Qubit dephasing by spectrally diffusing quantum two-level systems

Shlomi Matityahu,[1,*] Alexander Shnirman,[1,2,†] and Moshe Schechter[3]

[1] *Institut für Theorie der Kondensierten Materie, Karlsruhe Institute of Technology, Karlsruhe 76131, Germany*
[2] *Institut für Quantenmaterialien und Technologien, Karlsruhe Institute of Technology, Karlsruhe 76021, Germany*
[3] *Department of Physics, Ben-Gurion University of the Negev, Beer Sheva 84105, Israel*

We investigate the pure dephasing of a Josephson qubit due to the spectral diffusion of two-level systems that are close to resonance with the qubit. We identify the parameter regime in which this pure dephasing rate can be of the order of the energy relaxation rate and, thus, the relation $T_2 = 2T_1$ is violated for the qubit. This regime is reached if the dynamics of the thermal TLSs responsible for the spectral diffusion is sufficiently slower than the energy relaxation of the qubit. By adding periodic bias modulating the qubit frequency or TLS excitation energies we show that this pure dephasing mechanism can be mitigated, allowing enhancement of superconducting qubits' coherence time. Mitigating pure dephasing, even if it is subdominant, is of special significance in view of recent suggestions for converting the dominant relaxation process (T_1) into erasure errors, leaving pure dephasing as the bottleneck for efficient quantum computation.

I. INTRODUCTION

Superconducting Josephson qubits have made an impressive impact on the field of quantum artificial systems in general and on quantum computing, in particular [1]. Improving the coherence and stability of these qubits is necessary to make further progress. This would allow increasing the computational depth, reducing the unavoidable errors and making more reliable quantum simulations. In this paper we investigate one of the main sources of dissipation in Josepshon quantum circuits—atomic defects with a two-level structure, that act as parasitic "hidden" qubits and couple to the logical qubits via the circuit charge or flux degrees of freedom. Such environmental exciatations are commonly referred to as two-level systems (TLSs) in the literature.

The TLSs with energy splitting of the order of that of the qubit are considered to be one of the major sources of qubit energy relaxation (T_1 process). Namely, the qubit can exchange an excitation with such a TLS and the TLS, in turn, dissipates its energy to a thermal bath of phonons. This mechanism has been investigated in great detail [2–5]. For clarity we will call such TLSs with energy splitting close to that of the qubit "quantum TLSs."

In addition, it is known that frequently randomly looking "T_1 fingerprint," i.e., the dependence of T_1 on the qubit frequency, fluctuates in time [3,6–8]. This can be explained by the spectral diffusion of the TLSs, i.e., the random diffusive motion of the TLS energy splitting [9]. This diffusion is explained by coupling of the quantum TLSs to an ensemble of thermal TLSs, i.e., TLSs with energy splittings smaller or of the order of the temperature. These thermal TLSs switch randomly between their two states, thus shifting the energy of the quantum TLSs [Fig. 1(a)].

Coupling of the qubit to the quantum TLSs causes not only energy relaxation but also frequency shift, known as the Lamb shift. The two are intimately related via the Kramers-Kronig relations. Spectral diffusion gives rise to temporal fluctuations of the Lamb shift and, thus, to the pure dephasing of the qubit. In this paper we investigate this process in detail and identify regimes in which the pure dephasing time is of the order of the relaxation time T_1. Following our previous paper [10], we also investigate the effect of qubit frequency modulation on its pure dephasing time.

We index the quantum TLSs by an integer n. Each quantum TLS is characterized by several parameters: ϵ_n is the energy splitting of the TLS, g_n is the coupling strength of the TLS to the qubit, and γ_n is the relaxation rate of the TLS due to phonons. The spectral diffusion adds yet another parameter $\mu_{\mathrm{av},n}$, which gives the characteristic span of the spectral diffusion in the long-time limit. In other words, this is the width of the energy interval "visited" by $\epsilon_n(t)$ over long times. In addition, the ensemble of the quantum TLSs is characterized by the typical level spacing δ, namely the energy distance, $\epsilon_n - \epsilon_{n+1}$, between spectrally neighboring TLSs [see Fig. 1(a)]. Taking into account the

―――――――
*Corresponding author: shlomimatit@gmail.com
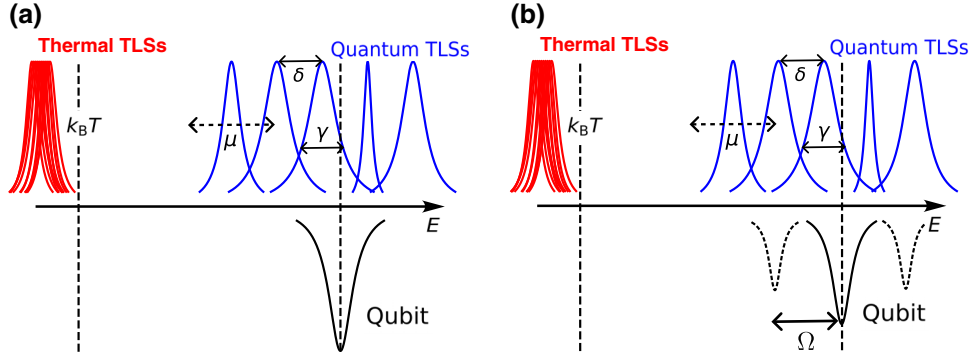†Corresponding author: alexander.shnirman@kit.edu

FIG. 1. Spectral schematics of the qubit, the quantum TLSs performing spectral diffusion, and the thermal TLSs responsible for it. Each quantum TLS is characterized by an absorption peak (blue) of width γ. The quantum TLSs perform spectral diffusion of span μ, which is caused by coupling to the thermal TLSs (red). The typical energy distance between adjacent quantum TLSs is δ. (a) Without periodic modulation; (b) with periodic modulation with frequency Ω. In this case the absorption peak of the qubit is split into multiple peaks at spectral distance Ω.

wide distribution of the coupling strengths g_n, it might be more prudent to consider separately the subsets of quantum TLSs with a specific coupling strength and introduce the g-dependent level spacing $\delta(g)$ as was done in Ref. [2].

Historically, the TLSs observed spectroscopically by Simmonds *et al.* in 2004 [11] were in the regime of very strong coupling $g \gg \gamma$ and spectrally rare $\delta \gg g$ (we drop the index n for brevity). The coupling strength g was observed as the avoided level crossing in qubit spectroscopy. The spectral diffusion did not play any significant role, and one could estimate $\mu_{av} \ll \delta$ and even $\mu_{av} \ll g$.

Similarly, spectrally rare ensembles of TLSs strongly interacting with the qubit were observed in later works [2,12]. In more recent studies [3,4,13], denser ensembles of quantum TLSs with weaker coupling to the qubit ($g < \gamma$) were observed. Such TLSs cause peaks in the relaxation rate of the qubit as a function of its energy splitting, denoted henceforth $T_{1,q}^{-1}$. The randomly looking dependence of the qubit relaxation rate on the qubit frequency suggested that $\delta \sim \gamma$, i.e., $T_{1,q}^{-1}$ was given by multiple partly overlapping Lorentzians. Moreover, following the time dependence of the positions of the strongest peaks [3], one could conclude that $\mu_{av} \gg \gamma$.

Here we propose a more differentiated approach. The experiments [3,4,13] show that the strongest peaks are still relatively rare, such that the level spacing of these peaks satisfies $\delta > \mu_{av} > \gamma$. These are relatively sparse TLSs that are strongly coupled to the qubit and at the same time are close to resonance with it. However, these peaks are not responsible for the background level of $T_{1,q}^{-1}$. We follow Ref. [2] and assume the existence of a much denser subset of quantum TLSs with an even weaker coupling to the qubit, a result of the abundance of TLSs with small tunneling amplitudes. These TLSs are so spectrally dense, such that the span of their spectral diffusion is larger than their level spacing and their inverse life time, i.e., $\mu_{av} > \gamma, \delta$. It is this subset that is mostly influenced by the spectral diffusion and can give the strongest contribution

to the pure dephasing of the qubit. In the other regimes, e.g., for $\mu_{av} < \gamma$, dephasing is suppressed as discussed in Appendix D.

To be concrete we assume, following Ref. [2], that the TLS density of states $dN/d\varepsilon$ satisfies

$$\frac{d^2N}{dg\,d\varepsilon} = \delta_{\mathrm{typ}}^{-1} \frac{\sqrt{1 - g^2/g_{\max}^2}}{g} \quad \text{for } g < g_{\max} \qquad (1)$$

and zero otherwise. Here δ_{typ} is a constant depending on the materials and the geometric shape of the qubit. It will play the role of the typical level spacing as we will see below. We conjecture this distribution to be valid down to very small values of g. Then it is guaranteed that we have a large number of TLSs such that $\mu_{av} > \gamma, \delta$ at high enough temperatures (note that $\mu_{av} \propto T$, see Ref. [14]). These are the TLSs that are of interest to us in this paper.

Analyzing the pure dephasing of the qubit due to the spectral diffusion, we will see that a major role is played by the typical relaxation time of the thermal TLSs, which we denote by $T_{1,T}$. This is the time scale at which the spectral diffusion of a quantum TLS saturates at the span of order μ_{av}. Our main result is the pure dephasing law $D(t)$, which reads

$$-2\ln D(t) \sim \frac{g_{\max}^4}{\delta_{\mathrm{typ}}}$$

$$\times \begin{cases} \dfrac{1}{\gamma}t^2, & t \ll \dfrac{\gamma}{\mu_{av}}T_{1,T} \\[2ex] \dfrac{T_{1,T}}{\mu_{av}}t\ln\left(\dfrac{\mu_{av}}{T_{1,T}\gamma}t\right), & T_{1,T} \gg t \gg \dfrac{\gamma}{\mu_{av}}T_{1,T}. \\[2ex] \dfrac{T_{1,T}}{\mu_{av}}t\ln\left(\dfrac{\mu_{av}}{\gamma}\right), & t \gg T_{1,T} \end{cases}$$

$$\qquad (2)$$

Here, γ and μ_{av} stand for typical values (see Appendix C). This result is valid at least as long as $-2\ln D(t) \lesssim 1$, thus

it can be safely used to estimate the effective dephasing rate (time). In Sec. IV we conclude that at low enough temperatures, when $T_{1,T}/T_{1,q} \gg \mu_{av}/\sqrt{\gamma \delta_{typ}}$, the effective pure dephasing time due to the spectral diffusion becomes comparable to $T_{1,q}$. This can explain numerous experimental observations, in which the qubit dephasing time $T_{2,q}$ is significantly shorter than its upper limit $2T_{1,q}$ (see, e.g., Refs. [6,7,15,16]). Below we reanalyze the dephasing law, Eq. (2), in the presence of periodic modulations of the qubit frequency and conclude that in this regime the modulations can strongly suppress the pure dephasing and, ultimately, restore the relation $T_{2,q} = 2T_{1,q}$.

