilde{t}^2$ ) order, the Josephson part  $\delta\omega_{E_J}$  is the sum of two contributions with distinct origins. The first one comes from the product of the terms proportional to  $A_{nm}^f$  and  $A_{nm}^r$  in  $\delta\hat{H}_T$  [Eq. (8)]. The second contribution is due to the terms we neglected in  $\delta\hat{H}$ . (The neglected terms are the pair tunneling and Josephson counterterm, as defined in Appendix A of Ref. 6.) Since we are studying decoherence effects in this work, we set  $A_{nm}^f = 0$  henceforth. Equations (4), (7), and (8) (with  $A_{nm}^f = 0$ ) constitute the starting point for the derivation of the master equation presented in the next section.

### III. QUBIT PHASE RELAXATION: THE MASTER EQUATION

The information on the time evolution of the qubit is contained in its density matrix  $\hat{\rho}(t)$ , which we decompose as

$$\hat{\rho} = \frac{1}{2} \left[ \hat{I} + \rho_z \hat{\sigma}^z \right] + \rho_+ \hat{\sigma}^- + \rho_+^* \hat{\sigma}^+ \quad (14)$$

In this section we present the final form of the master equation for the density matrix. The derivation can be found in Appendix A, where we start from the Hamiltonian of the system presented in the previous section and employ the standard Born-Markov and secular (rotating wave) approximations<sup>15</sup> to arrive at the expressions given here.

The diagonal component  $\rho_z$  of the density matrix obeys the equation

$$\frac{d\rho_z}{dt} = -[\Gamma_{1\rightarrow 0} + \Gamma_{0\rightarrow 1}] \rho_z + [\Gamma_{0\rightarrow 1} - \Gamma_{1\rightarrow 0}] \quad (15)$$

where, assuming equal gaps in the leads ( $\Delta^L = \Delta^R \equiv \Delta$ ),

$$\Gamma_{1\rightarrow 0} = \frac{2g_T}{\pi g_K} \int_{\Delta}^{+\infty} d\epsilon f(\epsilon) (1 - f(\epsilon + \omega_{10})) \left[ \frac{\epsilon(\epsilon + \omega_{10}) + \Delta^2}{\sqrt{\epsilon^2 - \Delta^2} \sqrt{(\epsilon + \omega_{10})^2 - \Delta^2}} |A_s^r|^2 + \frac{\epsilon(\epsilon + \omega_{10}) - \Delta^2}{\sqrt{\epsilon^2 - \Delta^2} \sqrt{(\epsilon + \omega_{10})^2 - \Delta^2}} |A_c^r|^2 \right] \quad (16)$$

and  $\Gamma_{0\rightarrow 1}$  is obtained by the replacement  $f \rightarrow 1 - f$ . Here  $g_K = e^2/h$  is the conductance quantum. The general solution to Eq. (15) is

$$\rho_z(t) = \rho_z(0) e^{-t/T_1} + \frac{\Gamma_{0\rightarrow 1} - \Gamma_{1\rightarrow 0}}{\Gamma_{0\rightarrow 1} + \Gamma_{1\rightarrow 0}} \quad (17)$$

where we introduced the relaxation time  $T_1$  as

$$\frac{1}{T_1} = \Gamma_{0\rightarrow 1} + \Gamma_{1\rightarrow 0}. \quad (18)$$

Equation (16) represents the generalization, valid for any  $\omega_{10} < 2\Delta$ , of the relaxation rate formula derived in Refs. 5,6 in the limit  $\omega_{10} \ll 2\Delta$  using Fermi's golden rule. Indeed, the assumption that quasiparticles have characteristic energies small compared to the gap enables us to approximately substitute  $\epsilon \rightarrow \Delta$  in the numerators in square brackets in Eq. (16), and neglecting terms of order  $\omega_{10}/\Delta$  we find

$$\Gamma_{1\rightarrow 0} \simeq |A_s^r|^2 S_{\text{qp}}(\omega_{10}), \quad (19)$$

where

$$S_{\text{qp}}(\omega) = \frac{16E_J}{\pi} \int_0^{+\infty} dx \frac{1}{\sqrt{x} \sqrt{x + \omega/\Delta}} f[(1+x)\Delta] \{1 - f[(1+x)\Delta + \omega]\} \quad (20)$$

and we remind that  $E_J = \Delta g_T / 8g_K$ . The agreement of Eq. (19) with the results of Refs. 5,6 validates the present approach. Since the relaxation rate is studied in detail in those references, we do not consider it here any further, except to note that the terms neglected in Eq. (19) can become important if the matrix element  $A_s^r$  is small,  $|A_s^r/A_c^r|^2 \lesssim \omega_{10}/\Delta$ . In fact,  $A_s^r$  can vanish at particular values of the external parameters used to tune

the qubit, for example in the flux qubit when the external flux equals half the flux quantum;<sup>5,6</sup> in such a case, one needs to retain the term proportional to  $A_c^r$  in Eq. (16) to evaluate the (non-vanishing) relaxation rate.

The master equation for the off-diagonal part of the density matrix is

$$\frac{d\rho_+}{dt} = i(\omega_{10} + \delta\omega)\rho_+ - \frac{1}{2T_1}\rho_+ - \Gamma_\phi\rho_+ \quad (21)$$

where  $\delta\omega$  is the quasiparticle-induced average frequency shift<sup>5,6</sup> discussed in the previous section,  $T_1$  is defined in Eq. (18), and the pure dephasing rate is

$$\Gamma_\phi = \frac{4g_T}{\pi g_K} \int_{\Delta^R}^{+\infty} d\epsilon f(\epsilon) [1 - f(\epsilon)] \left[ \frac{\epsilon^2 + \Delta^L \Delta^R}{\sqrt{\epsilon^2 - (\Delta^L)^2} \sqrt{\epsilon^2 - (\Delta^R)^2}} |A_s^d|^2 + \frac{\epsilon^2 - \Delta^L \Delta^R}{\sqrt{\epsilon^2 - (\Delta^L)^2} \sqrt{\epsilon^2 - (\Delta^R)^2}} |A_c^d|^2 \right] \quad (22)$$

where we assumed  $\Delta^R > \Delta^L$ . The general solution to Eq. (21) is

$$\rho_+(t) = \rho_+(0) e^{i(\omega_{10} + \delta\omega)t} e^{-t/T_2} \quad (23)$$

with

$$\frac{1}{T_2} = \frac{1}{2T_1} + \Gamma_\phi. \quad (24)$$

The pure dephasing rate defined in Eq. (22) has a structure similar to that of the relaxation rate, Eq. (16), if we substitute  $\omega_{10} \rightarrow 0$  and  $A_{s(c)}^r \rightarrow A_{s(c)}^d$  in the latter. Thus we recover the relationship between the power spectral density  $S(\omega)$  of a noise source and the decoherence rates discussed in the Introduction,  $\Gamma_{1 \rightarrow 0} \propto S(\omega_{10})$  and  $\Gamma_\phi \propto S(0)$ . However, in Eq. (22) we have explicitly assumed an asymmetric junction,  $\Delta^R > \Delta^L$ , and extension of this result to the typical case of a symmetric junction ( $\Delta^R = \Delta^L$ ) is problematic. Indeed, let us consider an almost symmetric junction,  $\Delta^R - \Delta^L \ll \Delta^R$ , with  $|A_s^d| \gtrsim |A_c^d|$  and a non-degenerate quasiparticle distribution [ $f(\epsilon) \ll 1$ ,  $\epsilon > \Delta^R$ ]; then we find, using from now on the notation  $\Delta = \Delta^R$ ,

$$\Gamma_\phi \simeq \frac{4g_T}{\pi g_K} |A_s^d|^2 \Delta \int_0^{+\infty} dx \frac{f[(1+x)\Delta]}{\sqrt{x}\sqrt{x + (\Delta - \Delta^L)/\Delta}} \simeq 2 |A_s^d|^2 S_{\text{qp}}(\Delta - \Delta^L) \quad (25)$$

In the symmetric junction limit  $\Delta^L \rightarrow \Delta$ ,  $\Gamma_\phi$  diverges logarithmically due to the singularity at  $x = 0$  of the integrand in Eq. (25); for example, in thermal equilibrium at temperature  $T \gg \Delta - \Delta^L$  we have

$$\Gamma_\phi \approx \frac{32E_J}{\pi} |A_s^d|^2 e^{-\Delta/T} \left[ \ln \frac{4T}{\Delta - \Delta^L} - \gamma_E \right] \quad (26)$$

Due to the logarithmic divergence, in general we cannot simply take  $\Gamma_\phi \propto S_{\text{qp}}(0)$ ; the correct procedure that leads to a finite dephasing rate is presented in the next section.

## A. Self-consistent dephasing rate

The terms in the right hand sides of the master equations (15) and (21) are proportional to the square of the tunneling amplitude via the tunneling conductance  $g_T \propto \tilde{t}^2$ ; this proportionality is a consequence of the lowest order perturbative treatment of the tunneling Hamiltonian [Eq. (8)], which enables us to neglect higher order (in  $\tilde{t}$ ) terms when evaluating certain correlation functions involving qubit and quasiparticle operators [see Appendix A for details]. This implies that those correlation functions oscillate but do not decay in time, which is a limitation of the used approximation: the inclusion of higher order effects introduces decaying factors of the form  $e^{-\gamma t}$  into the correlation functions, where at leading order the decay rate  $\gamma$  is itself proportional to the tunneling conductance. Here we discuss an Ansatz for  $\gamma$  whose validity is checked perturbatively in Appendix B. As we show there, a finite decay rate  $\gamma$  reflects itself into a smearing of the singularity for  $\Delta^L = \Delta$  of the integrand in Eq. (25),

$$\int_0^{+\infty} \frac{dx}{x} = \int_0^{+\infty} \frac{dx}{\sqrt{x}} \int_0^{+\infty} \frac{dy}{\sqrt{y}} \delta(x-y) \rightarrow \int_0^{+\infty} \frac{dx}{\sqrt{x}} \int_0^{+\infty} \frac{dy}{\sqrt{y}} \frac{1}{\pi} \frac{\gamma/\Delta}{(x-y)^2 + (\gamma/\Delta)^2} \quad (27)$$

In the problem at hand there are two inverse time scales which could serve as a low-energy cut-off to regularize the integral as in the above equation, the relaxation rate  $\Gamma_{1 \rightarrow 0}$  and the pure dephasing rate  $\Gamma_\phi$ . A finite relaxation rate means that the qubit excited level has a finite width; one could argue that this uncertainty in the energy will in turn reflect itself in an uncertainty of the energy exchanged between qubit and quasiparticles, thus smearing the singularity as in Eq. (27). However, relaxation rate and dephasing rate are determined by different matrix elements [cf. Eqs. (11)-(12)], so one can imagine, at least in principle, a limiting situation in which the relaxation rate vanishes, which would then cause the dephasing rate to diverge. Therefore, we expect that dephasing processes will themselves be the ultimate limiting factors for coherence, so that  $\gamma = \Gamma_\phi$ . With this identification, we arrive at the self-consistent expression for the pure dephasing rate

$$\Gamma_\phi = \frac{32E_J}{\pi} |A_s^d|^2 \int_0^{+\infty} \frac{dx}{\sqrt{x}} \int_0^{+\infty} \frac{dy}{\sqrt{y}} f[(1+x)\Delta] \times \{1 - f[(1+y)\Delta]\} \frac{1}{\pi} \frac{\Gamma_\phi/\Delta}{(x-y)^2 + (\Gamma_\phi/\Delta)^2} \quad (28)$$

Equation (28) is the central result of this paper. It is valid for symmetric junctions (or nearly symmetric,  $\Delta^R - \Delta^L \ll \Gamma_\phi$ ) and we show in Appendix B that it agrees with the result of the perturbative derivation of the master equation extended with logarithmic accuracy to the next to leading order in  $\tilde{t}^2$ .

Similarly to the relaxation rate, for some specific values of the qubit parameters the matrix element  $A_s^d$  can be small or even vanish exactly. Then one should take into account the second term in square brackets in Eq. (22) to get

$$\Gamma_\phi = \frac{32E_J}{\pi} |A_c^d|^2 \int_0^{+\infty} dx f[(1+x)\Delta] \{1 - f[(1+x)\Delta]\} \quad (29)$$

An estimate for the actual dephasing rate is given by the larger of the two rates calculated using Eq. (28) or Eq. (29).

### B. Non-equilibrium quasiparticles

The relaxation rate in Eq. (19) depends explicitly on the qubit properties via the matrix element  $A_s^d$ , while the spectral density  $S_{\text{qp}}$  accounts for the dynamics of quasiparticle tunneling. The same structure is present in the right hand sides of Eqs. (28)-(29) – a matrix element multiplies factors describing the tunneling dynamics. These factors can be further simplified under certain assumptions. Here we focus on Eq. (28) and distinguish two cases: first, let us assume that the quasiparticle energy is small compared to the dephasing rate,  $\delta E \ll \Gamma_\phi$ , and that quasiparticles are non-degenerate,  $f[(1+y)\Delta] \ll 1$ . Then integrating first over  $y$  and then over  $x$  we find

$$\Gamma_\phi \simeq \frac{16E_J}{\pi} |A_s^d|^2 \sqrt{\frac{\Delta}{\Gamma_\phi}} x_{\text{qp}}, \quad (30)$$

where

$$x_{\text{qp}} = \sqrt{2} \int_0^{+\infty} \frac{dx}{\sqrt{x}} f[(1+x)\Delta] \quad (31)$$

is the quasiparticle density normalized by the density of Cooper pairs. Indicating with  $f_0$  the typical occupation probability, we estimate<sup>16</sup>  $x_{\text{qp}} \sim f_0 \sqrt{\delta E/\Delta}$ . Then solving Eq. (30) for  $\Gamma_\phi$ , the requirement  $\Gamma_\phi \gg \delta E$  can be written as

$$\frac{16}{\pi} \frac{E_J}{\Delta} |A_s^d|^2 f_0 \gg \frac{\delta E}{\Delta} \quad (32)$$

This condition is in practice difficult to satisfy, since with our assumptions  $f_0 \ll 1$ , while  $|A_s^d| \leq 1$ ,  $E_J/\Delta \lesssim 1$ , and at the lowest experimental temperatures  $\delta E/\Delta \sim T/\Delta \gtrsim 0.01$ . Thus we conclude that for non-degenerate quasiparticles an upper bound for the dephasing rate is given by  $\Gamma_\phi \lesssim \delta E$ .

