Perturbative Diagonalization for Time-Dependent Strong Interactions

Z. Xiao,[1,*] E. Doucet,[1] T. Noh,[1,2] L. Ranzani,[3] R.W. Simmonds,[4] L.C.G. Govia,[3] and A. Kamal[1,†]

[1]*Department of Physics and Applied Physics, University of Massachusetts, Lowell, Massachusetts 01854, USA*

[2]*Associate of the National Institute of Standards and Technology, Boulder, Colorado 80305, USA*

[3]*Quantum Engineering and Computing, Raytheon BBN Technologies, Cambridge, Massachusetts 02138, USA*

[4]*National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305, USA*

We present a systematic method to implement a perturbative Hamiltonian diagonalization based on the time-dependent Schrieffer-Wolff transformation. Applying our method to strong parametric interactions we show how, even in the dispersive regime, full Rabi model physics is essential to describe the dressed spectrum. Our results unveil several qualitatively different results, including realization of large energy-level shifts, tunable in magnitude and sign with the frequency and amplitude of the pump mediating the parametric interaction. Crucially, Bloch-Siegert shifts, typically thought to be important only in the ultrastrong or deep-strong coupling regimes, can be rendered large even for weak dispersive interactions to realize points of exact cancelation of dressed shifts ("blind spots") at specific pump frequencies. The framework developed here highlights the rich physics accessible with time-dependent interactions and serves to significantly expand the functionalities for control and readout of strongly interacting quantum systems.

I. INTRODUCTION

Time-dependent interactions provide a powerful paradigm for the study of out-of-equilibrium quantum matter. Though originally motivated by the interest in the exploration of new phases in many-body quantum systems, recent developments in this domain have largely been triggered by studies of low-dimensional platforms, such as trapped ions and cavity and circuit QED, where a precise control of quantum dynamics is essential for realizing high-fidelity quantum information processing. However, the theoretical framework to describe the dynamics in the presence of such time-dependent interactions remains rudimentary, especially in the strong-coupling regime relevant to most applications.

One approach to studying strongly coupled systems is to derive a low-energy effective Hamiltonian by using a unitary transformation that decouples the high-frequency ("fast") subspace from the low-energy ("slow") subspace. A standard and widely used method to implement this is the Schrieffer-Wolff transformation (SWT) [1] that develops the diagonalized Hamiltonian as a perturbation series, whose radius of convergence is dictated by the ratio of interaction strength (g) and the gap between the low- and

high-energy subspaces (Δ). In contrast to the conventional Dyson series expansion, the SWT method gives direct access to the effective Hamiltonian at each order in the perturbation series; inferring the effective Hamiltonian from a time-ordered matrix exponential describing the propagator is typically nontrivial, especially because truncating the Dyson series does not preserve unitarity, and is shown to fail for degenerate ground states [2].

In this work, we present a generalization of the SWT for time-dependent strong interactions. While this exercise has been attempted before in a few sporadic examples [3–6], our approach clarifies how the inertial term generated due to the time-dependent generator of the SWT (equivalent to an extra dynamical rotation of the eigenbasis) can be accounted for at successive orders in a perturbation series; this is especially crucial since its inclusion magnifies the corrections arising from counter-rotating terms that are typically neglected in the dispersive regime ($g/\Delta \ll 1$).

We emphasize that the approach we describe is very general, and can be applied to a wide family of time-dependent Hamiltonians. Here we demonstrate it for the two archetypal examples of strong light-matter interaction: the Rabi model and a Kerr oscillator coupled to a linear oscillator by a parametrically modulated interaction. Such periodic driving is of interest in a myriad of applications, such as the realization of synthetic gauge fields

_____

*zhihao_xiao@uml.edu

†archana_kamal@uml.edu

in optical lattices [7], simulation of topological [8,9] and dynamically localized phases [10,11], implementation of fast entangling gates [12], tunable qubit readout [13], and state stabilization [14,15] and transfer [16,17].

The paper is organized as follows. We begin with a general description of the diagonalization procedure employing a series of sequential SWT generators and present the equation of motion for constructing the time-dependent SWT generator at a given order in perturbation in Sec. II. Then, in Sec. III, we apply this method to the time-dependent Rabi Hamiltonian and explicitly derive the condition for validity of the dispersive approximation in the presence of a parametrically mediated interaction in this system. We also derive analytical expressions for leading-order dispersive shifts, and comment on their unique features distinct from the "usual" (but time-dependent) Jaynes-Cummings (JC) case. We then discuss a generalization of the parametric QED system to include multilevel effects by considering the case of a Kerr oscillator (also known as a transmon) parametrically coupled to a linear oscillator and identify a "parametric straddling" regime by studying the induced frequency shifts as a function of the pump frequency in Sec. IV. Section V is devoted to the discussion of induced dissipation in the systems studied in Secs. III and IV. Finally, we conclude with a summary of results and present potential directions along which the current analysis can be extended in Sec. VI. Appendices A–E include additional mathematical details and proofs.

II. TIME-DEPENDENT SCHRIEFFER-WOLFF TRANSFORMATION

For a system described by a time-dependent Hamiltonian $H(t) = H_0 + \lambda V(t)$, an effective diagonal Hamiltonian at a given order M can be obtained by performing a sequence of $M - 1$ successive time-dependent SWTs. To this end, the aim is to thus construct a unitary operator $\hat{U}(t)$ perturbatively, as a series of successive time-dependent rotations

$$\hat{U}(t) = \prod_{j=1}^{M-1} \exp[\hat{S}^{(j)}(t)], \quad (1)$$

where $\hat{S}^{(n)}(t) \sim \mathcal{O}(\lambda^n)$ are the anti-Hermitian SWT generators. Under the action of $U(t)$, the lab frame Hamiltonian $\hat{H}(t) \equiv \hat{H}^{(1)}(t)$ is transformed into $\hat{H}^{(M)}(t)$,

$$\hat{H}^{(M)}(t) = \hat{H}_0 + \hat{V}^{(M)}(t). \quad (2)$$

Then, the effective diagonalized ("low-energy") Hamiltonian to order λ^M is given as

$$\hat{H}_{\text{eff}}^{(M)}(t) \equiv \hat{H}_0 + \mathcal{P}_0 \bullet \hat{V}^{(M)}(t) = \hat{H}_0 + \hat{V}_{\text{D}}^{(M)}(t), \quad (3)$$

where

$$\mathcal{P}_0 \bullet \hat{\Theta} = \sum_k |\psi_k\rangle\langle\psi_k|\hat{\Theta}|\psi_k\rangle\langle\psi_k| \quad (4)$$

is the pinching channel in the eigenbasis $\{|\psi_k\rangle\}$ of $\hat{H}_0$, which defines the "low-energy" subspace for the purpose of diagonalization.

The off-diagonal part of the interaction Hamiltonian contains terms of $\mathcal{O}(\lambda^M)$ or higher, i.e.,

$$\hat{V}_{\text{OD}}^{(M)}(t) \equiv \mathcal{Q}_0 \bullet \hat{V}^{(M)}(t) = \sum_{m=M}^{\infty} \lambda^m \hat{V}_{\text{OD},m}^{(M)}(t), \quad (5)$$

where

$$\mathcal{Q}_0 \bullet \hat{\Theta} = \sum_{j \neq k} |\psi_j\rangle\langle\psi_j|\hat{\Theta}|\psi_k\rangle\langle\psi_k| \quad (6)$$

projects onto the off-diagonal elements in the eigenbasis of $\hat{H}_0$. Such a construction is achievable by finding the operators $\hat{S}^{(j)}(t)$ using the following theorem.

Theorem 1. *At each order M in the perturbation series, if there exists $\hat{S}^{(M)}(t)$ ($M \geq 1$) satisfying the differential equation*

$$i\frac{\partial \hat{S}^{(M)}(t)}{\partial t} + [\hat{S}^{(M)}(t), \hat{H}_0] + \lambda^M \hat{V}_{\text{OD},M}^{(M)}(t) = 0 \quad (7)$$

then $\hat{U}_M(t) = \exp[-\hat{S}^{(M)}(t)]$ can be used to eliminate off-diagonal terms in $\hat{H}^{(M)}(t)$ up to $\mathcal{O}(\lambda^M)$, such that

$$\hat{H}^{(M+1)}(t) = \hat{H}_{\text{eff}}^{(M)}(t) + \mathcal{O}(\lambda^{M+1}), \quad (8)$$

where $\hat{H}_{\text{eff}}^{(M)}(t) \equiv \hat{H}_0 + \mathcal{P}_0 \bullet \hat{V}^{(M)}(t) = \hat{H}_0 + \hat{V}_{\text{D}}^{(M)}(t)$. Here $\hat{V}_{\text{OD},M}^{(M)}(t)$ denotes the leading off-diagonal term at $\mathcal{O}(\lambda^M)$ in the perturbation series, and includes the terms generated by the action of lower-order time-dependent transformations on $V(t)$. Note that, with $\hat{H}_{\text{eff}}^{(M)}(t)$ being diagonal, the off-diagonal part and the remaining diagonal part of $\hat{H}^{(M+1)}(t)$ are both $\mathcal{O}(\lambda^{M+1})$. In this way, the Mth order of the off-diagonal part of $\hat{H}^{(M)}(t)$ is eliminated as it is transformed into $\hat{H}^{(M+1)}(t)$, with

$$\hat{H}^{(M+1)}(t) = \hat{U}_M^\dagger(t)\hat{H}^{(M)}(t)\hat{U}_M(t) - i\hat{U}_M^\dagger(t)\frac{\partial \hat{U}_M(t)}{\partial t}. \quad (9)$$

See Appendix A for a detailed proof and a specific example.

It is worthwhile to note that typical methods of perturbative diagonalization develop corrections at different orders by expanding a single generator as a series in λ, i.e., $\hat{S}^{(M)}(t) = \exp[\sum_{n=1}^{M-1} \lambda^n G^{(n)}(t)]$, which leads to a
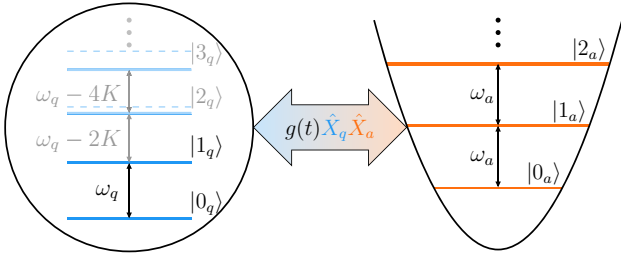
FIG. 1. Schematic representation of a bipartite system with strong time-dependent coupling. For the qubit case, we confine ourselves to the lowest two levels, while for a multilevel generalization, we consider the energy level ladder of a Kerr oscillator with anharmonicity K.

set of coupled differential equations for generators $G^{(n)}(t)$ ($n \geq 3$) for explicitly time-dependent interactions [5]. On the other hand, the construction presented here allows cancelation of the derivative term at each order in λ *independently* since the orders are tracked by updating the off-diagonal contribution at each step, which greatly simplifies the equations of motion for higher-order generators (see Appendix A).

III. PARAMETRIC LIGHT-MATTER INTERACTIONS

A. Rabi Hamiltonian

We now apply the sequential SWT construction described in the previous section to a time-dependent Rabi Hamiltonian describing a qubit-resonator system (Fig. 1). The system Hamiltonian with time-dependent transverse coupling is given by ($\hbar = 1$)

$$\hat{H}(t) = -\frac{\omega_q}{2}\hat{\sigma}_q^z + \omega_a\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) + g(t)\hat{\sigma}_q^x(\hat{a}^\dagger + \hat{a}). \quad (10)$$

For a block off-diagonal interaction of the form in Eq. (10), the first time-dependent unitary rotation in our SWT series is described by the generator

$$\hat{S}^{(1)}(t) = \xi_+^{(1)}(t)\hat{\sigma}_q^+\hat{a}^\dagger - \xi_-^{(1)}(t)\hat{\sigma}_q^-\hat{a}^\dagger - \text{H.c.}, \quad (11)$$

where the time dependence of $V^{(1)}(t) \equiv g(t)\hat{\sigma}_q^x(\hat{a}^\dagger + \hat{a})$ is ensconced in the solution of the coefficients $\xi_\pm^{(1)}(t)$. These can be evaluated using Eq. (7) for $M = 1$, which leads to the following set of linear differential equations,

$$\pm i\dot{\xi}_\pm^{(1)}(t) - (\omega_q \pm \omega_a)\xi_\pm^{(1)}(t) + g(t) = 0. \quad (12)$$

The time-dependent Hamiltonian correct to second order in g_p, $\hat{H}^{(2)}(t)$ in the interaction picture, defined with respect to the time independent $\hat{H}_0$, can then be obtained as

$$\hat{H}^{(2)}(t) = \frac{1}{2}[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)] + \mathcal{O}(g_p^3)$$

$$= -g(t)\left[\text{Re}\{\xi_-^{(1)}(t)\} + \text{Re}\{\xi_+^{(1)}(t)\}\right]\hat{\sigma}_q^z\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)$$

$$- \frac{g(t)}{2}\left[\left(\xi_-^{(1)*}(t) + \xi_+^{(1)*}(t)\right)e^{-2i\omega_a t}\hat{a}^2\hat{\sigma}_q^z + \text{H.c.}\right]$$

$$+ \frac{g(t)}{2}\left(\text{Re}\{\xi_-^{(1)}(t)\} - \text{Re}\{\xi_+^{(1)}(t)\}\right) + \mathcal{O}(g^3(t)). \quad (13)$$

