

Automated analysis of single-tone spectroscopic data of cQED systems

G.Fedorov^{1, 2, a)} and A. Ustinov^{1, 3}¹⁾*Russian Quantum Center, Skolkovo village, Russia*²⁾*Moscow Institute of Physics and Technology, Dolgoprudny, Russia*³⁾*Karlsruhe Institute of Technology, Karlsruhe, Germany*

(Dated: 24 December 2018)

FiXme Fatal: Is abstract supposed to be re-written completely? To build a full-scale quantum processor it is necessary to automate as many steps as possible on the physical, hardware level. Circuit quantum electrodynamics (cQED) is a contemporary architecture for dispersive readout and Purcell protection of superconducting qubits of various types, and thus it is necessary to develop software that is able to perform every kind of automatic calibration of such systems from scratch without any human participation. An important step towards this goal is to build a noise-insensitive and accurate computer vision tool to process three-dimensional spectroscopic data. In this work, we present and describe two scalable algorithms that are able to extract the Hamiltonian parameters of the cQED systems from spectroscopic data.

I. INTRODUCTION

Computation on a quantum computer involves operating large numbers of physical quantum bits (qubits). One of the most significant challenges is that qubit systems are not produced identical and cannot be operated identically^{1,2}: typically, control parameters for a device are determined by following a sequence of calibration experiments, and vary among the devices. It is already understood that manual tuning of these parameters is not a viable solution and operating a few dozens of qubits requires developing an autonomous system able to handle the physical device calibration on its own. Moreover, a lot of research is still being done to find the optimal chip designs and fabrication methods. This research includes gathering statistics on many different samples that should be measured and consistently evaluated. An automated tool not only will speed up this process, but as well exclude human error.

In this work, we are proposing and implementing several methods of data processing and computer vision that aid automatic calibration of circuit QED³ architectures used to read out superconducting qubits' states. Computer vision is understood here as an enterprise that uses statistical methods to disentangle data using models constructed with the aid of geometry, physics, and learning theory⁴. As it will be shown below the output of circuit calibration is almost always an image and it is hardly possible to automatically extract the required information from it without resorting to methods of computer vision.

The proposed methods are accurate, fast and robust to noise, applicable to superconducting qubits with a periodic spectrum described with a closed form model equations, and compatible with the paradigm of one readout resonator per one qubit^{5,6}.

The flow of an experiment that characterizes a single qubit-resonator cell of a quantum processor chip is depicted

schematically in Fig. 1(a). It is similar to the procedures described in Refs. [2 and 7] for measuring cQED samples with qubits of tunable frequency. For non-tunable qubits, a different algorithm of automatic calibration is described in the patent application Ref. [8].

First of all, it is necessary to locate the resonance peaks corresponding to the readout resonators. The resonator search step then provides the scan range for the probe frequency Δf_p . **FiXme Note: Δ should not be used to denote a range here. Furthermore in the article Δ is used at least in three different meanings: (1) a range (2) a reminder (3) a standardized sample(i.e. with zero mean - see note below) - this should be fixed!** in the next step: single-tone spectroscopy (STS). STS tracks the frequency shifts of the resonance peak while changing the magnetic field applied to the sample (controlled by some current I). By processing the results of the STS, one can get a coarse estimate of the qubit frequency depending on I denoted as $f_{ge}^{(0)}(I)$ and accurate estimates of the resonator frequency $f_r(I)$ and the coupling strength g . Based on these results, one may set the scan range for the two-tone spectroscopy (TTS) in excitation frequency Δf_{exc} and current ΔI . Finally, by processing TTS result, one can obtain the precise dependence of the qubit ge transition frequency on current $f_{ge}(I)$, and other transition frequencies ($f_{gf/2}(I)$, etc.) with accuracy sufficient for pulsed experiments. In overall, the process is structured so that the results of one measurement step define the parameters of the next step, successively acquiring more accurate information about the physical system.

The measurement outcomes that contain information about the system properties may be divided in two groups: some results contain curves (data are 1D manifolds), and others contain heatmap images (2D manifolds). The scope of this work is confined to the automatic analysis the latter results; precisely, we are developing a computer vision system that is capable of extracting analytic data from STS images. In Fig. 1(b) we illustrate two types of correlated measured data and the theoretical model: STS data has to be interpreted either using two discontinuous curves (upper row) or a single continuous curve (lower row) that are defined by the model.

^{a)}Electronic mail: gleb.fedorov@phystech.edu

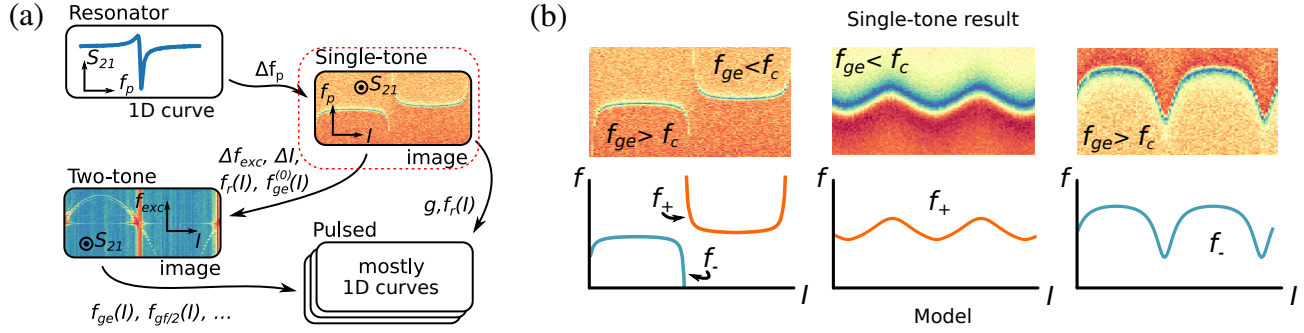


FIG. 1. (a) Common experiment flow for a tunable-frequency qubit: usually, single- and two-tone spectroscopic measurements are performed before pulsed experiments, and each next experiment is based on the results of the previous ones. In the red dashed frame, the scope of this work is shown. (b) Examples of the desired results after the detection for three types of qualitatively different STS outcomes. The heatmaps should be brought into correspondence with the 2D model curves. In the first column so called avoided crossing pattern is shown. Two model curves has to be fitted simultaneously since the model is discontinuous. Conversely, the second and third images may be described by a single continuous curve; however, the model equations are different for these two cases.