The second case we consider, for both degenerate and non-degenerate quasiparticles, is in fact that of small dephasing rate,  $\Gamma_\phi \ll \delta E$ . Then neglecting terms of order

$\Gamma_\phi/\delta E$ , Eq. (28) simplifies to

$$\begin{aligned} \Gamma_\phi &\simeq \frac{32E_J}{\pi} |A_s^d|^2 \int_0^{+\infty} \frac{dx}{\sqrt{x}} \text{Re} \frac{1}{\sqrt{x + i\Gamma_\phi/\Delta}} \\ &\quad \times f[(1+x)\Delta] \{1 - f[(1+x)\Delta]\} \\ &\sim \frac{32E_J}{\pi} |A_s^d|^2 f_0 (1 - f_0) \ln \frac{4\delta E}{\Gamma_\phi} \end{aligned} \quad (33)$$

We note that both Eqs. (30) and (33) can be written approximately in the form<sup>17</sup>  $\Gamma_\phi \propto |A_s^d|^2 S_{\text{qp}}(\Gamma_\phi)$ ; however, the proportionality coefficients are different in the two cases. Solving Eq. (33) for  $\Gamma_\phi$  by iterations gives

$$\Gamma_\phi \approx \frac{32E_J}{\pi} |A_s^d|^2 f_0 (1 - f_0) \ln \frac{\pi\delta E}{8E_J |A_s^d|^2 f_0 (1 - f_0)}. \quad (34)$$

As a specific example, we consider from now on a quasi-equilibrium distribution  $f(\epsilon) = e^{-\epsilon/T_e}$ , where  $T_e$  is the effective quasiparticle temperature.<sup>18</sup> In this case we have  $\delta E = T_e$  and  $f_0 = e^{-\Delta/T_e} \ll 1$ , so that the dephasing rate is

$$\Gamma_\phi(T_e) \approx \frac{32E_J}{\pi} |A_s^d|^2 e^{-\Delta/T_e} \left[ \frac{\Delta}{T_e} + \ln \frac{\pi T_e}{8E_J |A_s^d|^2} \right]. \quad (35)$$

To conclude this section, we note that the divergence for  $\Delta^R = \Delta^L$  in Eq. (22) is a consequence of the square root singularity of the BCS density of states at the gap edge. Therefore possible modifications of the density of states (e.g., broadening<sup>19</sup>) would in principle lead to different estimates of the dephasing rate; the effect of a small density of subgap states has been recently considered in Ref. 20. However, we argue in Appendix E that these potential modifications are not relevant to current experiments with Al-based qubits, which we focus on for the remainder of the paper.

## IV. PHASE RELAXATION OF SINGLE-JUNCTION QUBITS

In this section we consider the dephasing rate for two single-junction systems, the phase qubit and the transmon, under the assumption of small qubit frequency,  $\omega_{10} \ll \Delta$  (see Appendix C 1 for the flux qubit). The calculations of the matrix element entering the relaxation rate are described in detail in Ref. 6, whose result we briefly summarize. Here we use (without giving all the details) the same approach of that work to obtain the matrix elements for dephasing. Interestingly, in all cases the pure dephasing rate  $\Gamma_\Phi$  turns out to add at most a small correction to  $1/T_2$  in comparison with the relaxation term  $1/2T_1$ .

### A. Phase qubit

In a phase qubit, the charging energy  $E_C$  is small compared to the transition frequency  $\omega_{10}$ . The latter depends on the external flux via the position  $\varphi_0$  of a minimum in the potential energy of the Hamiltonian in Eq. (3), as determined by

$$E_J \sin \varphi_0 + E_L (\varphi_0 - 2\pi\Phi_e/\Phi_0) = 0 \quad (36)$$

Then the frequency is

$$\omega_{10} = \sqrt{8E_C(E_L + E_J \cos \varphi_0)}. \quad (37)$$

For a small effective temperature  $T_e \ll \omega_{10}$  the relaxation time is

$$\frac{1}{T_1} = \frac{1}{\pi} \frac{\omega_p^2}{\omega_{10}} e^{-\Delta/T_e} \sqrt{\frac{\pi T_e}{\omega_{10}}} (1 + \cos \varphi_0), \quad (38)$$

where

$$\omega_p = \sqrt{8E_C E_J} \quad (39)$$

is the plasma frequency of the junction.

Within the same approximations used to obtain the above formulas,<sup>21</sup> the matrix element for dephasing is

$$|A_s^d|^2 = \frac{1}{8} \left( \frac{E_C}{\omega_{10}} \right)^2 (1 - \cos \varphi_0) \quad (40)$$

and substituting into Eq. (35) we get

$$\Gamma_\phi = \frac{E_C}{2\pi} \frac{\omega_p^2}{\omega_{10}^2} e^{-\Delta/T_e} \left[ \frac{\Delta}{T_e} + \ln \frac{8\pi T_e \omega_{10}^2}{E_C \omega_p^2 (1 - \cos \varphi_0)} \right] \times (1 - \cos \varphi_0). \quad (41)$$

Note that the factor in front of  $e^{-\Delta/T_e}$  is smaller for  $\Gamma_\phi$  in comparison with that for  $1/T_1$  because the matrix element for dephasing is smaller than that for relaxation by a factor  $E_C/\omega_{10}$ . At low temperatures the terms in square brackets in Eq. (41) are dominated by  $\Delta/T_e$  and hence, neglecting factors  $\cos \varphi_0$  as they are small compared to unity, the condition  $2T_1\Gamma_\phi > 1$  can be written as

$$\frac{T_e}{\Delta} < \left( \frac{\omega_{10}}{\Delta} \right)^{1/3} \left( \frac{E_C}{\omega_{10}} \right)^{2/3} \quad (42)$$

Typically for a phase qubit the product on the right is of order  $10^{-2}$ , while  $T_e/\Delta \sim 10^{-1}$ . Therefore the pure dephasing contribution to  $T_2$  [Eq. (24)] can be neglected. Interestingly, for a quasiparticle temperature of the order of the base temperature,  $T/\Delta \sim 10^{-2}$ , relaxation and pure dephasing would have similar order of magnitudes, although both would be much smaller than at  $T_e/\Delta \sim 10^{-1}$  due to their common exponential suppression by the Boltzmann factor.

### B. Transmon

The Hamiltonian of transmon is given by Eq. (3) with  $E_L = 0$ , supplemented by a periodic boundary condition in phase.<sup>4</sup> For our purposes, the transmon can be considered as a particular case of the phase qubit with  $\varphi_0 = 0$  [see Eq. (36)]. With these parameters, one obtains from Eq. (38) the correct estimate for the relaxation time  $T_1$ ,

$$\frac{1}{T_1} = \frac{2}{\pi} \omega_p e^{-\Delta/T_e} \sqrt{\frac{\pi T_e}{\omega_p}} \quad (43)$$

However, the vanishing for  $\varphi_0 = 0$  of the matrix element in Eq. (40) is not the correct result for the transmon: careful evaluation of the matrix element, following the procedure outlined in Appendices B and C of Ref. 6, gives an exponentially small value,  $A_s^d \propto \exp[-\sqrt{8E_J/E_C}]$ . This exponential suppression is sufficient to ensure that the dephasing rate is dominated by the contribution in Eq. (29), since the matrix element entering that equation has no such suppression,

$$|A_c^d|^2 = \frac{1}{4} \left( \frac{E_C}{\omega_p} \right)^2 = \frac{1}{32} \frac{E_C}{E_J} \quad (44)$$

Substituting this expression into Eq. (29), for the quasi-equilibrium distribution function we find

$$\Gamma_\phi = \frac{1}{\pi} E_C e^{-\Delta/T_e} \frac{T_e}{\Delta} \quad (45)$$

Using Eqs. (43) and (45) it is easy to show that for the transmon  $2T_1\Gamma_\phi \ll 1$ ; therefore, as for the phase qubit, the pure dephasing contribution to  $T_2$  is negligible.

## V. PHASE RELAXATION OF MULTI-JUNCTION QUBITS

The results of Sec. III are readily generalized to multi-junction systems by following the same procedure as in Sec. V of Ref. 6. Assuming the same gaps and distribution functions in all superconducting elements, we simply need to substitute

$$E_J |A_{s(c)}^d|^2 \rightarrow \sum_{j=0}^M E_{Jj} |A_{s(c),j}^d|^2 \quad (46)$$

in Eqs. (28) and (29) (and hence in subsequent equations in Sec. III B). Here index  $j$  denotes the  $M + 1$  junctions with Josephson energy  $E_{Jj}$  and capacitance  $C_j$ , while the matrix elements are defined by

$$A_{s,j}^d = \frac{1}{2} \left( \langle 1 | \sin \frac{\hat{\varphi}_j}{2} | 1 \rangle - \langle 0 | \sin \frac{\hat{\varphi}_j}{2} | 0 \rangle \right) \quad (47)$$

with  $\varphi_j$  the phase difference across junction  $j$ . The similar definition for  $A_{c,j}^d$  is obtained by replacing sine with

cosine. We remind that the phases are not independent, as they are constrained by the flux quantization condition

$$\sum_{j=0}^M \varphi_j = 2\pi f, \quad f = \Phi_e / \Phi_0. \quad (48)$$

Below we consider explicitly the two-junction split transmon, while the many-junction fluxonium is analyzed in Appendix C 2.

### A. Split transmon

The split transmon single degree of freedom is governed by the same Hamiltonian of the single-junction transmon, but the SQUID loop has a flux-dependent effective Josephson energy

$$E_J(f) = (E_{J0} + E_{J1}) \cos(\pi f) \sqrt{1 + d^2 \tan^2(\pi f)} \quad (49)$$

with

$$d = \frac{|E_{J0} - E_{J1}|}{E_{J0} + E_{J1}} \quad (50)$$

quantifying the junction asymmetry. In quasi-equilibrium at the effective temperature  $T_e$ , the relaxation time is given by<sup>6</sup>

$$\frac{1}{T_1(f)} = \sqrt{\frac{T_e}{\pi \omega_p(f)}} e^{-\Delta/T_e} \frac{\omega_p^2(f) + \omega_p^2(0)}{\omega_p(f)} \quad (51)$$

where

$$\omega_p(f) = \sqrt{8E_C E_J(f)}, \quad E_C = \frac{e^2}{2(C_0 + C_1)} \quad (52)$$

We note that the smaller the asymmetry, the larger the tunability of the qubit, since  $\omega_p(0)/\omega_p(1/2) = 1/\sqrt{d}$ . However, this flexibility comes at the price of enhancing the relaxation rate,  $T_1(0)/T_1(1/2) = (1+d)/(2d^{3/4})$ . In Fig. 1 we plot the normalized relaxation rate  $T_1(0)/T_1(f)$  as a function of reduced flux  $f$  for three values of the asymmetry parameter. We note that the relaxation rate rises by about a factor 1.5 up to  $f \sim 0.4$ , but can increase sharply for small asymmetry as  $f \rightarrow 0.5$ .

The matrix elements for dephasing are [cf. Eq. (40)]

$$|A_{s,j}^d|^2 = \frac{1}{8} \left( \frac{E_C}{\omega_p(f)} \right)^2 [1 - \cos(\pi f \pm \vartheta)] \quad (53)$$

where the upper (lower) sign should be used for  $j = 1$  ( $j = 0$ ) and

$$\tan(\vartheta) = d \tan(\pi f) \quad (54)$$

Note that in contrast with the single junction transmon, the matrix elements in general do not vanish (except at  $f = 0$ ). Using Eqs. (49), (53), and (54) we find

$$\sum_{j=0}^1 E_{Jj} |A_{s,j}^d|^2 = \frac{1}{64} E_C \left( \frac{\omega_p^2(0)}{\omega_p^2(f)} - 1 \right) \quad (55)$$

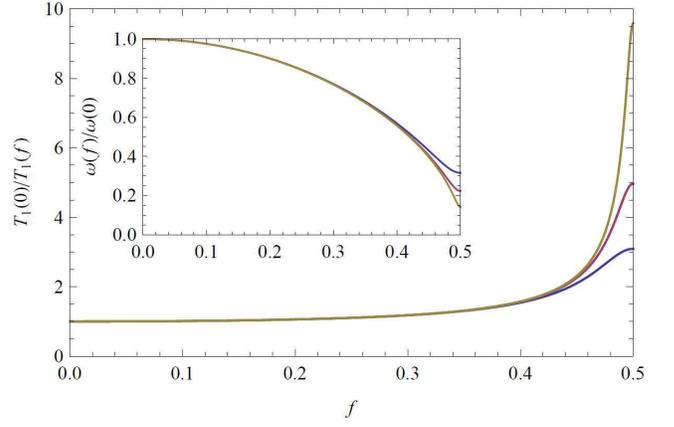


FIG. 1: Normalized relaxation rate  $T_1(0)/T_1(f)$  vs. reduced flux  $f$  for (top to bottom)  $d = 0.02, 0.05, 0.1$ . Inset: normalized frequency  $\omega_p(f)/\omega_p(0)$  vs. reduced flux for the same values of the asymmetry parameter (but decreasing top to bottom).