II. THEORETICAL FRAMEWORK

A. The model

We consider the same Hamiltonian as in our previous paper [10],

$$
\mathcal{H} = -\frac{1}{2} \left[E_0 + A \cos(\Omega t) \right] \sigma_z - \frac{1}{2} \sum_n \varepsilon_n(t) \tau_z^{(n)}
$$
$$
+ \frac{1}{2} \sum_n g_n \sigma_x \tau_x^{(n)} - \frac{i}{2} \sum_n \gamma_n (1 - \tau_z^{(n)}), \quad (3)
$$

where σ_α and $\tau_\alpha^{(n)}$ ($\alpha = x, y, z$) are the Pauli matrices that describe the qubit and the nth quantum TLS, respectively. The qubit energy splitting is modulated harmonically around E_0 with amplitude A and frequency Ω. In our previous paper [10] we assumed that the energy splittings ε_n of the TLSs are constant (but randomly distributed). Here we consider the effect of spectral diffusion due to interaction with thermal TLSs, giving rise to fluctuations in ε_n. The third term is the interaction between the qubit and the TLSs, characterized by the coupling constants g_n. Finally, the dissipation of the TLSs due to phonons can be described by the last non-Hermitian term with relaxation rates γ_n. The reason this description is sufficient and the full Linbladian is not needed is explained in Appendix A where the microscopic derivation is provided.

In our previous paper [10] we have studied the effect of the periodic modulation with frequency Ω on the energy relaxation rate of the qubit ($T_{1,q}^{-1}$). The main effect was the replication of the absorption peaks in $T_{1,q}^{-1}(E_0)$, at energies shifted from E_0 by multiples of Ω [see Fig. 1(b)]. We have assumed the condition $\Omega > \gamma_n, g_n$, meaning that the replications of each absorption peak extend beyond its width and demonstrated that the modulation conserves the average of the relaxation rate over E_0 but the fluctuations in the relaxation rate can be suppressed. In this work we investigate the effect of the spectral diffusion on the pure dephasing of the qubit, both in the absence and in the presence of the periodic modulation in the regime $\Omega > \gamma_n, g_n$.

B. Qubit relaxation and decoherence

We consider the initial state $|\psi(t=0)\rangle = (\alpha |0\rangle + \beta |1\rangle) |0, \dots, 0\rangle$, in which the qubit is prepared in a superposition of the ground ($|0\rangle$) and excited ($|1\rangle$) states, whereas all the quantum TLSs are in their ground states. Within the rotating-wave approximation only the component containing the excited state of the qubit ($\propto \beta$) evolves in time in a nontrivial manner, whereas the component $\propto \alpha$ stays intact. As in our previous paper [10], we analyze the model Hamiltonian (3) using the Weisskopf-Wigner approach [17]. We treat the third term as a perturbation, use the rotating-wave approximation, and consider the single-excitation ansatz $|\psi_I(t)\rangle = \alpha |0\rangle |0, \dots, 0\rangle + \beta \left[a(t) |1\rangle |0, \dots, 0\rangle + \sum_n b_n(t) |0\rangle |0, \dots, 0, 1_n, 0, \dots, 0\rangle \right]$ for the state in the interaction picture, with the initial conditions $a(0) = 1$, $b_n(0) = 0 \, \forall n$. The reduced density matrix of the qubit reads

$$
\rho = \begin{pmatrix} 1 - |\beta|^2 |a(t)|^2 & \alpha \beta^* a(t) \\ \alpha^* \beta a(t)^* & |\beta|^2 |a(t)|^2 \end{pmatrix}. \quad (4)
$$

This expression should be averaged over the realizations of the spectral diffusion $\varepsilon_n(t)$, which will be denoted by $\langle \, \rangle$ (see Sec. II C below). Below we solve the Schrödinger equation to find the amplitude of staying in the excited state of the qubit, $a(t)$, and then consider its average $\langle a(t) \rangle$, as well the average $\langle |a(t)|^2 \rangle$. From Eq. (4) we see that $\langle |a(t)|^2 \rangle$ describes the relaxation, which we discussed in Ref. [10], whereas $M(t) \equiv |\langle a(t) \rangle|$ contains also the pure dephasing, which is the subject of this paper.

The Schrödinger (Weisskopf-Wigner) equations read [10]

$$
\dot{a}(t) = -\frac{i}{2} \sum_n g_n e^{i \left[E_0 t + \phi(t) - \int_0^t \varepsilon_n(t') dt' + i \gamma_n t \right]} b_n(t),
$$
$$
\dot{b}_n(t) = -\frac{i}{2} g_n e^{-i \left[E_0 t + \phi(t) - \int_0^t \varepsilon_n(t') dt' + i \gamma_n t \right]} a(t), \quad (5)
$$

where $\phi(t) = \int_0^t A \cos(\Omega t') dt' = (A/\Omega) \sin(\Omega t)$. We integrate the second equation, employ the well-known expansion $e^{i\phi(t)} = \sum_m J_m(A/\Omega) e^{im\Omega t}$ [$J_m(x)$ being the Bessel functions of the first kind], and substitute the result into the first equation. We obtain the following integrodifferential equation:

$$
\dot{a}(t) = -\frac{1}{4} \sum_n \sum_{m_1, m_2} g_n^2 J_{m_1} J_{m_2} e^{i(m_1 - m_2)\Omega t}
$$
$$
\times \int_0^t e^{i \left[(E_0 + m_2 \Omega)\tau - \int_{t-\tau}^t \varepsilon_n(t') dt' \right]} e^{-\gamma_n \tau} a(t - \tau) d\tau. \quad (6)
$$

Here and in what follows we drop the argument A/Ω of the Bessel function.

In the weak coupling limit $g_n \ll \gamma_n$, we can employ the Markov approximation and replace $a(t - \tau)$ by $a(t)$ in the integrand. We can also extend the limit of the integration to ∞. The condition $\Omega \gg \gamma_n, g_n$ also implies that $e^{i(m_1-m_2)\Omega t}$ oscillates rapidly and averages to zero unless $m_1 = m_2$. If, in addition, $\varepsilon_n(t)$ varies slowly on the time scale of order $1/\gamma_n$, we can approximate $\int_{t-\tau}^{t} \varepsilon_n(t')dt' \approx \varepsilon_n(t)\tau$. This is justified if

$$\int_{t-\tau}^{t} \left[\varepsilon_n(t') - \varepsilon_n(t) \right] dt' \ll 1 \qquad (7)$$

for $\tau \leq 1/\gamma$, i.e., if the maximum deviation in $\varepsilon_n(t)$ during the time interval of order $1/\gamma_n$ is smaller than γ_n. This condition [Eq. (7)] will be examined for self-consistency in Sec. IV. For the time being, we assume this condition holds and obtain the time-local differential equation $\dot{a}(t) = -C(t)a(t)$, where $C(t) = \sum_n C_n(t)$ and

$$C_n(t) = \frac{1}{4} \sum_{m=-\infty}^{\infty} \frac{g_n^2 J_m^2}{\gamma_n - i(E_0 + m\Omega - \varepsilon_n(t))}. \qquad (8)$$

With the initial condition $a(0) = 1$, the solution reads $a(t) = \prod_n a_n(t)$, where

$$a_n(t) = e^{-\int_0^t C_n(t')dt'}. \qquad (9)$$

Since the spectral diffusion fluctuations of $\varepsilon_n(t)$ are independent, the averaging of $a(t)$ as well as of $|a(t)|^2$ can be performed separately for each n, i.e.,

$$M(t) \equiv |\langle a(t) \rangle| = \prod_n |\langle a_n(t) \rangle|, \qquad (10)$$

and similarly $\langle |a(t)|^2 \rangle = \prod_n \langle |a_n(t)|^2 \rangle$.

C. Spectral diffusion

Various protocols measuring qubit decoherence are sensitive to noise at different frequencies. Such protocols measure the occupation probability of a specific quantum state (typically the ground state $|0\rangle$) as a function of time between the protocol pulses. Each experimental data point is obtained by averaging over N_{runs}, where each run involves a free evolution period t. Some protocols (e.g., Ramsey oscillations) are sensitive to low-frequency noise, i.e., to quantities that fluctuate between different runs (on timescale $N_{\mathrm{runs}}t$), and some protocols (e.g., various dynamical decoupling schemes, such as the Hahn spin echo) are sensitive only to noise at higher frequencies, i.e., to quantities that fluctuate within a single run (on timescale t).

Motivated by this observation, we define

$$\varepsilon_n(t) = \varepsilon_{n,0} + x_n + y_n(t), \qquad (11)$$

where $\varepsilon_{n,0}$ is the TLS energy splitting in the absence of interaction with thermal TLSs, x_n is the "quasistatic" shift

at time $t = 0$ due to interaction with thermal TLSs being in a specific configuration, and $y_n(t)$ is the dynamical shift due to flipping of thermal TLSs [note that $y_n(t = 0) = 0$]. The initial shift x_n changes from run to run of the experiment but does not vary during a single run. For a classical Markov process, the average $\langle \rangle$ over the realizations of the stochastic process $\varepsilon_n(t)$ is uniquely determined by the probability distribution function $P(x_n)$ for the initial state x_n of the process and the conditional probability distribution function $P(y_n + x_n; t|x_n; 0)$ for the state at time t given the initial state.

According to Klauder and Anderson [9], at short times where $|y_n(t)|$ is much smaller than the ultimate span of the diffusion (this is introduced below and is called $\mu_{\mathrm{av},n}$), the probability distribution $P(y_n + x_n; t|x_n; 0)$ depends only on y_n and has the form

$$P(y_n + x_n; t|x_n; 0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\tau y_n - t f_n(\tau)} d\tau. \qquad (12)$$

Here the function $f_n(\tau)$ characterizes the diffusion process. One gets Gaussian diffusion if $f_n(\tau) \propto \tau^2$ and Lorentzian diffusion if $f_n(\tau) \propto |\tau|$. We consider the more detailed model of Klauder and Anderson [9], in which fluctuations in ε_n are due to the coupling of quantum TLSs to thermal TLSs [14], which fluctuate between two states with the typical switching rate $\kappa = T_{1,T}^{-1}$. Within an ensemble of thermal TLSs the switching rate varies over a wide range [14]. Yet, for the observables of interest to us, taking κ to be the typical switching rate preserves their qualitative behavior, and is assumed henceforth. The coupling is assumed to be dipolar, decaying as $1/r^3$ (r is the distance between the quantum TLS and the thermal TLS), with two possible values $\pm v_n/r^3$ depending on the state of the thermal TLS. The function $f_n(\tau)$ in this case is found to be $\propto |\tau|$ for $|\tau| > \rho_n \equiv \pi r_{\mathrm{min}}^3/(2v_n)$ (r_{min} is the short-distance cutoff, such that v_n/r_{min}^3 is the largest possible coupling) and $\propto \tau^2$ for $|\tau| < \rho_n$. These two limiting cases can be interpolated by the function $f_n(\tau) = m_n \left(\sqrt{\tau^2 + \rho_n^2} - \rho_n \right)$, where $m_n = (2/3)\pi^2 v_n \mathcal{N}\kappa \approx (v_n/\bar{r}^3)\kappa$, with $\bar{r}$ being the average distance between the thermal TLSs (their density is $\mathcal{N} \approx 1/\bar{r}^3$). The functional form of $P(y_n + x_n; t|x_n; 0)$ as a function of y_n is Lorentzian at short times $t \ll 1/(m_n \rho_n)$. In this regime the essential decay of the exponential function inside the integral in Eq. (12) happens at $\tau \gg \rho_n$, so that $f_n(\tau) \approx m_n|\tau|$. At longer times, $t \gg 1/(m_n \rho_n)$, essential decay happens already at $\tau \ll \rho_n$, so that $f_n \propto \tau^2$, and the functional form of $P(y_n + x_n; t|x_n; 0)$ is Gaussian.

