Efficient estimation of resonant coupling between quantum systems

Markku P.V. Stenberg,1 Yuval R. Sanders,2 and Frank K. Wilhelm1,2
1Theoretical Physics, Saarland University, 66123 Saarbrücken, Germany
2IQC and Dept. of Physics and Astronomy, University of Waterloo,
200 University Ave. W, Waterloo, ON, N2L 3G1, Canada

We present an efficient method for the characterization of two coupled discrete quantum systems, one of which can be controlled and measured. For two systems with transition frequencies \( \omega_q, \omega_r \) and coupling strength \( g \) we show how to obtain estimates of \( g \) and \( \omega_r \) whose error decreases exponentially in the number of measurement shots rather than a power law expected in simple approaches. Thereby the algorithm can identify \( g \) and \( \omega_r \) with high precision in few hundred measurement shots. This is achieved by adapting measurement settings by feeding back already obtained data. Both \( g \) and \( \omega_r \) are determined simultaneously. Our method works quite generally also in the presence of relaxation and typical noise. It is applicable to many candidate technologies for quantum computation, e.g., for the characterization of spurious two-level systems in superconducting qubits or stripline resonators.

Parameter estimation in many microscopic and some macroscopic systems inevitably involves quantum measurements. This implies that parameters cannot be identified with a single measurement shot since the outcome of such a measurement is generally random. Instead, the standard approach is to determine ensemble averages for many experiments and fit the parameters of certain quantitative model to those averages. The most common example for this is spectroscopy: it involves direct measurement of the energy splittings between quantum states in the form of resonances to incoming radiation. Typically, a large ensemble average is produced by gathering data from a large number of independent trials; either simultaneously on an ensemble of molecules (in nuclear magnetic resonance [1]) or from many repetitions of a specific experiment (in optical spectroscopy of single molecules, quantum dots, or superconducting qubits [2]).

While being reliable in many contexts, this approach is often too resource-intensive. Specifically, the error in the estimate of a single expectation value at a fixed measurement setting decreases in proportion to \( M_r^{-\frac{1}{2}} \) after \( M_r \) measurement shots. Moreover, many choices of measurement settings are usually required for complex measurement tasks. Such slowness of parameter estimation can also turn into imprecision in the estimate if the parameters of interest drift as a function of time, broadening spectroscopic signatures.

Imprecision in system characterization is particularly problematic for quantum information processing, i.e., for the initial calibration of the system and maintaining it. We use a controlled qubit to characterize the uncertain frequency of another mode \( \omega_r \) that is coupled to the qubit with uncertain coupling strength \( g \). Our algorithm employs modern Bayesian inference techniques to choose informative experimental settings while remaining computationally feasible. These modern inference techniques are able to use each bit of data obtained from experiment, instead of inferred expectation values only. This is both faster and more precise than the conventional approach involving ensemble averages. We demonstrate that our approach is robust against experimental imperfections.

Bayesian inference has been applied to identify qubit Hamiltonians in a standard tomographic setting using fixed, evenly spaced waiting times \( t \) between preparation and measurement [3]. In extracting an unknown frequency of a qubit, significant advantage can be achieved through an adaptive algorithm [4, 5] that updates the measurement setting during the experimental data collection and tends to choose exponentially increasing, rather

![Figure 1: (Color online) (a) Theoretically obtained swap-spectrum in frequency – waiting time plane. Here \( \delta = \omega_q - \omega_{r,0} \), with \( \omega_q \) the qubit frequency, \( \omega_{r,0} \) the resonator frequency, and \( g_0 \) the coupling strength. The color scale of the swap-spectrum represents the probability of the qubit being in its excited state. (b) Outcomes of a set of simulated single-shot measurements; blue: ground state, red: excited state.](image-url)
than evenly spaced, waiting times. Such a problem is equivalent to extracting \( g \) in a model discussed below when the frequency \( \omega_r \) is known. The assumption that \( \omega_r \) is known, however, limits the applicability of the model. The situation where both \( g \) and \( \omega_r \) are initially unknown is more widely and practically applicable. To solve the latter problem, we deliver a strategy that chooses both \( t \) and the qubit frequency \( \omega_q \) to achieve near-optimal scaling of errors in estimates.