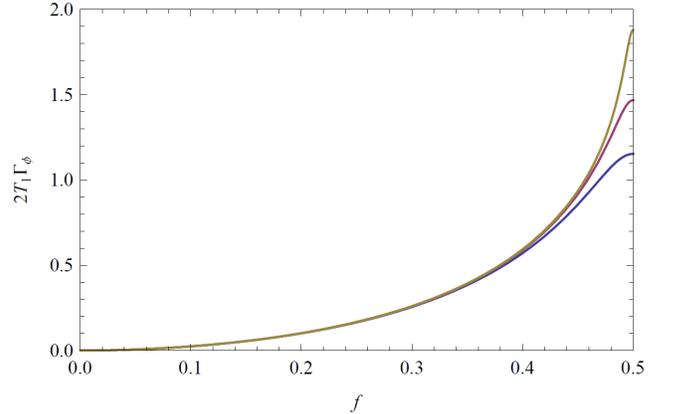


FIG. 2: Normalized dephasing rate  $2T_1\Gamma_\phi$  vs. reduced flux  $f$  for (top to bottom)  $d = 0.02, 0.05, 0.1$ . Other parameters are specified in the text after Eq. (56). The vanishing of  $\Gamma_\phi$  as  $f \rightarrow 0$  is an artifact of the approximations used to obtain Eq. (56); a finite dephasing rate at any flux would be obtained by including a subleading contribution analogous to Eq. (45).

and the above-described generalization to multi-junction systems of Eq. (35) gives

$$\Gamma_\phi = \frac{1}{2\pi} E_C \left( \frac{\omega_p^2(0)}{\omega_p^2(f)} - 1 \right) e^{-\Delta/T_e} \times \left[ \frac{\Delta}{T_e} + \ln \frac{8\pi T_e}{E_C (\omega_p^2(0)/\omega_p^2(f) - 1)} \right] \quad (56)$$

In Fig. 2 we show examples of the dependence of  $2T_1\Gamma_\phi$  on flux for different values of the asymmetry parameter  $d$  and typical values of the other dimensionless parameters ( $E_J(0)/E_C = 80$ ,  $\omega_p(0)/\Delta = 0.2$ ,  $T_e/\Delta = 0.06$ ); we note that near  $f = 1/2$  and for small asymmetry, pure dephasing dominates over relaxation,  $2T_1\Gamma_\phi > 1$ . Therefore the pure dephasing effect of quasiparticle tunneling

could be measured in a split transmon if other sources of dephasing (such as flux, photon, and charge noise) can be suppressed. Charge noise, in particular, can become the dominant dephasing mechanism as  $f \rightarrow 1/2$ , since the Cooper pair box regime of small  $E_J(f)/E_C$  is approached in this case for small asymmetry.<sup>4</sup> However, the contribution of  $\Gamma_\phi$  to  $1/T_2$  becomes relevant and thus potentially observable at values of reduced flux smaller than  $1/2$ , where the system is still in the transmon regime; for example, for  $f \sim 0.35$  where  $E_J(f)/E_C \sim 0.45E_J(0)/E_C$ , we estimate  $2T_1\Gamma_\phi \sim 0.4$ .

## VI. $T_2^*$ AND ANDREEV STATES IN A JOSEPHSON JUNCTION

In the previous sections we have considered the pure dephasing due to the interaction between tunneling quasiparticles and qubit. Here we study a different quasiparticle mechanism affecting the measured dephasing rate  $1/T_2^*$ : as discussed briefly in Sec. II and in more detail in Ref. 6, the quasiparticles renormalize the qubit frequency by shifting it by an amount  $\delta\omega$  which depends on the quasiparticle occupation. Therefore fluctuations in the occupation induce frequency fluctuations that can cause additional dephasing. In this section we focus on the phase qubit and show that this mechanism is not active during a single measurement, so that it does not contribute to the pure dephasing rate  $\Gamma_\phi$ ; however, it can contribute to the time  $T_2^*$  by changing the qubit frequency from measurement to measurement. In other words, this mechanism being slow on the scale of the qubit coherence time, its dephasing effect can be corrected by using echo techniques.

In a Josephson junction, weakly bound quasiparticles occupy the Andreev states that carry the dissipationless supercurrent.<sup>22</sup> Changes in the occupations of these states affect the value of the critical current (or equivalently of the Josephson energy) and in turn fluctuations in  $E_J$  lead to frequency fluctuations. As we show below, the parameter determining the relative magnitude of these fluctuations is the inverse square root of the (effective) number of transmission channels through the junction; therefore this fluctuation mechanism could be relevant in small junctions. For each transmission channel  $p$  ( $p = 1, \dots, N_{ch}$ ) with transmission probability  $T_p$  [defined after Eq. (6)], we find a corresponding Andreev bound state with binding energy [see Appendix D]

$$\omega_p^A = \Delta - E_p^A, \quad E_p^A = \Delta \left( 1 - \frac{1}{2} T_p \sin^2 \frac{\varphi_0}{2} \right) \quad (57)$$

This result is valid for  $T_p \ll 1$ ; the expression valid for arbitrary  $T_p$  can be found in Ref. 22. The (zero temperature) Josephson energy entering Eq. (3) is given by  $E_J = \Delta \sum_p T_p/4$ . To account for the occupations  $x_p^A$  of the Andreev states, due for example to finite tempera-

ture, in Eq. (37) we replace  $E_J$  by

$$E_J \rightarrow \frac{\Delta}{4} \sum_{p=1}^{N_{ch}} T_p (1 - 2x_p^A) \quad (58)$$

From this substitution we see that a change in the occupation of a single Andreev level can lead to a small change  $\delta E_J$  in the Josephson energy and hence in the qubit frequency, with a relative frequency shift of the order of  $\delta E_J/E_J \sim 1/N_{ch}$ . This effect could be measurable in small junction ( $N_{ch} \sim 10^5$ ) qubits and may have already been observed in a transmon, where slow frequency jumps of few parts per million magnitude have been measured.<sup>7</sup> More generally we find for the qubit frequency  $\omega_q$  at a given set of occupation numbers  $x_p^A$

$$\omega_q \simeq \omega_{10} - \frac{8E_C}{\omega_{10}} \cos \varphi_0 \sum_{p=1}^N \frac{\Delta}{4} T_p x_p^A \quad (59)$$

Here we assumed that on average the occupation numbers are small,  $x_{qp}^A = \langle x_p^A \rangle \ll 1$ ; in quasi-equilibrium the average takes the exponentially small value  $x_{qp}^A = e^{-\Delta/T_e}$ .<sup>23</sup> From this expression we see that fluctuations of the occupations of the Andreev states lead to frequency fluctuations. The mean square fluctuations of  $x_p^A$  are related to the average  $x_{qp}^A$  as<sup>24</sup>

$$\langle (\Delta x_p^A)^2 \rangle \equiv \langle (x_p^A - x_{qp}^A)^2 \rangle = x_{qp}^A (1 - x_{qp}^A) \quad (60)$$

Using this expression for the non-degenerate case  $x_{qp}^A \ll 1$ , we find for the root-mean-square frequency fluctuations

$$\frac{\sqrt{\langle (\Delta \omega_q)^2 \rangle}}{\omega_{10}} = |\cos \varphi_0| \frac{\omega_p^2}{\omega_{10}^2} \sqrt{x_{qp}^A} \frac{1}{\sqrt{N_e}} \quad (61)$$

where

$$N_e = \frac{\left( \sum_p T_p \right)^2}{\sum_p T_p^2} \quad (62)$$

is the effective number of channels;  $N_e$  coincides with  $N_{ch}$  if all the channels have equal transmission probabilities. The number  $N_e$  can be estimated independently by measuring the so called subgap structure due to Andreev reflections,<sup>25</sup>

$$N_e = \frac{\delta I_1}{\delta I_2} \frac{g_T}{2g_K}, \quad (63)$$

where the first factor in the right hand side is the ratio between the current step  $\delta I_1$  measured as the voltage increases from below to above  $2\Delta/e$  and the subgap current step  $\delta I_2$  at  $V \sim \Delta/e$ . This ratio is related to junction transparency and is of the order<sup>26,27</sup>  $\delta I_2/\delta I_1 \sim 10^{-5} - 10^{-3}$ , while depending on junction area

the ratio between junction conductance  $g_T$  and the conductance quantum  $g_K$  is  $g_T/g_K \sim 1-100$ , so we estimate  $N_e \sim 10^3$  to  $10^7$  for junction sizes from small to large.

The dephasing effect of the above frequency fluctuations gives observable contribution to  $T_2^*$  if

$$\langle (\Delta\omega_q)^2 \rangle^{1/2} T_2 \gtrsim 1. \quad (64)$$

Using Eq. (61) this condition is

$$T_2 \gtrsim \frac{\omega_{10}}{2\omega_p^2} \sqrt{\frac{N_e}{x_{\text{qp}}^A}} \sim \frac{1}{\omega_p} \sqrt{\frac{N_e}{x_{\text{qp}}^A}} \quad (65)$$

Assuming equilibrium between the occupation factors of Andreev states and free-quasiparticle states at the effective temperature  $T_e \approx 140$  mK (so that  $x_{\text{qp}}^A = e^{-\Delta/T_e}$ ), since  $\omega_p \sim 10^{11}$  s $^{-1}$  we find  $T_2 \gtrsim 10^{-6}$  s ( $10^{-4}$  s) for small (large) junctions. For phase qubits, which are fabricated with large junctions, this estimate is two to three orders of magnitude longer than the observed coherence time.<sup>2</sup> Therefore fluctuations in the occupations of Andreev levels do not contribute significantly to dephasing in current experiments with phase qubits.

The dephasing effect of the frequency fluctuations can be corrected using an echo pulse if the occupations do not change during a single measurement. In other words, if the rate at which the occupations change is small compared to  $1/T_2$ , then the fluctuations contribute to the decoherence time  $T_2^*$  rather than to  $T_2$ . Within our model Hamiltonian, Eq. (1), the only processes that can change the quasiparticle occupations are due to the interaction between qubit and quasiparticles; for an occupied Andreev level, this interaction leads to its ionization, with the qubit relaxing and giving its energy to a bound quasiparticle which is then excited into the continuum part of the spectrum. Since this process relaxes the qubit, it can in principle contribute to  $1/T_1$ . We show in Appendix D that this *intrinsic* contribution is small compared to the relaxation rate due to the interaction of the qubit with the bulk quasiparticles. There are of course *extrinsic* mechanisms that could affect the occupations of the Andreev states and hence the rate of frequency fluctuations. An example of such a mechanism is flux noise; we estimate that the ionization rate due to flux noise is in fact small compared to the experimental  $1/T_2$  – see Appendix D 2. Another mechanism is the quasiparticle recombination caused by the electron-phonon interaction. The recombination rate is  $\approx x_{\text{qp}}/\tau_0$ , with the characteristic time  $\tau_0 \sim 10^{-7} - 10^{-6}$  s in aluminum and  $\sim 10^{-10}$  s in niobium.<sup>28,29</sup> Since at low temperatures<sup>7,30</sup>  $x_{\text{qp}} \sim 10^{-7} - 10^{-8}$ , we find that the recombination rate is much smaller than  $1/T_2$ .

So far we have considered the effect of fluctuations of the Andreev levels occupations. Other mechanisms can in principle contribute to decoherence. For example, fluctuations of the order parameter  $\Delta$  in the vicinity of the junction also affect the Josephson energy, see Eq. (58); however, at low temperatures the typical time

scale over which  $\Delta$  changes in response to a sudden perturbation is very short, of order  $1/\Delta$ ,<sup>31</sup> so these fluctuations do not lead to additional decoherence. Another mechanism is associated with fluctuations in the number of free (rather than bound) quasiparticles. As discussed at the end of Sec. II, there are two contributions to the average frequency shift – the Josephson one,  $\delta\omega_{E_J}$ , and the quasiparticle one,  $\delta\omega_{\text{qp}}$ . Fluctuations of free quasiparticle occupations affect the latter, but their contribution to inhomogeneous broadening is small. Indeed, the average frequency shift can be obtained by considering the effect of quasiparticles on the junction impedance;<sup>5,6</sup> in quasiequilibrium the contribution of the normalized quasiparticle density  $x_{\text{qp}}$  to the quasiparticle part  $Y_{\text{qp}}$  of the junction impedance  $Y_J$  is smaller than the term in  $Y_J$  proportional to  $x_{\text{qp}}^A$  by the parameter  $\sqrt{T_e/\omega_{10}}$ . Moreover, the root mean square fluctuations of  $x_{\text{qp}}$  scale as the inverse square root of the volume of the electrodes<sup>24</sup> and can therefore be neglected for macroscopic electrodes.

## VII. SUMMARY

In this work we have studied decoherence caused by quasiparticles in superconducting qubits and obtained estimates for the pure dephasing rate  $\Gamma_\phi$  and for the contribution of inhomogeneous broadening to the decoherence rate  $1/T_2^*$ . We have presented a master equation approach that not only reproduces and generalizes the formula for the relaxation rate  $1/T_1$  of Refs. 5,6 [see Eq. (16)], but also gives a self-consistent expression for the pure dephasing rate  $\Gamma_\phi$ , Eq. (28). Moreover, in studying  $1/T_2^*$  we have derived a formula, Eq. (61), for the typical fluctuation of the qubit frequency due to change in the occupations of Andreev states. These two equations are our main results.

Application of Eq. (28) to single-junction qubits such as the phase qubit, the transmon (Sec. IV), and the flux qubit (Appendix C 1), and to the many-junctions fluxonium (Appendix C 2) shows that in these systems the pure dephasing rate is a small contribution to decoherence,  $2T_1\Gamma_\phi < 1$ . In the split transmon (Sec. V A), on the other hand, the quasiparticle dephasing rate can be larger than the relaxation rate when the external flux that tunes the qubit frequency approaches half the flux quantum, see Fig. 2; together with its temperature and flux dependence [Eq. (56)], the increased importance of  $\Gamma_\phi$  in this regime could permit its experimental measurement.