The implementation of our method depends on the theoretical description of the cQED systems and physical qubits; therefore, in Appendices A and B we derive all the necessary equations to form the model curves that are expected to appear in single-tone spectroscopy. Without loss of generality, we use the transmon⁹ with an asymmetric SQUID (superconducting quantum interference device) as a qubit; however, since our methods do not depend on the particular shape of the qubit transitions, for other types of qubits the logic will stay the same and there is no loss of generality.

For the single-tone spectroscopy, the model curves are (see Fig. 1(b)):

$$f_{\pm}(I) \equiv f_r(I) = \frac{f_c + f_{ge}(I)}{2} \pm \sqrt{g^2 + (f_{ge}(I) - f_c)^2/4}, \quad (1)$$

where f_c stands for current-independent bare cavity frequency. Here we note that for the avoided crossings pattern both curves are necessary; alternatively, for the cases when the qubit is entirely below (above) the bare cavity frequency only $f_+(I)$ ($f_-(I)$) is used. The equations showing the dependence of the qubit transitions $f_{ge}(I)$ on current are expressed as

$$f_{ge}(I) = f_{ge}^{max} \left[\cos^2 \left(\frac{\pi(I - I_{ss})}{\Pi} \right) + d^2 \sin^2 \left(\frac{\pi(I - I_{ss})}{\Pi} \right) \right]^{\frac{1}{4}}, \quad (2)$$

where Φ_0 is the flux quantum, d is the SQUID asymmetry and f_{ge}^{max} is the qubit frequency at $I = I_{ss}$. I_{ss} (sweet spot current) is the current exactly compensating the non-zero residual flux that is always present in experiment coming from the local magnetic fields on the chip. Finally, Π is the period of the spectrum in current.

In overall, we have 6 fitting parameters for the single-tone spectroscopy (f_c , g , Π , I_{ss} , f_{ge}^{max} , d). Since they are derived from the Hamiltonian for the compound cavity-qubit system as described in Appendix A hereafter the parameters are collectively called Hamiltonian parameters.

The structure of the rest of the paper is as follows. First, we describe our methods to extract the Hamiltonian parameters

from the STS data. Next, the accuracy, performance and reliability of the algorithm is addressed; finally, we discuss the limitations of the algorithm, further applications and future work.

II. METHODS

In this section, we describe the approaches towards extraction of Hamiltonian parameters from the two types of spectroscopic measurement results. Additionally, we describe important peculiarities of the data itself and some essential experimental details.

In Fig. 2(a), we show as an example a full experimental result obtained by STS for a notch-type (side-coupled) coplanar waveguide resonator coupled to a tunable transmon qubit. There, the absolute value of the complex scattering parameter S_{21} between ports 1 and 2 of the sample is shown in colour, and the axes show corresponding the probe frequency f_p and current values for which S_{21} was recorded. We note that depending on the sample configuration one may measure, in general, any scattering parameter S_{ij} . An important fact is that for other kinds of the readout resonators the phase $\angle S_{ij}(f_p)$ should be used instead of the amplitude $|S_{21}|$. While the amplitude has a local minimum at resonance, the phase would have a maximum of its derivative.

To extract the Hamiltonian parameters, we propose to fit the model (1) to the data in Fig. 2. **FiXme Error: Change y-axis label from f_p to f_r to on Fig. 2 (c) and x-label from coil current to I . It is enough to have in the figure comments that I denotes coil current(a) in the sense of maximum likelihood.** However, as one can see, the data are more complex than the model **FiXme Note: At this point of text the possible linkage between the experimental data and the model is not clear enough (apart from some visual similarities) as they contain an extra dimension: the model is a one-dimensional**

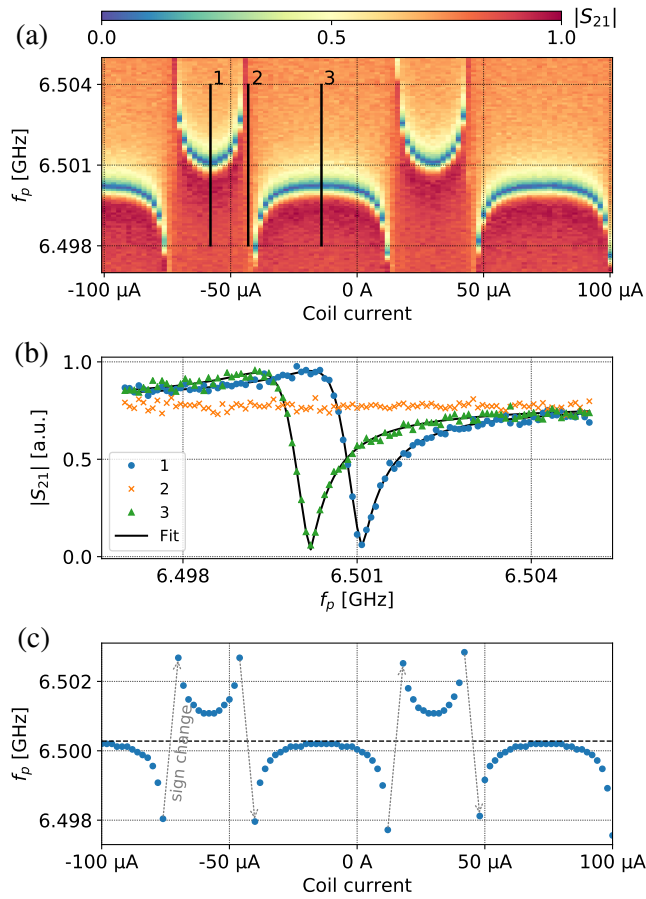


FIG. 2. (a) An experimental spectrum of a resonator strongly coupled to a transmon qubit depending on the coil current I . (b) Slices of the transmission from (a) showing two slices with fits (1,3) and a plateau with no dip (2) present. (c) Extracted data (blue dots) and mean frequency value over all points $\langle f_r \rangle_I$ (black dashed line). Grey arrows show where $f_r - \langle f_r \rangle_I$ changes sign.

manifold while the data have two dimensions, and it is then not clear how to compare them. We solve this problem by a special dimensionality reduction **FiXme Note: dimensionality reduction typically means either feature selection or feature extraction in machine learning. Neither meaning appears to be applicable here, so would it be possible to chose another name for the procedure?** procedure (described below), but even after that the fitting is not a trivial task because the maximum likelihood estimation implicates global optimization of an ill-defined $loss$ **FiXme Error: You should probably talk about "likelihood" function rather than "loss" function - see note in "Full model fitting". A more ore less arbitrary loss functions are used for example in decision theory. This is hardly the case here - you should justify the application of a particular loss function** function. It is so, firstly, due to the periodic dependence of the frequencies on current I and the unknown position of the qubit sweet spot I_{ss} and, secondly, due to the strongly non-linear dependence of the model on other parameters. Fortunately, it is possible to get a good initial guess for Π and I_{ss}

parameters and then brute force search the solution in other variables since we have well defined bounds for each of them. Finally, we can polish the result in all six parameters using a local optimizer.