A qubit coupled to a resonator is described by the Jaynes-Cummings Hamiltonian

\[
\hat{H}_{JC} = \frac{\hbar \omega_q}{2} \hat{\sigma}_z + \hbar \omega_r \left( \hat{a} \hat{a} + \frac{1}{2} \right) + \hbar g \left\{ \hat{\sigma}_+ \hat{a} + \hat{\sigma}_- \hat{a}^\dagger \right\},
\]

where \( g \ll \omega_r \). This model is broadly applicable to potential quantum computing technologies \[6\] and reproduces also the dynamics of spurious two-level systems \[7–9\], a notorious source of decoherence for superconducting qubits.

The standard method to estimate \( g \) and \( \omega_r \) is called swap-spectroscopy \[10\]. One starts by preparing the qubit in the excited state and the cavity in its ground state (typically by moving \( \omega_q \) far away from \( \omega_r \) and then exciting it with an external pulse). Then \( \omega_q \) is fixed to a chosen value that in Fig. 1 determines the horizontal coordinate (see the caption). The system is allowed to evolve a time \( t \) (vertical coordinate of Fig. 1) after which it is measured in the \( \hat{\sigma}_z \) basis. The system is then reset to its ground state before the next measurement.

As the Jaynes-Cummings Hamiltonian conserves the total number of excitations \( \hat{N} = \hat{a} \hat{a}^\dagger + \frac{1}{2} \) we can describe the resulting dynamics in the single-excitation subspace by the Hamiltonian \( \hat{H}' = \frac{\Delta \omega}{2} \hat{\eta}_z + g \hat{\eta}_x \) where \( \hat{\eta}_z \) are Pauli matrices and the detuning \( \Delta \omega = \omega_q - \omega_r \). This form of the relevant Hamiltonian occurs, in addition to the applications mentioned above, also in settings in magnetic resonance spectroscopy. For the Hamiltonian \( \hat{H}' \) the probability of the qubit being in its excited state is

\[
P_{t, \omega_q}(1|g, \omega_r) = \frac{1}{2} \left( 4 g^2 \frac{\cos \omega_R t}{\omega_R^2} + 1 + \frac{\Delta \omega^2}{\omega_R^2} \right),
\]

with \( \omega_R = \sqrt{\Delta \omega^2 + 4 g^2} \). The Supplementary Material generalizes this formula to account for qubit relaxation.

In conventional swap-spectroscopy, the measurement is repeated at a fixed setting \( (\omega_q, t) \) in order to establish an ensemble to calculate the relative frequency of the excited state and hence approximate its probability. To estimate \( g \), usual swap-spectroscopy first seeks \( \omega_q \) where \( \Delta \omega = 0 \), and where the Chevron pattern of the oscillating excitation probability has a maximum visibility, cf. Fig. 1. Measuring the angular frequency of these oscillations then yields \( 2g \) by Eq. \( 2 \).

Our algorithm, rather than establishing ensemble averages, chooses a new pair \( (\omega_q, t) \) after each measurement

\[
P(g, \omega_r|d) = \frac{P(d|g, \omega_r) P(g, \omega_r)}{P(d)}, \quad d = 0, 1
\]

for our system. This formula can be understood as a rule for iterative learning of the parameters. One starts with an initial probability distribution, or prior, \( P(g, \omega_r) \) that describes one’s \( a \ priori \) conception about the uncertain parameters. Based on the measurement outcome \( d \), c.f. Fig. \( 2b \), one can use Bayes’ theorem to update the probability distribution into a \( posterior \) based on the \( likelihood \) of the data \( P(d|g, \omega_r) \) according to conjectured model parameters \( g \) and \( \omega_r \). The normalization factor in the denominator can be calculated via the integral \( P(d) = \int P(d|g, \omega_r) P(g, \omega_r) dg d\omega_r \). The estimate is obtained from the mean value of the \( posterior \). The \( posterior \) is then identified as the prior for the next measurement. Thus each measurement outcome is immediately incorporated into our knowledge of the system.