Finally in Sec. VI we have considered the contribution to the decoherence rate  $1/T_2^*$  due to quasiparticles bound into Andreev states localized near the Josephson junction. Fluctuations of the occupations of these levels from measurement to measurement can in principle induce dephasing which can be corrected with an echo pulse. In practice, this mechanism gives negligible contributions to dephasing in current experiments with phase qubits: due to the short observed  $T_2$  time, Eq. (64) im-

plies that the fluctuations of the occupations would need to cause relative frequency fluctuations of the order  $10^{-3}$  to start affecting the coherence of the qubit.

### Acknowledgments

This research was funded by Yale University, the Swiss NSF, the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), through the Army Research Office, the American NSF (Contract DMR-1004406), and the DOE (Contract DE-FG02-08ER46482).

### Appendix A: Derivation of the master equation

In this Appendix we summarize the main steps of the derivation of the master equation presented in Sec. III. Our starting point is the von Neumann equation,<sup>15</sup> which we write for the two components of the qubit (i.e., reduced) density matrix as

$$\frac{d\rho_z}{dt} = -i\text{Tr} \left\{ \left[ \delta\hat{H}; \hat{\rho}_t \right] \hat{\sigma}^z \right\} \quad (\text{A1})$$

$$\frac{d\rho_{\pm}}{dt} = i\omega_{10}\rho_{\pm} - i\text{Tr} \left\{ \left[ \delta\hat{H}; \hat{\rho}_t \right] \hat{\sigma}^{\pm} \right\} \quad (\text{A2})$$

Here  $\rho_t$  is the total density matrix of the system, comprising both qubit and quasiparticles,  $[\cdot; \cdot]$  denotes the commutator and, as discussed in Sec. II, for our purposes the interaction Hamiltonian  $\delta\hat{H} = \hat{H}_T$  is given by Eq. (8) with  $A_{nm}^f = 0$ . More useful forms of the traces in the right hand sides of the above equations are

$$\begin{aligned} \text{Tr} \left\{ \left[ \hat{H}_T; \hat{\rho}_t \right] \hat{\sigma}^z \right\} &= \langle\langle \left[ \hat{\sigma}^z; \hat{H}_T \right] \rangle\rangle \\ &= 2\tilde{t} \langle\langle (\hat{\sigma}^+ - \hat{\sigma}^-) \sum_{n,m,\sigma} A_{nm}^r \hat{\alpha}_{n\sigma}^{L\dagger} \hat{\alpha}_{m\sigma}^R \rangle\rangle + \text{H.c.}' \end{aligned} \quad (\text{A3})$$

and similarly

$$\begin{aligned} \text{Tr} \left\{ \left[ \hat{H}_T; \hat{\rho}_t \right] \hat{\sigma}^+ \right\} &= \tilde{t} \langle\langle \hat{\sigma}^z \sum_{n,m,\sigma} A_{nm}^r \hat{\alpha}_{n\sigma}^{L\dagger} \hat{\alpha}_{m\sigma}^R \rangle\rangle \\ &\quad - 2\tilde{t} \langle\langle \hat{\sigma}^+ \sum_{n,m,\sigma} A_{nm}^d \hat{\alpha}_{n\sigma}^{L\dagger} \hat{\alpha}_{m\sigma}^R \rangle\rangle + \text{H.c.}' \end{aligned} \quad (\text{A4})$$

where angular brackets denote quantum statistical averaging with respect to the total density matrix and the prime denotes that Hermitian conjugation is not applied to qubit operators (i.e., Pauli matrices).

The averages in the right hand sides of Eqs. (A3)-(A4) can be found by solving the equations governing their

time evolution, such as

$$\begin{aligned} -i\partial_t \langle\langle \hat{\sigma}^{\pm} \hat{\alpha}_{n\sigma}^{\dagger L} \hat{\alpha}_{m\sigma}^R \rangle\rangle &= \langle\langle \left[ \hat{H}; \hat{\sigma}^{\pm} \hat{\alpha}_{n\sigma}^{\dagger L} \hat{\alpha}_{m\sigma}^R \right] \rangle\rangle \\ &= (\pm\omega_{10} + \epsilon_n^L - \epsilon_m^R) \langle\langle \hat{\sigma}^{\pm} \hat{\alpha}_{n\sigma}^{\dagger L} \hat{\alpha}_{m\sigma}^R \rangle\rangle \\ &\quad + \tilde{t} \left\{ \pm A_{nm}^{d*} \rho_{\pm}(t) [f_n^L(1-f_m^R) + (1-f_n^L)f_m^R] - \frac{1}{2} A_{nm}^{r*} \right. \\ &\quad \left. \times [(1 \pm \rho_z(t)) f_n^L(1-f_m^R) - (1 \mp \rho_z(t))(1-f_n^L)f_m^R] \right\} \end{aligned} \quad (\text{A5})$$

The terms in curly brackets originate from averages of one qubit operator times four quasiparticle operators evaluated in the Born approximation,<sup>15</sup> for example

$$\begin{aligned} \sum_{i,j,\rho} \langle\langle \hat{\sigma}^+ \left\{ \hat{\alpha}_{j\rho}^{R\dagger} \hat{\alpha}_{i\rho}^L; \hat{\alpha}_{n\sigma}^{L\dagger} \hat{\alpha}_{m\sigma}^R \right\} \rangle\rangle \\ = \rho_+(t) [f_n^L(1-f_m^R) + (1-f_n^L)f_m^R] \end{aligned} \quad (\text{A6})$$

where  $\{\cdot; \cdot\}$  is the anticommutator. The solution of Eq. (A5) is

$$\begin{aligned} \langle\langle \hat{\sigma}^{\pm} \hat{\alpha}_{n\sigma}^{\dagger L} \hat{\alpha}_{m\sigma}^R \rangle\rangle &= i\tilde{t} \int_0^t d\tau e^{i(\pm\omega_{10} + \epsilon_n^L - \epsilon_m^R + i0^+)(t-\tau)} \\ &\quad \left\{ \pm A_{nm}^{d*} \rho_{\pm}(\tau) [f_n^L(1-f_m^R) + (1-f_n^L)f_m^R] - \frac{1}{2} A_{nm}^{r*} \right. \\ &\quad \left. [(1 \pm \rho_z(\tau)) f_n^L(1-f_m^R) - (1 \mp \rho_z(\tau))(1-f_n^L)f_m^R] \right\} \end{aligned} \quad (\text{A7})$$

A similar expression can be derived for the average in Eq. (A4) that contains  $\hat{\sigma}^z$ . After substituting these expressions into Eqs. (A3)-(A4) and the results into Eqs. (A1)-(A2), we perform two additional approximations. First, we neglect fast rotating terms; this so-called secular (or rotating wave) approximation<sup>15</sup> is valid when the decoherence rate is small on the scale of the qubit frequency,  $1/T_2\omega_{10} \ll 1$ , and it amounts to keeping in the equation for  $\rho_z$  only the terms proportional to  $(1 \pm \rho_z)$  and in the equation for  $\rho_{\pm}$  only those proportional to  $\rho_{\pm}$ . With this approximation we find

$$\begin{aligned} \frac{d\rho_z(t)}{dt} &= -2\tilde{t}^2 \int_0^t d\tau \sum_{n,m} |A_{nm}^r|^2 \\ &\quad \times \left\{ \rho_z(\tau) [f_n^L(1-f_m^R) + (1-f_n^L)f_m^R] \right. \\ &\quad \times [e^{+++} + e^{-+-} + e^{+--} + e^{-++}] \\ &\quad + [f_n^L(1-f_m^R) - (1-f_n^L)f_m^R] \\ &\quad \left. \times [e^{+++} - e^{-+-} - e^{+--} + e^{-++}] \right\} \end{aligned} \quad (\text{A8})$$

and

$$\begin{aligned} \frac{d\rho_+(t)}{dt} &= i\omega_{10}\rho_+(t) \\ -2\tilde{t}^2 \int_0^t d\tau \sum_{n,m} \rho_+(\tau) & [f_n^L(1-f_m^R) + (1-f_n^L)f_m^R] \\ \times \left\{ 2|A_{nm}^d|^2 [e^{++-} + e^{+--}] + |A_{nm}^r|^2 [e^{0+-} + e^{0-+}] \right\} \end{aligned} \quad (\text{A9})$$

where we use the shorthand notation

$$e^{\alpha\beta\gamma} = e^{i(\alpha\omega_{10} + \beta\epsilon_n^L + \gamma\epsilon_m^R + i0^+)(t-\tau)} \quad (\text{A10})$$

Next we introduce the Markov approximation<sup>15</sup> by substituting in the integrands of Eqs. (A8)-(A9)  $\rho_z(\tau) \rightarrow \rho_z(t)$ ,  $\rho_+(\tau) \rightarrow e^{-i\omega_{10}(t-\tau)}\rho_+(t)$  and extending the lower integration limits from 0 to  $-\infty$ . Then the  $\tau$ -integrals can be performed using the identity

$$\int_{-\infty}^t d\tau e^{i(\omega+i0^+)(t-\tau)} = iP\frac{1}{\omega} + \pi\delta(\omega) \quad (\text{A11})$$

where  $P$  denotes the principal part. We note that in Eq. (A8) the contributions of the principal parts cancel out, while after rewriting the summations over  $n, m$  as integrals over the quasiparticle energies the  $\delta$ -functions can be used to eliminate one of these integrals. Assuming equal gaps in the leads, we finally arrive at Eq. (15).

Applying the above steps to Eq. (A9), we find that the principal parts cancel out in the term proportional to  $A_{nm}^d$ ; in that term we assume different gaps with  $\Delta_R > \Delta_L$  to get expression (22) for the pure dephasing rate  $\Gamma_\phi$ . On the other hand, we can take the gaps to be the same in the term proportional to  $A_{nm}^r$ ; then the  $\delta$ -functions give rise to the contribution  $-1/2T_1\rho_+$  in Eq. (21). As for the principal parts, they contribute a term  $i\delta\tilde{\omega}\rho_+(t)$  with

$$\delta\tilde{\omega} = |A_s^r|^2 [F_{\text{qp}}(-\omega_{10}) - F_{\text{qp}}(\omega_{10})] \quad (\text{A12})$$

The function  $F_{\text{qp}}$  is defined in Appendix A of Ref. 6; as in that work, we have neglected here contributions suppressed by the factor  $\omega_{10}/\Delta$ . We note that while  $\delta\tilde{\omega}$  has a structure similar to that of  $\delta\omega_{\text{qp}}$  in Ref. 6, due to the projection onto the qubit subspace described in Sec. II the expression in Eq. (A12) accounts for virtual transitions between the qubit states only and neglects those to other states of the full system. In systems with small anharmonicity (e.g., the transmon and phase qubit) these transitions cannot be neglected and the average frequency shift must be calculated using the formulas in Ref. 6. Finally, we remind that the total average frequency shift  $\delta\omega$  contains also a Josephson part  $\delta\omega_{E_J}$ , as discussed in Sec. II.

## Appendix B: Dephasing at next-to-leading order

The self-consistent equation (28) for  $\Gamma_\phi$  requires going beyond the lowest order (in the tunneling amplitude

$\tilde{t}$ ) perturbative considerations of Appendix A in order to regularize the logarithmic divergence in Eq. (22) for equal gaps. Here we focus on the next to leading order contributions to validate that equation. First, however, let us discuss briefly the smearing of the singularity, Eq. (27), which is obtained as follows: after the Markov approximation, the term in Eq. (A9) proportional to  $A_{nm}^d$  is explicitly

$$\begin{aligned} -4\tilde{t}^2\rho_+(t) \sum_{n,m} [f_n^L(1-f_m^R) + (1-f_n^L)f_m^R] |A_{nm}^d|^2 \\ \lim_{\gamma \rightarrow 0^+} \int_{-\infty}^t d\tau \left[ e^{i(\epsilon_n^L - \epsilon_m^R + i\gamma)(t-\tau)} + e^{i(-\epsilon_n^L + \epsilon_m^R + i\gamma)(t-\tau)} \right] \end{aligned} \quad (\text{B1})$$

Rather than taking the limit, we assume  $\gamma$  small but finite (in particular,  $\gamma \ll \omega_{10}$  for the rotating wave approximation to be valid). After integration the last line becomes

$$\frac{2\gamma}{(\epsilon_n^L - \epsilon_m^R)^2 + \gamma^2} \quad (\text{B2})$$

This explains the origin of the last factor in the second line of Eq. (27), with the other factors accounting for the square root singularity of the BCS density of states. We now want to show that the identification  $\gamma = \Gamma_\phi$  is correct at next to leading order. To do so, we initially assume that the left/right gaps are different, so that the logarithmic divergence is absent and the perturbative expansion in  $\tilde{t}$  is justified. Next, we keep only those terms that would become logarithmically divergent in the limit of equal gaps.