An outline of the method which uses Fig. 2(a) as an example is presented below. As it has already been mentioned, the chosen type of the resonator-qubit arrangement which yields the avoided crossings pattern is not unique: there are two other cases when the qubit spectrum lies entirely below or above the resonator frequency; however, they are treated exactly the same way and are simpler in terms of the loss function behaviour.

A. Extracting $f_r(I)$ from data

Firstly, we are going to reduce the dimensionality of the data, i.e. extract the resonance frequency for each I .

We do this by employing the *circfit*¹⁰ library which is capable of fitting various types of microwave resonators. For each I we fit the complex transmission $S_{21}(f_p)$ as in Fig. 2(b) (solid black lines) and extract the resonance frequency from the model. A possible caveat is that for some I values the resonance dip may disappear (see Fig. 2(b)) so the fit will fail. Therefore, such slices are excluded in advance via a threshold condition.

The resulting plot of the extracted resonator frequency versus current I is shown in Fig. 2(c) (blue dots). There are some current values located between the branches where the data points are missing, as expected, due to the absence of the resonance. Additionally, we plot here the mean value of the detected frequencies shown as a dashed black line. This parameter is important since it, will firstly, serve as an initial guess for the cavity frequency of the model (1): $f_c \approx \langle f_r \rangle_I$ and, secondly, will be used in the period and phase extraction algorithm which tracks the changes of the sign of the value $\Delta f_r = f_r - \langle f_r \rangle_I$ (marked as grey dashed lines in Fig. 2(c)). Locating these sign changes allows to find the qubit sweet spot without fitting the full model.

B. Extracting period, phase and sweet spot locations

As we have stated above, one of the serious obstacles for the fitting is the periodicity of the data on one of the fitting parameters. Along with the equally unknown phase of the signal, this leads to the presence of many local minima in the loss function which impede the progress of iterative optimization algorithms. In other words, the unknown parameters Π and I_{ss} are preventing us from finding the global minimum. Fortunately, it is possible to determine the period and the sweet spot location (or, the phase of the data) without fitting the full model (1) which alleviates the stated problem. In the absence of noise, this would be a trivial task. However, in the conditions of experiment, the noise to some extent is always present, so below we suggest a solution that is insensitive to local sporadic perturbations of the data.

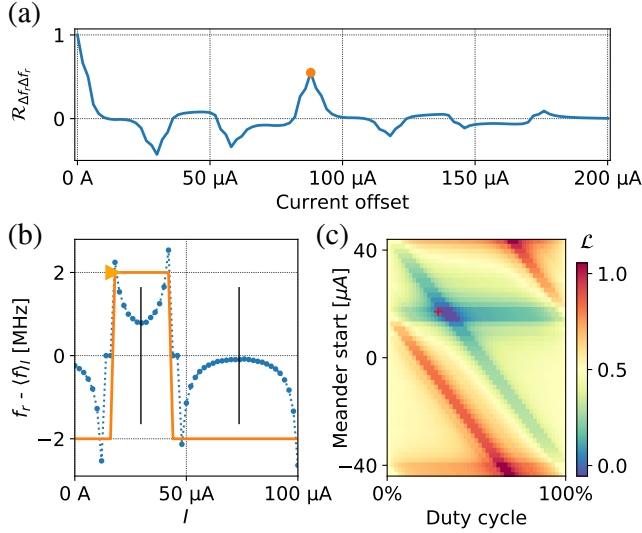


FIG. 3. Period and phase extraction procedure. (a) Autocorrelation function depending on the current offset ΔI shows a prominent local maximum at $88 \mu\text{A}$ (orange dot). (b) Construction for phase estimation; $\Delta f_r = f_r - \langle f_r \rangle_I$ (blue) is fitted with a square wave (orange, start marked with a triangle). Vertical bars mark candidate sweet spots. (c) Loss function (normalized) for the square wave fitting procedure from (b); red cross indicates the parameters of the meander shown there.

A powerful tool for finding the period in a given dataset y (especially when it contains just a few periods) is the autocorrelation function $\mathcal{R}_{yy}(l) = \sum_n y_n y_{n-l}$. The location of the largest of its local maxima (except for the $l = 0$) equals exactly the sought period. **FiXme Note: Probably close to be true but the reference to the proof should be provided** However, this is only true when the mean value of the function is zero; otherwise, $\mathcal{R}_{yy}(l)$ would be linear with a steep slope that may smear out all the extrema. Therefore, instead of directly calculating the autocorrelation of $f_r(I)$ we consider the function $\Delta f_r(I)$ introduced above which has zero mean and the same period. **FiXme Note: It is a standard procedure called ... "standardizing the data" (make the sample has zero mean and unit variance) - you should refer to it** For this function, which looks exactly like one in Fig. 2(c) but centered vertically around zero, the autocorrelation function $\mathcal{R}_{\Delta f_r \Delta f_r}(\Delta I)$ depending on the current offset ΔI is shown in Fig. 3(a). As one can see, ΔI spans $200 \mu\text{A}$ just as the data itself. This means that to calculate $\mathcal{R}_{\Delta f_r \Delta f_r}(\Delta I)$ the data is being zero-padded at all ΔI except for $\Delta I = 0$, and this is why we get diminishing correlation peaks at Π , 2Π , etc. The orange dot in the plot shows the highest local extremum of $\mathcal{R}_{\Delta f_r \Delta f_r}(\Delta I)$ at $88 \mu\text{A}$. It is a very prominent peak and can be easily distinguished among all others. There is as well a small peak at $176 \mu\text{A}$ which corresponds to $\Delta I = 2\Pi$. Note that the autocorrelation function is not ideally smooth and has some abrupt bends on the sides of the peaks. This happens because there were some missing points in the data (corresponding to the plateaus of Fig. 2(b)) that were replaced by zeros to ensure the correct mapping between the current indices and current

values. This zero padding may be noticed as well in Fig. 3(b).