We now want to optimize the measurement settings based on the current knowledge about \( \omega_r \) and \( g \). In prin-

![Figure 2](image-url)
principle, for each adaptation step one can maximize utility (e.g., a negative trace of the posterior covariance matrix or information gain $I$) of the next shot as if it were the last in the series, a so-called “greedy” algorithm. Computationally however, maximizing utility between the shots is quite unwieldy. Moreover, to optimize the whole series of measurement shots, i.e., to optimize globally, it is not sufficient to optimize greedily, i.e., locally. We have therefore analyzed offline exemplary instances of information gain landscapes $I(t,\omega)$ to develop a global strategy that we then tested. These considerations suggest a strategy in which $t \sim 1/\sigma_g$ and $\omega_q \sim \mu_\omega + \sigma_\omega$. With respect to the prior, $\mu_\omega$ is the mean of $\omega_r$ and $\sigma_\omega$ is the standard deviation of $g(\omega_r)$. Our measurement strategy chooses the $M^{th}$ measurement setting $(t,\omega_q)$ according to the rule

$$
\begin{align*}
t &= \begin{cases} 
\frac{a(r_1 - \frac{1}{2})}{|a + b|} & \text{if } M \leq M_0 \\
\frac{\mu_\omega + c(r_2 - \frac{1}{2})\sigma_\omega}{\sigma_g} & \text{if } M > M_0,
\end{cases} \\
\omega_q &= \begin{cases} 
\mu_\omega + \frac{1}{2} & \text{if } M \leq M_0 \\
\mu_\omega + c(r_2 - \frac{1}{2})\sigma_\omega & \text{if } M > M_0.
\end{cases}
\end{align*}
$$

Here, $a = 1.57$, $b = 0.518$, and $c = 3.0$ are numerical constants that we have found to yield a robust and efficient strategy. Furthermore, $z$ is a standard normal deviate and $r_{1,2}$ are uniform random variables on the interval $[0,1]$. This strategy is the central result of our paper. The measurement settings with $M \leq M_0$ are chosen more uniformly to obtain a unimodal posterior in the beginning of the series which makes the strategy more effective. We choose $M_0 = 15$ which we found to be sufficient especially for our prior and the parameters considered in Figs. 3, 4a, and 4b. Figure 2b illustrates an exemplary trace of measurement shots chosen adaptively according to Eq. (4).

A numerical challenge with Bayesian inference techniques is that each application of Bayes’ theorem requires the evaluation of the computationally expensive integral $P(d)$. To calculate the integrals we adopt a sequential Monte-Carlo approach [12,13] with moving grid points or “particles” whose density describes the probability distribution of interest [10]. In evaluating a probability distribution in fixed grid points, the density of the grid points would limit the precision on the estimates, which is mitigated by the adaptive grid. We perform the calculations with 50 000 particles, so for instance in computing $P(d)$ one evaluates 50 000 times the likelihood function.

To consider the performance of our algorithm, we have applied it to ensembles of 10 000 simulated samples with randomly chosen parameters $(g_0,\omega_r,0)$. Here the subscript 0 denotes a specific fixed true value, in contrast to the symbols naming a quantity. The values $g_0$ have been chosen from a log-normal distribution with the mean $\mu_g$ and the standard deviation $\sigma_g$ while $\omega_r,0$ have been chosen from the normal distribution $\mathcal{N}(\mu_\omega,\sigma_\omega)$. For each sample we have chosen the initial prior of the $(g,\omega_r)$ estimate to coincide with the probability distribution from which the true values $(g_0,\omega_r,0)$ are randomly picked. Unless specified otherwise we have chosen $\sigma_g = 0.25 \mu_g$ and $\sigma_\omega = \mu_\omega$.

Figure 3 exhibits the relative median squared error as a function of measurement shots for $g$ (solid) and $\omega_r$ (dotted). The changes in parameters represent the effects of relaxation and noise: (i) $T_1 = \infty$, $P_c = 0$ (black), (ii) $T_1 = \infty$, $P_c = 0.1$ (blue), (iii) $T_1 = 2000\pi$, $P_c = 0.1$ (red), and (iv) $T_1 = 40\pi/P_\sigma$, $P_c = 0.1$ (brown). Here $T_1$ is the relaxation time, $P_c$ is the probability for a readout error, and $\mu_g$ the mean of $g$ over the ensemble.