To begin our derivation, we note that in Eq. (A9) the first term in square brackets multiplying  $A_{nm}^d$  originates from  $\langle\langle \hat{\sigma}^+ \hat{\alpha}_{n\sigma}^{\dagger L} \hat{\alpha}_{m\sigma}^R \rangle\rangle$ , as explained in Appendix A. Together with the other term in square brackets, they give rise to the pure dephasing rate term in the master equation (A9) via the equality

$$\begin{aligned} 2\tilde{t} \sum_{n,m,\sigma} \left[ A_{nm}^d \langle\langle \hat{\sigma}^+ \hat{\alpha}_{n\sigma}^{\dagger L} \hat{\alpha}_{m\sigma}^R \rangle\rangle + A_{nm}^{d*} \langle\langle \hat{\sigma}^+ \hat{\alpha}_{m\sigma}^R \hat{\alpha}_{n\sigma}^L \rangle\rangle \right] \\ = i\Gamma_\phi \rho_+(t) \end{aligned} \quad (\text{B3})$$

In what follow we first consider in some detail the next order contributions to  $\langle\langle \hat{\sigma}^+ \hat{\alpha}_{n\sigma}^{\dagger L} \hat{\alpha}_{m\sigma}^R \rangle\rangle$  and then discuss briefly the contributions to other averages. Without invoking the lowest order Born approximation, the equation of motion for  $\langle\langle \hat{\sigma}^+ \hat{\alpha}_{n\sigma}^{\dagger L} \hat{\alpha}_{m\sigma}^R \rangle\rangle$  is obtained by adding to the right hand side of Eq. (A5) the terms

$$\begin{aligned} \tilde{t} \sum_{i,j,\rho} \left[ A_{ij}^d N_{nm,ij}^{\sigma,\rho} + A_{ij}^{d*} M_{nm,ij}^{\sigma,\rho} - \frac{1}{2} A_{ij}^r Q_{nm,ij}^{\sigma,\rho} \right. \\ \left. - \frac{1}{2} A_{ij}^{r*} P_{nm,ij}^{\sigma,\rho} + \frac{1}{2} A_{ij}^{r*} S_{nm,ij}^{\sigma,\rho} + \frac{1}{2} A_{ij}^{r*} R_{nm,ij}^{\sigma,\rho} \right] \end{aligned} \quad (\text{B4})$$

with the definitions

$$M_{nm,ij}^{\sigma,\rho} = \langle\langle \hat{\sigma}^+ \left\{ \hat{\alpha}_{j\rho}^{R\dagger} \hat{\alpha}_{i\rho}^L; \hat{\alpha}_{n\sigma}^{L\dagger} \hat{\alpha}_{m\sigma}^R \right\} \rangle\rangle \quad (\text{B5})$$

$$N_{nm,ij}^{\sigma,\rho} = \langle\langle \hat{\sigma}^+ \left\{ \hat{\alpha}_{i\rho}^{L\dagger} \hat{\alpha}_{j\rho}^R; \hat{\alpha}_{n\sigma}^{L\dagger} \hat{\alpha}_{m\sigma}^R \right\} \rangle\rangle \quad (\text{B6})$$

$$P_{nm,ij}^{\sigma,\rho} = \langle\langle \hat{\sigma}^z \left\{ \hat{\alpha}_{j\rho}^{R\dagger} \hat{\alpha}_{i\rho}^L; \hat{\alpha}_{n\sigma}^{L\dagger} \hat{\alpha}_{m\sigma}^R \right\} \rangle\rangle \quad (\text{B7})$$

$$Q_{nm,ij}^{\sigma,\rho} = \langle\langle \hat{\sigma}^z \left\{ \hat{\alpha}_{i\rho}^{L\dagger} \hat{\alpha}_{j\rho}^R; \hat{\alpha}_{n\sigma}^{L\dagger} \hat{\alpha}_{m\sigma}^R \right\} \rangle\rangle \quad (\text{B8})$$

$$R_{nm,ij}^{\sigma,\rho} = \langle\langle \left[ \hat{\alpha}_{j\rho}^{R\dagger} \hat{\alpha}_{i\rho}^L; \hat{\alpha}_{n\sigma}^{L\dagger} \hat{\alpha}_{m\sigma}^R \right] \rangle\rangle \quad (\text{B9})$$

$$S_{nm,ij}^{\sigma,\rho} = \langle\langle \left[ \hat{\alpha}_{i\rho}^{L\dagger} \hat{\alpha}_{j\rho}^R; \hat{\alpha}_{n\sigma}^{L\dagger} \hat{\alpha}_{m\sigma}^R \right] \rangle\rangle \quad (\text{B10})$$

In introducing these definitions we have subtracted out the lowest order contributions already appearing in Eq. (A5). Then in that equation and in Eqs. (B5) and (B7) the density matrix should be understood as the lowest (zeroth) order one. In other words, by construction the quantities defined in Eqs. (B5)-(B10) account for higher order (in  $\tilde{t}$ ) contributions; these can be found by considering the equations of motions for those quantities, such as

$$\begin{aligned} -i\partial_t M_{nm,ij}^{\sigma,\rho} &= (\omega_{10} + \epsilon_n^L - \epsilon_m^R + \epsilon_j^R - \epsilon_i^L) M_{nm,ij}^{\sigma,\rho} \\ &+ \tilde{t} \sum_{k,l,\mu} \langle\langle \hat{\sigma}^+ \left\{ A_{kl}^d \hat{\alpha}_{k\mu}^{L\dagger} \hat{\alpha}_{l\mu}^R + A_{kl}^{d*} \hat{\alpha}_{l\mu}^{R\dagger} \hat{\alpha}_{k\mu}^L; \mathcal{A}_{nm,ij}^{\sigma,\rho} \right\} \rangle\rangle \\ &- \frac{1}{2} \hat{\sigma}^z \left\{ A_{kl}^r \hat{\alpha}_{k\mu}^{L\dagger} \hat{\alpha}_{l\mu}^R + A_{kl}^{r*} \hat{\alpha}_{l\mu}^{R\dagger} \hat{\alpha}_{k\mu}^L; \mathcal{A}_{nm,ij}^{\sigma,\rho} \right\} \\ &+ \frac{1}{2} \left[ A_{kl}^r \hat{\alpha}_{k\mu}^{L\dagger} \hat{\alpha}_{l\mu}^R + A_{kl}^{r*} \hat{\alpha}_{l\mu}^{R\dagger} \hat{\alpha}_{k\mu}^L; \mathcal{A}_{nm,ij}^{\sigma,\rho} \right] \end{aligned} \quad (\text{B11})$$

where  $\mathcal{A}_{nm,ij}^{\sigma,\rho}$  stands for the anticommutator

$$\mathcal{A}_{nm,ij}^{\sigma,\rho} = \left\{ \hat{\alpha}_{j\rho}^{R\dagger} \hat{\alpha}_{i\rho}^L; \hat{\alpha}_{n\sigma}^{L\dagger} \hat{\alpha}_{m\sigma}^R \right\} \quad (\text{B12})$$

At lowest order, all the averages in the right hand side of Eq. (B11) vanish; non-vanishing contributions can in principle be found by considering once again the equation of motions for those averages. As it is well known, proceeding in this manner we would obtain a hierarchy of coupled equations.<sup>32</sup> Here we make two approximations: first, we truncate the hierarchy at this level; second, as explained above we keep only those terms that in the limit of equal gaps would give logarithmically divergent contributions to the master equation. As a first step, this amounts to performing a mean-field like approximation in which the averages in the right hand side of Eq. (B11) are written in terms of product of averages as in the following example:

$$\begin{aligned} \langle\langle \hat{\sigma}^+ \left\{ \hat{\alpha}_{k\mu}^{L\dagger} \hat{\alpha}_{l\mu}^R; \mathcal{A}_{nm,ij}^{\sigma,\rho} \right\} \rangle\rangle &= 2 \langle\langle \hat{\sigma}^+ \hat{\alpha}_{k\mu}^{L\dagger} \hat{\alpha}_{l\mu}^R \rangle\rangle \langle\langle \mathcal{A}_{nm,ij}^{\sigma,\rho} \rangle\rangle \\ &+ 2 \langle\langle \hat{\sigma}^+ \hat{\alpha}_{n\sigma}^{L\dagger} \hat{\alpha}_{m\sigma}^R \rangle\rangle \langle\langle \mathcal{A}_{kl,ij}^{\mu,\rho} \rangle\rangle \end{aligned} \quad (\text{B13})$$

where

$$\langle\langle \mathcal{A}_{nm,ij}^{\sigma,\rho} \rangle\rangle = \delta_{ni} \delta_{mj} \delta_{\sigma\rho} [f_n^L (1 - f_m^R) + (1 - f_n^L) f_m^R] \quad (\text{B14})$$

Similar expressions can be written for the other averages appearing in Eq. (B11). In the second step we check which of the terms obtained in this way are logarithmically divergent in the limit of equal gaps and discard those that are finite (here we employ again the Born-Markov<sup>33</sup> and rotating wave approximations).

Applying the above procedure to Eq. (B11) we find that the terms in the last two lines can be neglected, while in terms originating from the second line we use Eq. (B3) as well as Eq. (A7) (in the rotating wave approximation, we only need to keep the term in the right hand side of that equation that contains  $\rho_+$ ). Solving the equation for  $M_{nm,ij}^{\sigma,\rho}$  so obtained we finally arrive at

$$\begin{aligned} M_{nm,ij}^{\sigma,\rho}(t) &= -t \Gamma_{\phi\rho_+}(t) \delta_{ni} \delta_{mj} \delta_{\sigma\rho} [f_n^L (1 - f_m^R) + (1 - f_n^L) f_m^R] - 2\tilde{t}^2 \rho_+(t) A_{ij}^d A_{nm}^{d*} [f_i^L (1 - f_j^R) + (1 - f_i^L) f_j^R] \\ &\times [f_n^L (1 - f_m^R) + (1 - f_n^L) f_m^R] \int_0^t du e^{i(\epsilon_n^L - \epsilon_m^R + \epsilon_j^R - \epsilon_i^L + i0^+)(t-u)} \int_0^u d\tau \left( e^{i(\epsilon_n^L - \epsilon_m^R + i0^+)(u-\tau)} + e^{i(-\epsilon_i^L + \epsilon_j^R + i0^+)(u-\tau)} \right) \end{aligned} \quad (\text{B15})$$

We then use the same approach to find the expression for  $N_{nm,ij}^{\sigma,\rho}$  [Eq. (B6)], which has the structure similar to that of the last term in Eq. (B15). Using these results

we get

$$\begin{aligned} &\sum_{i,j,\rho} (A_{ij}^d N_{nm,ij}^{\sigma,\rho} + A_{ij}^{d*} M_{nm,ij}^{\sigma,\rho}) \\ &= -\Gamma_{\phi\rho_+}(t) A_{nm}^{d*} [f_n^L (1 - f_m^R) + (1 - f_n^L) f_m^R] \\ &\times \left\{ t + \int_0^t d\tau e^{i(\epsilon_n^L - \epsilon_m^R + i0^+)(t-\tau)} \right\} \end{aligned} \quad (\text{B16})$$

To obtain the last term in curly brackets we used the identity

$$\int_0^t du \int_0^u d\tau h(\tau, u) = \int_0^t d\tau \int_0^t du h(\tau, u) - \int_0^t du \int_0^u d\tau h(u, \tau) \quad (\text{B17})$$

to combine contributions coming from  $M_{nm,ij}^{\sigma,\rho}$  and  $N_{nm,ij}^{\sigma,\rho}$  in a compact form.

Using the same procedure one can find the expressions for the quantities defined in Eqs. (B7)-(B10). Those quantities, however, do not contribute to the master equation within the approximations we are employing (in particular, we remind that in the rotating wave approximation we neglect by assumption terms small by the factor  $\Gamma_\phi/\omega_{10}$ ). Therefore, we obtain the following next-to-leading order equation of motion for  $\langle\langle \hat{\sigma}^+ \hat{\alpha}_{n\sigma}^{\dagger L} \hat{\alpha}_{m\sigma}^R \rangle\rangle$  by substituting Eq. (B16) into Eq. (B4) and adding the result to the left hand side of Eq. (A5):

$$\begin{aligned} -i\partial_t \langle\langle \hat{\sigma}^+ \hat{\alpha}_{n\sigma}^{\dagger L} \hat{\alpha}_{m\sigma}^R \rangle\rangle &= (\omega_{10} + \epsilon_n^L - \epsilon_m^R) \langle\langle \hat{\sigma}^+ \hat{\alpha}_{n\sigma}^{\dagger L} \hat{\alpha}_{m\sigma}^R \rangle\rangle \\ + \tilde{t} A_{nm}^{d*} \rho_+ &\left( 1 - \Gamma_\phi t - \Gamma_\phi \int_0^t d\tau e^{i(\epsilon_n^L - \epsilon_m^R + i0^+)(t-\tau)} \right) \\ &\times [f_n^L(1 - f_m^R) + (1 - f_n^L)f_m^R] \\ - \frac{\tilde{t}}{2} A_{nm}^{r*} &[(1 + \rho_z) f_n^L(1 - f_m^R) - (1 - \rho_z)(1 - f_n^L)f_m^R] \end{aligned} \quad (\text{B18})$$

As explained at the beginning of this Appendix, we want to show that this equation agrees at next to leading order with the smearing obtained by introducing a finite decay rate in the terms responsible for dephasing, with the decay rate given by  $\Gamma_\phi$  itself. Indeed, introducing this decay in Eq. (A7) we find

$$\begin{aligned} \langle\langle \hat{\sigma}^+ \hat{\alpha}_{n\sigma}^{\dagger L} \hat{\alpha}_{m\sigma}^R \rangle\rangle &= i\tilde{t} A_{nm}^{d*} \int_0^t d\tau e^{i(\omega_{10} + \epsilon_n^L - \epsilon_m^R + i\Gamma_\phi)(t-\tau)} \\ \rho_+(\tau) &[f_n^L(1 - f_m^R) + (1 - f_n^L)f_m^R] \\ - \frac{1}{2} i\tilde{t} A_{nm}^{r*} &\int_0^t d\tau e^{i(\omega_{10} + \epsilon_n^L - \epsilon_m^R + i0^+)(t-\tau)} \\ &[(1 + \rho_z(\tau)) f_n^L(1 - f_m^R) - (1 - \rho_z(\tau))(1 - f_n^L)f_m^R] \end{aligned} \quad (\text{B19})$$