Assuming the interaction time dependence to be sinusoidal and decomposing it into distinct frequency components, i.e.,

$$g(t) = \sum_{p=1}^{K}[g_p \exp(-i\omega_p t) + g_p^* \exp(+i\omega_p t)], \quad (14)$$

leads to the solution for Eq. (12) as

$$\xi_\pm^{(1)}(t) = \sum_{p=1}^{K}\left[\frac{g_p}{\omega_q \pm \omega_a \mp \omega_p}\exp(-i\omega_p t)\right.$$

$$\left. + \frac{g_p^*}{\omega_q \pm \omega_a \pm \omega_p}\exp(+i\omega_p t)\right]. \quad (15)$$

For the special case of monochromatic parametric driving ($K = 1$), Eq. (13) simplifies to

$$\hat{H}^{(2)}(t) = -2|g_p|^2\sum_\pm\left[\left(\frac{1}{\omega_\pm + \omega_p} + \frac{1}{\omega_\pm - \omega_p}\right)\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)\hat{\sigma}_q^z + \frac{1}{2}\left(\frac{1}{\omega_\pm + \omega_p} + \frac{1}{\omega_\pm - \omega_p}\right)\right]\cos^2(\omega_p t + \phi_p)$$

$$- \left[\left\{e^{-2i\omega_a t}|g_p|^2\sum_\pm\left(\frac{1}{\omega_\pm + \omega_p} + \frac{1}{\omega_\pm - \omega_p}\right) + e^{-2i(\omega_p + \omega_a)t}g_p^2\left(\frac{1}{\omega_- - \omega_p} + \frac{1}{\omega_+ + \omega_p}\right)\right.\right.$$

$$\left.\left. + e^{+2i(\omega_p - \omega_a)t}(g_p^*)^2\left(\frac{1}{\omega_- + \omega_p} + \frac{1}{\omega_+ - \omega_p}\right)\right\}\hat{a}^2 + \text{H.c.}\right]\frac{\hat{\sigma}_q^z}{2} + \mathcal{O}(g_p^3), \quad (16)$$

where $\phi_p = \arg(g_p)$ and $\omega_\pm = \omega_q \pm \omega_a$, with $+, -$ denoting the sum and difference of the qubit and the resonator frequencies, respectively. The squeezing terms in Eq. (16) are fast rotating for pump frequencies near $\omega_\pm$ and their effect can hence be neglected as their time-average vanishes. On the other hand, when $\omega_p \sim \pm\omega_a$ and the squeezing terms are nearly resonant, the effective detuning is comparable to ω_q, which is in gigahertz; this makes the effect of these terms negligible since $|g_p|/\omega_q \ll 1$. By a similar argument, cubic contributions towards χ_p from g_p^3 can be neglected under the rotating-wave approximation. This order necessarily involves terms of the form $[\hat{S}^{(1)}, [\hat{S}^{(1)}, \hat{V}^{(1)}]]$, and each term of this form contains an asymmetric number of raising and lowering operators $(\hat{a}^\dagger, \hat{a})$ such that it can only be rendered resonant for large detunings from parametric resonance.

Consequently, the diagonal part of the Hamiltonian in Eq. (16) to order $\mathcal{O}(g_p^2)$, $\hat{H}_{\text{eff}}^{(2)}(t) = \mathcal{P}_0 \bullet \frac{1}{2}[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)]$ simplifies to its time-averaged form as

$$\hat{H}_{\text{eff}}^{(2)} = \chi_p^{(2)} \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) \hat{\sigma}_q^z + \Omega_p^{(2)} \qquad (17)$$

with

$$\chi_p^{(2)} = -2|g_p|^2 \left(\frac{\omega_-}{\omega_-^2 - \omega_p^2} + \frac{\omega_+}{\omega_+^2 - \omega_p^2} \right) \qquad (18)$$

denoting the leading-order dispersive shift and

$$\Omega_p^{(2)} = |g_p|^2 \left(\frac{\omega_-}{\omega_-^2 - \omega_p^2} - \frac{\omega_+}{\omega_+^2 - \omega_p^2} \right) \qquad (19)$$

being the overall energy shift.

To eliminate $\mathcal{O}(g_p^2)$ off-diagonal terms in Eq. (16), we develop the next-order SWT generator as $\hat{S}^{(2)}(t) = \xi^{(2)}(t)$ $\hat{\sigma}_q^z(\hat{a}^\dagger)^2 - \text{H.c.}$, where $\xi^{(2)}(t)$ is again evaluated using Eq. (7) for $M = 2$. The resultant effective Hamiltonians are (see Appendix D for detailed expressions of higher-order shifts)

$$\hat{H}_{\text{eff}}^{(3)}(t) = 0, \qquad (20a)$$

$$\hat{H}_{\text{eff}}^{(4)}(t) = \mathcal{P}_0 \bullet \frac{1}{8} \left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t) \right] \right] \right] \qquad (20b)$$

$$+ \mathcal{P}_0 \bullet \frac{1}{2} \left[\hat{S}^{(2)}(t), \mathcal{Q}_0 \bullet \frac{1}{2} \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t) \right] \right]. \qquad (20c)$$

Figure 2 shows a comparison of dressed shifts predicted from Eqs. (20) with exact shifts estimated from a numerical simulation of the parametric Rabi Hamiltonian. Note
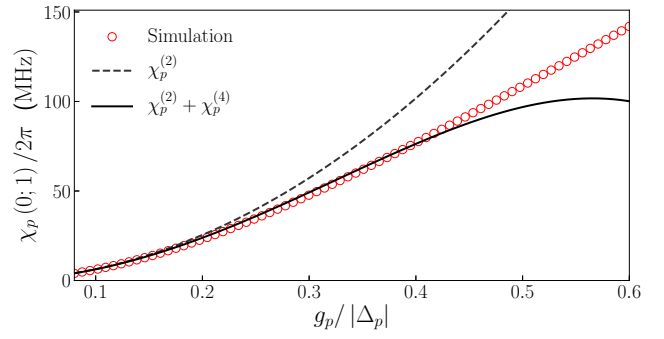

FIG. 2. Comparison of analytical calculations for the linear dispersive shift, obtained from the $\chi_p \hat{a}^\dagger \hat{a} \hat{\sigma}_q^z$ term in $\hat{H}_{\text{eff}}^{(2)}(t)$ and $\hat{H}_{\text{eff}}^{(4)}(t)$, respectively, with exact results obtained from the numerically simulated resonator spectrum. All simulations are performed with $\omega_q = 2\pi \times 5$ GHz, $\omega_a = 2\pi \times 3$ GHz, $\omega_p = 2\pi \times 1.5$ GHz, and $\kappa = 2\pi \times 1.5$ MHz.

that, for this specific case, the presence of the inertial term significantly modifies the magnitude of $|\xi_\pm^{(1)}|$, and hence the condition for validity of the perturbative expansion $|\xi_\pm^{(1)}| \ll 1$ leads to

$$|g_p| \ll \Delta_p, \quad \Delta_p = \min\{|\omega_\pm \pm \omega_p|\}. \qquad (21)$$

This implies that the interaction strength needs to be small compared to the detunings measured in a rotating frame defined with respect to the pump frequency ω_p. This presents an opportunity to realize large dressed shifts, in excess of 10 MHz with $|g_p|$ of only a few tens of megahertz, as evident from the results presented in Fig. 2. For comparison, achieving a similar value of the shift with static couplings would require an interaction strength of magnitude comparable to the qubit-resonator detuning (i.e., $|g_p| \sim \omega_-$). Both (i) the improvement of the radius of convergence to as large as $|g_p/\Delta_p| \sim 0.4$ upon including the quartic contribution, and (ii) the reliable prediction of dispersive shifts in tens of megahertz from analytical theory, confirm how the generalized SWT method developed here provides a powerful framework to capture the effect of time-dependent strong interactions.

While the example discussed here involves a monochromatic pump, the sequential SWT method can accommodate "colored" pumps, arbitrary forms of $g(t)$ via a decomposition into a Fourier series, and frequency modulation of system frequencies (see Appendix B). A specific case, relevant to experimental situations, is inclusion of residual or parasitic static coupling present due to practical design limitations. The effect of such a static interaction can be captured by including a zero-frequency component in Eq. (14):

$$g(t) = \sum_{j=p,s} [g_j \exp(-i\omega_j t) + g_j^* \exp(+i\omega_j t)]. \qquad (22)$$

Following the same procedure as outlined earlier, the time-averaged dispersive shift is the sum of different frequency components of the interaction, $\chi_p^{(2)} = \sum_j \chi_j^{(2)}$. This provides a means to cancel out the net combined shift by canceling the shift caused by the static coupling against that caused by the parametric coupling. Note that this can be accomplished even when $|g_p| < |g_s|$, so long as a viable ω_p can be found.

B. Validity of the rotating-wave approximation

A crucial consideration for strong interactions is the validity of the Hamiltonian derived under the rotating-wave approximation (RWA). This is especially relevant for periodically driven systems where, depending on the choice of ω_p, any (or even a combination) of the time-dependent terms can be rendered resonant. Here we elucidate how it is imperative to make the RWA after deriving the transformed $H_{\mathrm{eff}}(t)$ to a given order, by contrasting the results obtained with the full Rabi interaction to the predictions for a parametrically driven Jaynes-Cummings (JC) interaction.

To delineate the importance of the counter-rotating terms, $\hat{\sigma}_q^+ \hat{a}^\dagger + \mathrm{H.c.}$, in the parametric dispersive regime [Eq. (21)], we now derive the dressed shifts for a qubit-resonator system coupled via a parametrically mediated JC interaction of the form $\hat{V}_{\mathrm{JC}}^{(1)}(t) = g(t)(\hat{\sigma}_q^+ \hat{a} + \hat{\sigma}_q^- \hat{a}^\dagger)$. In this case, unlike Eq. (11), the time-dependent SWT generator includes only the difference frequency components in line with the JC Hamiltonian,

$$\hat{S}_{\mathrm{JC}}^{(1)}(t) = -\xi_-^{(1)}(t)\hat{\sigma}_q^- \hat{a}^\dagger - \mathrm{H.c.} \tag{23}$$

The resultant leading-order effective Hamiltonian in the interaction picture is given by

$$\hat{H}_{\mathrm{JC}}^{(2)}(t) = \chi_{\mathrm{JC}}^{(2)}(t)\hat{\sigma}_q^z\left(\hat{a}^\dagger \hat{a} + \frac{1}{2}\right) + \Omega_{\mathrm{JC}}^{(2)}(t), \tag{24}$$

where $\chi_{\mathrm{JC}}^{(2)}(t) = -g(t)\mathrm{Re}[\xi_-^{(1)}(t)]$ and the overall energy shift $\Omega_{\mathrm{JC}}^{(2)}(t) = \frac{1}{2}g(t)\mathrm{Re}[\xi_-^{(1)}(t)]$. Note that the qubit-state-dependent squeezing component of the effective Hamiltonian present in the Rabi model case does not appear for JC interaction. Considering a single-frequency pump as before, it leads to a simple expression for the time-averaged parametric dispersive shift:

$$\chi_{\mathrm{JC}}^{(2)} = -2|g_p|^2 \frac{\omega_-}{\omega_-^2 - \omega_p^2}. \tag{25}$$

As evident, the dispersive shift for the JC Hamiltonian lacks the sum frequency contribution as a result of the missing counter-rotating terms in the Hamiltonian, in contrast to the dispersive shift for the Rabi Hamiltonian in Eq. (18).

Figure 3 presents a detailed comparison of analytically calculated qubit-induced dispersive shifts on the resonator frequency, as a function of the modulation frequency ω_p, against the backdrop of the exact resonator transmission spectrum obtained from a numerical simulation for both the Rabi and JC Hamiltonians. As evident, parametric interactions support a unique "blind spot" where $\chi_p = 0$, even for large g_p/Δ_p, for an appropriate pump frequency at the geometric average of the sum and the difference frequencies $\omega_{\mathrm{BS}} = (\omega_+\omega_-)^{1/2} = (\omega_q^2 - \omega_a^2)^{1/2}$. (Note that no blind spot exists for $\omega_q < \omega_a$.) This decoupling is caused due to equal and opposite contributions to the dressed energy-level shifts from the difference and the sum frequency components of the Rabi Hamiltonian at ω_{BS}, respectively. Note that no blind spot exists for $\omega_q < \omega_a$. Such a cancelation at the two-level approximation is not possible with shifts estimated in Eq. (25) with JC interaction, of which the profile of qubit-induced shifts on the resonator frequency for various values of ω_p against the backdrop of the exact resonator transmission spectrum obtained from a numerical simulation is shown in Fig. 3(b).

In the same vein, the JC Hamiltonian fails to capture the pole near the sum frequency ω_+ entirely. This is clear from Eq. (25), which unlike Eq. (18), lacks the sum frequency contribution. In fact, the contribution of the missing sum frequency terms can be significant even at pump frequencies far detuned from ω_+; this is most strikingly evident at ω_{BS}, where a nonzero value of χ_{JC} persists due to the absence of Bloch-Siegert contributions. Thus, unlike the Rabi model, the JC interaction does not support zero-contrast points with monochromatic driving. Both these aspects highlight how making the rotating-wave approximation in the lab frame, before the SWT is performed, can be problematic and lead to erroneous results for time-dependent interactions.