Table I: Number of outliers per 10 000 simulated samples with $T_1 = \infty$, $P_c = 0$. Rows correspond to the number of outliers with squared error larger than $\hat{E}_2^g$ after a given number of measurement shots (indicated by the columns).

| $\hat{E}_2^g$/Shots | 150 | 300 | 600 | 1200 | 3600 | 6600 |
|---------------------|-----|-----|-----|------|------|------|
| $10^{-6}$           | 533 | 466 | 276 | 25   | 3    | 0    |
| $10^{-5}$           | 265 | 251 | 111 | 18   | 2    | 0    |
| $10^{-4}$           | 118 | 116 | 25  | 14   | 1    | 0    |
robust against such outliers through repetition as follows. After 300 measurement shots we set the prior widths back to their original values but keep the mean of the probability distribution unchanged. Another 300 measurement shots are performed thereafter. We then compare the estimates after 300 and 600 measurement shots. If their difference is smaller than a set threshold, we conclude that we have found a correct estimate, otherwise we start a new search of the estimate. For the new search we choose a prior whose mean values for $g$ and $\omega_r$ are randomly chosen from the original prior while the prior widths equal those of the original prior. Table I summarizes the performance of our outlier correction scheme. Outliers are defined as the samples with the squared error of the $g$ estimate larger than threshold $\bar{\xi}_g^2$. Our scheme appears to reduce the number of outliers with an acceptable overhead.

Figures 4 exhibit the average number of measurement shots required to meet the desired level of relative mean squared error $E_g^2 = 10^{-4}$ (black), $10^{-7}$ (red), and $10^{-10}$ (blue). The number of measurements presented as a function of (a) probability of readout error $P_e$, (b) relaxation time $T_1$, (c) standard deviation $\sigma_g$ over the initial prior, and (d) standard deviation $\sigma_\omega$ over the initial prior.

In conclusion, by adaptively focusing measurements on the regions of high information gain we obtained a globally efficient measurement strategy for two control parameters. Our algorithm makes advanced spectroscopy drastically more efficient.

We thank D. Sank, J. Kelly, and J.M. Martinis for numerous discussions. F.W.M. acknowledges early discussions with C. Bender, H. Zähle, and M. Hein. Work supported by the European Union through ScaleQIT as well as by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), through the Army Research Office. All statements of fact, opinion or conclusions contained herein are those of the authors and should not be construed as representing the official views or policies of IARPA, the ODNI, or the U. S. Government.

[markku.stenber@iki.fi]

[1] M. H. Levitt, *Spin Dynamics: Basics of Nuclear Magnetic Resonance* (Wiley, Chichester, 2008), 2nd ed.

[2] W. Demtröder, *Laser Spectroscopy*, Vol. 1 (Springer, Berlin, 2008), 4th ed.

[3] S. G. Schirmer and D. K. L. Oi, Phys. Rev. A 80, 022333 (2009).
[4] A. Sergeevich et al., Phys. Rev. A 84, 052315 (2011).
[5] C. Ferrie, C. E. Granade, and D. G. Cory, Quant. Inf. Proc. 12, 611 (2013).
[6] S. Haroche and J.-M. Raimond, Exploring the Quantum: Atoms, Cavities, and Photons (Oxford University Press, New York, 2006); L. S. Bishop et al., Nat. Phys. 5, 105 (2009); A. Imamoğlu et al., Phys. Rev. Lett. 83, 4204 (1999).
[7] M. Neeley et al., Nat. Phys. 4, 523 (2008).
[8] T. Palomäki et al., Phys. Rev. B 81, 144503 (2010).
[9] Y. Shalibo et al., Phys. Rev. Lett. 105, 177001 (2010).
[10] M. Mariantoni et al., Science 334, 61 (2011).
[11] D. Sivia and J. Skilling, Data Analysis: A Bayesian Tutorial (Oxford University Press, New York, 2006).
[12] M. West, J. R. Statist. Soc. B 55, 409 (1993).
[13] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, Radar and Signal Processing IEE Proc. F 140, 107 (1993).
[14] J. Liu and M. West in, Sequential Monte Carlo Methods in Practice, edited by A. Doucet, N. Freitas, and N. Gordon (Springer, New York, 2001).
[15] C. E. Granade et al., New J. Phys. 14, 103013 (2012).
[16] We modify the algorithm, e.g., by resampling particles from log-normal rather than normal [c.f. Eq. (3.6) in [14]] distribution to avoid negative values of $g$.
[17] D. Sank, J. Kelly, and J. Martinis, private communication.
**Supplementary Information**