It should be emphasized that the distribution function Eq. (12) does not represent a stationary distribution since it does not approach a limiting distribution for $t \to \infty$. As discussed by Klauder and Anderson [9], the generalization of the distribution function to longer times has the form

$$P(y_n + x_n; t|x_n; 0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\tau \, e^{i[y_n + x_n(1 - e^{-R_n t})]\tau}$$
$$\times \, e^{-\int_0^t dt' f_n(\tau e^{-R_n t'})}. \quad (13)$$

The form of Eq. (13) guarantees that this probability distribution has a stationary form at times $t \gg 1/R_n$, corresponding to the timescale for which most of the thermal TLSs have flipped at least once. The distribution function in this limit depends only on the final position $x_n + y_n$ irrespective of the initial position. This stationary limit is used below to average over the "quasistatic" shift x_n between different runs of a dephasing protocol (e.g., Ramsey or spin echo), i.e., we define $P_\infty(x_n) = P(x_n; t \to \infty|x_n; 0)$, which can be written approximately as

$$P_\infty(x_n) \approx \frac{1}{2\pi} \int_{-\infty}^{\infty} d\tau \, e^{ix_n\tau} e^{-f_n(\tau)/R_n}. \quad (14)$$

We take $R_n \equiv R$ to be the typical switching rate of the thermal TLSs defined above $R = \kappa = T_{1,T}^{-1}$. We also define $\mu_{\mathrm{av},n} = v_n/\bar{r}^3$ and $\mu_{\mathrm{max},n} = v_n/r_{\mathrm{min}}^3$. Then m_n defined above is $\mu_{\mathrm{av},n}R$ and $\rho_n = 1/\mu_{\mathrm{max},n}$. We therefore obtain

$$P_\infty(x_n) \approx \frac{1}{2\pi} \int_{-\infty}^{\infty} d\tau \, e^{ix_n\tau} e^{-\mu_{\mathrm{av},n}\left(\sqrt{\tau^2 + \rho_n^2} - \rho_n\right)}. \quad (15)$$

Assuming $\mu_{\mathrm{av},n} \ll \mu_{\mathrm{max},n}$, the distribution function $P_\infty(x_n)$ has the Lorentzian form $P_\infty(x_n) = (\mu_{\mathrm{av},n}/\pi)/(\mu_{\mathrm{av},n}^2 + x_n^2)$ for $|x_n| < \mu_{\mathrm{max},n}$, and should decay faster than $1/x_n^3$ for $|x_n| > \mu_{\mathrm{max},n}$, such that the second moment $\langle x_n^2 \rangle$ is finite.

III. ANALYSIS

A. Expansion at short times

To further investigate Eq. (10), we limit ourselves to the short-time expansion of Eq. (9), up to the second order in C_n. This approximation may be justified by two quite different arguments:

(1) In many cases multiple resonant TLSs contribute simultaneously. Then, even if each individual $a_n(t)$ has only slightly deviated from 1 and can be expanded, the product $a(t) = \prod_n a_n(t)$ has already substantially decayed. This consideration is a variation of the central limit theorem in statistics.

(2) The usual definition of the dephasing time as the time at which $|\langle a(t) \rangle|$ (averaging over the realizations of the spectral diffusion) has decayed by a factor of $1/e$ is, nowadays, of only indicative significance in the field of quantum computing. Much more relevant is the time at which, e.g., $|\langle a(t) \rangle|$ reaches an error threshold allowed by the requirements of, e.g., quantum error correction. Thus, we might be more interested to calculate the time at which $|\langle a(t) \rangle| \sim 1 - \epsilon$ for some $\epsilon \ll 1$. To answer this question the short-time expansion is usually sufficient.

Thus we expand

$$a_n(t) = e^{-\int_0^t C_n(t')dt'} \approx 1 - \int_0^t C_n(t_1)dt_1$$
$$+ \frac{1}{2} \int_0^t \int_0^t C_n(t_1)C_n(t_2)dt_1 dt_2 + \dots \quad (16)$$

and perform the averaging described in Appendix B. For the dephasing of the off-diagonal elements of the density matrix we obtain

$$|\langle a(t) \rangle| = \sqrt{\langle |a(t)|^2 \rangle} \times D(t), \quad (17)$$

where

$$D(t) \equiv \exp\left[-\frac{1}{2}\sum_n \int_0^t \int_0^t \langle \delta C_n(t_1)\delta C_n^*(t_2)\rangle \, dt_1 dt_2\right] \quad (18)$$

is the pure dephasing decay factor. In the simplest case one would expect $\sqrt{\langle |a(t)|^2 \rangle} \sim \exp[-\Gamma_1 t/2]$, $|\langle a(t)\rangle| = \exp[-\Gamma_2 t]$, $D(t) \sim \exp[-\Gamma_\varphi t]$. Thus, Eq. (17) stands in clear correspondence to the well-known Bloch's relation $\Gamma_2 = \Gamma_1/2 + \Gamma_\varphi$. Of course, in reality, these decay laws are not necessarily exponential. The quantity $D(t)$ is the central subject of our study.

B. The correlators

Next we employ the spectral diffusion propagators, Eqs. (12) and (15). Using Eqs. (8) and (11), the first moment $\langle C_n(t) \rangle$ can be written as

$$\langle C_n(t) \rangle = \frac{g_n^2}{4} \sum_{m=-\infty}^{\infty} J_m^2 \int_{-\infty}^{\infty} dx_n P_\infty(x_n)$$
$$\times \int_{-\infty}^{\infty} \frac{P(y_n + x_n; t|x_n; 0)}{\gamma_n - i(\Delta_n - x_n - y_n + m\Omega)} dy_n, \quad (19)$$

where $\Delta_n \equiv E_0 - \varepsilon_{n,0}$. Here we have averaged over the spectral diffusion realizations x_n and y_n using the distribution functions $P_\infty(x_n)$ and $P(y_n + x_n; t|x_n; 0)$ discussed above. The correlator $\langle C_n(t_1)C_n^*(t_2)\rangle$ can be written similarly as

$$\langle C_n(t_1)C_n^*(t_2)\rangle$$
$$= \frac{g_n^4}{16} \sum_{m,m'=-\infty}^{\infty} J_m^2 J_{m'}^2 \int_{-\infty}^{\infty} dx_n P_\infty(x_n)$$
$$\times \int_{-\infty}^{\infty} \frac{P(y_{n2} + x_n; t_2|x_n; 0)}{\gamma_n + i(\Delta_n - x_n - y_{n2} + m\Omega)} dy_{n2}$$
$$\times \int_{-\infty}^{\infty} \frac{P(y_{n1} + y_{n2} + x_n; t_1|y_{n2} + x_n; t_2)}{\gamma_n - i(\Delta_n - x_n - y_{n1} - y_{n2} + m'\Omega)} dy_{n1}, \quad (20)$$

where we assumed $t_1 > t_2$.

C. Estimation of the pure dephasing factor $D(t)$

The estimation of $D(t)$, defined in Eq. (18), is performed in Appendix C separately for $t \ll 1/R \equiv T_{1,T}$ and for $t \gg T_{1,T}$. This is because the spectral diffusion undergoes a qualitative change at times of order $T_{1,T}$ as described in Sec. II C. At $t \ll T_{1,T}$ the distribution function $P(y_n + x_n; t|x_n; 0) \sim (W_n(t)/W_n^2(t) + y_n^2)$ is a Lorentzian depending on the relative deviation y_n only and the width $W_n(t)$ grows linearly in time (for details see Sec. II C and Appendix C). At $t \gg T_{1,T}$ the distribution function $P(y_n + x_n; t|x_n; 0)$ saturates at $P_\infty(y_n + x_n)$.

One finally ends up with

$$
-2 \ln D(t) \sim \frac{g_{\max}^4}{\delta_{\text{typ}}} \left(\sum_{m=-\infty}^{\infty} J_m^4 \right)
$$
$$
\times
\begin{cases}
\dfrac{1}{\gamma} t^2, & t \ll \dfrac{\gamma}{\mu_{\text{av}}} T_{1,T} \\[2ex]
\dfrac{T_{1,T}}{\mu_{\text{av}}} t \ln\left(\dfrac{\mu_{\text{av}}}{T_{1,T}\gamma} t \right), & T_{1,T} \gg t \gg \dfrac{\gamma}{\mu_{\text{av}}} T_{1,T} \\[2ex]
\dfrac{T_{1,T}}{\mu_{\text{av}}} t \ln\left(\dfrac{\mu_{\text{av}}}{\gamma} \right), & t \gg T_{1,T}.
\end{cases}
$$
$$(21)$$

Here, γ and μ_{av} stand for typical values (see Appendix C). The case $t \ll T_{1,T}$ is subdivided into the regime of quasistatic dephasing ($t \ll (\gamma/\mu_{\text{av}})T_{1,T}$), in which the random phase is dominated by the distribution of the initial spectral positions x_n, and the regime of dynamical dephasing, in which the spectral diffusion during the time interval t matters (see more details in Appendix C). The validity of this result is insured at least as long as $-2\ln[D(t)] \lesssim 1$ as discussed in Appendix C. Thus it can be safely used in order to estimate the effective dephasing rate.

IV. DISCUSSION

Here we analyze the pure dephasing law given by Eq. (21). We first start the discussion in the absence of modulations: $A = 0$, in which case $\sum_{m=-\infty}^{\infty} J_m^4 = 1$ as $J_m(0) = \delta_{m,0}$. First, we notice that the condition (7) assumed in the course of the derivation and required for the validity of the Markovian approximation is equivalent to $T_{1,T}\gamma(\gamma/\mu_{\text{av}}) > 1$. This follows from $W_n(t) \approx (\mu_{\text{av},n}/T_{1,T})t$ and the requirement that the maximum deviation in $\varepsilon_n(t)$ during the time interval of order $1/\gamma_n$ is smaller than γ_n.