IV. MULTILEVEL EFFECTS

To investigate the effect of higher excitations in time-dependent settings, we next consider a multilevel generalization of the Rabi model by replacing the qubit with a Kerr oscillator. The Kerr oscillator provides an economical way to study the perturbative corrections for the case of a multi-frequency spectrum, while introducing only one additional parameter (anharmonicity K, with $K > 0$ assumed) to the problem. Furthermore, it attracts considerable theoretical and practical interest since, besides being a canonical example of a nonlinear quantum optical system, it forms the cornerstone of many quantum information platforms using Josephson junction-based superconducting circuits [18–20].

The system of a Kerr oscillator coupled to a linear oscillator via a time-dependent transverse coupling can be
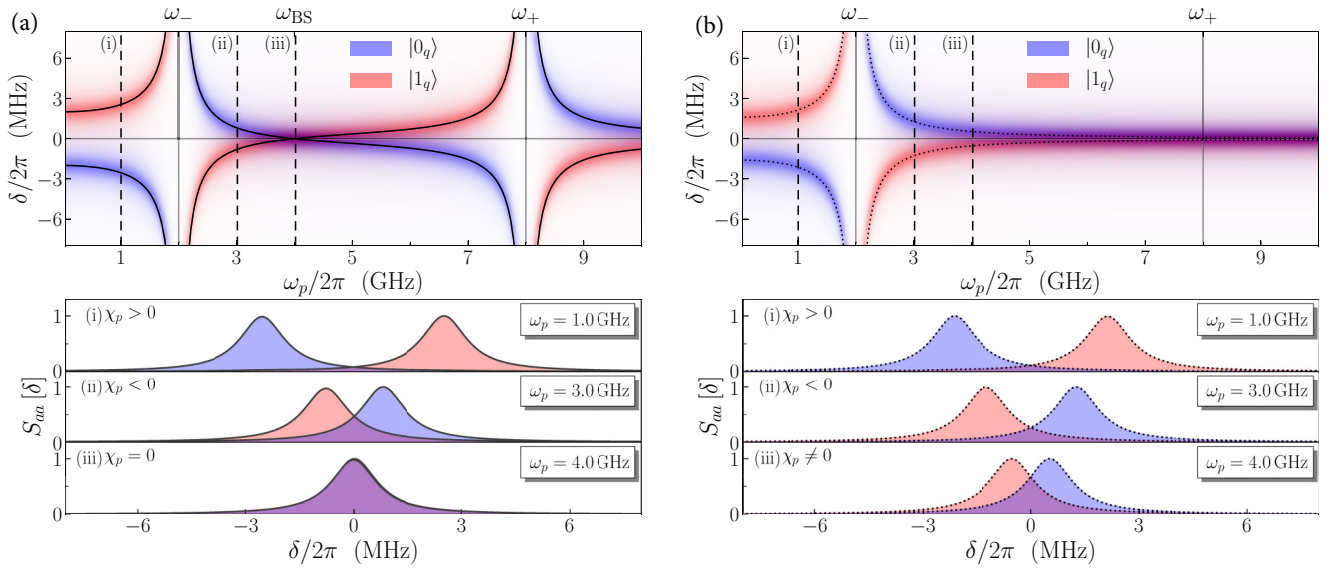
FIG. 3. (a) Comparison of analytical results for the resonator frequency shifts calculated using Eq. (18) (black) with the resonator spectrum $S_{aa}[\delta]$, for δ the detuning from ω_a, obtained from exact numerical simulations (color) of the time-dependent Rabi Hamiltonian. The resonator spectral density is shown as a function of the pump frequency ω_p for a qubit in the ground (blue) and excited (red) states. Bottom panel shows the normalized resonator spectra for selected values of ω_p, demonstrating positive, negative, and zero χ_p ("blind spot") for the Rabi model. (b) Analogous comparison for the Jaynes-Cummings Hamiltonian. Note that no blind spot occurs in the absence of the Bloch-Siegert contribution mediated by the sum frequency terms. All simulations are performed with $\omega_q = 2\pi \times 5$ GHz, $\omega_a = 2\pi \times 3$ GHz, $g_p = 2\pi \times 40$ MHz, and $\kappa = 2\pi \times 1.5$ MHz.

described by the Hamiltonian

$$\hat{H}(t) = \omega_b \hat{b}^\dagger \hat{b} - K(\hat{b}^\dagger \hat{b})^2 + \omega_a \hat{a}^\dagger \hat{a} + g(t)(\hat{b}^\dagger + \hat{b})(\hat{a}^\dagger + \hat{a}),$$

where $\hat{b}^\dagger$ ($\hat{a}^\dagger$), $\hat{b}$ ($\hat{a}$) are creation and annihilation operators, ω_b (ω_a) is the frequency of the Kerr (linear) oscillator, and $g(t) = 2g_p \cos(\omega_p t)$ as before. Following the same procedure as done for the qubit, we can define a SWT generator S_1 of the form

$$\hat{S}^{(1)}(t) = g_p \sum_{\pm} e^{\pm i\omega_p t}[(\hat{\Omega}_+[\hat{n}_b] \pm \omega_p)^{-1}\hat{b}^\dagger \hat{a}^\dagger$$

$$+ (\hat{\Omega}_-[\hat{n}_b] \pm \omega_p)^{-1}\hat{b}^\dagger \hat{a}] - \text{H.c.}, \quad (26)$$

where $\hat{\Omega}_\pm[\hat{n}_b] = (\omega_b + K)\hat{\mathbb{I}} - 2K\hat{n}_b \pm \omega_a\hat{\mathbb{I}}$ with $\hat{n}_b = \hat{b}^\dagger \hat{b}$.

To leading order in g_p, we find the resultant cross-Kerr shift corresponding to the transition between the Fock states $\{|n_b - 1\rangle, |n_b\rangle\}$ of the Kerr oscillator as

$$\chi_p^{(2)}(n_b - 1; n_b)$$

$$= g_p^2 \sum_{\pm} \left(\frac{n}{\Omega_-(n) \pm \omega_p} + \frac{n}{\Omega_+(n) \pm \omega_p} \right.$$

$$\left. - \frac{n+1}{\Omega_-(n+1) \pm \omega_p} - \frac{n+1}{\Omega_+(n+1) \pm \omega_p} \right)\Bigg|_{n_b-1}^{n_b}, \quad (27)$$

where $\Omega_\pm(n_b) = \langle n_b | \hat{\Omega}_\pm[\hat{n}_b] | n_b \rangle$ denotes the number-dependent sum and difference frequencies for the system of coupled oscillators. Several observations are in order.

(i) The expression for the shift in Eq. (27) involves eight terms as compared to the two terms in Eq. (18) for the two-level qubit case; this is because each state $|n_b\rangle$ is coupled to neighboring states $|n_b \pm 1\rangle$, each of which contribute to the net shift on the energy of a given state.

(ii) Given the distinct number-dependent transition frequencies for the Kerr oscillator, the regime for validity of the perturbative diagonalization now depends on the number of excitations, i.e., $g_p/|\Delta_p(n_b)| < 1$, where $\Delta_p(n_b) = \min\{\Omega_\pm(n_b) \pm \omega_p\}$.

(iii) We develop the perturbation series with $g_p/|\Delta_p(n_b)|$ as the small parameter. This is a crucial improvement over some previous analyses, which treat the Kerr term perturbatively [21,22], since now the shifts calculated using Eq. (27) hold for any value of anharmonicity K relative to detuning from the parametric resonance $\Delta_p(n_b)$.

To elucidate the last point, Fig. 4 shows the detailed profile of calculated $\chi_p^{(2)}(0; 1)$ (the "transmon" limit [23]) in two distinct regimes based on the magnitude of the K relative to the detuning $\Omega_-(0)$; in both cases we assume weak anharmonicity such that $K < \omega_{a,b}$.
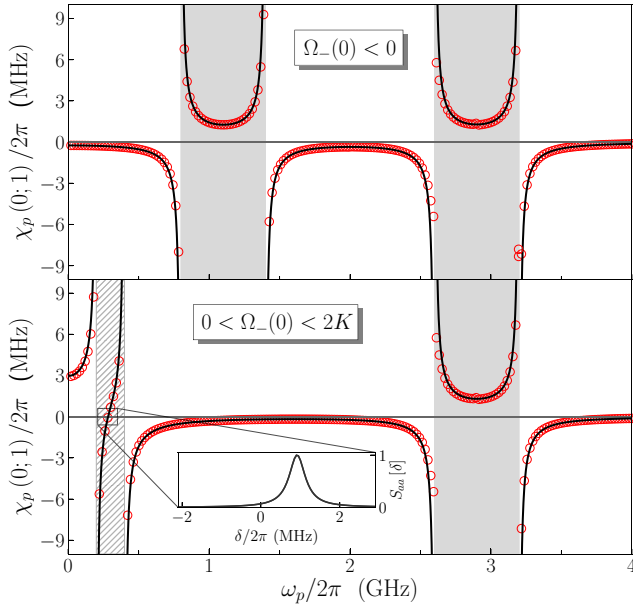
FIG. 4. Parametric dispersive shifts of a resonator coupled to a Kerr oscillator for two distinct regimes of anharmonicity: $\omega_a = 2\pi \times 2.0$ GHz, $\omega_b = 2\pi \times 1.5$ GHz (top) and $\omega_a = 2\pi \times 1.5$ GHz, $\omega_b = 2\pi \times 2.0$ GHz (bottom). Both simulations use $K = 2\pi \times 300$ MHz, $g_p = 2\pi \times 10$ MHz, and $\kappa = 2\pi \times 0.5$ MHz. The red circles are results from simulations of a truncated three-level system, and the black line is the result of the analytical calculation based on Eq. (27). In both panels, the gray regions correspond to the parametric straddling regimes, where $\chi_p(0;1) > 0$ is realized. The hatched area in the bottom panel corresponds to the region enclosing the blind spot. The inset shows the normalized resonator spectrum at the blind spot; note that, while there is no state-dependent shift, there is an overall shift of the resonator frequency at ω_{BS} unlike the qubit case (see Appendix C).

Case I: $\Omega_-(0) < 0$ *or* $0 < 2K < \Omega_-(0)$. As shown in Fig. 4, either positive or negative shifts can be realized depending on whether the ratio $K/\Delta_p(n_b)$ is larger or smaller than unity. In contrast, the sign of the static shift $\chi_s^{(2)}(0;1)$, $\omega_p = 0$ in Eq. (27), is fixed by the sign of the anharmonicity K alone. Specifically, $\chi_p^{(2)}(0;1)$ reverses sign when $\Omega_-(2) < \omega_p < \Omega_-(1)$ or $\Omega_+(2) < \omega_p < \Omega_+(1)$. This *parametric straddling regime* [13] is reminiscent of the static straddling regime achieved with fixed couplings for $\omega_b - 3K < \omega_a < \omega_b - K$ [24], except now the sign reversal can be achieved by tuning the frequency ω_p of the coupling while keeping its magnitude g_p fixed.

Case II: $0 < \Omega_-(0) < 2K$. Here, in addition to the sign reversal of induced frequency shifts between parametric straddling and dispersive regimes, a blind spot $\chi_p^{(2)}(0;1) = 0$ is realized for a pump frequency $\omega_{BS} \simeq (\Omega_-(1)|\Omega_-(2)|)^{1/2}$. Note that, unlike the two-level qubit

case, this does not result from the cancelation of Bloch-Siegert (sum) and JC (difference) contributions, but rather from the difference-frequency contributions to $\chi_p(0;1)$ corresponding to the lowest two transition frequencies of the Kerr oscillator.

V. INDUCED DISSIPATION

In the presence of parametric interactions, the dissipative interaction of the qubit with the environment (through the resonator) also becomes tunable with the pump frequency and amplitude, in sharp contrast with the static scenario where the induced dissipation is fixed [25]. Using a formal master equation construction (see Appendix E), we find that, for $|\omega_p - \omega_-| \sim g_p$, the dominant process is qubit relaxation ($\hat{\sigma}_q^-$) at rate $\gamma_-^\downarrow$, while for $|\omega_p - \omega_+| \sim g_p$, induced qubit heating ($\hat{\sigma}_q^+$) at rate $\gamma_+^\uparrow$ dominates, with the respective rates given by

$$\gamma_-^\downarrow \approx \sum_\pm \kappa(\omega_q \pm \omega_p)\left(\frac{g_p}{\omega_q - \omega_a \pm \omega_p}\right)^2, \quad (28a)$$

$$\gamma_+^\uparrow \approx \sum_\pm \kappa(-\omega_q \pm \omega_p)\left(\frac{g_p}{\omega_q + \omega_a \pm \omega_p}\right)^2. \quad (28b)$$

Here $\kappa(\omega)$ denotes the rate proportional to the noise spectral density of the environment at frequency ω. A similar calculation for the case of the Kerr resonator leads to an n_b-dependent relaxation (heating) rate when the interaction is driven near the difference (sum) frequency, $\Omega_\pm(n_b)$, for the corresponding number-dependent transition (see Appendix E). Note that the "quantum heating" rate $\gamma_+^\uparrow$ is nonzero even when the resonator remains in its ground state, since it is mediated through coupling-induced amplification of vacuum fluctuations [26]. Crucially, the induced dissipation rates do *not* null at the blind spots unlike the coherent shifts; they remain highly suppressed since the corresponding pump frequency is far detuned from both the sum and difference frequencies, i.e., $|\omega_\pm \pm \omega_{BS}| \gg g_p$.