The dynamics of the qubit-resonator system is governed by the Master equation
\[ \dot{\hat{\rho}}_S = -i \left[ \hat{H}, \hat{\rho}_S \right] + \mathcal{D} \left[ \hat{A} \right] \hat{\rho}_S, \] (1)
with \( \hat{\rho}_S \) the density matrix in the Schrödinger picture. In the absence of relaxation the system is described by the Jaynes-Cummings Hamiltonian
\[ \hat{H} = \hat{H}_0 + \hat{H}_1, \]
\[ \hat{H}_0 = \frac{E_r}{2} \hat{\sigma}_z + \frac{E_q}{2} \hat{\tau}_z, \quad \hat{H}_1 = \hbar g (\hat{\sigma}_+ \hat{\tau}_- + \hat{\sigma}_- \hat{\tau}_+). \] (2)
Here \( E_q = \hbar \omega_q \) and \( E_r = \hbar \omega_r \), are the bare qubit and resonator energies, respectively, and the qubit-resonator coupling is characterized by \( g \). Furthermore, \( \hat{\sigma}_z (\hat{\tau}_z) \) are Pauli matrices in qubit (resonator) subspace, whereas \( \hat{\sigma}_z \) and \( \hat{\tau}_z \) are raising and lowering operators. We focus on energy relaxation of the qubit as the main channel of loss, described by a Lindblad form
\[ \mathcal{D} \left[ \hat{A} \right] \hat{\rho}_S \equiv \hat{A} \hat{\rho}_S \hat{A}^\dagger - \frac{1}{2} \left\{ \hat{A}^\dagger \hat{A}, \hat{\rho}_S \right\} \] (3)
with the amplitude damping Lindblad operator
\[ \hat{A} = \sqrt{\Gamma} \hat{\sigma}_- \] (4)
where \( \Gamma \) is the relaxation rate.

We would first like to show that we can drop the state \(|11\rangle\) in the equations from the very beginning. We switch to the interaction picture with the Schwinger-Tomonaga equation
\[ \dot{\hat{\rho}}_I = -i \left[ \hat{H}_I, \hat{\rho}_I \right] + \mathcal{D} \left[ \hat{A}_I \right] \hat{\rho}_I. \] (5)
It is easy to show that \( \hat{H}_I |11\rangle = 0 \). Moreover, \( \hat{A}_I = e^{-i E_r t} \hat{A} \) implies \( \hat{A}_I \hat{\rho}_I \hat{A}_I^\dagger |11\rangle = 0 \). Thus all the time-dependence left is
\[ \langle 11 | \dot{\hat{\rho}}_I | 11 \rangle = -\Gamma \langle 11 | \hat{\rho}_I | 11 \rangle. \] (6)
Therefore if the initial state does not include a \(|11\rangle\) component, that state stays unoccupied.

Below we assume that the initial state at \( t = 0 \) is \(|10\rangle\). As the Jaynes-Cummings Hamiltonian preserves the number of excitations we can describe the resulting dynamics in the single-excitation subspace by the Hamiltonian \( \hat{H}' = \frac{\Delta \omega}{2} \hat{\eta}_z + \hbar \omega_R \). Here \( \hat{\eta}_z \) are Pauli matrices and \( \Delta \omega = \omega_q - \omega_r \) is the detuning frequency. The eigenstates of \( \hat{H}' \) are
\[ |e\rangle = \cos \frac{\theta}{2} |10\rangle + \sin \frac{\theta}{2} |01\rangle, \]
\[ |g\rangle = -\sin \frac{\theta}{2} |10\rangle + \cos \frac{\theta}{2} |01\rangle, \] (7)
where \( \tan \theta = \frac{2 \hbar \omega_R}{\Delta \omega} \). The eigenenergies are \( E = \pm \hbar \omega_R \), with \( \omega_R = \sqrt{\omega^2 + 4 g^2} \). Note that \( \Delta \omega = 0 \) corresponds to \( \theta = \pi/2 \).