Now, having found the period it is possible to precisely determine the phase of the signal. This is done via finding a global maximum of the zero-lag correlation function **FiXme Note: Why would finding of a global maximum of the zero-lag correlation function .. give the required parameter value? It appears to be an interesting idea and should be elaborated further** $\mathcal{R}_{\Delta f_r S}(0)$ between $\Delta f_r(I)$ and a square wave signal $S(I, \Pi, \phi, D)$ **FiXme Note: What is square wave signal?** having the same period but unknown phase ϕ and duty cycle D . The high and low levels of the square must be opposite in value, i.e. 1 and -1 **FiXme Note: Is the above mentioned wave signal complex? Is it $S_{12} \text{ squared?}$, and the absolute value does not matter.** An illustration of a square wave function satisfying the optimal condition is presented in Fig. 3(b) in orange. The phase ϕ (in μA) denotes the x-coordinate of the first point after the rising edge, and is marked with a triangle. Generally, the idea behind this is to robustly detect sign changes that were shown back in Fig. 2(c). We can not rely on a more simple algorithm that walks through the points and marks where the function changes its sign because such algorithm may fail in the presence of noise. Additionally, it will still work correctly even if the mean value $\langle f_r \rangle_I$ does not lie exactly between the branches and intersects one of them.

The global optimal ϕ and D are found using a brute force algorithm. It calculates the loss function $\mathcal{L} = -\mathcal{R}_{\Delta f_r S}(0)$ **FiXme Note: It should be explained why does minimization of this loss function lead to optimal parameters values.** on a 50×50 grid of (ϕ, D) and takes the minimal value of all. This method is stable and universal due to the evident boundaries on $\phi \in [-\Pi/2, \Pi/2)$ and $D \in [0, 1]$. The loss function topography for the avoided crossing patterns is nicely structured and for our example is shown in Fig. 3(c). One peculiarity is that instead of a single minimum it has an area of the same minimal value. Again, this effect comes from the missing zero-padded f_r points at some I values. However, any value from this valley suits well enough for our purposes, and the algorithm finds no difficulty in locating it.

Having found values for Π , ϕ and D we now can calculate the currents of the transmon sweet spots in the case of the avoided crossings pattern and the smooth patterns:

$$I_{ss} = \begin{cases} \phi + \Pi(1 + D)/2, & \text{for the avoided crossings} \\ \phi + \Pi D/2, & \text{otherwise} \end{cases} \quad (3)$$

To distinguish between these two cases when the noise is not too large, one may calculate the maximal absolute differential of the frequency data $\max_{i>0} |f_{r,i} - f_{r,i-1}|$ and compare it to the peak-to-peak amplitude $\max_{i,j} |f_{r,i} - f_{r,j}|$. For the avoided crossings these values are close and for the smooth dependencies they are not. However, in the presence of noise this indicator may fail, and one will have to check both current values to be I_{ss} by fitting the full model two times, and then choose among the two possibilities based on the loss value.

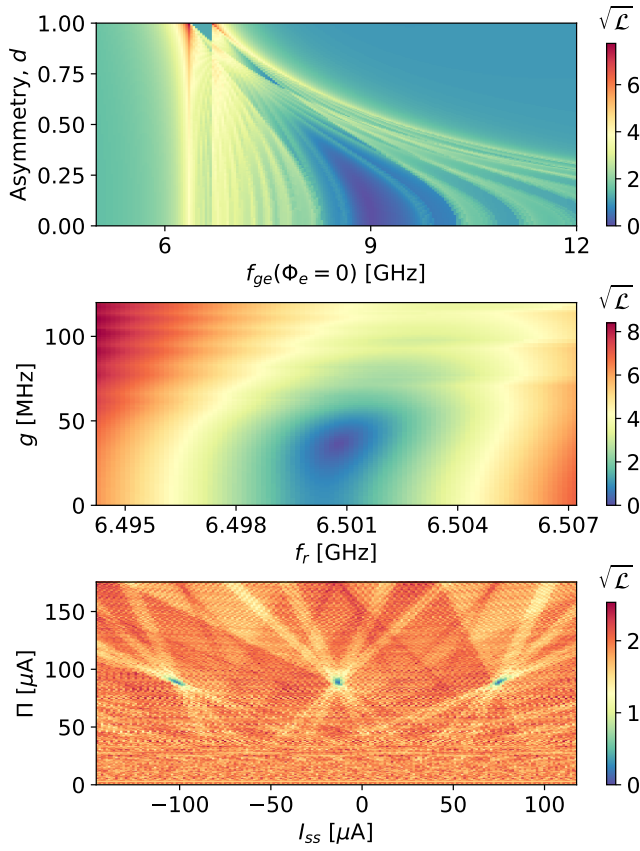


FIG. 4. Slices of the loss function (4) for the full model and experimental data from the example made around the optimal point. Root-mean-square per-point loss $\sqrt{\mathcal{L}}$ is in MHz. (a) f_{ge}^{max} and d are varied, other parameters optimal. A large valley is located near 9 GHz, and some smaller locally minimal ones are present all around. (b) g and f_r are varied, others optimal. The loss function for these two parameters is well-conditioned near the optimum. (c) Π and I_{ss} varied, others optimal. This subplot illustrates a complex structure of local minima around the true one which we find analytically.