As evident from Eq. (28), a fixed frequency qubit parametrically coupled to an environment can be employed as a novel broadband noise sensor. The resultant qubit relaxation or heating rates sample the environmental density of states over a set bandwidth (set by a Lorentzian profile of the resonator in QED setups), with the center of this "filter function" set by the choice of pump frequency mediating the interaction. Such pump-mediated noise spectroscopy can be thought of as a continuous-wave analogue of spin locking with the baseband frequency set by ω_p in gigahertz, making it immune to low-frequency fluctuations that typically limit conventional protocols relying on variation of the drive amplitude [27]. It is also an attractive complement to dynamical decoupling protocols that characterize

slow or quasistatic noise [28], and can be used for sensing high-frequency quantum noise that was shown to limit coherence in recent qubit designs [29].

VI. CONCLUSIONS

We present an *ab-initio* framework for diagonalizing strong time-dependent interactions, based on a time-dependent generalization of the SWT. The excellent quantitative agreement between exact numerical simulations and our analytical calculations for perturbative shifts and decay of dressed states demonstrates how the SWT can be used for developing a well-controlled perturbation series in time-dependent settings. The time-dependent SWT method presents a complementary approach to exact methods such as Floquet theory. Though exact, the complexity of Floquet-based diagonalization grows exponentially with the number of driving fields, making it computationally demanding. On the other hand, while perturbative, SWT-based expansion is more amenable in terms of analytical tractability of the effective Hamiltonian, thereby offering physically relevant insights. In fact, recent works have used SWT-like techniques for block diagonalization of the Floquet quasienergy operator in order to gain a more intuitive picture about validity of high-frequency approximations employed while truncating the Floquet Hilbert space [30].

Applying our analytical technique to parametric interactions, we predict several unique features accessible with parametric cavity- and circuit-QED systems. Most notably, tunable dispersive shifts in tens of megahertz can be realized with modest interaction strengths by simply tuning the pump frequency sufficiently close to the sum or difference frequency, while maintaining large detuning between the physical subsystems. This is a particularly attractive functionality for multiqubit circuit-QED architectures where frequency crowding becomes a limiting issue [12]. Achieving such tunability in architectures based on static coupling and tunable qubits necessarily comes with trade-offs, such as flux-noise-limited qubit coherence and enhanced crosstalk, leading to limited control flexibility. Furthermore, our results show how the physics of the Rabi model can be crucial even in the dispersive regime $g_p \ll |\omega_-|$ and leads to significant corrections beyond the usual rotating-wave Jaynes-Cummings model; these are crucial for predicting exact cancelation points ("blind spots") of the induced energy shifts. We note that these corrections to the dressed shift are distinct from non-RWA corrections in the ultrastrong coupling ($g < \omega_- \lesssim \omega_a$) [31] or the deep-strong coupling ($\omega_- < g \sim \omega_a$) [32] regime. In addition to their theoretical appeal, identifying and engineering such blind spots is a topic of active interest in several applications, such as protecting qubits against photon shot-noise-induced dephasing [33,34] and mitigating ZZ-induced crosstalk in two-qubit

gates [35–38] in circuit-QED architectures. Extending our analysis to weakly anharmonic systems, we show how frequency-tunable interactions support a rich structure of multilevel shifts, including switching between parametric dispersive and parametric straddling regimes via the choice of pump frequency. Previously, such effects have been restricted to the purview of highly nonlinear multilevel atoms with complicated selection-rule engineering [39].

The parametric QED regime presented here enables several expeditious applications of the burgeoning parametric quantum toolbox. In addition, they provide a natural starting point for several theoretical extensions, such as multipump generalizations, the multiphoton Rabi model [40], and Hamiltonian engineering for quantum simulation of nontrivial gauge structures [41], that we hope to investigate in future work.

APPENDIX A: TIME-DEPENDENT SCHRIEFFER-WOLFF TRANSFORMATION

The essence of constructing a sequence of successive time-dependent SWTs that perturbatively diagonalizes a general time-dependent Hamiltonian is to find an appropriate choice for the Mth-order SWT generator $\hat{S}^{(M)}(t)$ such that the off-diagonal part of the interaction Hamiltonian only contains terms of $\mathcal{O}(\lambda^M)$ or higher, i.e.,

$$\hat{V}_{\mathrm{OD}}^{(M)}(t) = \sum_{m=M}^{\infty} \lambda^m \hat{V}_{\mathrm{OD},m}^{(M)}(t). \qquad (A1)$$

The form of this generator can be calculated using Eq. (7) presented in the main text; in this section we present a detailed proof of this construction.

We construct this proof by following a process similar to mathematical induction. Without loss of generality, we begin by assuming that the lab frame interaction Hamiltonian $\hat{V}(t) \sim \mathcal{O}(\lambda)$, such that $\hat{H}^{(1)}(t)$ satisfies Eq. (5) by construction. Then, assuming that $\hat{H}^{(M)}(t)$ satisfies Eq. (5), we now prove that $\hat{H}^{(M+1)}(t)$ also satisfies Eq. (5) as long as $\hat{S}^{(M)}(t)$ satisfies Eq. (7).

Utilizing the Baker-Campbell-Hausdorff expansion and the derivative of the exponential map, the unitary transformation from $\hat{H}^{(M)}(t)$ to $\hat{H}^{(M+1)}(t)$ can be expressed as, with $\hat{U}_M(t) = \exp[-\hat{S}^{(M)}(t)]$,

$$\hat{H}^{(M+1)}(t) = \hat{U}_M^\dagger(t)\hat{H}^{(M)}(t)\hat{U}_M(t) - i\hat{U}_M^\dagger(t)\frac{\partial \hat{U}_M(t)}{\partial t}$$

$$= \hat{H}_0 + \sum_{n=0}^{\infty}\frac{1}{(1+n)!}\left(\mathrm{ad}_{\hat{S}^{(M)}(t)}\right)^n [\hat{S}^{(M)}(t), \hat{H}_0] + \sum_{n=0}^{\infty}\frac{1}{n!}\left(\mathrm{ad}_{\hat{S}^{(M)}(t)}\right)^n \hat{V}^{(M)}(t)$$

$$+ \sum_{n=0}^{\infty}\frac{1}{(1+n)!}\left(\mathrm{ad}_{\hat{S}^{(M)}(t)}\right)^n \left(i\frac{\partial \hat{S}^{(M)}(t)}{\partial t}\right)$$

$$= \hat{H}_0 + \frac{1}{1!}\hat{W}^{(M)}(t) + \left(\frac{1}{1!} - \frac{1}{2!}\right)[\hat{S}^{(M)}(t), \hat{V}^{(M)}(t)] + \frac{1}{2!}[\hat{S}^{(M)}(t), \hat{W}^{(M)}(t)]$$

$$+ \left(\frac{1}{2!} - \frac{1}{3!}\right)[\hat{S}^{(M)}(t), [\hat{S}^{(M)}(t), \hat{V}^{(M)}(t)]] + \frac{1}{3!}[\hat{S}^{(M)}(t), [\hat{S}^{(M)}(t), \hat{W}^{(M)}(t)]]$$

$$+ \left(\frac{1}{3!} - \frac{1}{4!}\right)[\hat{S}^{(M)}(t), [\hat{S}^{(M)}(t), [\hat{S}^{(M)}(t), \hat{V}^{(M)}(t)]]]$$

$$+ \frac{1}{4!}[\hat{S}^{(M)}(t), [\hat{S}^{(M)}(t), [\hat{S}^{(M)}(t), \hat{W}^{(M)}(t)]]] + \cdots$$

$$= \hat{H}_0 + \hat{W}^{(M)}(t) + \sum_{n=1}^{\infty}\frac{1}{(n+1)!}\left(\mathrm{ad}_{\hat{S}^{(M)}(t)}\right)^n \hat{W}^{(M)}(t)$$

$$+ \sum_{n=1}^{\infty}\frac{n}{(n+1)!}\left(\mathrm{ad}_{\hat{S}^{(M)}(t)}\right)^n \hat{V}^{(M)}(t), \tag{A2}$$

where $\mathrm{ad}_{\hat{S}^{(M)}(t)}\bullet \equiv [\hat{S}^{(M)}(t), \bullet]$ denotes the adjoint action, and

$$\hat{W}^{(M)}(t) = i\frac{\partial \hat{S}^{(M)}(t)}{\partial t} + [\hat{S}^{(M)}(t), \hat{H}_0] + \hat{V}^{(M)}(t). \tag{A3}$$

Next, following the line of inductive reasoning we assume that $\hat{H}^{(M)}$ satisfies Eq. (5). Therefore, the interaction Hamiltonian can be written as

$$\hat{V}^{(M)}(t) = \hat{V}_{\mathrm{D}}^{(M)}(t) + \sum_{m=M}^{\infty} \lambda^m \hat{V}_{\mathrm{OD},m}^{(M)}(t)$$

$$= \hat{V}_{\mathrm{D}}^{(M)}(t) + \lambda^M \hat{V}_{\mathrm{OD},M}^{(M)}(t) + \sum_{m=M+1}^{\infty} \lambda^m \hat{V}_{\mathrm{OD},m}^{(M)}(t), \tag{A4}$$

where we have singled out the lowest-order off-diagonal term $\lambda^M \hat{V}_{\mathrm{OD},M}^{(M)}(t)$ that we aim to eliminate. Given Eq. (7) for $\hat{S}^{(M)}(t)$, Eq. (A3) reduces to

$$\hat{W}^{(M)}(t) = i\frac{\partial \hat{S}^{(M)}(t)}{\partial t} + [\hat{S}^{(M)}(t), \hat{H}_0] + \hat{V}_{\mathrm{D}}^{(M)}(t)$$

$$+ \lambda^M \hat{V}_{\mathrm{OD},M}^{(M)}(t) + \sum_{m=M+1}^{\infty} \lambda^m \hat{V}_{\mathrm{OD},m}^{(M)}(t)$$

$$= \hat{V}^{(M)}(t) - \lambda^M \hat{V}_{\mathrm{OD},M}^{(M)}(t)$$

$$= \hat{V}_{\mathrm{D}}^{(M)}(t) + \sum_{m=M+1}^{\infty} \lambda^m \hat{V}_{\mathrm{OD},m}^{(M)}(t), \tag{A5}$$

which is exactly the remainder of $\hat{V}^{(M)}$ without its lowest-order off-diagonal component. Note that $\hat{V}^{(M)}(t), \hat{W}^{(M)}(t)$ are both at least of the order of $\mathcal{O}(\lambda^{n\geq 1})$, while $\hat{S}^{(M)}(t) \sim \mathcal{O}(\lambda^M)$ according to Eq. (7). Thus,

$$\sum_{n=1}^{\infty}\frac{1}{(n+1)!}\left(\mathrm{ad}_{\hat{S}^{(M)}(t)}\right)^n \hat{W}^{(M)}(t) \sim \mathcal{O}(\lambda^{n\geq M+1}), \tag{A6a}$$

$$\sum_{n=1}^{\infty}\frac{n}{(n+1)!}\left(\mathrm{ad}_{\hat{S}^{(M)}(t)}\right)^n \hat{V}^{(M)}(t) \sim \mathcal{O}(\lambda^{n\geq M+1}). \tag{A6b}$$

Substituting Eqs. (A6) into Eq. (A2), we arrive at the desired expression for $\hat{H}^{(M+1)}(t)$:

$$\hat{H}^{(M+1)}(t) = \hat{H}_0 + \hat{V}_{\mathrm{D}}^{(M)}(t) + \mathcal{O}(\lambda^{M+1}). \tag{A7}$$

Note that $\hat{H}^{(M+1)}(t)$ lacks any $\mathcal{O}(\lambda^M)$-order off-diagonal term. Thus, by constructing a time-dependent Schrieffer-Wolff generator using Eq. (7), $\hat{H}^{(M)}(t)$ can be transformed

into $\hat{H}^{(M+1)}(t)$ with the off-diagonal component at the Mth order eliminated. Since this property of the transformation is valid for any $M \geq 1$, a sequence of successive SWTs can be applied to eliminate the off-diagonal terms up to any desired order in λ.

Note that, although a formally infinite number of terms are involved in the expressions of $\hat{W}^{(M)}(t)$, $\hat{V}^{(M)}(t)$, $\hat{W}^{(M+1)}(t)$, $\hat{V}^{(M+1)}(t), \ldots,$ in practice none of these calculations need to be done beyond the desired order of diagonalization in λ, since they are neglected anyway in the final outcome.