Taking the time derivative of this equation we get

$$\begin{aligned} -i\partial_t \langle\langle \hat{\sigma}^+ \hat{\alpha}_{n\sigma}^{\dagger L} \hat{\alpha}_{m\sigma}^R \rangle\rangle &= (\omega_{10} + \epsilon_n^L - \epsilon_m^R) \langle\langle \hat{\sigma}^+ \hat{\alpha}_{n\sigma}^{\dagger L} \hat{\alpha}_{m\sigma}^R \rangle\rangle \\ + \tilde{t} A_{nm}^{d*} \rho_+ &[f_n^L(1 - f_m^R) + (1 - f_n^L)f_m^R] \\ - \frac{\tilde{t}}{2} A_{nm}^{r*} &[(1 + \rho_z) f_n^L(1 - f_m^R) - (1 - \rho_z)(1 - f_n^L)f_m^R] \\ - \Gamma_\phi \tilde{t} A_{nm}^{d*} &\int_0^t d\tau \rho_+(\tau) e^{i(\omega_{10} + \epsilon_n^L - \epsilon_m^R + i\Gamma_\phi)(t-\tau)} \\ &\times [f_n^L(1 - f_m^R) + (1 - f_n^L)f_m^R] \end{aligned} \quad (\text{B20})$$

At next-to-leading order, one should expand the exponentially decaying part of  $\rho_+$  [cf. Eqs. (23)-(24)] in the second line of Eq. (B20) and hence substitute there, with logarithmic accuracy,  $\rho_+ \rightarrow \rho_+(1 - \Gamma_\phi t)$ . The last term in Eq. (B20) is explicitly of higher order, so one can use  $\rho_+(\tau) \simeq e^{i\omega_{10}\tau}$  and drop  $\Gamma_\phi$  in the exponent. In this way we recover Eq. (B18), thus showing for  $\langle\langle \hat{\sigma}^+ \hat{\alpha}_{n\sigma}^{\dagger L} \hat{\alpha}_{m\sigma}^R \rangle\rangle$  the validity of our Ansatz. To complete the proof, we repeat the above steps for other averages, such as  $\langle\langle \hat{\sigma}^+ \hat{\alpha}_{m\sigma}^{\dagger R} \hat{\alpha}_{n\sigma}^L \rangle\rangle$  and  $\langle\langle \hat{\sigma}^z \hat{\alpha}_{n\sigma}^{\dagger L} \hat{\alpha}_{m\sigma}^R \rangle\rangle$ . The latter contributes to the  $1/2T_1$  term in the master equation (21) and at next-to-leading order the only correction we find is that corresponding to the expansion of the exponentially decaying part of  $\rho_+$ , as discussed above for the second line in Eq. (B20).

## Appendix C: Phase relaxation in flux qubit and fluxonium

### 1. Flux qubit

In a flux qubit, the external flux threading the superconducting loop is tuned to half the flux quantum,  $f = \Phi_e/\Phi_0 \simeq 1/2$ , and the potential energy takes the form of a double well. Then the qubit states  $|\pm\rangle$  are the two lowest tunnel-split states in this potential with energy difference

$$\omega_{10}(f) = \sqrt{\bar{\epsilon}^2 + [(2\pi)^2 \bar{E}_L (f - 1/2)]^2} \quad (\text{C1})$$

where for  $\bar{E}_J \gg \bar{E}_C$  we have

$$\bar{\epsilon} = 2\sqrt{\frac{2}{\pi}} \sqrt{8\bar{E}_C \bar{E}_J} \left( \frac{8\bar{E}_J}{\bar{E}_C} \right)^{1/4} e^{-\sqrt{8\bar{E}_J/\bar{E}_C}} \quad (\text{C2})$$

Expressions for the renormalized parameters  $\bar{E}_C$ ,  $\bar{E}_J$ , and  $\bar{E}_L$  in terms of the bare parameters of the Hamiltonian (3) can be found in Sec. IV.B of Ref. 6. It was shown there that the matrix element  $A_s^r$  vanishes at  $f = 1/2$  because of symmetry considerations, thus leading to a minimum for the relaxation rate. Here we focus on the case  $f = 1/2$  and therefore we need to evaluate the contribution to relaxation originating from the last line in Eq. (16). The relevant matrix element is

$$|A_c^r| = \frac{\bar{\epsilon}}{\omega_{10}(f)} \quad (\text{C3})$$

which equals unity at  $f = 1/2$ . Then from Eq. (16) we obtain

$$\frac{1}{T_1} = \frac{8}{\pi} E_J \sqrt{\frac{\omega_{10}}{2\Delta}} x_{\text{qp}} = \frac{8}{\pi} E_J \sqrt{\frac{\pi \bar{\epsilon} T_e}{\Delta^2}} e^{-\Delta/T_e} \quad (\text{C4})$$

Turning now to the dephasing rate, we find at  $f = 1/2$  the following expression for the matrix element

$$|A_s^d| = \frac{D}{2\sqrt{2}} \frac{\bar{\epsilon}}{\bar{E}_J} \left( \frac{\bar{E}_J}{\bar{E}_C} \right)^{1/3}, \quad (\text{C5})$$

where  $D \approx 1.45$  is a numerical coefficient.<sup>6</sup> Using Eq. (35) and (C4), after straightforward algebra we arrive at

$$2T_1\Gamma_\phi = \frac{D^2}{\sqrt{\pi}} \sqrt{\frac{\Delta}{E_C}} \sqrt{\frac{\Delta}{T_e}} \left(\frac{\bar{\epsilon}}{\bar{E}_C}\right)^{3/2} \left(\frac{\bar{E}_C}{\bar{E}_J}\right)^{4/3} \times \left\{ \frac{\Delta}{T_e} + \ln \left[ \frac{\pi}{D^2} \frac{T_e}{\Delta} \frac{\Delta}{\bar{E}_C} \left(\frac{\bar{E}_C}{\bar{\epsilon}}\right)^2 \left(\frac{\bar{E}_J}{\bar{E}_C}\right)^{1/3} \right] \right\} \quad (\text{C6})$$

Due to the exponential suppression of the splitting, Eq. (C2), this quantity is in general small. Indeed, for  $\bar{E}_C/\Delta, T_e/\Delta > 0.01$  and  $\bar{E}_J/\bar{E}_C \gtrsim 15$  we find  $2T_1\Gamma_\phi \lesssim 0.01$ .

## 2. Fluxonium

In the fluxonium an array of  $M \gg 1$  identical junctions (each with Josephson energy  $E_{J1} \gg E_{C1}$  large compared to their charging energy) acts as a lossy inductor connected to a weaker junction with  $E_{J0} < E_{J1}$ . The inductive energy of the array is  $E_L = E_{J1}/M$  and the losses are due to quasiparticle tunneling through the array junctions. In fact, for external flux near half the flux quantum the relaxation time is determined by this loss mechanism,<sup>6</sup>

$$\frac{1}{T_1} = 4\pi E_L \sqrt{\frac{\pi T_e}{\omega_{10}(f)}} e^{-\Delta/T_e} \left(\frac{\omega_{10}(1/2)}{\omega_{10}(f)}\right)^2, \quad (\text{C7})$$

since as discussed above for the flux qubit the contribution of the weaker junction is suppressed at  $f = 1/2$  [cf. Eq. (C4)]. Note that at  $f = 1/2$  the rate in Eq. (C7) is larger than that in Eq. (C4) by the factor  $(\Delta/E_J)(E_L/\omega_{10}(1/2))$ .

To calculate the dephasing rate, we note that at  $f = 1/2$  the matrix element for the weak junction is the same as for the flux qubit, Eq. (C5),

$$|A_{s,0}^d| = \frac{D}{2\sqrt{2}} \frac{\omega_{10}(1/2)}{\bar{E}_{J0}} \left(\frac{\bar{E}_{J0}}{\bar{E}_{C0}}\right)^{1/3} \quad (\text{C8})$$

while for each array junction we get

$$|A_{s,1}^d| = \frac{\pi}{2M} |A_{s,0}^d| \quad (\text{C9})$$

Then the coefficient containing the sum over all junctions is

$$\sum_{j=0}^M E_{Jj} |A_{s,j}^d|^2 = \left(E_{J0} + \frac{\pi^2}{4} E_L\right) |A_{s,0}^d|^2, \quad (\text{C10})$$

which in the limit  $\bar{E}_J/\bar{E}_C \gg 1$  is exponentially suppressed, see Eq. (C2). Therefore at  $f = 1/2$  the dephasing rate  $\Gamma_\phi$  has the same exponential suppression as in the flux qubit. Since as discussed above the fluxonium relaxation rate is parametrically larger than the

flux qubit one,<sup>34</sup> we find again that the pure dephasing rate is small compared to the relaxation rate for large  $\bar{E}_J/\bar{E}_C$ . The latter condition is not satisfied experimentally, since typically<sup>35</sup>  $E_J/E_C \lesssim 5$ , and numerical calculations beyond the scope of the present work may be needed to address this parameter regime. However, we note that in all cases studied here decreasing the ratio  $E_J/E_C$  increases the relative contribution of pure dephasing to  $1/T_2$ .

## Appendix D: Andreev bound states and ionization rate

The goals of this Appendix are to derive Eqs. (57) starting from the model defined by Eqs. (1)-(6), and to estimate the ionization rates due to qubit-quasiparticles interaction and flux noise. In the low energy limit where the characteristic energy of the quasi-particles  $\delta E$  as well as the qubit transition frequency  $\omega_{10}$  are small compared to the superconducting gap  $\Delta$ , we approximate the BCS coherence factors as  $u_n^j \approx v_n^j \approx 1/\sqrt{2}$ . Then considering for now a single channel junction, Eq. (6) takes the form<sup>6</sup>

$$\hat{H}_T = i\tilde{t} \sin(\hat{\varphi}/2) \sum_{n,m,\sigma} \hat{\alpha}_{n\sigma}^{L\dagger} \hat{\alpha}_{m\sigma}^R + \text{H.c.} \quad (\text{D1})$$

Assuming for simplicity identical left/right leads, we perform a canonical rotation into a new quasiparticle basis defined by the operators

$$\hat{\gamma}_{\pm n\sigma} = \frac{1}{\sqrt{2}} (\hat{\alpha}_{n\sigma}^L \pm i\hat{\alpha}_{n\sigma}^R). \quad (\text{D2})$$

In this basis we have [cf. Eq. (4)]

$$\hat{H}_{\text{qp}} = \hat{H}_{\text{qp}+} + \hat{H}_{\text{qp}-}, \quad \hat{H}_{\text{qp}\pm} = \sum_{n,\sigma} \epsilon_n \hat{\gamma}_{\pm n\sigma}^\dagger \hat{\gamma}_{\pm n\sigma} \quad (\text{D3})$$

$$\hat{H}_T = \tilde{t} \sin \frac{\hat{\varphi}}{2} \sum_{n,m,\sigma} \left( \hat{\gamma}_{-n\sigma}^\dagger \hat{\gamma}_{-m\sigma} - \hat{\gamma}_{+n\sigma}^\dagger \hat{\gamma}_{+m\sigma} \right). \quad (\text{D4})$$

Denoting with  $|j\rangle$  and  $\mathcal{E}_j$  the eigenstates and eigenenergies of  $\hat{H}_\varphi$  [Eq. (3)], the total Hamiltonian  $\hat{H}$  can then be split into parts that are respectively diagonal and non-diagonal in the qubit subspace,  $\hat{H} = \hat{H}_d + \hat{H}_{\text{nd}}$ , with the diagonal part defined as

$$\hat{H}_d = \sum_j \mathcal{E}_j |j\rangle\langle j| + \sum_j |j\rangle\langle j| \left( \hat{H}_{j+} + \hat{H}_{j-} \right), \quad (\text{D5})$$

where

$$\hat{H}_{j\pm} = \hat{H}_{\text{qp}\pm} \mp \tilde{t} s_{jj} \sum_{n,m,\sigma} \hat{\gamma}_{\pm n\sigma}^\dagger \hat{\gamma}_{\pm m\sigma}. \quad (\text{D6})$$

and we have used the definition (10) for the matrix elements  $s_{ij}$ . The non-diagonal part is given by

$$\hat{H}_{\text{nd}} = \tilde{t} \sum_{i \neq j} s_{ij} |i\rangle\langle j| \sum_{m,n,\sigma} \left( \hat{\gamma}_{-n\sigma}^\dagger \hat{\gamma}_{-m\sigma} - \hat{\gamma}_{+n\sigma}^\dagger \hat{\gamma}_{+m\sigma} \right) \quad (\text{D7})$$

It describes real transitions in which qubit and quasiparticles exchange energy. The term proportional to  $\tilde{t}$  in the diagonal part, on the other hand, accounts for virtual transitions that renormalize the spectrum. Indeed, as we show next, for  $s_{jj} > 0$  ( $s_{jj} < 0$ ) there exists a sub-gap Andreev bound state in the  $\gamma_+$  ( $\gamma_-$ ) subspace. Because the two subspaces are uncoupled, we can restrict ourselves to either one of those; in the following we consider the  $\gamma_+$  subspace.