C. Full model fitting

Having performed the aforementioned preliminary steps, it is now possible to fit the full model to the extracted points. To do this, we employ brute force optimization combined with the Nelder-Mead simplex downhill algorithm¹¹ to polish the brute force result. For both methods, we use a common loss function defined as follows. For the known probe frequency span of the data Δf_p (Fig. 2(a), whole y-axis) and the set of N extracted points $\{p_i\} = \{(I_i, f_{r,i})\}$, we calculate the loss function **FiXme Note: Essentially, you are minimizing log likelihood assuming gaussian noise - it should be explained**

Parameter	Value range	Steps
f_c	$\langle f_r \rangle_I \pm 1$ MHz	3
g	20 - 40 MHz	5
f_{ge}^{max}	4 - 12 GHz	80
d	0 - 0.9	9

TABLE I. Grid specifications for the brute force algorithm for STS detection.

in this way as

$$\mathcal{L} = \frac{1}{N} \sum_{i=0}^N [f_{r,i} - \mathcal{M}(I_i, \Pi, I_{ss}, f_c, g, f_{ge}^{max}, d)]^2, \quad (4)$$

$$\mathcal{M} = \begin{cases} f_+, & |f_+ - f_c| < \Delta f_p / 2 \\ f_-, & \text{otherwise,} \end{cases} \quad (5)$$

where, as in the previous section,

$$f_{\pm}(I_i) = f_{\pm}(I_i, \Pi, I_{ss}, f_c, g, f_{ge}^{max}, d)$$

The condition of Eq. (5) means that we choose only the model frequencies that lie within a window Δf_p around the model f_c parameter. This ensures, firstly, that in the optimum we will not take any excess points outside the frequency scan, and, secondly, that we have a single model value for each current.

To substantiate the choice of the optimization algorithms, we present in Fig. 4 three visualizations of the the defined loss function. The plots show how the function behaves if a certain pair of 6 model parameters is varied while others are optimal. From the plots it is obvious that the loss function is ill-defined and has a lot of local minima. Moreover, it may not be smooth everywhere because of the condition (5). The f_{ge}^{max} and d parameters present the most significant difficulty in terms of false minima as can be seen from Fig. 4(a). In contrast, f_c presents the least difficulty, as can be seen from Fig. 4(b). The last plot Fig. 4(c) shows a very complex structure of the loss function and narrow optimal valleys and serves as an illustration of why the period and phase extraction algorithm is important.

The brute force algorithm acts on the grid specified in Table I. The ranges in the grid are based on the usual design parameters of the qubit samples in our database and the number of steps is chosen so that the algorithm reliably finds the optimal valley. After the coarse brute force optimization is done and the optimal valley is located, we apply the Nelder-Mead search on all 6 parameters.

III. RESULTS

In this section, we discuss the performance and accuracy of the two algorithms described in the previous section. Firstly, we test the resonator spectrum recognition, and then the qubit spectrum detection algorithm.

The algorithm was implemented in Python using the *brute* and *minimize* routines of the SciPy¹² library. We present here three examples of the detection for all possible qubit-resonator

Parameter	(a)				(b)				(c)			
	Brute	N-M	\mathbb{H}^{-1}	Correct	Brute	N-M	\mathbb{H}^{-1}	Correct	Brute	N-M	\mathbb{H}^{-1}	Correct
f_c , GHz	6.5003	6.5007	3×10^{-5}	n/a	6.962	6.9631	1.2×10^{-3}	n/a	6.47	6.465	1.4×10^{-3}	n/a
g , MHz	24	35.8	1.13	n/a	36	64.9	111	n/a	36	86.1	7.3	n/a
f_{ge}^{max} , GHz	8	8.97	0.09	9.04	10.3	11.3	9.66	9.08	6.3	5.89	0.07	5.9
d	0.5	0.09	0.08	0.26	0.5	0.49	0.52	0.6	0.1	0.25	0.4	0.42
Loss, kHz	251	20			338	51			2038	149		

TABLE II. Analysis of optimal parameters found for the three cases in Fig. 5 after the brute and Nelder-Mead optimizations. \mathbb{H}^{-1} stands for the standard deviations found from the diagonal of the covariance matrix which was calculated using the Hessian \mathbb{H} of the loss function at the optimum.

dispositions. The resulting fits are presented in Fig. 5 along with the original data and the error plots; we will further reference them as (a), (b) and (c).

The chosen algorithm order works as expected, as one can see from Table II comparing the brute estimation and the final result after Nelder-Mead is performed. The significant improvement in the loss value of the polished result compared to the brute estimation is due to the more accurate determination of the cavity frequency f_c which leads as well to major shifts in optimal g , f_{ge}^{max} and d . However, even these optimal qubit parameters and coupling strength may be inaccurate even on GHz scale. For example, in case (b) for f_{ge}^{max} we see a 300 MHz difference between the optimal and the correct value found with TTS. For (a) and (c), the differences in f_{ge}^{max} are much smaller (about 10-70 MHz) but instead we see significant errors in d . Inaccuracy in the qubit parameters is caused by low sensitivity of the resonator frequency to the qubit frequency when they are far away and by

Finally, The match is almost perfect, and the noise in the residuals is random and thus shows no systematic errors. In Fig. 6 we also present the fitting results for different powers of added noise. The signal-to-noise ratio (SNR) is defined for our data in a similar manner as the SNR in the resonator fitting tests¹⁰; since the data for a resonator lie on a circle in the complex plane, we take its radius r as the signal amplitude. Then, if the noise is of the form $\frac{\xi_1 + i\xi_2}{\sqrt{2}}$, where ξ_1, ξ_2 distributed normally with zero mean and variance σ , the SNR is defined as r/σ . As one can see from the graph, the algorithm stays robust even at very low SNRs that are even lower than the low limit SNRs for the *circleft* stability. This is possible because a fallback strategy of taking the smallest amplitude point as the resonance is applied when *circleft* fails. Ultimately, the key point of this stress test is to show that the algorithm can be applied in real-life scenarios when the data may be of low quality.

The fitting example above runs for 3.2 seconds on a 5-year old Intel Core i5-3337U CPU, from which 2.7 seconds are for the brute and 0.4 of a second is for the Nelder-Mead minimization. On a contemporary Intel Core i7-7700 CPU, it takes about one second to run the whole procedure. The time costs mostly come from the square root calculation necessary for Eqs. (A3) and (B3) and because that the brute routine is not parallelized. However, this is still fast, firstly, because record-

ing a resonator spectrum such as one in the example takes around 20 seconds, and, secondly, it is much faster than a human would do.