1. Example for $M_{\text{max}} = 5$

As an illustrative example, we detail the procedure of obtaining a diagonalized Hamiltonian correct to fifth order (i.e., $M_{\text{max}} = 5$), obtained by eliminating off-diagonal interaction terms to $\mathcal{O}(\lambda^4)$. To this end, without loss of generality, we begin by assuming that $\hat{V}_{\text{D}}^{(1)} = 0$ and write

$$\hat{V}^{(1)}(t) = \lambda \hat{V}_{\text{OD}}^{(1)}(t). \tag{A8}$$

Using Eq. (A3), this leads to the following differential equation for $\hat{S}^{(1)}(t)$:

$$i\frac{\partial \hat{S}^{(1)}(t)}{\partial t} + [\hat{S}^{(1)}(t), \hat{H}_0] + \lambda \hat{V}_{\text{OD}}^{(1)}(t) = 0. \tag{A9}$$

Therefore,

$$\hat{W}^{(1)}(t) = i\frac{\partial \hat{S}^{(1)}(t)}{\partial t} + [\hat{S}^{(1)}(t), \hat{H}_0] + \hat{V}^{(1)}(t) = 0. \tag{A10}$$

According to Eq. (A2),

$$\hat{V}^{(2)}(t) = \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right] + \frac{1}{3}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right]$$
$$+ \frac{1}{8}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right]\right] + \mathcal{O}(\lambda^5). \tag{A11}$$

Next, following this procedure the second-order generator $\hat{S}^{(2)}(t)$ can be evaluated from

$$i\frac{\partial \hat{S}^{(2)}(t)}{\partial t} + [\hat{S}^{(2)}(t), \hat{H}_0] + \lambda^2 \hat{V}_{\text{OD},2}^{(2)}(t) = 0, \tag{A12}$$

where $\hat{V}_{\text{OD},2}^{(2)}(t)$ is entirely generated from the off-diagonal contribution due to action of the previous SWT on the interaction, i.e.,

$$\lambda^2 \hat{V}_{\text{OD},2}^{(2)}(t) = \mathcal{Q}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]. \tag{A13}$$

Substituting this into the expression for $\hat{W}^{(2)}$, we obtain

$$\hat{W}^{(2)} = i\frac{\partial \hat{S}^{(2)}(t)}{\partial t} + [\hat{S}^{(2)}(t), \hat{H}_0] + \hat{V}^{(2)}(t)$$
$$= i\frac{\partial \hat{S}^{(2)}(t)}{\partial t} + [\hat{S}^{(2)}(t), \hat{H}_0] + \mathcal{Q}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right] + \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]$$
$$+ \frac{1}{3}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right] + \frac{1}{8}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}, \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right]\right] + \mathcal{O}(\lambda^5)$$
$$= \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right] + \frac{1}{3}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right] + \frac{1}{8}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right]\right] + \mathcal{O}(\lambda^5). \tag{A14}$$

Next, according to Eq. (A2),

$$\hat{V}^{(3)}(t) = \hat{W}^{(2)}(t) + \frac{1}{2}\left[\hat{S}^{(2)}(t), \hat{W}^{(2)}(t)\right] + \frac{1}{6}\left[\hat{S}^{(2)}(t), \left[\hat{S}^{(2)}(t), \hat{W}^{(2)}(t)\right]\right]$$
$$+ \frac{1}{2}\left[\hat{S}^{(2)}(t), \hat{V}^{(2)}(t)\right] + \frac{1}{3}\left[\hat{S}^{(2)}(t), \left[\hat{S}^{(2)}(t), \hat{V}^{(2)}(t)\right]\right] + \mathcal{O}(\lambda^8)$$
$$\simeq \hat{W}^{(2)}(t) + \frac{1}{2}\left[\hat{S}^{(2)}(t), \hat{W}^{(2)}(t)\right] + \frac{1}{2}\left[\hat{S}^{(2)}(t), \hat{V}^{(2)}(t)\right] + \mathcal{O}(\lambda^5)$$
$$= \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right] + \frac{1}{3}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right]$$
$$+ \frac{1}{8}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right]\right] + \frac{1}{2}\left[\hat{S}^{(2)}(t), \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right] + \frac{1}{2}\left[\hat{S}^{(2)}(t), \hat{V}^{(2)}(t)\right] + \mathcal{O}(\lambda^5),$$

where in the second step we drop the terms $[\hat{S}^{(2)}(t), [\hat{S}^{(2)}(t), \hat{W}^{(2)}(t)]], [\hat{S}^{(2)}(t), [\hat{S}^{(2)}(t), \hat{V}^{(2)}(t)]] \sim \mathcal{O}(\lambda^6)$, since we are only interested in elimination of off-diagonal terms up to $\mathcal{O}(\lambda^4)$. This leads to the following differential equation for third-order generator $\hat{S}^{(3)}(t)$:

$$i\frac{\partial \hat{S}^{(3)}(t)}{\partial t} + [\hat{S}^{(3)}(t), \hat{H}_0] + \lambda^3 \hat{V}_{\text{OD},3}^{(3)}(t) = 0. \tag{A15}$$

Here, now the only off-diagonal terms at cubic order are obtained as

$$\lambda^3 \hat{V}_{\text{OD},3}^{(3)}(t) = \mathcal{Q}_0 \bullet \frac{1}{3}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right]. \tag{A16}$$

This leads to

$$\begin{aligned}
\hat{W}^{(3)} &= i\frac{\partial \hat{S}^{(3)}(t)}{\partial t} + [\hat{S}^{(3)}(t), \hat{H}_0] + \hat{V}^{(3)}(t) \\
&= i\frac{\partial \hat{S}^{(3)}(t)}{\partial t} + [\hat{S}^{(3)}(t), \hat{H}_0] + \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right] \\
&\quad + \mathcal{Q}_0 \bullet \frac{1}{3}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right] + \mathcal{P}_0 \bullet \frac{1}{3}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right] \\
&\quad + \frac{1}{8}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right]\right] + \frac{1}{2}\left[\hat{S}^{(2)}(t), \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right] \\
&\quad + \frac{1}{2}\left[\hat{S}^{(2)}(t), \hat{V}^{(2)}(t)\right] + \mathcal{O}(\lambda^5) \\
&= \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right] + \mathcal{P}_0 \bullet \frac{1}{3}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right] \\
&\quad + \frac{1}{8}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right]\right] \\
&\quad + \frac{1}{2}\left[\hat{S}^{(2)}(t), \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right] + \frac{1}{2}\left[\hat{S}^{(2)}(t), \hat{V}^{(2)}(t)\right] + \mathcal{O}(\lambda^5).
\end{aligned}$$

Noting that $\hat{V}^{(3)}(t), \hat{W}^{(3)}(t) \sim \mathcal{O}(\lambda^2)$, and $\hat{S}^{(3)}(t) \sim \mathcal{O}(\lambda^3)$, we can ignore terms of the type $[\hat{S}^{(3)}(t), \hat{W}^{(3)}(t)]$, $[\hat{S}^{(3)}(t), \hat{V}^{(3)}(t)]$, which are $\mathcal{O}(\lambda^5)$, since we are interested in obtaining the diagonalized Hamiltonian to $\mathcal{O}(\lambda^4)$. Therefore,

$$\begin{aligned}
\hat{V}^{(4)}(t) &= \hat{W}^{(3)}(t) + \frac{1}{2}\left[\hat{S}^{(3)}(t), \hat{W}^{(3)}(t)\right] + \mathcal{O}(\lambda^9) + \frac{1}{2}\left[\hat{S}^{(3)}(t), \hat{V}^{(3)}(t)\right] + \mathcal{O}(\lambda^8) \\
&\simeq \hat{W}^{(3)}(t) + \mathcal{O}(\lambda^5) \\
&= \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right] + \mathcal{P}_0 \bullet \frac{1}{3}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right] \\
&\quad + \frac{1}{8}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}\right]\right]\right] \\
&\quad + \frac{1}{2}\left[\hat{S}^{(2)}(t), \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right] + \frac{1}{2}\left[\hat{S}^{(2)}(t), \hat{V}^{(2)}(t)\right] + \mathcal{O}(\lambda^5). \tag{A17}
\end{aligned}$$

Next, the differential equation for $\hat{S}^{(4)}(t)$ becomes

$$i\frac{\partial \hat{S}^{(4)}(t)}{\partial t} + [\hat{S}^{(4)}(t), \hat{H}_0] + \lambda^4 \hat{V}_{\text{OD},4}^{(4)}(t) = 0, \tag{A18}$$

where the three off-diagonal contributions at $\mathcal{O}(\lambda^4)$ order are given as

$$
\begin{aligned}
\lambda^4 \hat{V}_{\mathrm{OD},4}^{(4)}(t) = \mathcal{Q}_0 \bullet \Bigg(& \frac{1}{8}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right]\right] \\
& + \frac{1}{2}\left[\hat{S}^{(2)}(t), \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right] + \frac{1}{2}\left[\hat{S}^{(2)}(t), \hat{V}^{(2)}(t)\right]\Bigg).
\end{aligned}
\tag{A19}
$$

Therefore,

$$
\begin{aligned}
\hat{W}^{(4)}(t) = {} & \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right] + \mathcal{P}_0 \bullet \frac{1}{3}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right] \\
& + \mathcal{P}_0 \bullet \frac{1}{8}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right]\right] \\
& + \mathcal{P}_0 \bullet \left(\frac{1}{2}\left[\hat{S}^{(2)}(t), \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right]\right) + \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(2)}(t), \hat{V}^{(2)}(t)\right] \\
& + \mathcal{O}(\lambda^5)
\end{aligned}
\tag{A20}
$$

and

$$
\begin{aligned}
\hat{V}^{(5)}(t) = {} & \hat{W}^{(4)}(t) + \frac{1}{2}\left[\hat{S}^{(4)}(t), \hat{W}^{(4)}(t)\right] + \frac{1}{2}\left[\hat{S}^{(4)}(t), \hat{V}^{(4)}(t)\right] + \mathcal{O}(\lambda^{10}) \\
\simeq {} & \hat{W}^{(4)}(t) + \mathcal{O}(\lambda^6) \\
= {} & \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right] + \mathcal{P}_0 \bullet \frac{1}{3}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right] \\
& + \mathcal{P}_0 \bullet \frac{1}{8}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right]\right] \\
& + \mathcal{P}_0 \bullet \left(\frac{1}{2}\left[\hat{S}^{(2)}(t), \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right]\right) + \mathcal{P}_0 \bullet \frac{1}{2}\left[\hat{S}^{(2)}(t), \hat{V}^{(2)}(t)\right] \\
& + \mathcal{O}(\lambda^5).
\end{aligned}
\tag{A21}
$$

In the examples discussed in the main text, $[\hat{S}^{(1)}(t), [\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)]]$ and $[\hat{S}^{(2)}(t), \mathcal{P}_0 \bullet [\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)]]$ generate no diagonal contributions; therefore,

$$
\begin{aligned}
\hat{V}^{(5)}(t) = \mathcal{P}_0 \bullet \Bigg(& \frac{1}{2}\left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right] + \frac{1}{8}\left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \left[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)\right]\right]\right] \\
& + \frac{1}{2}\left[\hat{S}^{(2)}(t), \hat{V}^{(2)}(t)\right]\Bigg) + \mathcal{O}(\lambda^5).
\end{aligned}
\tag{A22}
$$

APPENDIX B: TIME-DEPENDENT SYSTEM FREQUENCY

Our framework can fully accommodate the case where the qubit frequency is time dependent (TD) due to, say, parametric flux variation. To demonstrate this, we take the Rabi model as an example and calculate only up to $\hat{S}^{(1)}(t)$. The total Hamiltonian now becomes $\hat{H}(t) = \hat{H}_0^{\mathrm{TD}}(t) + \hat{V}(t)$, where $\hat{H}_0^{\mathrm{TD}}(t) = -(\omega_q(t)/2)\hat{\sigma}_q^z + \omega_a(\hat{a}^\dagger \hat{a} + 1/2)$. It is more convenient to obtain the SWT generator by first transforming into the interaction picture, where

$$
\hat{V}_{\mathrm{I}}(t) = \exp\left[i\int_0^t \hat{H}_0^{\mathrm{TD}}(t')dt'\right]\hat{V}(t)\exp\left[-i\int_0^t \hat{H}_0^{\mathrm{TD}}(t')dt'\right]
\tag{B1}
$$

and $i\partial\hat{S}_{\mathrm{I}}^{(1)}/\partial t + \hat{V}_{\mathrm{I}}(t) = 0$. We can then solve for

$$
\hat{S}_{\mathrm{I}}^{(1)}(t) = \tilde{\xi}_+(t)\hat{\sigma}_q^+\hat{a}^\dagger - \tilde{\xi}_-(t)\hat{\sigma}_q^-\hat{a}^\dagger - \mathrm{H.c.},
\tag{B2}
$$

where

$$\tilde{\xi}_{\pm}(t) = \pm i \int_0^t g(t') \exp\left\{ i \int_0^{t'} \left[\pm \omega_q(t'') + \omega_a \right] dt'' \right\} dt'. \tag{B3}$$