Since  $\hat{H}_d$  is diagonal in the qubit space, to find the spectrum we only need to calculate the eigenvalues of the quasiparticle Hamiltonians  $\hat{H}_{j\pm}$ . We denote with  $|A_j\rangle$  the wavefunction of the Andreev state when the qubit is in state  $|j\rangle$ ; to solve the Schrödinger equation  $\hat{H}_{j+}|A_j\rangle = E|A_j\rangle$  we make the Ansatz  $|A_j\rangle = \sum_{n\sigma} a_{jn} \gamma_{+n\sigma}^\dagger |\emptyset\rangle$ , where  $|\emptyset\rangle$  denotes the quasiparticle vacuum state,  $\gamma_{\pm n\sigma} |\emptyset\rangle = 0$ , and obtain the following system of linear equations

$$a_{jn} = \tilde{t} s_{jj} \frac{1}{\epsilon_n - E} \sum_m a_{jm}. \quad (\text{D8})$$

To find the eigenenergy  $E$ , we sum both sides over  $n$  and in the low energy limit we write  $\epsilon_n \approx \Delta + \xi_n^2/(2\Delta)$ ; then the sum over  $n$  in the right hand side can be approximated by an integral,  $\sum_n \approx \nu_0 \int d\xi$ , and we arrive at

$$1 = \pi \nu_0 \tilde{t} s_{jj} \sqrt{\frac{2\Delta}{\Delta - E}}. \quad (\text{D9})$$

A solution with energy  $E < \Delta$  exists if and only if  $s_{jj} > 0$  (the opposite holds in the  $\gamma_-$  subspace where a bound state exists if and only if  $s_{jj} < 0$ ). The corresponding bound state energy is

$$E_j^A = \Delta [1 - 2(\pi \nu_0 \tilde{t})^2 s_{jj}^2], \quad (\text{D10})$$

This energy depends on the state of the qubit via the matrix element  $s_{jj}$ . However, for the low-energy states of the phase qubit this matrix element is the same at leading order in  $E_C/\omega_{10} \ll 1$ , since the square of the matrix element is<sup>6</sup>

$$s_{ij}^2 = \delta_{i,j} \left[ 1 - 2 \frac{E_C}{\omega_{10}} \left( i + \frac{1}{2} \right) \right] \sin^2 \frac{\varphi_0}{2} + \frac{E_C}{\omega_{10}} [j \delta_{i,j-1} + (j+1) \delta_{i,j+1}] \cos^2 \frac{\varphi_0}{2} \quad (\text{D11})$$

up to higher order terms  $\propto (E_C/\omega_{10})^2$ . Keeping only the leading term in this equation, introducing the transmission probability  $\mathbb{T} = (2\pi \nu_0 \tilde{t})^2$  in Eq. (D10), and generalizing it to multiple channels, we arrive at Eq. (57). (In that equation the subscript  $p$  denotes the transmission channel, and we have dropped the qubit state index  $j$  since, as explained above, the leading order expression is independent of  $j$ .)

For later use, we note that the normalization condition  $\sum_n (a_{jn})^2 = 1/2$ , which accounts for spin degeneracy, together with the square of Eq. (D8), leads to the

amplitudes

$$a_{jn} = \frac{1}{\sqrt{\pi \nu_0}} \frac{(2\Delta \omega_j^A)^{3/4}}{\xi_n^2 + 2\Delta \omega_j^A}, \quad (\text{D12})$$

where  $\omega_j^A = \Delta - E_j^A$  is the binding energy.

### 1. Ionization rate

The ionization of the Andreev level can be caused by quantum fluctuations of the phase difference across the junction induced by the finite charging energy  $E_C$ ; the ionization rate can be calculated using Fermi's golden rule by treating the non-diagonal part (D7) of the Hamiltonian as a perturbation. For a qubit initially in the state  $|i\rangle$ , the ionization rate  $\Gamma_i^A$  is given by

$$\Gamma_i^A = 2\pi \sum_{nj} \left| \langle j, \epsilon_{jn} | \hat{H}_{\text{nd}} | i, A_i \rangle \right|^2 \times \delta(\mathcal{E}_j + \epsilon_n - \mathcal{E}_i - E_i^A) (1 - f(\epsilon_n)). \quad (\text{D13})$$

Here  $|\epsilon_{jn}\rangle$  is a scattering state in the continuum part of the quasiparticle spectrum and the factor  $(1 - f(\epsilon_n))$  gives the probability that this state is empty. The matrix element in Eq. (D13) is the product of the off-diagonal matrix element  $s_{ji}$  times the overlap of the wavefunctions of bound and scattering states at the junction,

$$\langle j, \epsilon_{jn} | \hat{H}_{\text{nd}} | i, A_i \rangle = -\tilde{t} s_{ji} \sum_m \psi_{jm}^*(\epsilon_n) \sum_{n'} a_{in'}, \quad (\text{D14})$$

where  $\psi_{jm}(\epsilon_n) = \langle \epsilon_m | \epsilon_{jn} \rangle$  and  $|\epsilon_m\rangle$  are the eigenstates of  $\hat{H}_{\text{qp}+}$ , see Eq. (D3). Next we calculate the wavefunctions for the continuum states by solving the scattering problem in the standard  $T$ -matrix approach.<sup>36</sup>

We focus again on the  $\gamma_+$  subspace and write  $\hat{H}_{j+} = \hat{H}_{\text{qp}+} + \hat{H}_{j1}$  with  $\hat{H}_{j1} = -\tilde{t} s_{jj} \sum_{nm,\sigma} \hat{\gamma}_{+n\sigma}^\dagger \hat{\gamma}_{+m\sigma}$  [cf. Eq. (D6)]. From the Schrödinger equation, we have for the scattering states  $|\epsilon_{jn}\rangle$

$$|\epsilon_{jn}\rangle = |\epsilon_n\rangle + \frac{1}{\epsilon_n - \hat{H}_{\text{qp}+} + i0^+} \hat{H}_{1j} |\epsilon_{jn}\rangle = \left[ \hat{1} + \frac{1}{\epsilon_n - \hat{H}_{\text{qp}+} + i0^+} \mathbb{T}_j(\epsilon_n) \right] |\epsilon_n\rangle, \quad (\text{D15})$$

where we have defined the  $T$ -matrix as

$$\mathbb{T}_j(\epsilon_n) = \hat{H}_{1j} + \hat{H}_{1j} [\epsilon_n - \hat{H}_{\text{qp}+} + i0^+]^{-1} \hat{H}_{1j} + \dots \quad (\text{D16})$$

The  $T$ -matrix is related to the quasiparticle Green's function  $G_j$  via

$$G_j = g + g \mathbb{T}_j g, \quad (\text{D17})$$

where  $g$  is the (diagonal in momentum) bare quasiparticle Green's function  $g_n(\omega) = 1/(\omega - \epsilon_n + i0^+)$ . Using the

inverse of this equation:  $\mathbb{T}_j = g^{-1}G_j g^{-1} - g^{-1}$ , we find upon projecting Eq. (D15) onto  $|\epsilon_m\rangle$

$$\psi_{jm}(\epsilon_n) = \lim_{\omega \rightarrow \epsilon_n} G_{j,mn}(\omega)(g_n(\omega))^{-1}. \quad (\text{D18})$$

The Green's function, as obtained from the equations of motion for  $\gamma_{+n\sigma}$ , is given by

$$G_{j,nm}(\omega) = \delta_{nm}g_n(\omega) - \frac{\tilde{t}s_{jj}g_n(\omega)g_m(\omega)}{1 + \tilde{t}s_{jj}\sum_p g_p(\omega)}. \quad (\text{D19})$$

Hence the continuum states are

$$\psi_{jm}(\epsilon_n) = \delta_{nm} - \frac{\tilde{t}s_{jj}g_m(\epsilon_n)}{1 - i\pi\nu_0\tilde{t}s_{jj}\sqrt{\frac{2\Delta}{\epsilon_n - \Delta}}}, \quad (\text{D20})$$

where we have used that in the low energy limit  $\sum_p g_p(\epsilon_n) \approx -i\pi\nu_0\sqrt{2\Delta}/(\epsilon_n - \Delta)$ .

Using Eqs. (D12) and (D20) we find

$$\sum_n a_{in} = 2^{1/4}\pi\nu_0\sqrt{\tilde{t}\Delta}s_{ii}, \quad (\text{D21})$$

$$\sum_m \psi_{jm}^*(\epsilon_n) = \frac{1}{1 + i\pi\nu_0\tilde{t}s_{jj}\sqrt{2\Delta}/(\epsilon_n - \Delta)}. \quad (\text{D22})$$

Substituting these expressions into Eq. (D13), and considering explicitly the case of a phase qubit, using the expressions for the matrix elements  $s_{ij}$  given in Eq. (D11) finally yields for the ionization rate of a single-channel junction

$$\Gamma_j^A = j\frac{\omega_p^2}{8\omega_{10}}(1 + \cos\varphi_0)\frac{\sqrt{\frac{2\omega_j^A}{\mathcal{E}_j - \mathcal{E}_{j-1} - \omega_j^A}}}{1 + \frac{\omega_{j-1}^A}{\mathcal{E}_j - \mathcal{E}_{j-1} - \omega_j^A}} \times (1 - f(\mathcal{E}_j - \mathcal{E}_{j-1} + E_j^A)), \quad (\text{D23})$$

with  $\varphi_0$ ,  $\omega_p$ , and  $\omega_{10}$  defined in Sec. IV A and we used that for a single-channel junction  $E_J = \Delta(\pi\nu_0\tilde{t})^2$ .

The above result can be easily generalized to the case of  $N_{ch}$  independent channels. Assuming for simplicity identical transmission amplitudes, the Andreev binding energy can be written as  $\omega_j^A = 2E_Js_{jj}^2/N_{ch}$  which for a phase qubit reduces approximately to  $\omega^A \approx E_J/N_{ch}$ . We assume  $N_{ch} \gg 1$  sufficiently large so that  $\omega^A \ll \omega_{10}$  and obtain for the ionization rate of each occupied channel

$$\Gamma_1^A \approx \frac{1}{4N_{ch}}\frac{\omega_p^2}{\omega_{10}}\sqrt{\frac{2\omega^A}{\omega_{10}}}\frac{1 + \cos\varphi_0}{2}. \quad (\text{D24})$$

A single ionization event is sufficient to relax the qubit energy, and the probability of at least one ionization event taking place during time  $t$ , when initially  $N_{occ} \leq N_{ch}$  Andreev levels are occupied, is given by  $p = 1 - e^{-N_{occ}\Gamma_1^A t}$ . Introducing the total ionization rate  $\Gamma_{tot}^A = N_{occ}\Gamma_1^A$  and using Eq. (38), leads to the estimate

$$T_1\Gamma_{tot}^A \approx \frac{1}{4\sqrt{2\pi}}\frac{N_{occ}}{N_{ch}^{3/2}}e^{\Delta/T_e}\sqrt{\frac{E_J}{T_e}}, \quad (\text{D25})$$

Defining the frequency shift  $\delta\omega_q = \omega_{10} - \omega_q$  and using Eq. (59) to estimate its value, we may eliminate  $N_{occ}$  and rewrite the above as

$$T_1\Gamma_{tot}^A \sim e^{\Delta/T_e}\sqrt{\frac{E_J}{N_{ch}T_e}}\frac{\delta\omega_q}{\omega_{10}} \sim 4 \times 10^4 \frac{\delta\omega_q}{\omega_{10}} \quad (\text{D26})$$

where we used  $E_J/N_{ch} \sim 10^{-5}\Delta$  and  $T_e \approx 140$  mK. Thus when  $\delta\omega_q \gtrsim 10^{-4}\omega_{10}$ , the qubit relaxation is likely dominated by the ionization process, rather than by quasiparticle transitions within the continuum. However, we note that the typical shift is much smaller than this,  $\delta\omega_q/\omega_{10} \sim e^{-\Delta/T_e} \sim 3 \times 10^{-7}$ , i.e.  $p \approx 0.012$ , so the contribution of ionization to qubit relaxation is negligible unless  $N_{occ}$  is anomalously large.

## 2. Ionization by flux noise

As an example of an *extrinsic* ionization mechanism, we consider here low frequency ( $\ll \omega_{10}$ ) flux noise. Small fluctuations  $\delta\Phi_e(t) \ll \Phi_0$  of the external flux induce small fluctuations  $\varphi_1(t)$  of the phase difference  $\varphi_0$ ,

$$\varphi_1(t) = 2\pi\frac{\delta\Phi_e(t)}{\Phi_0}\frac{E_L}{E_L + E_J\cos\varphi_0} \quad (\text{D27})$$

[see Eq. (36)]. Since the low-frequency fluctuations do not induce qubit transitions, their effect is accounted for by substituting  $\varphi_0 \rightarrow \varphi_0 + \varphi_1(t)$  into the diagonal matrix element  $s_{jj}$  in Eq. (D6). At linear order in  $\varphi_1$  we thus obtain the time-dependent perturbation (in the  $\gamma_+$  subspace)

$$\hat{V}(t) = -\tilde{t}\frac{\varphi_1(t)}{2}\cos\left(\frac{\varphi_0}{2}\right)\sum_{n,m,\sigma}\gamma_{+n\sigma}^\dagger\gamma_{+m\sigma}, \quad (\text{D28})$$

Using Fermi's golden rule and following similar steps as in the previous section, the total ionization rate can be expressed as

$$\Gamma_{tot}^A = N_{occ}\left(\frac{E_J}{N_{ch}}\right)^{3/2}\left|\sin\frac{\varphi_0}{2}\right|\frac{1 + \cos\varphi_0}{2} \times \int_{\omega_A}^{\infty} d\omega S_{\varphi\varphi}(\omega)\frac{\sqrt{\omega - \omega_A}}{\omega}(1 - f(\omega + E_A)). \quad (\text{D29})$$

where  $S_{\varphi\varphi}(\omega) = 1/(2\pi)\int e^{i\omega t}\langle\varphi_1(t)\varphi_1(0)\rangle dt$  is the phase fluctuation spectrum and the binding energy introduces a natural low-frequency cutoff. For non-degenerate quasiparticles,  $f(\omega_A + E_A) \ll 1$ , and a power-law spectrum of the form<sup>2,3,10</sup>  $S_{\varphi\varphi}(\omega) = (\delta\varphi)^2/(2\pi\omega^\alpha)$ , we obtain

$$\Gamma_{tot}^A = \frac{(\delta\varphi)^2}{2\pi}N_{occ}\left(\frac{E_J}{N_{ch}}\right)^{3/2}\left|\sin\frac{\varphi_0}{2}\right|\frac{1 + \cos\varphi_0}{2} \times \omega_A^{\frac{1}{2}-\alpha}\int_1^{\infty} dx\frac{\sqrt{x-1}}{x^{\alpha+1}} \quad (\text{D30})$$

For  $\alpha = 1$  (pure  $1/f$  noise), the remaining integral is equal to  $\pi/2$  and since  $\omega_A = 2(E_J/N_{ch})\sin^2(\varphi_0/2)$ , we arrive at

$$\Gamma_{tot}^A = \frac{(\delta\varphi)^2}{4\sqrt{2}} N_{occ} \left( \frac{E_J}{N_{ch}} \right) \frac{1 + \cos\varphi_0}{2}. \quad (\text{D31})$$

The measured<sup>2,3</sup> magnitude of the fluctuations is small,  $\delta\varphi \sim 10^{-6}$ ; since  $E_J/N_{ch} \sim 10^{-5}\Delta$  and  $N_{occ} \ll N_{ch} \lesssim 10^7$ , we estimate this rate to be much smaller than 1 Hz.