We have tested the algorithm on more than 100 real spectra that we have in our experimental database, and it finds an optimum for any kind of the dependence of the resonance on the flux caused by various dispositions of the qubit and the resonator, as in Fig. 9. Example fitting results for the variants in Fig. 9(b,c) are provided in ???. There is a caveat, however, in some of such cases. It turns out, that for large qubit-cavity detunings two different sets of parameters minimizing the loss function equally well are possible. An illustration for this statement can be seen in Fig. 7(a),(b). The algorithm was launched using different grid ranges for f_q^{max} . In one case (orange) the search was done below, and in the other (green) above the mean resonance frequency $\langle f_c \rangle_I$. The resulting fits are equally accurate, and without other information about the system it would not be possible to choose between the two. Therefore, some hints should be passed to the algorithm to avoid ambiguity; for instance, whether to look for the qubit above or below the resonator. Otherwise, it would be necessary to check both possibilities with other methods, i.e. via two-tone spectroscopy.

IV. DISCUSSION

Appendix A: Transmon Hamiltonian

The simplest version of this qubit consists of a Josephson junction shunted with a large capacitor. Flux tunability of the frequency is attained by replacing a single Josephson junction with a SQUID) as in Fig. 8(a) and applying external magnetic flux Φ_e to its loop. This configuration can be equivalently represented with a shunted junction of tunable energy, Fig. 8(b). The Hamiltonian for such equivalent circuit is as follows:

$$\hat{H}_{tr} = 4E_C \hat{n}^2 - E_J(\Phi_e) \cos \hat{\phi}, \quad (\text{A1})$$

where $E_{J1,2}$ are the Josephson energies, $E_C = e^2/2C_\Sigma$, $C_\Sigma = C_s + C_1 + C_2$, is the charging energy, \hat{n} and $\hat{\phi}$ are the operators for the Cooper pair number and the phase of the qubit island. For the equivalent Josephson energy E_J one obtains

$$E_J(\Phi_e) = E_{J\Sigma} \cos(\pi\Phi_e/\Phi_0) \sqrt{1 + d^2 \tan^2(\pi\Phi_e/\Phi_0)}, \quad (\text{A2})$$

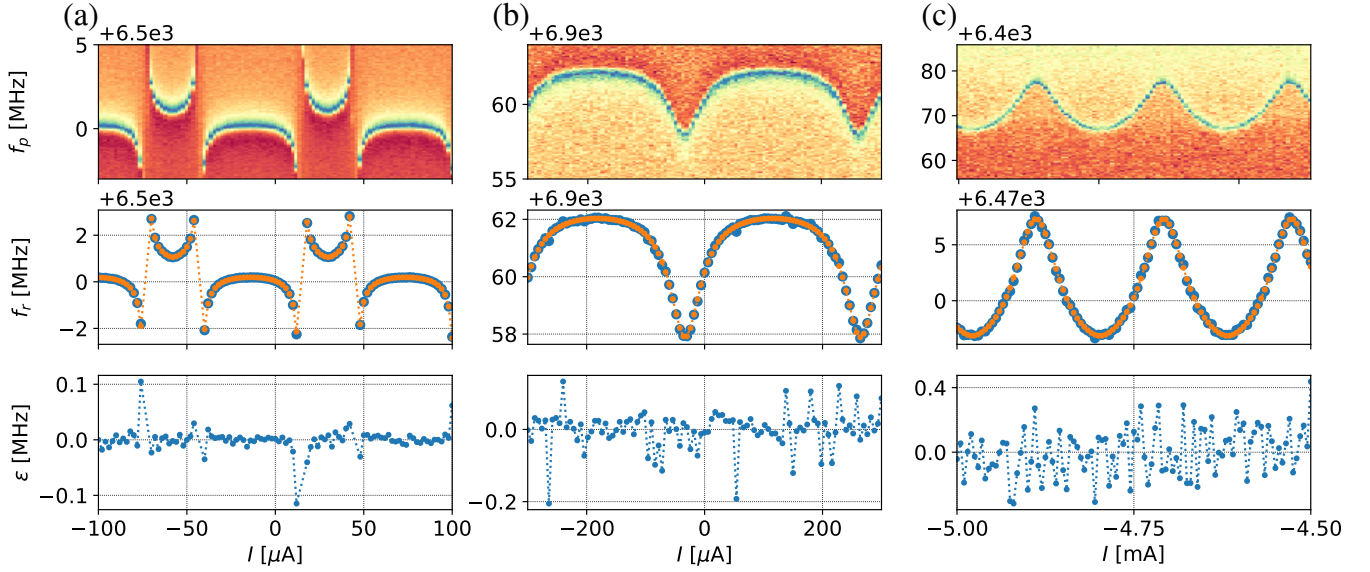


FIG. 5. The results of the algorithm execution on three experimental examples from our database. In the upper row, original images are presented. In the middle row, the data (blue dots) and the fitted model (orange connected dots) are shown. In the lower row are errors ϵ . (a) Avoided crossings pattern; per-point RMS error around 30 kHz. (b) The qubit entirely above the resonator; per-point RMS error around 60 kHz. (c) The qubit below the resonator; per-point RMS error around 150 kHz.