As a concrete example, we consider the frequently encountered case of parametric flux variation of the qubit frequency, $\Phi(t) = \Phi_0 + \delta\Phi \cos(\omega_d t)$ with $|\delta\Phi| \ll \Phi_0$. Such nonlinear modulation of the qubit frequency can be written as

$$\omega_q[\Phi(t)] = \omega_q \cos[\Phi_0] + \delta\Phi \cos(\omega_d t)] \approx \omega_{q0} + \nu \cos(\omega_d t),$$

where $\omega_{q0} = \omega_q \cos[\Phi_0]$ and $\nu = \delta\Phi \times \omega_q \sin[\Phi_0]$. Substituting this into Eq. (B3) and making use of the Jacobi-Anger expansion to resolve the exponent of the sinusoidal modulation, we obtain

$$\begin{aligned}
\tilde{\xi}_{\pm}(t) &= \pm i \int_0^t g(t') \exp\left\{ i \left[\left(\pm \omega_{q0} + \omega_a \right) t' \pm \frac{\nu}{\omega_d} \sin(\omega_d t') \right] \right\} dt' \\
&= \pm \sum_{k=-\infty}^{+\infty} J_k\left(\frac{\nu}{\omega_d} \right) \sum_p \left\{ \frac{g_p}{\pm\omega_{q0} \pm k\omega_d + \omega_a - \omega_p} \exp\left[i \left(\pm\omega_{q0} \pm k\omega_d + \omega_a - \omega_p \right) t \right] \right. \\
&\quad \left. + \frac{g_p^*}{\pm\omega_{q0} \pm k\omega_d + \omega_a + \omega_p} \exp\left[i \left(\pm\omega_{q0} \pm k\omega_d + \omega_a + \omega_p \right) t \right] \right\},
\end{aligned} \tag{B4}$$

where we have assumed sinusoidal coupling, as in Eq. (14) of the main text. Here $J_k(\nu/\omega_d)$ is the Bessel function of the first kind, and k is an integer. The expression for the corresponding SWT generator in the Schrödinger picture,

$$\begin{aligned}
\hat{S}^{(1)}(t) &= \tilde{\xi}_+(t) \exp\left\{ -i \int_0^t \left[+\omega_q(t') + \omega_a \right] dt' \right\} \hat{\sigma}_q^+ \hat{a}^\dagger \\
&\quad - \tilde{\xi}_-(t) \exp\left\{ -i \int_0^t \left[-\omega_q(t') + \omega_a \right] dt' \right\} \hat{\sigma}_q^- \hat{a}^\dagger - \text{H.c.}
\end{aligned} \tag{B5}$$

leads to

$$\begin{aligned}
\xi_{\pm}(t) &= \pm \sum_{k,k'=-\infty}^{+\infty} J_k(\nu/\omega_d) J_{k'}(\nu/\omega_d) \\
&\quad \times \sum_p \left\{ \frac{g_p}{\pm\omega_{q0} \pm k\omega_d + \omega_a - \omega_p} e^{i\left[\pm(k-k')\omega_d - \omega_p \right]t} \right. \\
&\quad \left. + \frac{g_p^*}{\pm\omega_{q0} \pm k\omega_d + \omega_a + \omega_p} e^{i\left[\pm(k-k')\omega_d + \omega_p \right]t} \right\}.
\end{aligned} \tag{B6}$$

This can now be used in $\hat{H}^{(2)}(t) = \frac{1}{2}[\hat{S}^{(1)}(t), \hat{V}^{(1)}(t)] + \mathcal{O}(g_p^3)$ [Eq. (13)] to calculate the leading-order effective Hamiltonian correct to $\mathcal{O}(g^2(t))$. For a single pump frequency ω_p, the corresponding dispersive shift, when $r \equiv 2\omega_p/\omega_d$ is an integer, is

$$\begin{aligned}
\chi_p^{(2)} &= -|g_p|^2 \sum_{k=-\infty}^{+\infty} J_k^2(\nu/\omega_d) \left(\frac{1}{\omega_{k,-} - \omega_p} + \frac{1}{\omega_{k,-} + \omega_p} + \frac{1}{\omega_{k,+} - \omega_p} + \frac{1}{\omega_{k,+} + \omega_p} \right) \\
&\quad - \frac{1}{2}(g_p^2 + g_p^{*2}) \sum_{k=-\infty}^{+\infty} \left(\frac{J_k(\nu/\omega_d) J_{k-r}(\nu/\omega_d)}{\omega_{k,-} - \omega_p} + \frac{J_k(\nu/\omega_d) J_{k+r}(\nu/\omega_d)}{\omega_{k,-} + \omega_p} \right. \\
&\quad \left. + \frac{J_k(\nu/\omega_d) J_{k-r}(\nu/\omega_d)}{\omega_{k,+} - \omega_p} + \frac{J_k(\nu/\omega_d) J_{k+r}(\nu/\omega_d)}{\omega_{k,+} + \omega_p} \right),
\end{aligned} \tag{B7}$$

where $\omega_{\pm,k} = (\omega_{q0} + k\omega_d) \pm \omega_a$. If r is not an integer,

$$\chi_p^{(2)} = -|g_p|^2 \sum_{k=-\infty}^{+\infty} J_k^2(\nu/\omega_d) \left(\frac{1}{\omega_{k,-} - \omega_p} + \frac{1}{\omega_{k,-} + \omega_p} \right.$$

$$\left. + \frac{1}{\omega_{k,+} - \omega_p} + \frac{1}{\omega_{k,+} + \omega_p} \right). \tag{B8}$$

Note that now, in addition to the qubit frequency ω_{q0}, we get contributions to the dressed shift at every order from sidebands at $\omega_{q0} + k\omega_d$, weighted by $J_k(\nu/\omega_d)$.

APPENDIX C: MULTILEVEL SHIFTS

The leading-order diagonalized Hamiltonian for the system of a Kerr oscillator parametrically coupled to a linear oscillator is given by

$$\hat{H}_D^{(2)} = |g_p|^2 \hat{n}_a \sum_{\pm,\pm} [(\hat{\Omega}_\pm[\hat{n}_b] \pm \omega_p)^{-1} \hat{n}_b$$

$$- (\hat{\Omega}_\pm[\hat{n}_b + 1] \pm \omega_p)^{-1}(\hat{n}_b + 1)], \tag{C1}$$

where $\hat{\Omega}_\pm[\hat{n}_b] = (\omega_b + K)\hat{\mathbb{I}} - 2K\hat{n}_b \pm \omega_a \hat{\mathbb{I}}$.

In addition to the cross-Kerr shift reported in the main text, the dispersive shift induced by the Kerr oscillator's "ground state" can be calculated from the matrix element

$$\chi_p^{(2)}(0) = \langle n_a, n_b = 0|\hat{H}_D^{(2)}|n_a, n_b = 0\rangle. \tag{C2}$$

Figure 5 shows a profile of $\chi_p(0)$ for the two regimes of nonlinearity discussed in the main text. It is interesting to note that, even at the "blind spot" realized in the regime $0 < \Omega_-(0) < 2K$, $\chi_p(0) \neq 0$ (see the inset of Fig. 4 in the main text); this is in contrast to the qubit case where the cancelation happens at zero detuning from the bare resonance of the linear oscillator.

APPENDIX D: HIGHER-ORDER CORRECTIONS

The higher-order contributions in λ can be developed using the prescription given in Appendix A. Here we report the next-order contribution in the systems we consider in the main text, which appears at quartic order in the coupling g_p^4 for the purely block off-diagonal interactions considered. Notably, there are two contributions at quartic order: (i) the diagonal term generated by the action of the first-order generator on the first-order off-diagonal interaction $\frac{1}{8}[\hat{S}^{(1)}(t), [\hat{S}^{(1)}(t), [\hat{S}^{(1)}(t), \hat{V}(t)]]]$, and (ii) the diagonal term generated by the second-order generator on the second-order off-diagonal interaction $\frac{1}{2}[\hat{S}^{(2)}(t), \hat{V}_{\text{OD}}^{(2)}(t)]$ [see Eq. (A13)]. It is worth noting that the contribution of type (ii) does not appear for the JC interaction, as the source of the nonzero $\hat{V}_{\text{OD}}^{(2)}(t)$ is the squeezing terms. Here we report the terms generated for both these interactions post-RWA for both the qubit and the Kerr resonator cases.
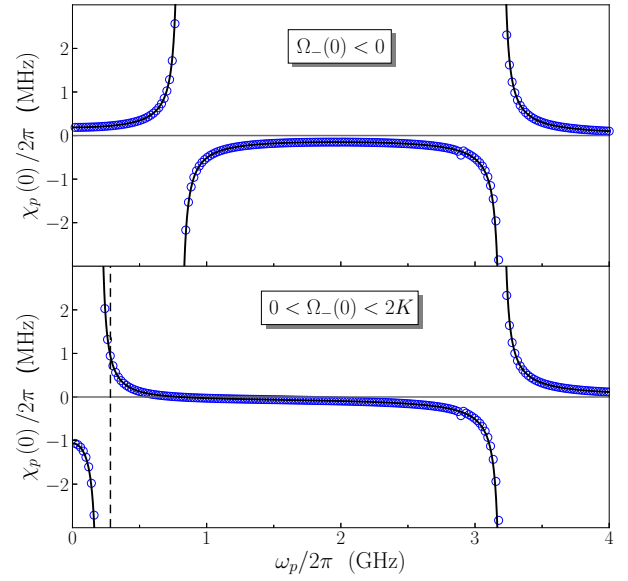

FIG. 5. (Top) Shift of a resonator coupled to a three-level system in its ground state, with $\omega_a = 2\pi \times 1.5$ GHz, $\omega_b = 2\pi \times 2.0$ GHz, $K = 2\pi \times 300$ MHz, $g_p = 2\pi \times 10$ MHz, and $\kappa = 2\pi \times 0.5$ MHz. The blue circles are results from simulations, and the black line is from Eq. (C2). (Bottom) Shift with $\omega_a = 2\pi \times 2.0$ GHz, $\omega_b = 2\pi \times 1.5$ GHz, and all other parameters as before.

1. Qubit case

The two terms that appear in the diagonalized Hamiltonian at quartic order are

$$\frac{1}{8}[\hat{S}^{(1)}(t), [\hat{S}^{(1)}(t), [\hat{S}^{(1)}(t), \hat{V}(t)]]]$$

$$= |g_p|^4 \left[\sum_\pm \left(\frac{1}{\omega_\pm + \omega_p} + \frac{1}{\omega_\pm - \omega_p} \right)^3 \right.$$

$$- \sum_\pm \left(\frac{1}{\omega_\pm + \omega_p} \right)^2 \left(\frac{1}{\omega_\pm - \omega_p} \right)$$

$$- \sum_\pm \left(\frac{1}{\omega_\pm - \omega_p} \right)^2 \left(\frac{1}{\omega_\pm + \omega_p} \right)$$

$$+ 2\sum_\pm \left(\frac{1}{\omega_\pm + \omega_p} + \frac{1}{\omega_\pm - \omega_p} \right)$$

$$\times \left(\frac{1}{\omega_\mp + \omega_p} + \frac{1}{\omega_\mp - \omega_p} \right)^2$$

$$- 2\sum_\pm \left(\frac{1}{\omega_\pm + \omega_p} \right)\left(\frac{1}{\omega_\mp + \omega_p} \right)\left(\frac{1}{\omega_\mp - \omega_p} \right)$$

$$\left. - 2\sum_\pm \left(\frac{1}{\omega_\pm - \omega_p} \right)\left(\frac{1}{\omega_\mp - \omega_p} \right)\left(\frac{1}{\omega_\mp + \omega_p} \right) \right] \hat{\sigma}_q^z$$

$$\times \left[(\hat{a}^\dagger \hat{a})^2 + \hat{a}^\dagger \hat{a} + \frac{1}{2} \right] \tag{D1}$$

and

$$\frac{1}{2}[\hat{S}^{(2)}(t), \hat{V}_{\mathrm{OD}}^{(2)}(t)]$$

$$= -|g_p|^4 \left[\frac{1}{\omega_a} \left(\sum_{\pm,\pm} \frac{1}{\omega_b \pm \omega_a \pm \omega_p} \right)^2 + \frac{1}{\omega_a + \omega_p} \left(\frac{1}{\omega_b + \omega_a + \omega_p} + \frac{1}{\omega_b - \omega_a - \omega_p} \right)^2 \right.$$

$$\left. + \frac{1}{\omega_a - \omega_p} \left(\frac{1}{\omega_b - \omega_a + \omega_p} + \frac{1}{\omega_b + \omega_a - \omega_p} \right)^2 \right] \frac{2\hat{a}^\dagger \hat{a} - 1}{4}. \tag{D2}$$

2. Kerr resonator case

For the Kerr resonator case, recall that $\chi_p^{(2)}(0;1)$, which is quadratic in g_p, exhibits poles at detunings of $-K$ and $-3K$ from the difference frequency ω_- (Fig. 3 of the main text); these correspond to the "single-photon" parametric resonances for the $|0\rangle \to |1\rangle$ and $|1\rangle \to |2\rangle$ transitions of the Kerr resonator, respectively. The next-order

corrections, quartic in g_p, describe the "two-photon" parametric resonances corresponding to the $|1\rangle \to |3\rangle$ transition of the Kerr resonator, thus manifesting as poles in $\chi_p^{(4)}(0;1)$ at $-5K$ and $+K$ detunings from the difference frequency. The resultant dispersive shifts can be read off from the prefactor of the $\hat{n}_a \hat{n}_b$ term in the SWT Hamiltonian diagonalized to fourth order, as

$$\chi_p^{(4)}(0;1) = -\frac{1}{4}|g_p|^4 \sum_{\pm} \left[-18 \left(\frac{1}{\Omega_-(3) \pm \omega_p} \right) \left(\frac{1}{\Omega_-(3) \pm \omega_p} \right) \left(\frac{1}{\Omega_-(2) \pm \omega_p} \right) \right.$$

$$+ 36 \left(\frac{1}{\Omega_-(2) \pm \omega_p} \right) \left(\frac{1}{\Omega_-(3) \pm \omega_p} \right) \left(\frac{1}{\Omega_+(2) \pm \omega_p} \right)$$

$$+ 54 \left(\frac{1}{\Omega_+(2) \pm \omega_p} \right) \left(\frac{1}{\Omega_-(3) \pm \omega_p} \right) \left(\frac{1}{\Omega_+(2) \pm \omega_p} \right)$$

$$+ 6 \left(\frac{1}{\Omega_-(0) \pm \omega_p} \right) \left(\frac{1}{\Omega_-(2) \pm \omega_p} \right) \left(\frac{1}{\Omega_+(1) \pm \omega_p} \right)$$

$$\left. + 4 \left(\frac{1}{\Omega_+(0) \pm \omega_p} \right) \left(\frac{1}{\Omega_-(1) \pm \omega_p} \right) \left(\frac{1}{\Omega_+(1) \pm \omega_p} \right) \right], \tag{D3}$$

where we have considered the pump to be near the difference frequency. Note that the dressed shifts to cubic order in g_p are zero under the RWA, since diagonal contributions to shifts require a balanced number of $\hat{b}^\dagger$ and $\hat{b}$.