To decide which contribution is the relevant one for the actual qubit dephasing, we evaluate the "quasistatic" dephasing at the crossover time

$$
\tilde{t} = \frac{\gamma}{\mu_{\text{av}}} T_{1,T}.
$$
$$(22)$$

If it is larger than some threshold ϵ, then the "quasistatic" contribution is the relevant one. Otherwise, the dephasing occurs only at longer times where the dynamic contribution is significant. From Eq. (21) we find

$$
-2 \ln D(\tilde{t}) \sim \frac{g_{\max}^4}{\delta_{\text{typ}}} \frac{\gamma \, T_{1,T}^2}{\mu_{\text{av}}^2} \sim \left(\Gamma_{1,q} T_{1,T} \right)^2 \frac{\gamma \delta_{\text{typ}}}{\mu_{\text{av}}^2}, \quad (23)
$$

where we have used Fermi's golden rule estimate for the qubit energy relaxation rate, $\Gamma_{1,q} \sim g_{\max}^2/\delta_{\text{typ}}$ (see, e.g., Ref. [10]). We are working in the regime $\gamma \ll \mu_{\text{av}}$. It is also natural to assume $\delta_{\text{typ}} \sim \gamma \ll \mu_{\text{av}}$ because in this case the randomly looking "T_1 fingerprint" [3,6,7] is reproduced. Then the number of contributing resonant TLSs $N \sim \mu_{\text{av}}/\delta_{\text{typ}} \log(\cdots)$ is large and the short-time expansion for each n is sufficient. Let us, first, consider the limit $-2 \ln D(\tilde{t}) = \left(\Gamma_{1,q} T_{1,T} \right)^2 \left(\gamma \delta_{\text{typ}}/\mu_{\text{av}}^2 \right) \gg 1$. That would necessarily (but not sufficiently) require $\Gamma_{1,q} T_{1,T} \gg 1$, i.e., the thermal TLSs relax slower than the qubit. Since $\Gamma_{1,q}$ is typically smaller than γ, the condition $\left(\Gamma_{1,q} T_{1,T} \right)^2 \left(\gamma \delta_{\text{typ}}/\mu_{\text{av}}^2 \right) \gg 1$ is more restrictive than the Markovian condition $T_{1,T}\gamma(\gamma/\mu_{\text{av}}) > 1$ discussed above. Thus, the validity of the Markovian condition (7) can be safely assumed. In this regime the substantial dephasing happens already during the period dominated by the initial (quasistatic) fluctuations, i.e., $-2 \ln D(t) \sim g_{\max}^4 t^2/(\gamma \delta_{\text{typ}}) \sim \left(\Gamma_{1,q} t \right)^2 (\delta_{\text{typ}}/\gamma)$. Thus, depending on the ratio $\delta_{\text{typ}}/\gamma$ the pure dephasing rate can become comparable or even somewhat larger than the relaxation rate $\Gamma_{1,q}$. The dephasing law Eq. (21) in this regime is sketched in Fig. 2, where we compare the behavior of the pure dephasing exponent $-\ln D(t)$ (thin solid black curve) with the relaxation exponent $\Gamma_{1,q}t/2$ (thick solid blue curve) for realistic parameters for modern superconducting qubits [3, 18,19]. The pure dephasing exponent is comparable to the relaxation exponent [they become equal at $t \approx (\gamma/\delta)\Gamma_{1,q}^{-1}$ as indicated by the dotted magenta vertical line] in the quasistatic regime $t < \tilde{t}$ where the dephasing process occurs ($\tilde{t}$ is indicated by a dash-dotted green vertical line).

In the other limit $\left(\Gamma_{1,q} T_{1,T} \right)^2 \left(\gamma \delta_{\text{typ}}/\mu_{\text{av}}^2 \right) \lesssim 1$ [but still $T_{1,T}\gamma(\gamma/\mu_{\text{av}}) > 1$], the dephasing happens at longer times, $t \gg \tilde{t}$. In this case, with logarithmic accuracy, the pure dephasing rate can be estimated to be approximately $(g_{\max}^4/\delta_{\text{typ}}) (T_{1,T}/\mu_{\text{av}}) \sim \Gamma_{1,q} \left(\Gamma_{1,q} T_{1,T} \right) \left(\delta_{\text{typ}}/\mu_{\text{av}} \right)$. Thus it is necessarily much smaller than $\Gamma_{1,q}$ and the pure dephasing due to the spectral diffusion is subdominant. Both limits could be relevant depending on the qubit relaxation rate and the relaxation rate of the thermal TLSs, the latter being strongly temperature dependent.

Our results are obtained for the limit where $\mu_{\text{av}} > \gamma$. In the opposite limit, i.e., for $\mu_{\text{av}} < \gamma$, pure dephasing is suppressed and dominated by the quasistatic contribution. In the ultimate regime of $\mu_{\max} < \gamma$ the effective dephasing rate is smaller by a factor of $\mu_{\max}/\gamma$ compared to its
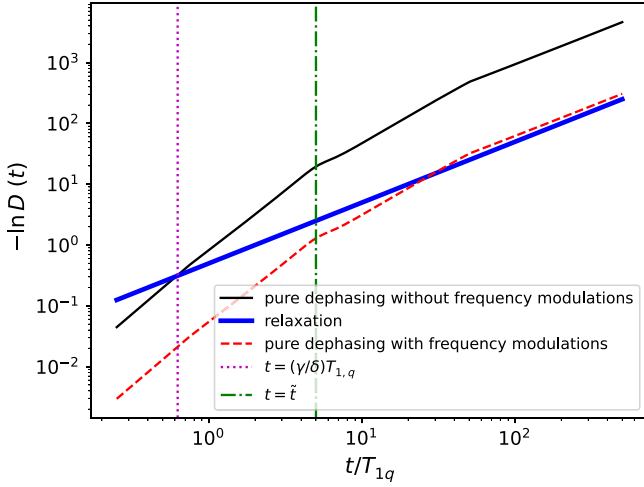
FIG. 2. The pure dephasing exponent $-\ln D(t)$ Eq. (21) due to spectral diffusion of quantum TLSs as a function of time (in units of the qubit relaxation time $T_{1,q} = \Gamma_{1,q}^{-1}$) in the regime $\gamma < \mu_{av}$ and $\left(\Gamma_{1,q}T_{1,T}\right)^2 \left(\gamma\delta_{typ}/\mu_{av}^2\right) \gg 1$ (the specific parameters are $\delta_{typ}/2\pi = 0.8$ MHz, $\gamma/2\pi = 0.5$ MHz, $\mu_{av}/2\pi = 5$ MHz, $T_{1,q} = \Gamma_{1,q}^{-1} = 20$ μs, and $T_{1,T} = 1$ ms [3,18,19]). In this regime the pure dephasing exponent without frequency modulations (thin solid black) is comparable in magnitude to the relaxation exponent $\Gamma_{1,q}t/2$ (thick solid blue) at short times $t < \tilde{t}$ [indicated by the dash-dotted green vertical line, see Eq. (22)], with crossing time $t \approx (\gamma/\delta)\Gamma_{1,q}^{-1}$ indicated by the dotted magenta vertical line. The red dashed curve is the pure dephasing exponent in the presence of frequency modulations with $A/\Omega = 10$, which suppresses the pure dephasing process and allows restoration of the relaxation-limited decoherence rate $\Gamma_2 = \Gamma_{1,q}/2$.

value in the limit of $\mu_{av} > \gamma$, as given in Eq. (21) (see Appendix D for derivation of pure dephasing in the limit of $\mu_{max} < \gamma$).

Our results reflect the dephasing process, in which maximal dephasing of the qubit is related to a shift of a resonant TLS by an energy which is of the order of its width γ. For $\mu_{av} < \gamma$ such processes are not allowed, and thus qubit dephasing is small. For $\mu_{av} > \gamma$ quasistatic dephasing dominates for $t < \tilde{t}$, which marks the border of the t^2 behavior of the dephasing. At longer times dynamical processes shift the TLS energies randomly, typically by more than their width γ. The dephasing process becomes a dynamical one and the effective dephasing rate is reduced as compared with the static regime [cf. Eq. (21)].

The dependence of the pure dephasing rate on μ_{av} and $T_{1,T}$ dictates its dependence on temperature. For very low temperatures $\mu_{av} < \gamma$, and pure dephasing is weak. Yet, once the temperature is large enough so that $\mu_{av} > \gamma$, we find an intriguing decrease in pure dephasing with increased temperature. This decrease is a consequence of the diminishing of the regime $t < \gamma T_{1,T}/\mu_{av}$ in which dephasing is quasistatic, as $T_{1,T}^{-1} \propto T^3$ (Refs. [14,20,21]), and $\mu_{av} \propto T$ (Ref. [14]). With increased temperature,

dynamical effects, which slow the growth of dephasing enter at earlier times, causing relative reduction in dephasing rate, which reaches a factor of $1/T^4$ at long times. As a result, we find a re-entrant temperature dependence of the pure dephasing rate, which peaks at the temperature where $\mu_{av} = \gamma$.

Let us now discuss the effect of qubit frequency modulations on the pure dephasing decay. We observe from Eq. (21) that the factor $\sum_{m=-\infty}^{\infty} J_m^4 \sim \Omega/A$ reduces effectively the dephasing rate for $\Omega \ll A$ [the fact that $\sum_{m=-\infty}^{\infty} J_m^4(A/\Omega) \sim \Omega/A$ results from the typical value $|J_m(x)| \sim \sqrt{\Omega/A}$ of the Bessel functions for $|m| < x$, see Ref. [10]]. From Eq. (21) one sees that the effective rate is reduced by a factor of $\sqrt{\Omega/A}$ in the case that the quasistatic contribution dominates pure dephasing, i.e., for $t \ll \tilde{t}$. For longer times, when the dynamical contribution dominates, the modulations reduce the pure dephasing rate by a factor of order Ω/A. In addition, the estimate of the dephasing at the crossover between the static and the dynamic regimes (23) is also modified and reads now

$$-2\ln D(\tilde{t}) \sim \frac{g_{max}^4}{\delta_{typ}} \left(\sum_{m=-\infty}^{\infty} J_m^4\right) \frac{\gamma T_{1,T}^2}{\mu_{av}^2}$$

$$\sim \left(\Gamma_{1,q}T_{1,T}\right)^2 \left(\sum_{m=-\infty}^{\infty} J_m^4\right) \frac{\gamma\delta_{typ}}{\mu_{av}^2}. \quad (24)$$

The quasistatic contribution at $t = \tilde{t}$ is reduced by the modulations, thus shrinking the parameter domain where its contribution dominates. On the other hand, in the dynamical regime, $t > \tilde{t}$, pure dephasing is subdominant even in the absence of modulations. The reduction of the pure dephasing exponent Eq. (21) due to frequency oscillations with $A/\Omega = 10$ is illustrated by the dashed red curve in Fig. 2.

We note that in contrast to our previous analysis of the effect of the modulations on the qubit relaxation rate [10], in which the modulations did not affect the average relaxation rate but only reduced its temporal fluctuations, here we obtain a reduction of the average pure dephasing rate. Such reduction of pure dephasing is certainly significant in the regime where pure dephasing rate is comparable of larger than relaxation-induced dephasing rate $\Gamma_{1,q}/2$. However, given recent results showing that relaxation-induced dephasing can be mitigated algorithmically [22,23], mitigation of pure dephasing, as shown here, is of utmost significance also in the regime where relaxation-induced dephasing dominates. Similar results of stabilization of relaxation time and reduction of pure dephasing rate were recently obtained by engineering the TLS noise spectral densities [24].

In this paper we considered the pure dephasing process defined by the Ramsey protocol. We have observed that periodic modulations can mitigate this dephasing.

In the regime where the pure dephasing dominates over relaxation-induced dephasing, it is also quasistatic. Thus, it can be mitigated by the echo technique as well. Remarkably, periodic modulations mitigate the nonstatic (exponential) dephasing as well, whereas the echo technique would be inefficient in this regime. Although the nonstatic dephasing is always subdominant, its suppression is very significant in view of the ideas of Refs. [22,23].