To treat relaxation, we now express the qubit lowering operator in these eigenstates as
\[ \hat{\sigma}_- = |0\rangle \langle 1| \otimes \hat{1} \]
\[ = |00\rangle \left( \cos \frac{\theta}{2} |e\rangle - \sin \frac{\theta}{2} |g\rangle \right) + \]
\[ + \left( \sin \frac{\theta}{2} |e\rangle + \cos \frac{\theta}{2} |g\rangle \right) \] (8)
Our approach is to perform a rotating wave approximation (RWA), recognizing that \( \omega_R/2 \geq \hbar g \gg \Gamma \). To carry out the transformation, we first introduce a further interaction picture where the perturbation now includes only the relaxation
\[ \dot{\hat{\rho}}_I = \mathcal{D} \left[ \hat{A}_r \right] \hat{\rho}_I. \] (9)
In this picture the lowering operator in the qubit subspace is
\[ \hat{\sigma}_-^{(I')} = e^{-i (E_q + E_r) t/2 \hbar} |00\rangle \left( \cos \frac{\theta}{2} e^{-i \omega_R t/2} |e\rangle \right. \]
\[ \left. - \sin \frac{\theta}{2} e^{i \omega_R t/2} |g\rangle \right) + e^{-i (E_q + E_r) t/2 \hbar} \]
\[ \times \left( \sin \frac{\theta}{2} e^{i \omega_R t/2} |e\rangle + \cos \frac{\theta}{2} e^{-i \omega_R t/2} |g\rangle \right); \] (10)
We can now drop all the terms involving the state \(|11\rangle\) using the same argument as in the context of Eq. (6). Hence we obtain
\[ \langle e | \hat{\rho}_I | e \rangle = -\cos^2 \frac{\theta}{2} \langle e | \hat{\rho}_I | e \rangle, \]
\[ \langle g | \hat{\rho}_I | g \rangle = -\sin^2 \frac{\theta}{2} \langle g | \hat{\rho}_I | g \rangle, \]
\[ \langle e | \hat{\rho}_I | g \rangle = -\frac{\Gamma}{2} \langle e | \hat{\rho}_I | g \rangle. \] (11)
We would now like to calculate the evolution of the population of \(|10\rangle\) starting in that state. The initial density matrix is
\[ \hat{\rho}_r(t = 0) = |10\rangle \langle 10| \]
\[ = \cos^2 \frac{\theta}{2} |e\rangle \langle e| + \sin^2 \frac{\theta}{2} |g\rangle \langle g| - \frac{1}{2} \sin \theta \langle e| \langle g| + |g\rangle \langle e|. \]
(13)
Using Eqs. (9) and (12) we find that the state decays according to
\[ \hat{\rho}_r(t) = \cos^2 \frac{\theta}{2} e^{-\frac{\Gamma}{2} \cos^2 \theta/2} |e\rangle \langle e| + \sin^2 \frac{\theta}{2} e^{-\frac{\Gamma}{2} \sin^2 \theta/2} |g\rangle \langle g| \]
\[ - \frac{1}{2} \sin \theta e^{-\frac{\Gamma}{2} \sin^2 \theta/2} \langle e| \langle g| + |g\rangle \langle e|. \] (14)
The projection operator $\hat{\Pi} = |10\rangle \langle 10|$ corresponding to the measurement is in the interaction picture

$$\hat{\Pi}_I(t) = \cos^2 \frac{\theta}{2} |e\rangle \langle e| + \sin^2 \frac{\theta}{2} |g\rangle \langle g|$$

$$- \frac{1}{2} \sin \theta \left( |e\rangle \langle g| e^{i\omega_R t} + |g\rangle \langle e| e^{-i\omega_R t} \right).$$

(15)

Since the trace of an operator is the same in the interaction and the Schrödinger pictures we find the occupation probability of the state $|10\rangle$

$$P_{t,\omega_R} (1|g,\omega) = \text{Tr} \left( \hat{\Pi} \hat{\rho} \right) = \frac{(1 + \cos \theta)^2}{4} e^{-\Gamma t(1+\cos \theta)/2}$$

$$+ \frac{(1 - \cos \theta)^2}{4} e^{-\Gamma t(1-\cos \theta)/2} + \frac{\sin^2 \theta}{2} e^{-\Gamma t/2} \cos \omega_R t.$$  

(16)

Using $\tan \theta = \frac{2g}{\Delta \omega}$ and $T_1 = \frac{1}{\Gamma}$ this can be recast as

$$P_{t,\omega_R} (1|g,\omega) = \left( \frac{\omega_R + \Delta \omega}{2\omega_R} \right)^2 e^{-(\omega_R + \Delta \omega)t/2\omega_R T_1}$$

$$+ \left( \frac{\omega_R - \Delta \omega}{2E} \right)^2 e^{-(\omega_R - \Delta \omega)t/2\omega_R T_1} + \frac{2g^2}{\omega_R} e^{-t/2T_1} \cos \omega_R t.$$  

(17)

This generalizes Eq. (2) in the main text for finite relaxation.