Crucially, the degree of each pole in the expression for the dispersive shift is connected to the order of the n-photon process. For instance, for $\omega_p = \Omega_-(3)$, the pole near $5K$ manifests as a peak in $\chi_p^{(4)}(0;1)$ that is symmetric in detuning, since it corresponds to a pole of degree 2 [first line of Eq. (D3)]. This can be distinguished from the quadratic contribution to the dispersive shift $\chi_p^{(2)}(1;2)$ due to the $|2\rangle \to |3\rangle$ transition, which leads to a pole at the same value of ω_p, but with a peak lineshape that is asymmetric in detuning as it is a pole of degree 1.

APPENDIX E: PARAMETRICALLY INDUCED PURCELL DECAY AND QUANTUM HEATING

To investigate the induced dissipation on the qubit or Kerr oscillator in the presence of a time-dependent interaction with a dissipative linear oscillator, in this section we construct a Lindblad-form quantum master equation. To this end, we transform the system-bath interaction using the same generator as that used to implement the time-dependent Schrieffer-Wolff transformation on the system Hamiltonian. Since this master equation construction is based on a second-order Dyson series expansion of the system-bath interaction, we restrict ourselves to the lowest-order operator transformations $\mathcal{O}(g_p/\Delta_p(n_b))$.

We present the results for the Kerr oscillator here since the qubit case is a specific limit of the multilevel physics of

this system. Specifically, our focus is obtaining the rate of induced, or Purcell, decay on a given transition of the Kerr oscillator, as well as the onset of dissipative terms mediated by the parametric coupling such as quantum heating. To this end, we assume that the linear resonator is coupled to an environment modeled as a collection of harmonic oscillators, $\hat{H}_E = \sum_\alpha \nu_\alpha \hat{\gamma}_\alpha^\dagger \hat{\gamma}_\alpha$, with an interaction of the

form

$$\hat{H}_{SE} = \sum_\alpha \mu_\alpha (\hat{\gamma}_\alpha^\dagger + \hat{\gamma}_\alpha)(\hat{a}^\dagger + \hat{a}). \tag{E1}$$

Under the action of the SWT, $\hat{H}_{SE}$ in the interaction frame transforms as

$$\hat{H}_{SE}^{\mathrm{SW}}(t) = e^{\hat{S}(t)} \hat{H}_{SE} e^{-\hat{S}(t)} \simeq \sum_\alpha \mu_\alpha \left(\hat{\gamma}_\alpha^\dagger e^{i\nu_\alpha t} + \hat{\gamma}_\alpha e^{-i\nu_\alpha t} \right) \left[e^{-i\omega_a t} \left\{ \hat{a} - g_p \left(\hat{\Omega}_{+-}^{-1} e^{i\hat{\Omega}_{+-}t} \hat{b}^\dagger - \hat{b}\, \hat{\Omega}_{-+}^{-1} e^{-i\hat{\Omega}_{-+}t} \right) \right. \right.$$

$$\left. \left. -g_p^* \left(\hat{\Omega}_{++}^{-1} e^{i\hat{\Omega}_{++}t} \hat{b}^\dagger - \hat{b}\, \hat{\Omega}_{--}^{-1} e^{-i\hat{\Omega}_{--}t} \right) \right\} + \mathrm{H.c.} \right], \tag{E2}$$

where we have introduced the operator $\hat{\Omega}_{\pm\pm}[\hat{n}_b] = \hat{\Omega}_\pm[\hat{n}_b] \pm \omega_p \mathbb{I}$, with $\Omega_{\pm\pm}(n_b) \equiv \langle n_b | \hat{\Omega}_{\pm\pm} | n_b \rangle$, for brevity of notation. Note that if $g_p/|\Omega_{\pm\pm}(n_b)|\rangle \ll 1$, the corresponding terms in Eq. (E2) are highly suppressed. Taking the zero-temperature limit and considering a continuous density of states $\Gamma(\nu)$ for the bath modes leads to the following form of the master equation:

$$\frac{\partial}{\partial t}\hat{\rho}_I(t) = (\hat{a}\hat{\rho}_I(t)\hat{a}^\dagger - \hat{a}^\dagger \hat{a}\hat{\rho}_I(t)) \int_0^\infty d\nu \int_0^\infty dt'\, \Gamma(\nu)|\mu|^2 e^{i\nu(t'-t)} e^{-i\omega_a(t'-t)}$$

$$+ (\hat{a}^\dagger \hat{\rho}_I(t)\hat{a} - \hat{a}\hat{a}^\dagger \hat{\rho}_I(t)) \int_0^\infty d\nu \int_0^\infty dt'\, \Gamma(\nu)|\mu|^2 e^{i\nu(t'-t)} e^{+i\omega_a(t'-t)}$$

$$+ (\hat{b}\hat{\rho}_I(t)\hat{b}^\dagger - \hat{b}^\dagger \hat{b}\hat{\rho}_I(t)) \sum_\pm \frac{g_p^2}{|\Omega_{-\pm}(n_b+1)|^2}$$

$$\times \int_0^\infty d\nu \int_0^\infty dt'\, \Gamma(\nu)|\mu|^2 e^{i\nu(t'-t)} e^{-i(\omega_a+\Omega_{-\pm}(n_b+1))(t'-t)} + (\hat{b}^\dagger \hat{\rho}_I(t)\hat{b} - \hat{b}\hat{b}^\dagger \hat{\rho}_I(t)) \sum_\pm \frac{g_p^2}{|\Omega_{-\pm}(n_b)|^2}$$

$$\times \int_0^\infty d\nu \int_0^\infty dt'\, \Gamma(\nu)|\mu|^2 e^{i\nu(t'-t)} e^{+i(\omega_a+\Omega_{-\pm}(n_b))(t'-t)} + (\hat{b}\hat{\rho}_I(t)\hat{b}^\dagger - \hat{b}^\dagger \hat{b}\hat{\rho}_I(t)) \sum_\pm \frac{g_p^2}{|\Omega_{+\pm}(n_b+1)|^2}$$

$$\times \int_0^\infty d\nu \int_0^\infty dt'\, \Gamma(\nu)|\mu|^2 e^{i\nu(t'-t)} e^{+i(\omega_a-\Omega_{+\pm}(n_b+1))(t'-t)} + (\hat{b}^\dagger \hat{\rho}_I(t)\hat{b} - \hat{b}\hat{b}^\dagger \hat{\rho}_I(t)) \sum_\pm \frac{g_p^2}{|\Omega_{+\pm}(n_b)|^2}$$

$$\times \int_0^\infty d\nu \int_0^\infty dt'\, \Gamma(\nu)|\mu|^2 e^{i\nu(t'-t)} e^{-i(\omega_a-\Omega_{+\pm}(n_b))(t'-t)} + \mathrm{H.c.} \tag{E3}$$

Before proceeding, it is worth clarifying the shorthand notation of Eq. (E3). As an example, we explicitly write out the following term in Eq. (E3):

$$\hat{b}\hat{\rho}_I(t)\hat{b}^\dagger \frac{g_p^2}{|\Omega_{-+}(n_b)|^2}$$

$$= \sum_{n_b} \frac{n_b g_p^2}{|\Omega_{-+}(n_b)|^2} |n_b - 1\rangle \langle n_b | \hat{\rho}_I(t) | n_b \rangle \langle n_b - 1|. \tag{E4}$$

This shows that there are distinct dissipative rates for each Fock-state transition $n_b \to n_b - 1$. To obtain the "diagonal" form of Eq. (E4), we have made a RWA to drop terms of the form $|n_b - 1\rangle \langle n_b | \hat{\rho}_I(t) | m_b \rangle \langle m_b -$

$1|$ for $n_b \neq m_b$ during the master equation derivation. These terms would have a time-dependent phase factor $\exp\{-i[\Omega_{-+}(n_b) - \Omega_{-+}(m_b)] t\}$ in the interaction frame, which for $\Omega_{-+}(n_b) - \Omega_{-+}(m_b) = 2K(m_b - n_b)$, is fast oscillating. Note that we have a shifted transition frequency corresponding to $(n_b + 1)$ excitations in the Kerr oscillator for $\hat{b} \bullet \hat{b}^\dagger$ due to noncommutation of $\hat{\Omega}_{\pm\pm}$ with $(\hat{b}, \hat{b}^\dagger)$, namely,

$$\hat{b}^\dagger \hat{\Omega}_{\pm\pm} = (\hat{\Omega}_{\pm\pm} + 2K\,\mathbb{I})\hat{b}^\dagger, \quad \hat{b}\hat{\Omega}_{\pm\pm} = (\hat{\Omega}_{\pm\pm} - 2K\,\mathbb{I})\hat{b}. \tag{E5}$$

Finally, the master equation can be written in a more compact form as

$$
\frac{\partial}{\partial t}\hat{\rho}_I(t) = \left[\kappa(\omega_a)\mathcal{D}[\hat{a}] \right.
$$
$$
+ \sum_{n_b}\left(\left[\gamma_-^\downarrow(n_b) + \gamma_+^\downarrow(n_b)\right]\mathcal{D}\left[|n_b - 1\rangle\langle n_b|\right] \right.
$$
$$
\left.\left. + \left[\gamma_-^\uparrow(n_b) + \gamma_+^\uparrow(n_b)\right]\mathcal{D}\left[|n_b\rangle\langle n_b - 1|\right]\right)\right]\hat{\rho}_I(t),
$$
$$
\tag{E6}
$$

where $\mathcal{D}[\hat{O}]\bullet = \hat{O}\bullet\hat{O}^\dagger - \frac{1}{2}\hat{O}^\dagger\hat{O}\bullet - \frac{1}{2}\bullet\hat{O}^\dagger\hat{O}$ and $\kappa(\omega) = 2\pi\Gamma(\omega)|\mu(\omega)|^2$, with $\gamma^\downarrow$ and $\gamma^\uparrow$ representing relaxation-type and heating-type dissipators, respectively. The Kerr oscillator dissipative rates are given by

$$
\gamma_-^\downarrow(n_b) = \sum_\pm \kappa[\omega_a + \Omega_{-\pm}(n_b)]\frac{g_p^2}{|\Omega_{-\pm}(n_b)|^2}, \tag{E7a}
$$

$$
\gamma_-^\uparrow(n_b) = \sum_\pm \kappa[-\omega_a - \Omega_{-\pm}(n_b)]\frac{g_p^2}{|\Omega_{-\pm}(n_b)|^2}, \tag{E7b}
$$

$$
\gamma_+^\downarrow(n_b) = \sum_\pm \kappa[-\omega_a + \Omega_{+\pm}(n_b)]\frac{g_p^2}{|\Omega_{+\pm}(n_b)|^2}, \tag{E7c}
$$

$$
\gamma_+^\uparrow(n_b) = \sum_\pm \kappa[+\omega_a - \Omega_{+\pm}(n_b)]\frac{g_p^2}{|\Omega_{+\pm}(n_b)|^2}. \tag{E7d}
$$

Typically, under the assumption that the bath modes remain in vacuum, only the relaxation process is present; however, for parametric pumping, heating is possible even with a zero-temperature environment due to amplification of the zero-point fluctuations. Note that if the pump detuning from either the sum or difference frequency is large compared to g_p, i.e., $\Omega_{\pm\pm}(n_b) \gg g_p$, then the corresponding dissipative rates are highly suppressed due to a large denominator in the respective prefactor. We focus on two distinct cases, depending on the choice of pump frequency to be near the sum or difference frequency of the transmon-resonator system.