V. CONCLUSIONS

In this paper we have analyzed the pure dephasing of a qubit due to the spectral diffusion of a quantum TLS, which in the first place are responsible for the energy relaxation time of the qubit, $T_{1,q}$. We have identified a regime, in which the contribution of the spectral diffusion to the pure dephasing rate is comparable to $T_{1,q}^{-1}$. This leads to an appreciable deviation from the relation $T_{2,q} = 2T_{1,q}$, as is indeed the case in numerous experiments (see, e.g., Refs. [6,7,15,16]).

This result should be compared to the alternative mechanisms that could violate the relation $T_{2,q} = 2T_{1,q}$. We first mention that the direct coupling of the thermal fluctuators to the transmon qubit does not contribute to the pure dephasing since transmon qubits are exponentially decoupled from the low-frequency charge noise [25]. On the other hand, the fluctuations of the critical current [26,27] or the noise of the photon occupation number in the nearby readout resonator [28] could be relevant. The latter mechanism can be suppressed in devices with tunable coupling [29].

To distinguish our pure dephasing mechanism from that caused by the photon shot noise, we note that the latter is characterized by more or less white noise spectrum. Thus the pure dephasing due to the photon shot noise is Markovian and the time decay is strictly exponential. In contrast, in our case the dominant part of the time decay is that given by the quasistatic noise, thus the time decay takes roughly the form $e^{-\Gamma_\varphi^2 t^2}$.

On the other hand, the pure dephasing due to the critical current noise is caused by thermal TLSs, thus giving a similar quasistatic time decay. Distinguishing the pure dephasing resulting from critical current fluctuations from the mechanism discussed in this paper would be possible via a response to an external electric field. The effect of such field on the TLSs causing the critical current fluctuations would be significantly weaker, since these are located in the Josephson junction itself [4].

We further analyze the effect of the spectral diffusion in the presence of periodic modulations of the qubit frequency and conclude that in this regime the modulations can strongly suppress the pure dephasing and, ultimately, restore the relation $T_{2,q} = 2T_{1,q}$, therefore enhancing the performance of superconducting qubits.

Recently it was suggested that relaxation-induced dephasing can be overcome in a certain type of error-correction scheme (erasure errors) using "dual rail" qubits [22,23]. This leaves pure dephasing as the bottleneck for quantum computation even in the regime where it is subdominant. Thus, mitigation of pure dephasing by periodic modulations as suggested here can be of considerable value also when pure dephasing is weak.

Analysis of periodic bias modulation in this work considers the application of the periodic biasing to the qubit's frequency. In the same way, the periodic bias field can be applied to modulate the excitation energies of the TLSs [30–33]. In this case detailed analysis of the dependence of the modulation on individual TLS parameters has to be performed. Yet, we expect the results presented in this paper to be applicable, up to quantitative corrections, to the case where the periodic bias field is applied to the TLSs.

The phenomenon of pure dephasing considered in this paper is related to the problem of frequency noise in resonators [34–36]. In both systems, thermally excited TLSs interact with resonant TLSs, and the energy shifts of the latter cause, via the Lamb shift, fluctuations in the resonance energy of the resonator, or in the excitation energy of the qubit. Theory for frequency noise in resonators requires the application of spectral diffusion theory [9,14] to this specific problem. This was done directly by Burin *et al.* [36], and via separation of the different contributions of strongly and weakly interacting thermal TLSs in the work of Faoro and Ioffe [35] (see Ref. [37]). Here we follow an approach similar to that of Burin *et al.* [36], directly applying the theory of Klauder and Anderson [9] and treating all thermal TLSs on the same footing. Yet, in similarity to Faoro and Ioffe [35], we are able to point to the source of strong pure dephasing being the spectral diffusion over energy shifts of the resonant TLSs of order of their resonance width, which corresponds to "strong interactions" in Ref. [35]. Re-entrant temperature dependence, as is found for the frequency noise in resonators [35,36], is also found here for the pure dephasing. This effect is stronger here, as temperature dictates not only the width of spectral diffusion, but also the timescale for thermal TLS relaxation, which is dominant in dictating pure dephasing.

APPENDIX A: DERIVATION OF THE NON-HERMITIAN HAMILTONIAN

The full Hamiltonian of our system reads $\mathcal{H}(t) = \mathcal{H}_{\text{qubit-TLS}}(t) + \mathcal{H}_{\text{TLS-phonons}} + \mathcal{H}_{\text{phonons}}$, where $\mathcal{H}_{\text{qubit-TLS}}(t)$

is given by Eq. (1) of the main text and

$$\mathcal{H}_{\text{TLS-phonons}} = \frac{1}{2}\sum_{n,k}\tau_x^{(n)}(v_{nk}a_k + v_{nk}^*a_k^\dagger),$$

$$\mathcal{H}_{\text{phonons}} = \sum_k \omega_k a_k^\dagger a_k. \tag{A1}$$

Here $a_k^\dagger$ and a_k are the phonon creation and annihilation operators, ω_k are the phonon frequencies, and v_{nk} describes the coupling between the nth TLS and the phonon mode with wave vector k. We now consider the interaction picture with respect to the perturbation $V = (1/2)\sum_n g_n\sigma_x\tau_x^{(n)} + (1/2)\sum_{n,k}v_{nk}\tau_x^{(n)}(a_k + a_k^\dagger)$ and employ the rotating-wave approximation. Within the single-excitation subspace spanned by the states $|1\rangle\,|\{0\}_{\text{TLS}}\rangle\,|\{0\}_{\text{phonons}}\rangle$, $|0\rangle\,|\{1_n\}_{\text{TLS}}\rangle\,|\{0\}_{\text{phonons}}\rangle$ and $|0\rangle\,|\{0\}_{\text{TLS}}\rangle\,|\{1_k\}_{\text{phonons}}\rangle$, in which either the qubit, a single TLS, or a single phonon is excited (we use the notation $|\{1_n\}_{\text{TLS}}\rangle = |0,\ldots,0,1_n,0,\ldots,0\rangle$ for the state in which the nth TLS is excited and all other TLSs are at their ground states and $|\{1_k\}_{\text{phonons}}\rangle = |0,\ldots,0,1_k,0,\ldots,0\rangle$ for the state with a single phonon occupying the mode with wave vector k). Consider the general state $|\psi_I(t)\rangle = a(t)\,|1\rangle\,|\{0\}_{\text{TLS}}\rangle\,|\{0\}_{\text{phonons}}\rangle + \sum_n b_n(t)\,|0\rangle\,|\{1_n\}_{\text{TLS}}\rangle\,|\{0\}_{\text{phonons}}\rangle + \sum_k c_k(t)\,|0\rangle\,|\{0\}_{\text{TLS}}\rangle\,|\{1_k\}_{\text{phonons}}\rangle$. Note that the qubit-reduced density matrix Eq. (4) is obtained by tracing over the TLS and phonon states, and using the normalization condition $|a(t)|^2 + \sum_n|b_n(t)|^2 + \sum_k|c_k(t)|^2 = 1$. The Schrödinger equation for $|\psi_I(t)\rangle$ reads

$$\dot{a}(t) = -\frac{i}{2}\sum_n g_n e^{i(E_0t+\phi(t)-\varepsilon_nt)}b_n(t),$$

$$\dot{b}_n(t) = -\frac{i}{2}\Bigg[g_n e^{-i(E_0t+\phi(t)-\varepsilon_nt)}a(t)$$

$$+ \sum_k v_{nk}e^{i(\varepsilon_n-\omega_k)t}c_k(t)\Bigg],$$

$$\dot{c}_k(t) = -\frac{i}{2}\sum_l v_{lk}^*e^{-i(\varepsilon_l-\omega_k)t}b_l(t), \tag{A2}$$

with initial conditions $a(0) = 1$, $b_n(0) = c_k(0) = 0\ \forall\, n, k$. Integrating the last equation and substituting into the second, we obtain

$$\dot{a}(t) = -\frac{i}{2}\sum_n g_n e^{i(E_0t+\phi(t)-\varepsilon_nt)}b_n(t),$$

$$\dot{b}_n(t) = -\frac{i}{2}g_n e^{-i(E_0t+\phi(t)-\varepsilon_nt)}a(t)$$

$$- \frac{1}{4}\sum_l e^{i(\varepsilon_n-\varepsilon_l)t}\int_0^t K_{nl}(t-t')b_l(t')\,dt', \tag{A3}$$

where $K_{nl}(\tau) = \sum_k v_{nk}v_{lk}^*e^{i(\varepsilon_l-\omega_k)\tau}$. We now apply the Weisskopf-Wigner (Markov) approximation [17]. This is justified because the functions $K_{nl}(\tau)$ decay rapidly on the characteristic time scale $1/\gamma_l$ (to be defined below) describing the temporal variation of $b_l(t)$. One then replaces $b_l(t')$ by $b_l(t)$ in the integral and extends the upper limit of the integration to ∞ to obtain the following equations:

$$\dot{a}(t) = -\frac{i}{2}\sum_n g_n e^{i(E_0t+\phi(t)-\varepsilon_nt)}b_n(t),$$

$$\dot{b}_n(t) = -\frac{i}{2}g_n e^{-i(E_0t+\phi(t)-\varepsilon_nt)}a(t) - \frac{1}{4}\sum_l C_{nl}e^{i(\varepsilon_n-\varepsilon_l)t}b_l(t), \tag{A4}$$

with $C_{nl} = \int_0^\infty K_{nl}(\tau)d\tau$. The nondiagonal terms ($n \neq l$) couple different TLSs. These could become relevant if two or more TLSs are close to degeneracy, $|\varepsilon_n - \varepsilon_l| \ll \gamma_l, \gamma_n$. Then, cooperative super-radiant and subradiant dynamics of this group of TLSs could emerge [38,39]. In our case, this possibility can be neglected, as the average level spacing Δ between the TLSs is much larger than the typical width γ_{typ}. Even if such a degeneracy happens, but the two TLSs are not spatially close to each other, the phases in $K_{nl}(\tau) = \sum_k v_{nk}v_{lk}^*e^{i(\varepsilon_l-\omega_k)\tau}$ will average this coupling to zero.

We therefore retain only diagonal terms, which gives

$$\dot{a}(t) = -\frac{i}{2}\sum_n g_n e^{i(E_0t+\phi(t)-\varepsilon_nt)}b_n(t),$$

$$\dot{b}_n(t) = -\frac{i}{2}g_n e^{-i(E_0t+\phi(t)-\varepsilon_nt)}a(t) - \frac{1}{4}C_{nn}b_n(t). \tag{A5}$$

Finally, we define $\tilde{b}_n(t) = e^{C_{nn}t/4}b_n(t)$ and obtain

$$\dot{a}(t) = -\frac{i}{2}\sum_n g_n e^{i[E_0t+\phi(t)-(\varepsilon_n-iC_{nn}/4)t]}\tilde{b}_n(t),$$

$$\dot{\tilde{b}}_n(t) = -\frac{i}{2}g_n e^{-i[E_0t+\phi(t)-(\varepsilon_n-iC_{nn}/4)t]}a(t). \tag{A6}$$

By absorbing the imaginary part of C_{nn} (Lamb shift) into the definition of ε_n, and defining $\gamma_n \equiv \text{Re}(C_{nn})/4$, one ends up with Eqs. (4) of the main text [with $\tilde{b}_n(t)$ replaced by $b_n(t)$]. By integrating out the phonons, one thus finds an exact description of the system in terms of the non-Hermitian Hamiltonian (2) of the main text.