### Appendix E: Modifications of the density of states

The logarithmic divergence of the dephasing rate and its regularization discussed in Sec. III are a consequence of the square root singularity of the BCS density of states at the gap edge. Here we discuss two other mechanisms that also can regularize the divergence and show that for Al-based qubits used at present they do not modify the results in the main text.

To begin with we consider the broadened density of states introduced by Dynes<sup>19</sup> to interpret experimental tunneling data. This phenomenological density of states is characterized by a broadening parameter  $\Gamma_D \ll \Delta$  and a finite density of subgap states. These states give rise to an additional contribution to the dephasing rate which we denote with  $\Gamma_\phi^{sg}$ ; assuming quasi-equilibrium, it is given by<sup>20</sup>

$$\Gamma_\phi^{sg}(T_e) = \frac{16E_J}{\pi} |A_s^d|^2 \left( \frac{\Gamma_D}{\Delta} \right)^2 \frac{T_e}{\Delta}. \quad (\text{E1})$$

and it is always smaller than the broadening,  $\Gamma_\phi^{sg} \ll \Gamma_D$ . Comparing Eqs. (35) and (E1), we see that a small broadening in the latter can compensate for the exponential suppression of the quasiparticle occupation in the former. Then we can distinguish three regimes: 1. at ‘‘high’’ temperatures, the dephasing rate is given by Eq. (35), since the broadening can be neglected in calculating  $\Gamma_\phi$ . The high-temperature regime is defined by the condition  $\Gamma_D \lesssim \Gamma_\phi(T_e)$ ; 2. at intermediate temperatures, when  $\Gamma_\phi^{sg}(T_e) \lesssim \Gamma_\phi(T_e) \lesssim \Gamma_D$ , the broadening of the density of states cannot be neglected. With logarithmic accuracy, this amounts to substitute  $\Gamma_\phi \rightarrow \Gamma_D$  in the last term in Eq. (33) [and hence replace the square bracket in Eq. (35) with  $\ln(T_e/\Gamma_D)$ ]; we note that since this substitution affects only the logarithm, use of Eq. (35) still

gives a correct order-of-magnitude estimate]. 3. at low temperatures, such that  $\Gamma_\phi(T_e) \lesssim \Gamma_\phi^{sg}(T_e)$  the subgap contribution becomes dominant.

In recent measurements<sup>30</sup> the intrinsic value of the broadening parameter in aluminum was found to be small,  $\Gamma_D/\Delta < 2 \times 10^{-7}$ . Using this value and the results of the next section, our estimates show that the low-temperature regime is entered for  $T_e \lesssim 60$  mK. In experiments with superconducting resonators<sup>37</sup> as well as qubits<sup>7,38</sup> the quasiparticle effective temperature is larger,  $T_e \sim 140$  mK, so we can neglect the subgap contribution to the dephasing rate for Al-based qubits, which we focus on in this paper. However, the subgap contribution may be relevant in other systems, such as qubits fabricated with niobium.<sup>20</sup>

While the above consideration are based on a phenomenological model, an *intrinsic* modification of the continuum part of the density of states near the junction is due to the presence of Andreev bound states. They modify the square root singularity into a square root threshold,

$$\sqrt{\frac{2\Delta}{\omega - \Delta}} \rightarrow \frac{\sqrt{2\Delta}\sqrt{\omega - \Delta}}{\omega - E^A} \quad (\text{E2})$$

with  $E^A$  the energy of the bound state defined in Eq. (57) (here we consider for simplicity the single channel case). The above substitution can be obtained using Eq. (D19) for the Green’s function to calculate the density of states. Assuming the binding energy  $\omega^A = \Delta - E^A$  to be small compared to the typical quasiparticle energy,  $\omega^A \ll \delta E$ , we find that the substitution (E2) would lead to the replacement of  $\Gamma_\phi$  with  $\omega^A$  in the right hand side of Eq. (33). In quasi-equilibrium this amounts to replacing the square brackets in Eq. (35) with

$$\ln \frac{T_e}{\omega^A} \sim \ln \frac{T_e}{E_J} + \ln N_{ch} \quad (\text{E3})$$

where  $N_{ch} \gg 1$  is the number of channels in the junction. We note that the tunneling limit we are considering consists in taking the transmission amplitude  $\tilde{t} \rightarrow 0$  at finite  $E_J$ , which implies  $N_{ch} \rightarrow \infty$ . Then in this limit the self-consistent approach is justified with logarithmic accuracy as explained in Appendix B. .

<sup>1</sup> D. P. DiVincenzo, Fortschr. Phys. **48**, 771 (2000).

<sup>2</sup> R. C. Bialczak, R. McDermott, M. Ansmann, M. Hofheinz, N. Katz, E. Lucero, M. Neeley, A. D. O’Connell, H. Wang, A. N. Cleland, and J. M. Martinis, Phys. Rev. Lett. **99**, 187006 (2007).

<sup>3</sup> F. Yoshihara, K. Harrabi, A. O. Niskanen, Y. Nakamura, and J. S. Tsai, Phys. Rev. Lett. **97**, 167001 (2006).

<sup>4</sup> J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, Alexandre Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. A **76**, 042319 (2007).

<sup>5</sup> G. Catelani, J. Koch, L. Frunzio, R. J. Schoelkopf, M. H. Devoret, and L. I. Glazman, Phys. Rev. Lett. **106**, 077002 (2011).

- <sup>6</sup> G. Catelani, R. J. Schoelkopf, M. H. Devoret, and L. I. Glazman, Phys. Rev. B **84**, 064517 (2011).
- <sup>7</sup> 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).
- <sup>8</sup> M. Lenander, H. Wang, R. C. Bialczak, E. Lucero, M. Mariantoni, M. Neeley, A. D. O'Connell, D. Sank, M. Weides, J. Wenner, T. Yamamoto, Y. Yin, J. Zhao, A. N. Cleland, and J. M. Martinis, Phys. Rev. B **84**, 024501 (2011).
- <sup>9</sup> L. Sun, L. DiCarlo, M. D. Reed, G. Catelani, L. S. Bishop, D. I. Schuster, B. R. Johnson, Ge A. Yang, L. Frunzio, L. I. Glazman, M. H. Devoret, and R. J. Schoelkopf, Phys. Rev. Lett. **108**, 230509 (2012).
- <sup>10</sup> G. Ithier, E. Collin, P. Joyez, P. J. Meeson, D. Vion, D. Esteve, F. Chiarello, A. Shnirman, Y. Makhlin, J. Schrieffer, and G. Schön, Phys. Rev. B **72**, 134519 (2005).
- <sup>11</sup> J. M. Martinis, S. Nam, J. Aumentado, K. M. Lang, and C. Urbina, Phys. Rev. B **67**, 094510 (2003).
- <sup>12</sup> A. P. Sears, A. Petrenko, G. Catelani, L. Sun, H. Paik, G. Kirchmair, L. Frunzio, L. I. Glazman, S. M. Girvin, and R. J. Schoelkopf, arXiv:1206.1265 (2012).
- <sup>13</sup> M. H. Devoret and J. M. Martinis, Quantum Inf. Process. **3**, 163 (2004).
- <sup>14</sup> We will reintroduce the channel index where necessary, e.g. in Sec. VI
- <sup>15</sup> H.-P. Breuer and F. Petruccione, *The theory of open quantum systems* (OUP, Oxford, 2002)
- <sup>16</sup> Here and in subsequent formulas we take the distribution function to be a smooth function of energy over an interval of order  $\delta E$  above the gap and to vanish quickly at higher energies.
- <sup>17</sup> A. Shnirman, G. Schön, I. Martin, and Y. Makhlin, in *Electron Correlation in New Materials and Nanosystems* (Springer Netherlands, 2007)
- <sup>18</sup> It was shown in Ref. 6 that the quasiparticle distribution function  $f$  should be interpreted as the energy mode  $f_E$ ; the latter is characterized by an effective temperature, in contrast with the charge mode disequilibrium whose presence requires a shift in the chemical potential, see, e.g., M. Tinkham, *Introduction to superconductivity*, 2nd ed., Ch. 11 (McGraw-Hill, New York, 1975).
- <sup>19</sup> R. C. Dynes, J. P. Garno, G. B. Hertel, and T. P. Orlando, Phys. Rev. Lett. **53**, 2437 (1984).
- <sup>20</sup> J. Leppäkangas and M. Marthaler, Phys. Rev. B **85**, 144503 (2012).
- <sup>21</sup> The main approximation used here is the harmonic one for the potential energy of the phase variable in Eq. (3), justified by the smallness of the fluctuations of the phase around a minimum. However, in Eq. (40) the dependance on  $(E_C/\omega_{10})^2$  is due to the cancellation of the leading terms in the difference  $s_{11} - s_{00}$  [cf. Eq. (11)]. Then a more careful analysis shows that corrections to the wavefunctions due to the cubic anharmonicity of the potential should be considered when evaluating  $A_s^d$ . Since the corresponding corrections are of order unity for typical phase qubit parameters, they do not affect our conclusions in the text after Eq. (42).
- <sup>22</sup> C. W. J. Beenakker, in *Transport Phenomena in Mesoscopic Systems*, edited by H. Fukuyama and T. Ando (Springer, Berlin, 1992).
- <sup>23</sup> This exponential factor coincides with the average occupation probability  $f(\Delta)$  at the gap edge; in Refs. 5,6 two contributions to the frequency, one coming from the Andreev states and the other from the bulk quasiparticles with energy at the gap edge, were combined together since for tunnel junctions  $x_{\text{qp}}^A = f(\Delta)$ . More generally, the two contributions must be considered separately. For example, for junctions with highly transparent channels we have  $x_{\text{qp}}^A = f(E_A) \neq f(\Delta)$ , and when considering fluctuations of the occupations their relative magnitudes scale differently for bound and free quasiparticles (as  $1/\sqrt{N}$  and inverse square root of the volume, respectively).
- <sup>24</sup> L. D. Landau and E. M. Lifshitz, *Statistical physics* (Course of Theoretical Physics, Vol. 5, Pergamon Press, 1980).
- <sup>25</sup> see, e.g., N. Agraït, A. Levy Yeyati, and J. M. van Ruitenbeek, Phys. Rep. **377**, 81 (2003), and references therein.
- <sup>26</sup> V. F. Maisi, O.-P. Saira, Yu. A. Pashkin, J. S. Tsai, D. V. Averin, and J. P. Pekola, Phys. Rev. Lett. **106**, 217003 (2011).
- <sup>27</sup> J. Teufel, *Superconducting Tunnel Junctions as Direct Detectors for Submillimeter Astronomy* (PhD thesis, Yale University, 2008).
- <sup>28</sup> S. B. Kaplan, C. C. Chi, D. N. Langenberg, J. J. Chang, S. Jafarey, and D. J. Scalapino, Phys. Rev. B **14**, 4854 (1976).
- <sup>29</sup> C. M. Wilson and D. E. Prober, Phys. Rev. B **69**, 094524 (2004).
- <sup>30</sup> O.-P. Saira, A. Kemppinen, V. F. Maisi, and J. P. Pekola, Phys. Rev. B **85**, 012504 (2012).
- <sup>31</sup> A. F. Volkov and Sh. M. Kogan, Sov. Phys. JETP **38**, 1018 (1974).
- <sup>32</sup> P. C. Martin and J. Schwinger, Phys. Rev. **115**, 1342 (1959).
- <sup>33</sup> More precisely, we here perform the first step of the Markov approximation – the substitution  $\rho_+(\tau) \rightarrow e^{-i\omega_{10}(t-\tau)}\rho_+(t)$ . The second step (the extension of the lower integration limit) will be taken at the end of the calculation.
- <sup>34</sup> We note that in the flux qubit the inductor is assumed to be lossless, while in the fluxonium the effective inductor formed by the junction array has finite losses.
- <sup>35</sup> V. E. Manucharyan, N. A. Masluk, A. Kamal, J. Koch, L. I. Glazman, and M. H. Devoret, Phys. Rev. B **85**, 024521 (2012).
- <sup>36</sup> S. Doniach and F. H. Sondheimer, *Green's functions for solid state physicists* (Imperial College Press, 1998).
- <sup>37</sup> 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).
- <sup>38</sup> A. D. Corcoles, J. M. Chow, J. M. Gambetta, C. Rigetti, J. R. Rozen, G. A. Keefe, M. B. Rothwell, M. B. Ketchen, and M. Steffen, Appl. Phys. Lett. **99**, 181906 (2011).