Case I: $\omega_p \approx |\omega_-|$. In this limit, depending on the detuning of the pump frequency, either $\Omega_{--}(n_b)$ or $\Omega_{-+}(n_b)$ is comparable in magnitude to g_p, such that the corresponding rate dominates. Furthermore, $\Omega_{-\pm}(n_b) \ll \omega_a$, such that the corresponding $\kappa(-\omega_a - \langle\hat{\Omega}_{-\pm}\rangle) \sim \kappa(-\omega_a) = 0$, since the bath spectrum has no negative frequency components. The result of this is that the prefactor for the corresponding heating term, $\mathcal{D}[\hat{b}^\dagger]$, is zero. Thus, whichever decay term is made non-negligible by the pump will have its counterpart heating term exactly zero. On the other hand, for whichever of $\Omega_{-\pm}(n_b)$ is not comparable to g_p, we have $\Omega_{-\pm} \gg g_p$, such that the corresponding decay term

is highly suppressed, while its counterpart heating term can be potentially nonzero but nonetheless still highly suppressed. For both the sum frequency components, we have $\Omega_{+\pm}(n_b) \gg g_p$ and, thus, both the corresponding decay and heating rates are negligible. In summary, the dominant contribution to the dissipation comes from the parametrically induced relaxation rate, which leads to the master equation

$$
\frac{\partial}{\partial t}\hat{\rho}_I(t) \simeq \left[\kappa(\omega_a)\mathcal{D}[\hat{a}] + \sum_{n_b}\gamma_-^\downarrow(n_b)\mathcal{D}[|n_b - 1\rangle\langle n_b|] \right]
$$
$$
\times \hat{\rho}_I(t). \tag{E8}
$$

Case II: $\omega_p \approx \omega_+$. Following a similar analysis to the preceding discussion, in this regime parametrically induced quantum heating dominates the dissipative dynamics, with the master equation taking the form

$$
\frac{\partial}{\partial t}\hat{\rho}_I(t) \simeq \left[\kappa(\omega_a)\mathcal{D}[\hat{a}] + \sum_{n_b}\gamma_+^\uparrow(n_b)\mathcal{D}[|n_b\rangle\langle n_b - 1|] \right]
$$
$$
\times \hat{\rho}_I(t). \tag{E9}
$$

As a final point, we emphasize that in either the qubit or the Kerr oscillator case, if the pump frequency is at the blind-spot frequency, the system still experiences induced decay and heating, even though the effective coherent coupling is exactly canceled. Nonetheless, since the blind-spot frequency, $\omega_p = \omega_{\text{BS}}$, is far detuned from either the sum or difference frequency, both Purcell decay and heating effects, though nonzero, are highly suppressed due to $|g_p/\Omega_p| \ll 1$, where $\Omega_p = \min\{\Omega_{\pm\pm}(n_b)\}$.

[1] Sergey Bravyi, David P. DiVincenzo, and Daniel Loss, Schrieffer–Wolff transformation for quantum many-body systems, Ann. Phys. (N. Y.) **326**, 2793 (2011).

[2] Christian Brouder, Gabriel Stoltz, and Gianluca Panati, Adiabatic approximation, Gell-Mann and low theorem, and degeneracies: A pedagogical example, Phys. Rev. A **78**, 042102 (2008).

[3] Marin Bukov, Michael Kolodrubetz, and Anatoli Polkovnikov, Schrieffer-Wolff Transformation for Periodically Driven Systems: Strongly Correlated Systems with Artificial Gauge Fields, Phys. Rev. Lett. **116**, 125301 (2016).

[4] Alexandru Petrescu, Moein Malekakhlagh, and Hakan E. Türeci, Lifetime renormalization of driven weakly anharmonic superconducting qubits. II. The readout problem, Phys. Rev. B **101**, 134510 (2020).

[5] Moein Malekakhlagh, Easwar Magesan, and David C. McKay, First-principles analysis of cross-resonance gate operation, Phys. Rev. A **102**, 042605 (2020).

[6] Michael A. Sentef, Jiajun Li, Fabian Künzel, and Martin Eckstein, Quantum to classical crossover of Floquet engineering in correlated quantum systems, Phys. Rev. Res. **2,** 033033 (2020).

[7] Luca Barbiero, Christian Schweizer, Monika Aidelsburger, Eugene Demler, Nathan Goldman, and Fabian Grusdt, Coupling ultracold matter to dynamical gauge fields in optical lattices: From flux attachment to $\mathbb{Z}_2$ lattice gauge theories, Sci. Adv. **5,** 7444 (2019).

[8] P. Roushan *et al.*, Chiral ground-state currents of interacting photons in a synthetic magnetic field, Nat. Phys. **13,** 146 (2016).

[9] Vittorio Peano, Martin Houde, Christian Brendel, Florian Marquardt, and Aashish A. Clerk, Topological phase transitions and chiral inelastic transport induced by the squeezing of light, Nat. Commun. **7,** 10779 (2016).

[10] Anina Leuch, Luca Papariello, Oded Zilberberg, Christian L. Degen, R. Chitra, and Alexander Eichler, Parametric Symmetry Breaking in a Nonlinear Resonator, Phys. Rev. Lett. **117,** 214101 (2016).

[11] Steven Mathey and Sebastian Diehl, Absence of Criticality in the Phase Transitions of Open Floquet Systems, Phys. Rev. Lett. **122,** 110602 (2019).

[12] Matthew Reagor *et al.*, Demonstration of universal parametric entangling gates on a multi-qubit lattice, Sci. Adv. **4,** 3603 (2018).

[13] T. Noh, Z. Xiao, K. Cicak, X. Y. Jin, E. Doucet, J. Teufel, J. Aumentado, L. C. G. Govia, L. Ranzani, A. Kamal, and R. W. Simmonds, arXiv:2103.09277 [quant-ph] (2021).

[14] Yao Lu, S. Chakram, N. Leung, N. Earnest, R. K. Naik, Ziwen Huang, Peter Groszkowski, Eliot Kapit, Jens Koch, and David I. Schuster, Universal Stabilization of a Parametrically Coupled Qubit, Phys. Rev. Lett. **119,** 150502 (2017).

[15] E. Doucet, F. Reiter, L. Ranzani, and A. Kamal, High fidelity dissipation engineering using parametric interactions, Phys. Rev. Res. **2,** 023370 (2020).

[16] A. J. Sirois, M. A. Castellanos-Beltran, M. P. DeFeo, L. Ranzani, F. Lecocq, R. W. Simmonds, J. D. Teufel, and J. Aumentado, Coherent-state storage and retrieval between superconducting cavities using parametric frequency conversion, Appl. Phys. Lett. **106,** 172603 (2015).

[17] X. Li, Y. Ma, J. Han, Tao Chen, Y. Xu, W. Cai, H. Wang, Y. P. Song, Zheng-Yuan Xue, Zhang-qi Yin, and Luyan Sun, Perfect Quantum State Transfer in a Superconducting Qubit Chain with Parametrically Tunable Couplings, Phys. Rev. Appl. **10,** 054009 (2018).

[18] Jens Koch, Terri M. Yu, Jay Gambetta, A. A. Houck, D. I. Schuster, J. Majer, Alexandre 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).

[19] Wenyuan Zhang, W. Huang, M. E. Gershenson, and M. T. Bell, Josephson Metamaterial with a Widely Tunable Positive or negative Kerr Constant, Phys. Rev. Appl. **8,** 051001 (2017).

[20] Christian Kraglund Andersen, Archana Kamal, Nicholas A. Masluk, Ioan M. Pop, Alexandre Blais, and Michel H. Devoret, Quantum versus Classical Switching Dynamics of Driven Dissipative Kerr Resonators, Phys. Rev. Appl. **13,** 044017 (2020).

[21] Simon E. Nigg, Hanhee Paik, Brian Vlastakis, Gerhard Kirchmair, S. Shankar, Luigi Frunzio, M. H. Devoret, R. J. Schoelkopf, and S. M. Girvin, Black-Box Superconducting Circuit Quantization, Phys. Rev. Lett. **108,** 240502 (2012).

[22] Nicolas Didier, Eyob A. Sete, Marcus P. da Silva, and Chad Rigetti, Analytical modeling of parametrically modulated transmon qubits, Phys. Rev. A **97,** 022330 (2018).

[23] J. Bourassa, F. Beaudoin, Jay M. Gambetta, and A. Blais, Josephson-junction-embedded transmission-line resonators: From Kerr medium to in-line transmon, Phys. Rev. A **86,** 013814 (2012).

[24] S. J. Srinivasan, A. J. Hoffman, J. M. Gambetta, and A. A. Houck, Tunable Coupling in Circuit Quantum Electrodynamics Using a Superconducting Charge Qubit with a v-Shaped Energy Level Diagram, Phys. Rev. Lett. **106,** 083601 (2011).

[25] Luke C. G. Govia and Frank K. Wilhelm, Unitary-Feedback-Improved Qubit Initialization in the Dispersive Regime, Phys. Rev. Appl. **4,** 054001 (2015).

[26] M. I. Dykman, M. Marthaler, and V. Peano, Quantum heating of a parametrically modulated oscillator: Spectral signatures, Phys. Rev. A **83,** 052115 (2011).

[27] D. H. Slichter, R. Vijay, S. J. Weber, S. Boutin, M. Boissonneault, J. M. Gambetta, A. Blais, and I. Siddiqi, Measurement-Induced Qubit State Mixing in Circuit QED from Up-Converted Dephasing Noise, Phys. Rev. Lett. **109,** 153601 (2012).

[28] Jonas Bylander, Simon Gustavsson, Fei Yan, Fumiki Yoshihara, Khalil Harrabi, George Fitch, David G. Cory, Yasunobu Nakamura, Jaw-Shen Tsai, and William D. Oliver, Noise spectroscopy through dynamical decoupling with a superconducting flux qubit, Nat. Phys. **7,** 565 (2011).

[29] Fei Yan, Simon Gustavsson, Archana Kamal, Jeffrey Birenbaum, Adam P. Sears, David Hover, Ted J. Gudmundsen, Danna Rosenberg, Gabriel Samach, S. Weber, Jonilyn L. Yoder, Terry P. Orlando, John Clarke, Andrew J. Kerman, and William D. Oliver, The flux qubit revisited to enhance coherence and reproducibility, Nat. Commun. **7,** 12964 (2016).

[30] André Eckardt and Egidijus Anisimovas, High-frequency approximation for periodically driven quantum systems from a Floquet-space perspective, New J. Phys. **17,** 093039 (2015).

[31] David Zueco, Georg M. Reuther, Sigmund Kohler, and Peter Hänggi, Qubit-oscillator dynamics in the dispersive regime: Analytical theory beyond the rotating-wave approximation, Phys. Rev. A **80,** 033846 (2009).

[32] J. Casanova, G. Romero, I. Lizuain, J. J. García-Ripoll, and E. Solano, Deep Strong Coupling Regime of the Jaynes-Cummings Model, Phys. Rev. Lett. **105,** 263603 (2010).

[33] A. P. Sears, A. Petrenko, G. Catelani, L. Sun, Hanhee Paik, G. Kirchmair, L. Frunzio, L. I. Glazman, S. M. Girvin, and R. J. Schoelkopf, Photon shot noise dephasing in the strong-dispersive limit of circuit QED, Phys. Rev. B **86,** 180504 (2012).

[34] Gengyan Zhang, Yanbing Liu, James J. Raftery, and Andrew A. Houck, Suppression of photon shot noise

dephasing in a tunable coupling superconducting qubit, Npj Quantum Inf. **3**, 1 (2017).

[35] Pranav Mundada, Gengyan Zhang, Thomas Hazard, and Andrew Houck, Suppression of Qubit Crosstalk in a Tunable Coupling Superconducting Circuit, Phys. Rev. Appl. **12**, 054023 (2019).

[36] Jaseung Ku, Xuexin Xu, Markus Brink, David C. McKay, Jared B. Hertzberg, Mohammad H. Ansari, and B. L. T. Plourde, Suppression of Unwanted *zz* Interactions in a Hybrid Two-Qubit System, Phys. Rev. Lett. **125**, 200504 (2020).

[37] M. Ganzhorn, G. Salis, D. J. Egger, A. Fuhrer, M. Mergenthaler, C. Müller, P. Müller, S. Paredes, M. Pechal, M. Werninghaus, and S. Filipp, Benchmarking the noise sensitivity of different parametric two-qubit gates in a single superconducting quantum computing platform, Phys. Rev. Res. **2**, 033447 (2020).

[38] Alexandru Petrescu, Camille Le Calonnec, Catherine Leroux, Agustin Di Paolo, Pranav Mundada, Sara Sussman, Andrei Vrajitoarea, Andrew A. Houck, and Alexandre Blais, Accurate methods for the analysis of strong-drive effects in parametric gates, arXiv:2107.02343 [quant-ph] (2021).

[39] Guanyu Zhu, David G. Ferguson, Vladimir E. Manucharyan, and Jens Koch, Circuit QED with fluxonium qubits: Theory of the dispersive regime, Phys. Rev. B **87**, 024510 (2013).

[40] P. Bertet, S. Osnaghi, P. Milman, A. Auffeves, P. Maioli, M. Brune, J. M. Raimond, and S. Haroche, Generating and Probing a Two-Photon Fock State with a Single Atom in a Cavity, Phys. Rev. Lett. **88**, 143601 (2002).

[41] N. Goldman and J. Dalibard, Periodically driven quantum systems: Effective Hamiltonians and engineered gauge fields, Phys. Rev. X **4**, 031027 (2014).