APPENDIX B: PURE DEPHASING AND RELAXATION IN TERMS OF THE CORRELATION FUNCTIONS OF $C_n(t)$

Expanding Eq. (9) up to second order in C_n and averaging, we obtain

$$\langle a_n(t)\rangle = \langle e^{-\int_0^t C_n(t')dt'}\rangle \approx 1 - \int_0^t \langle C_n(t_1)\rangle \, dt_1 + \frac{1}{2}\int_0^t\int_0^t \langle C_n(t_1)C_n(t_2)\rangle \, dt_1 dt_2. \tag{B1}$$

Denoting $C_n(t) = C'_n(t) + iC''_n(t)$ and taking the absolute value, we obtain up to second order in $C_n(t)$

$$|\langle a_n(t)\rangle| \approx 1 - \int_0^t \langle C'_n(t_1)\rangle \, dt_1 + \frac{1}{2}\int_0^t\int_0^t \left(\langle C'_n(t_1)C'_n(t_2)\rangle - \langle C''_n(t_1)C''_n(t_2)\rangle\right) dt_1 dt_2 + \frac{1}{2}\left(\int_0^t \langle C''_n(t_1)\rangle \, dt_1\right)^2. \tag{B2}$$

Similarly, up to the second order in $C_n(t)$

$$\langle |a_n(t)|^2\rangle = \langle e^{-2\int_0^t C'_n(t')dt'}\rangle \approx 1 - 2\int_0^t \langle C'_n(t_1)\rangle \, dt_1 + 2\int_0^t\int_0^t \langle C'_n(t_1)C'_n(t_2)\rangle \, dt_1 dt_2, \tag{B3}$$

and therefore

$$\sqrt{\langle |a_n(t)|^2\rangle} \approx 1 - \int_0^t \langle C'_n(t_1)\rangle \, dt_1 + \int_0^t\int_0^t \langle C'_n(t_1)C'_n(t_2)\rangle \, dt_1 dt_2 - \frac{1}{2}\left(\int_0^t \langle C'_n(t_1)\rangle \, dt_1\right)^2. \tag{B4}$$

As we are interested in the product $\prod_n a_n(t)$, we calculate the logarithms of Eqs. (B2) and (B4):

$$\ln|\langle a_n(t)\rangle| \approx -\int_0^t \langle C'_n(t_1)\rangle \, dt_1 + \frac{1}{2}\int_0^t\int_0^t \left(\langle C'_n(t_1)C'_n(t_2)\rangle - \langle C''_n(t_1)C''_n(t_2)\rangle\right) dt_1 dt_2$$

$$+ \frac{1}{2}\left(\int_0^t \langle C''_n(t_1)\rangle \, dt_1\right)^2 - \frac{1}{2}\left(\int_0^t \langle C'_n(t_1)\rangle \, dt_1\right)^2$$

$$\ln\sqrt{\langle |a_n(t)|^2\rangle} \approx -\int_0^t \langle C'_n(t_1)\rangle \, dt_1 + \int_0^t\int_0^t \langle C'_n(t_1)C'_n(t_2)\rangle \, dt_1 dt_2 - \left(\int_0^t \langle C'_n(t_1)\rangle \, dt_1\right)^2. \tag{B5}$$

Up to the second order in $C_n(t)$, we can thus write

$$\ln|\langle a_n(t)\rangle| \approx \ln\sqrt{\langle |a_n(t)|^2\rangle} - \frac{1}{2}\int_0^t\int_0^t \left(\langle C'_n(t_1)C'_n(t_2)\rangle + \langle C''_n(t_1)C''_n(t_2)\rangle\right) dt_1 dt_2$$

$$+ \frac{1}{2}\left(\int_0^t \langle C'_n(t_1)\rangle \, dt_1\right)^2 + \frac{1}{2}\left(\int_0^t \langle C''_n(t_1)\rangle \, dt_1\right)^2, \tag{B6}$$

or equivalently,

$$\ln\sqrt{\langle |a_n(t)|^2\rangle} \approx -\int_0^t \langle C'_n(t_1)\rangle \, dt_1 + \int_0^t\int_0^t \langle \delta C'_n(t_1)\delta C'_n(t_2)\rangle \, dt_1 dt_2, \tag{B7}$$

$$\ln|\langle a_n(t)\rangle| \approx \ln\sqrt{\langle |a_n(t)|^2\rangle} - \frac{1}{2}\int_0^t\int_0^t \langle \delta C_n(t_1)\delta C_n^*(t_2)\rangle \, dt_1 dt_2, \tag{B8}$$

where $\delta C_n(t) \equiv C_n(t) - \langle C_n(t)\rangle$ and $\delta C'_n(t) \equiv C'_n(t) - \langle C'_n(t)\rangle$.

We are now in a position to perform the product over n and exponentiate. This gives for the energy relaxation (T_1):

$$\sqrt{\langle |a(t)|^2\rangle} \approx \exp\left[-\sum_n \int_0^t \langle C'_n(t')\rangle \, dt' + \sum_n \int_0^t\int_0^t \langle \delta C'_n(t_1)\delta C'_n(t_2)\rangle \, dt_1 dt_2\right]. \tag{B9}$$

The quite unexpected second term in the exponent reflects the effect of the fluctuations. Exponentiating Eq. (B7) we obtain Eqs. (17) and (18) of the main text.

APPENDIX C: ESTIMATION OF $D(t)$

1. The short-time limit $t \ll T_{1,T}$

At short times $t \ll T_{1,T}$, the relevant distribution function is given by Eq. (12). In this regime, the condition $t \ll (m_n \rho_n)^{-1} = (\mu_{\max,n}/\mu_{av,n})T_{1,T}$ holds, and the diffusion is Lorentzian [as argued following Eq. (12)] with distribution function

$$P(y_n + x_n; t|x_n; 0) = \frac{1}{\pi} \frac{W_n(t)}{W_n^2(t) + y_n^2}, \tag{C1}$$

which is homogeneous, i.e.,

$$P(y_n + y_n' + x_n; t|y_n' + x_n; t') = \frac{1}{\pi} \frac{W_n(t - t')}{W_n^2(t - t') + y_n^2}. \tag{C2}$$

Here the width $W_n(t)$ is given by [14]

$$W_n(t) \propto \mu_{av,n} \int_0^\infty \frac{dy}{\cosh^2 y} \int_0^1 \frac{1 - e^{-x^2 t/T_1(y)}}{x} dx, \tag{C3}$$

where $y = E/(2T)$ and $x = \Delta_0/E$. Here Δ_0 and E are the tunneling amplitude and energy splitting of the fluctuating thermal TLSs, respectively. These are broadly distributed according to the standard tunneling model [20,21]. In Eq. (C3) $T_1(y) \propto T_{1,\min}(T)y^3 \coth y$, with $T_{1,\min}(T) \propto T^3$ being the relaxation time of the thermal TLSs with the largest value of Δ_0, and thus with the fastest decay rate to the phonon bath. According to Black and Halperin [14], the function $W_n(t)$ is linear in time for $t \ll T_{1,\min}(T)$ and approaches $\mu_{av,n}$ logarithmically at times $t \gg T_{1,\min}(T)$. In what follows we treat thermal TLSs as having a single relaxation time $T_{1,T} \equiv 1/R \sim T_{1,\min}(T)$.

Performing the integrations we obtain

$$\langle C_n(t) \rangle = \frac{g_n^2}{4} \sum_{m=-\infty}^{\infty} J_m^2 \int_{-\infty}^{\infty} \frac{P_\infty(x_n)}{\gamma_n + W_n(t) - i(\Delta_n - x_n + m\Omega)} dx_n,$$

$$\langle C_n(t_1)C_n^*(t_2) \rangle = \frac{g_n^4}{16} \sum_{m,m'=-\infty}^{\infty} J_m^2 J_{m'}^2 \frac{2\gamma_n + 2W_n(t_2) + W_n(t_1 - t_2) + i(m - m')\Omega}{2\gamma_n + W_n(t_1 - t_2) + i(m - m')\Omega}$$

$$\times \int_{-\infty}^{\infty} \frac{P_\infty(x_n)}{[\gamma_n + W_n(t_2) + i(\Delta_n - x_n + m\Omega)][\gamma_n + W_n(t_2) + W_n(t_1 - t_2) - i(\Delta_n - x_n + m'\Omega)]} dx_n. \tag{C4}$$

We assume $\mu_{\max,n} \gg \mu_{av,n}$, such that $P_\infty(x_n)$ can be taken as a Lorentzian function of width $\mu_{av,n}$ up to $|x_n| = \mu_{\max,n}$. Since $\mu_{av,n} \gg W_n(t)$ for $t \ll T_{1,T}$, in the regime $\mu_{\max,n} \gg \gamma_n$ the non-Lorentzian tails of $P_\infty(x_n)$ at $|x_n| > \mu_{\max,n}$ can be neglected, and we obtain

$$\langle C_n(t) \rangle \approx \frac{g_n^2}{4} \sum_{m=-\infty}^{\infty} \frac{J_m^2}{\mu_{av,n} + \gamma_n + W_n(t) - i(\Delta_n + m\Omega)},$$

$$\langle C_n(t_1)C_n^*(t_2) \rangle \approx \frac{g_n^4}{16} \sum_{m,m'=-\infty}^{\infty} J_m^2 J_{m'}^2 \frac{2\mu_{av,n} + 2\gamma_n + 2W_n(t_2) + W(t_1 - t_2) + i(m - m')\Omega}{2\gamma_n + W_n(t_1 - t_2) + i(m - m')\Omega}$$

$$\times \frac{1}{[\mu_{av,n} + \gamma_n + W_n(t_2) + i(\Delta_n + m\Omega)][\mu_{av,n} + \gamma_n + W_n(t_2) + W_n(t_1 - t_2) - i(\Delta_n + m'\Omega)]}. \tag{C5}$$

We note that if $\mu_{av,n}$ and $\mu_{\max,n}$ are comparable, then $P_\infty(x_n)$ is not a Lorentzian function, but the results are qualitatively similar. We next assume $\Omega \gg \gamma_n, W_n(t)$ and therefore only terms $m = m'$ contribute to the sum in the second equation of

Eqs. (C4). The results (C5) hold for an arbitrary relation between $\mu_{\text{av},n}$ and γ_n. From now on we use the assumed above regime $\mu_{\text{av},n} \gg \gamma_n, W_n$ (the other regimes are discussed in Appendix D). In this regime we otain

$$\langle C_n(t) \rangle \approx \frac{g_n^2}{4} \sum_{m=-\infty}^{\infty} \frac{J_m^2}{\mu_{\text{av},n} - i(\Delta_n + m\Omega)},$$

$$\langle C_n(t_1) C_n^*(t_2) \rangle \approx \frac{g_n^4 \mu_{\text{av},n}}{8\left(2\gamma_n + W_n(t_1 - t_2)\right)} \sum_{m=-\infty}^{\infty} \frac{J_m^4}{\mu_{\text{av},n}^2 + (\Delta_n + m\Omega)^2}. \tag{C6}$$