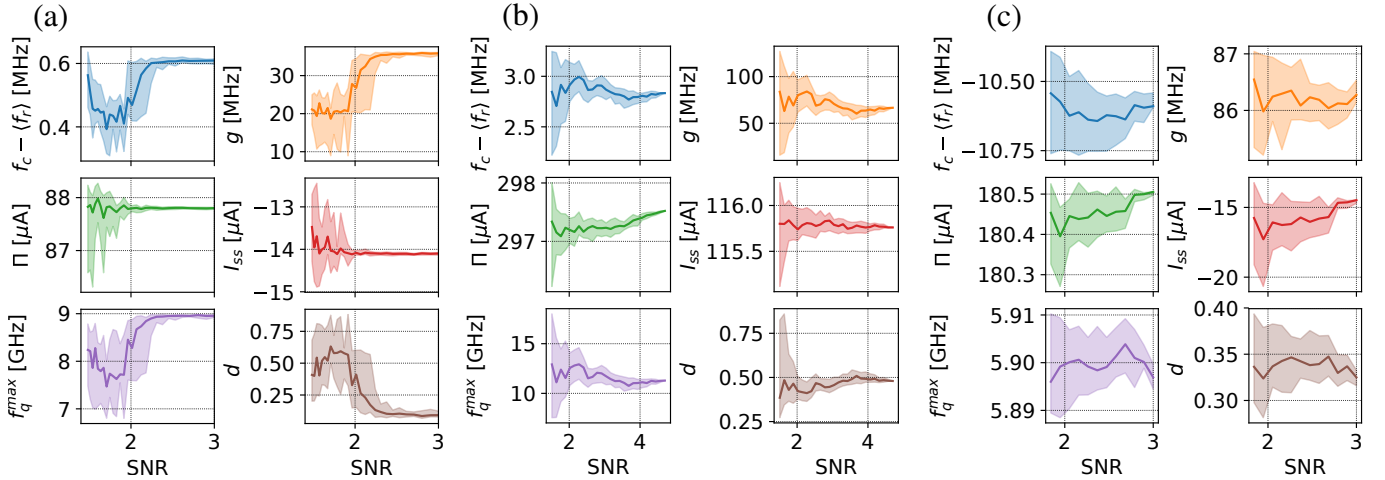


FIG. 6. Behaviour of the algorithm on the real data from Fig. 5 with added gaussian noise of varying power (i.e., with artificially reduced SNR). The clouds show the 25th and 75th percentiles of the optimal parameter samples after 50 tries, solid lines show median values. (a) Detecting avoided crossings; the data used are from Fig. 5(a). The algorithm is accurate above SNR=3 but is still robust at detecting avoided crossings down to SNR=2 but with reduced accuracy in f_{ge}^{max} , g and d . (b) Detection for the case of smooth dependence of the resonator frequency on current; the data used are from Fig. 5(c). Due to the problems with the resonator fitting, the stability is reached only above SNR=5. Correct qubit parameters $f_q^{max}=5.91$ GHz and $d=0.42$ (from TTS); f_q^{max} from the fit in the stable region is 5.89 GHz. However, d still can not be determined accurately on the data with noise; it is estimated to be 0.35 on the original data.

where $E_{J\Sigma} = E_{J1} + E_{J2}$, $d = \frac{E_{J1} - E_{J2}}{E_{J1} + E_{J2}}$ is the asymmetry of the SQUID. As one can notice, the dependence is periodic in Φ_e .

It is also possible to derive analytical expressions for the energy levels and transition frequencies for this type of qubits. The energy of the m^{th} level is⁹

$$E_m = m\sqrt{8E_J(\Phi_e)E_C} - \frac{E_C}{12}(6m^2 + 6m), \quad (\text{A3})$$

Level	Energy	Transition	Frequency
$g (E_0)$	0	ge	ω_{ge}
$e (E_1)$	$\sqrt{8E_JE_C} - E_C$	$gf/2$	$\omega_{ge} - 0.5E_C$
$f (E_2)$	$2\sqrt{8E_JE_C} - 3E_C$	$ef, gd/3$	$\omega_{ge} - E_C$
$d (E_3)$	$3\sqrt{8E_JE_C} - 6E_C$	$ed/2$	$\omega_{ge} - 1.5E_C$
E_4	$4\sqrt{8E_JE_C} - 10E_C$	$fd, eE_4/3$	$\omega_{ge} - 2E_C$

TABLE III. Energies and some transition (single and multi-photon) frequencies for the first 5 levels of the transmon calculated with (A3).

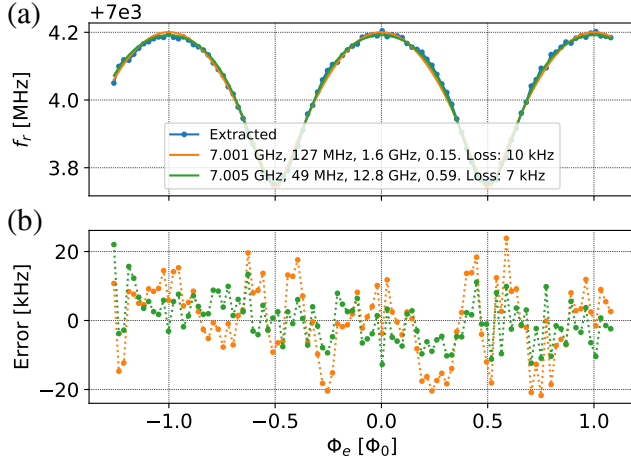


FIG. 7. (a) Two alternative fits (orange and green lines) for the same data (blue dots). The corresponding parameters f_c , g , f_{ge}^{max} , d for the two alternative fits are shown in the legend along with the loss values. (b) Residuals for the two fits, in orange (qubit below) and green (qubit above the resonator).

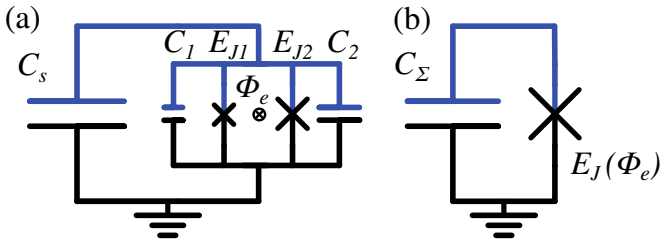


FIG. 8. (a) A tunable transmon circuit with an asymmetric SQUID, $E_{J1} \neq E_{J2}$. (b) Equivalent transmon with tunable energy $E_J(\Phi_e)$ and unified capacitance C_Σ . The qubit island containing its single degree of freedom is in blue.

and some of the transition frequencies are presented in Table III. The qubit frequency may be approximated as

$$f_{ge}(\Phi_e) \approx \sqrt{8E_J(\Phi_e)E_C} = f_{ge}^{max} \sqrt{\cos(\pi\Phi_e/\Phi_0) \sqrt{1 + d^2 \tan^2(\pi\Phi_e/\Phi_0)}}, \quad (\text{A4})$$

where $f_{ge}^{max} = \sqrt{8E_J(0)E_C}$. This simplifies the expression for the frequency since now it depends on two parameters instead of three.

One final note is that in real-life applications is not possible to know directly the flux Φ_e that is threaded through the SQUID. The experimenter usually knows only the current I (or voltage) which he applies to some coil that is connected inductively to the SQUID. Then $\Phi_e = MI + \Phi_r$, where M stands for the mutual inductance of the coil and the SQUID, and Φ_r is some residual flux inherent to the sample.