Thus, near the resonance ($\Delta_n + m\Omega \approx 0$) we can estimate $|\langle C_n(t) \rangle|^2 \propto 1/\mu_{\text{av},n}^2 \ll \langle C_n(t_1) C_n^*(t_2) \rangle$. Therefore, we obtain $\langle \delta C_n(t_1) \delta C_n^*(t_2) \rangle \approx \langle C_n(t_1) C_n^*(t_2) \rangle$. For $t \ll T_{1,T}$, the diffusion width $W_n(t)$ is a linear function of time, $W_n(t) \approx (\mu_{\text{av},n}/T_{1,T})t \equiv \alpha_n t$. Therefore, using the relation

$$\int_0^t \int_0^t k(t_1 - t_2) dt_1 dt_2 = \int_0^{2t} (2t - \tau) k(\tau) d\tau, \tag{C7}$$

we obtain

$$\int_0^t \int_0^t \langle \delta C_n(t_1) \delta C_n^*(t_2) \rangle \, dt_1 dt_2 \approx \frac{g_n^4 \mu_{\text{av},n}}{8} \sum_{m=-\infty}^{\infty} \frac{J_m^4}{\mu_{\text{av},n}^2 + (\Delta_n + m\Omega)^2} \int_0^{2t} \frac{2t - \tau}{2\gamma_n + \alpha_n \tau} d\tau$$

$$= \frac{g_n^4 \mu_{\text{av},n}}{4} \left[-\frac{t}{\alpha_n} + \frac{\alpha_n t + \gamma_n}{\alpha_n^2} \ln\left(1 + \frac{\alpha_n t}{\gamma_n}\right) \right] \sum_{m=-\infty}^{\infty} \frac{J_m^4}{\mu_{\text{av},n}^2 + (\Delta_n + m\Omega)^2}. \tag{C8}$$

At times $t \ll \gamma_n/\alpha_n = (\gamma_n/\mu_{\text{av},n})T_{1,T}$, we obtain

$$\int_0^t \int_0^t \langle \delta C_n(t_1) \delta C_n^*(t_2) \rangle \, dt_1 dt_2 \approx \frac{g_n^4 \mu_{\text{av},n}}{8\gamma_n} \sum_{m=-\infty}^{\infty} \frac{J_m^4}{\mu_{\text{av},n}^2 + (\Delta_n + m\Omega)^2} \, t^2, \tag{C9}$$

which corresponds to the "quasistatic" limit. On the other hand, at longer times $(\gamma_n/\mu_{\text{av},n})T_{1,T} \ll t \ll T_{1,T}$, we have

$$\int_0^t \int_0^t \langle \delta C_n(t_1) \delta C_n^*(t_2) \rangle \, dt_1 dt_2 \approx \frac{g_n^4}{4} T_{1,T} \sum_{m=-\infty}^{\infty} \frac{J_m^4}{\mu_{\text{av},n}^2 + (\Delta_n + m\Omega)^2} \, t \ln\left(\frac{\mu_{\text{av},n}}{T_{1,T}\gamma_n} t\right). \tag{C10}$$

To make a rough estimate, we replace all $\mu_{\text{av},n}$ by a typical value μ_{av} and similarly for γ_n. Then, for each value of m, we sum the contributions of the TLSs that are resonant at this m, i.e., such that $\Delta_n + m\Omega \lesssim \mu_{\text{av}}$. This gives us $d\epsilon = \mu_{\text{av}}$ in the distribution (1). Using this distribution we estimate

$$\sum_n \frac{g_n^4}{\mu_{\text{av}}^2 + (\Delta_n + m\Omega)^2} \sim \frac{d\epsilon}{\delta_{\text{typ}}\mu_{\text{av}}^2} \int_0^{g_{\text{max}}} g^3 dg \sim \frac{g_{\text{max}}^4}{\delta_{\text{typ}}\mu_{\text{av}}}. \tag{C11}$$

Since our theory holds only for the TLSs weakly coupled to the qubit, i.e., $g_n < \gamma_n$, it would be reasonable to put $g_{\text{max}} \sim \gamma$. We, thus, arrive at

$$-2\ln[D(t)] = \sum_n \int_0^t \int_0^t \langle \delta C_n(t_1) \delta C_n^*(t_2) \rangle \, dt_1 dt_2 \sim \frac{g_{\text{max}}^4}{\delta_{\text{typ}}} \left(\sum_{m=-\infty}^{\infty} J_m^4 \right) \begin{cases} \dfrac{1}{\gamma} t^2, & t \ll \dfrac{\gamma}{\mu_{\text{av}}} T_{1,T} \\[2ex] \dfrac{T_{1,T}}{\mu_{\text{av}}} t \ln\left(\dfrac{\mu_{\text{av}}}{T_{1,T}\gamma} t\right), & T_{1,T} \gg t \gg \dfrac{\gamma}{\mu_{\text{av}}} T_{1,T}. \end{cases} \tag{C12}$$

One might wonder if this result is valid at all relevant times in view of the short-time expansions (B1) and (B3) used in the calculation. Since the number of contributing TLSs is always large, approximately $\mu_{\text{av}}/\delta_{\text{typ}}$, the dephasing law (C12) is valid at least as long as $-2\ln[D(t)] \lesssim 1$. Thus we can safely use (C12) in order to estimate the effective dephasing rate (time). At much longer times (C12) might be inaccurate.

2. The long-time limit $t \gg T_{1,T}$

Next we discuss the long time limit, $t \gg 1/R \equiv T_{1,T}$. We approximate

$$\int_0^t \int_0^t \langle \delta C_n(t_1) \delta C_n^*(t_2) \rangle \, dt_1 dt_2 \approx \int_{T_{1,T}}^t \int_{T_{1,T}}^t \langle \delta C_n(t_1) \delta C_n^*(t_2) \rangle \, dt_1 dt_2, \tag{C13}$$

and note that $P(y_n + x_n; t|x_n; 0) \to P_\infty(x_n + y_n)$ for $t \gg T_{1,T}$ and similarly $P(y_{n1} + y_{n2} + x_n; t_1|y_{n2} + x_n; t_2) \to P_\infty(x_n + y_{n1} + y_{n2})$ for $t_1, t_2 \gg T_{1,T}$, except for the region $t_1 - t_2 < T_{1,T}$, where Eq. (C2) holds. Then for $t_1, t_2 \gg T_{1,T}$, we use

$$\int_{-\infty}^\infty \frac{P_\infty(x_n)}{\gamma_n - i(\Delta_n - x_n + m\Omega)} dx_n = \frac{1}{\gamma_n + \mu_{\text{av},n} - i(\Delta_n + m\Omega)}, \tag{C14}$$

to obtain

$$\langle C_n(t) \rangle = \frac{g_n^2}{4} \sum_{m=-\infty}^\infty \frac{J_m^2}{\gamma_n + \mu_{\text{av},n} - i(\Delta_n + m\Omega)}. \tag{C15}$$

The second moment is

$$\langle C_n(t_1) C_n^*(t_2) \rangle = \frac{g_n^4}{16} \sum_{m,m'=-\infty}^\infty \frac{J_m^2 J_{m'}^2}{\left[\gamma_n + \mu_{\text{av},n} + i(\Delta_n + m\Omega) \right] \left[\gamma_n + \mu_{\text{av},n} - i(\Delta_n + m'\Omega) \right]} \tag{C16}$$

for $t_1 - t_2 \gtrsim T_{1,T}$ and

$$\langle C_n(t_1) C_n^*(t_2) \rangle = \frac{g_n^4}{16} \sum_{m,m'=-\infty}^\infty \frac{J_m^2 J_{m'}^2}{\left[\gamma_n + \mu_{\text{av},n} + i(\Delta_n + m\Omega) \right] \left[\gamma_n + W_n(t_1 - t_2) - i(\Delta_n + m'\Omega) \right]} \tag{C17}$$

for $t_1 - t_2 \ll T_{1,T}$. The contribution to the second integral in the right-hand side of Eq. (C13) vanishes for $t_1 - t_2 \gtrsim T_{1,T}$, whereas for $t_1 - t_2 \ll T_{1,T}$ the integral is the same as that calculated in Eq. (C8).

One finally ends up with

$$-2 \ln D(t) \sim \frac{g_{\text{max}}^4}{\delta_{\text{typ}}} \left(\sum_{m=-\infty}^\infty J_m^4 \right) \frac{T_{1,T}}{\mu_{\text{av}}} t \ln \left(\frac{\mu_{\text{av}}}{\gamma} \right), \quad t \gg T_{1,T}. \tag{C18}$$

The validity of this result is again insured at least as long as $-2 \ln[D(t)] \lesssim 1$ as discussed right after Eq. (C12). Thus it can be safely used in order to estimate the effective dephasing rate. Combining Eqs. (C12) and (C18) we obtain Eq. (21) of the main text.

APPENDIX D: THE LIMIT $\mu_{\text{av}} \ll \gamma$

As the average span of the diffusion $\mu_{\text{av},n}$ becomes smaller than the width γ_n, the results for the pure dephasing change. Analyzing Eq. (C5) we observe that the inequality $| \langle C_n(t) \rangle |^2 \ll \langle C_n(t_1) C_n^*(t_2) \rangle$ does not hold anymore. Thus, the fluctuations $\langle \delta C_n(t_1) \delta C_n^*(t_2) \rangle$ are suppressed and the dephasing is reduced. In the intermediate regime $\mu_{\text{av},n} < \gamma_n < \mu_{\text{max},n}$ the analysis is complicated and we concentrate on the ultimate case $\mu_{\text{max},n} < \gamma_n$.

In the limit $\mu_{\text{max},n} \ll \gamma_n$ the center of the Lorentzian functions appearing in Eqs. (19) and (20) does not fluctuate significantly on the energy scale compared to the width γ_n of the Lorentzian. The contributions of the dynamic shifts is thus small and to obtain the contribution of the "quasistatic" dephasing we can take the conditional probabilities to be δ functions of the dynamic shift. We thus end up with

$$\langle C_n(t)\rangle \approx \frac{g_n^2}{4} \sum_{m=-\infty}^{\infty} J_m^2 \int_{-\infty}^{\infty} \frac{P_\infty(x_n)}{\gamma_n - i(\Delta_n - x_n + m\Omega)} dx_n,$$

$$\langle C_n(t_1)C_n^*(t_2)\rangle \approx \frac{g_n^4}{16} \sum_{m,m'=-\infty}^{\infty} J_m^2 J_{m'}^2 \int_{-\infty}^{\infty} \frac{P_\infty(x_n)}{[\gamma_n + i(\Delta_n - x_n + m\Omega)][\gamma_n - i(\Delta_n - x_n + m'\Omega)]} dx_n. \tag{D1}$$

We next expand in the small quantity x_n/γ_n to second order and keep only terms with $m = m'$. This gives

$$\langle C_n(t)\rangle \approx \frac{g_n^2}{4} \sum_{m=-\infty}^{\infty} \frac{J_m^2}{\gamma_n - i(\Delta_n + m\Omega)} \left[1 - \frac{\mu_{\max,n}^2}{[\gamma_n - i(\Delta_n + m\Omega)]^2}\right],$$

$$\langle C_n(t_1)C_n^*(t_2)\rangle \approx \frac{g_n^4}{16} \sum_{m=-\infty}^{\infty} \frac{J_m^4}{\gamma_n^2 + (\Delta_n + m\Omega)^2} \left[1 - \mu_{\max,n}^2 \frac{\gamma_n^2 - 3(\Delta_n + m\Omega)^2}{[\gamma_n^2 + (\Delta_n + m\Omega)^2]^2}\right], \tag{D2}$$

where we have used the fact that $\langle x_n\rangle = 0$ [since $P_\infty(x_n)$ is a symmetric function] and $\langle x_n^2\rangle \approx \mu_{\max,n}^2$. Hence to leading order in $\mu_{\max,n}/\gamma_n$ we obtain

$$\langle \delta C_n(t_1)\delta C_n^*(t_2)\rangle \approx \frac{g_n^4}{16} \sum_{m=-\infty}^{\infty} \frac{J_m^4 \mu_{\max,n}^2}{[\gamma_n^2 + (\Delta_n + m\Omega)^2]^2}. \tag{D3}$$

Assuming $\delta < \gamma$ and taking into account the resonant contributions we arrive at

$$-2\ln D(t) \sim (\Gamma_{1,q}t)^2 \frac{\mu_{\max}^2}{\gamma^2} \frac{\delta}{\gamma} \sum_m J_m^4. \tag{D4}$$

The effective dephasing rate, even in the absence of the modulation, is always smaller than the relaxation rate $\Gamma_{1,q}$ and is further reduced by the modulations.

[1] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, Superconducting qubits: Current state of play, Annu. Rev. Condens. Matter Phys. **11**, 369 (2020).

[2] J. M. Martinis, K. B. Cooper, R. McDermott, M. Steffen, M. Ansmann, K. D. Osborn, K. Cicak, S. Oh, D. P. Pappas, R. W. Simmonds, and C. C. Yu, Decoherence in Josephson qubits from dielectric loss, Phys. Rev. Lett. **95**, 210503 (2005).

[3] P. V. Klimov, *et al.*, Fluctuations of energy-relaxation times in superconducting qubits, Phys. Rev. Lett. **121**, 090502 (2018).

[4] J. Lisenfeld, A. Bilmes, A. Megrant, R. Barends, J. Kelly, P. Klimov, G. Weiss, J. M. Martinis, and A. V. Ustinov, Electric field spectroscopy of material defects in transmon qubits, npj Quantum Inf. **5**, 105 (2019).

[5] L. V. Abdurakhimov, I. Mahboob, H. Toida, K. Kakuyanagi, Y. Matsuzaki, and S. Saito, Identification of different types of high-frequency defects in superconducting qubits, PRX Quantum **3**, 040332 (2022).

[6] S. Schlör, J. Lisenfeld, C. Müller, A. Bilmes, A. Schneider, D. P. Pappas, A. V. Ustinov, and M. Weides, Correlating decoherence in transmon qubits: Low frequency noise by single fluctuators, Phys. Rev. Lett. **123**, 190502 (2019).

[7] J. J. Burnett, A. Bengtsson, M. Scigliuzzo, D. Niepce, M. Kudra, P. Delsing, and J. Bylander, Decoherence benchmarking of superconducting qubits, npj Quantum Inf. **5**, 54 (2019).

[8] M. Carroll, S. Rosenblatt, P. Jurcevic, I. Lauer, and A. Kandala, Dynamics of superconducting qubit relaxation times, npj Quantum Inf. **8**, 132 (2022).

[9] J. R. Klauder and P. W. Anderson, Spectral diffusion decay in spin resonance experiments, Phys. Rev. **125**, 912 (1962).

[10] S. Matityahu, A. Shnirman, and M. Schechter, Stabilization of qubit relaxation rates by frequency modulation, Phys. Rev. Appl. **16**, 044036 (2021).

[11] R. W. Simmonds, K. M. Lang, D. A. Hite, S. Nam, D. P. Pappas, and J. M. Martinis, Decoherence in Josephson phase qubits from junction resonators, Phys. Rev. Lett. **93**, 077003 (2004).

[12] G. J. Grabovskij, T. Peichl, J. Lisenfeld, G. Weiss, and A. V. Ustinov, Strain tuning of individual atomic tunneling systems detected by a superconducting qubit, Science **338**, 232 (2012).

[13] R. Barends, J. Kelly, A. Megrant, D. Sank, E. Jeffrey, Y. Chen, Y. Yin, B. Chiaro, J. Mutus, C. Neill, P. O'Malley, P. Roushan, J. Wenner, T. C. White, A. N. Cleland, and J.

M. Martinis, Coherent Josephson qubit suitable for scalable quantum integrated circuits, Phys. Rev. Lett. **111**, 080502 (2013).

[14] J. L. Black and B. I. Halperin, Spectral diffusion, phonon echoes, and saturation recovery in glasses at low temperatures, Phys. Rev. B **16**, 2879 (1977).

[15] 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, Observation of high coherence in Josephson junction qubits measured in a three-dimensional circuit QED architecture, Phys. Rev. Lett. **107**, 240501 (2011).

[16] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Error mitigation extends the computational reach of a noisy quantum processor, Nature **567**, 491 (2019).

[17] V. Weisskopf and E. Wigner, Berechnung der natürlichen linienbreite auf grund der diracschen lichttheorie, Z. Phys. **63**, 54 (1930).

[18] J. Lisenfeld, A. Bilmes, S. Matityahu, S. Zanker, M. Marthaler, M. Schechter, G. Schön, A. Shnirman, G. Weiss, and A. V. Ustinov, Decoherence spectroscopy with individual two-level tunneling defects, Sci. Rep. **6**, 23786 (2016).

[19] S. Matityahu, A. Shnirman, G. Schön, and M. Schechter, Decoherence of a quantum two-level system by spectral diffusion, Phys. Rev. B **93**, 134208 (2016).

[20] P. W. Anderson, B. I. Halperin, and C. M. Varma, Anomalous low-temperature thermal properties of glasses and spin glasses, Philos. Mag.: A J. Theor. Exp. Appl. Phys. **25**, 1 (1972).

[21] W. A. Phillips, Tunneling states in amorphous solids, J. Low Temp. Phys. **7**, 351 (1972).

[22] A. Kubica, A. Haim, Y. Vaknin, H. Levine, F. Brandão, and A. Retzker, Erasure qubits: Overcoming the T_1 limit in superconducting circuits, Phys. Rev. X **13**, 041022 (2023).

[23] H. Levine, A. Haim, J. S. C. Hung, N. Alidoust, M. Kalaee, L. DeLorenzo, E. A. Wollack, P. A. Arriola, A. Khalajhedayati, R. Sanil, Y. Vaknin, A. Kubica, A. A. Clerk, D. Hover, F. Brandão, A. Retzker, and O. Painter, Demonstrating a long-coherence dual-rail erasure qubit using tunable transmons, arXiv:2307.08737.

[24] X. You, Z. Huang, U. Alyanak, A. Romanenko, A. Grassellino, and S. Zhu, Stabilizing and improving qubit coherence by engineering the noise spectrum of two-level systems, Phys. Rev. Appl. **18**, 044026 (2022).

[25] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Charge-insensitive qubit design derived from the Cooper pair box, Phys. Rev. A **76**, 042319 (2007).

[26] D. J. Van Harlingen, T. L. Robertson, B. L. T. Plourde, P. A. Reichardt, T. A. Crane, and J. Clarke, Decoherence in Josephson-junction qubits due to critical-current fluctuations, Phys. Rev. B **70**, 064517 (2004).

[27] C. D. Nugroho, V. Orlyanchik, and D. J. Van Harlingen, Low frequency resistance and critical current fluctuations in Al-based Josephson junctions, Appl. Phys. Lett. **102**, 142602 (2013).

[28] J. Gambetta, A. Blais, D. I. Schuster, A. Wallraff, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Qubit-photon interactions in a cavity: Measurement-induced dephasing and number splitting, Phys. Rev. A **74**, 042318 (2006).

[29] G. Zhang, Y. Liu, J. J. Raftery, and A. A. Houck, Suppression of photon shot noise dephasing in a tunable coupling superconducting qubit, npj Quantum Inf. **3**, 1 (2017).

[30] M. S. Khalil, S. Gladchenko, M. J. A. Stoutimore, F. C. Wellstood, A. L. Burin, and K. D. Osborn, Landau-Zener population control and dipole measurement of a two-level-system bath, Phys. Rev. B **90**, 100201 (2014).

[31] Y. J. Rosen, M. S. Khalil, A. L. Burin, and K. D. Osborn, Random-defect laser: Manipulating lossy two-level systems to produce a circuit with coherent gain, Phys. Rev. Lett. **116**, 163601 (2016).

[32] S. Matityahu, H. Schmidt, A. Bilmes, A. Shnirman, G. Weiss, A. V. Ustinov, M. Schechter, and J. Lisenfeld, Dynamical decoupling of quantum two-level systems by coherent multiple Landau–Zener transitions, npj Quantum Inf. **5**, 114 (2019).

[33] L. Yu, S. Matityahu, Y. J. Rosen, C.-C. Hung, A. Maksymov, A. L. Burin, M. Schechter, and K. D. Osborn, Experimentally revealing anomalously large dipoles in the dielectric of a quantum circuit, Sci. Rep. **12**, 16960 (2022).

[34] J. Burnett, L. Faoro, I. Wisby, V. L. Gurtovoi, A. V. Chernykh, G. M. Mikhailov, V. A. Tulin, R. Shaikhaidarov, V. Antonov, P. J. Meeson, A. Y. Tzalenchuk, and T. Lindström, Evidence for interacting two-level systems from the 1/f noise of a superconducting resonator, Nat. Commun. **5**, 4119 (2014).

[35] L. Faoro and L. B. Ioffe, Interacting tunneling model for two-level systems in amorphous materials and its predictions for their dephasing and noise in superconducting microresonators, Phys. Rev. B **91**, 014201 (2015).

[36] A. L. Burin, S. Matityahu, and M. Schechter, Low-temperature $1/f$ noise in microwave dielectric constant of amorphous dielectrics in Josephson qubits, Phys. Rev. B **92**, 174201 (2015).

[37] The two approaches are in principle equivalent, the requirement in Ref. [35] for an energy-dependent DOS is a consequence of their neglect of the logarithmic expansion in time of the spectral diffusion width and its temperature dependence.

[38] R. H. Dicke, Coherence in spontaneous radiation processes, Phys. Rev. **93**, 99 (1954).

[39] K. Hepp and E. H. Lieb, On the superradiant phase transition for molecules in a quantized radiation field: The Dicke maser model, Ann. Phys. (N. Y) **76**, 360 (1973).