Appendix B: Circuit QED

The readout of the superconducting qubits is now predominantly done using an ancilla system which is usually implemented as a superconducting microwave resonator which acts as an electromagnetic cavity in the standard cavity QED. Truncating the qubit to two levels, one may obtain the following Hamiltonian for the compound cavity-qubit system (in RWA):

$$\hat{H}/\hbar = \frac{f_q}{2} \hat{\sigma}_z + f_c \hat{a}^\dagger \hat{a} + g(\hat{\sigma}^- \hat{a}^\dagger + \hat{\sigma}^+ \hat{a}), \quad (\text{B1})$$

where f_q is the qubit frequency, f_c is the cavity frequency and g is the coupling strength. As long as the RWA is done, this Hamiltonian may be diagonalized analytically¹³:

$$E_{g,0}/\hbar = \frac{f_c - f_q}{2}, \quad (\text{B2})$$

$$E_{\pm,n}/\hbar = (n+1)f_c \pm \frac{1}{2} \sqrt{4g^2(n+1) + (f_q - f_c)^2}. \quad (\text{B3})$$

This is very convenient for our purposes. By substituting the dependence of the qubit frequency $f_q \equiv f_{ge}(\Phi_e)$ into these equations, we can get straightforwardly the full system spectrum in dependence on the magnetic flux. In Fig. 9 we have used the equations (A3), (B2) and (B3) to model a tunable transmon interacting with a cavity for various Φ_e and various f_{ge}^{max} , d . In the lower row of the figure, one can see that it is possible to extract the dependence of the cavity frequency f_c on Φ_e ; for example, the well-known avoided crossing pattern can be directly observed in Fig. 9(a), and the other two possible behaviours for the qubit entirely above or below the resonator in Fig. 9(b),(c). To shorten the notation, in the following we will define the corresponding branch frequencies of (B3) as $f_{\pm} = (E_{\pm,0} - E_{g,0})/2\pi$.

In Fig. 9(a) it is also possible to see the entire spectrum of a transmon ge transition predicted by equations (A2), (A3). It has a cosine-like periodic shape with a period of one flux quantum Φ_0 . Consequently, it has two extrema, the upper and the lower which are called “sweet spots” due to the first-order insensitivity to Φ_e , and thus to possible flux noise. In the following, by saying sweet spot we will assume the upper one whose frequency is $f_{ge}(\Phi_e = 0) \equiv f_{ge}^{max}$.

We will use the model (B2), (B3) to fit the resonator frequency that we can find in an experiment. The only conceptual problem for the fitting that is left now is that the function we want to use as a model is not single-valued. Indeed, in Fig. 9(a, top) for each value of magnetic flux we always find two frequency points corresponding to the qubit and to the resonator, respectively. However, in practice only a narrow scan around the resonator frequency such as in Fig. 9(a, bottom) is required, and thus no ambiguity occurs.

¹J. Kelly, P. O’Malley, M. Neeley, H. Neven, and J. M. Martinis, “Physical qubit calibration on a directed acyclic graph,” arXiv preprint arXiv:1803.03226 (2018).

²S. M. Smith, *Metrology of Quantum Control and Measurement in Superconducting Qubits*, Ph.D. thesis, University of California Santa Barbara (2018).

³A. Blais, J. Gambetta, A. Wallraff, D. Schuster, S. Girvin, M. Devoret, and R. Schoelkopf, “Quantum-information processing with circuit quantum electrodynamics,” *Physical Review A* **75**, 032329 (2007).

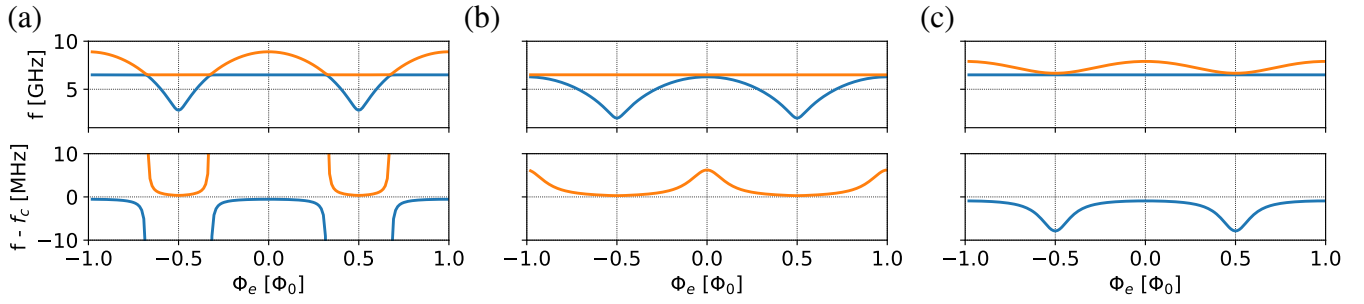


FIG. 9. Frequency spectrum of the transmon-resonator system. Parameters used: $f_{ge}(0) \approx \sqrt{8E_C E_J(0)}/2\pi = 8.5$ GHz, $d = 0.3$, $f_r = 6.4$ GHz, $g = 30$ MHz. For each subplot two transition branches $f_{\pm} = (E_{\pm,0} - E_{g,0})/2\pi$ are shown (orange and blue, respectively) both forming the resonator and qubit lines. As one can notice, there are three qualitatively different cases of the resonator-qubit disposition. Lower row shows a zoomed area around f_c that looks differently in each case.

⁴D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach* (Pearson, 2011).

⁵R. Versluis, S. Poletto, N. Khammassi, B. Tarasinski, N. Haider, D. Michalak, A. Bruno, K. Bertels, and L. DiCarlo, "Scalable quantum circuit and control for a superconducting surface code," *Physical Review Applied* **8**, 034021 (2017).

⁶J. Kelly, R. Barends, A. G. Fowler, A. Megrant, E. Jeffrey, T. C. White, D. Sank, J. Y. Mutus, B. Campbell, Y. Chen, *et al.*, "State preservation by repetitive error detection in a superconducting quantum circuit," *Nature* **519**, 66 (2015).

⁷M. Jerger, *Experiments on superconducting qubits coupled to resonators*, Ph.D. thesis, Karlsruhe Institute of Technology (2013).

⁸B. J. Bloom, S. A. Caldwell, M. Reagor, and C. Rigetti, U.S. Patent No. US 20180260732A (13 Sep. 2018).

⁹J. Koch, M. Y. Terri, J. Gambetta, A. A. Houck, D. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, "Charge-insensitive qubit design derived from the cooper pair box," *Physical Review A* **76**, 042319 (2007).

¹⁰S. Probst, F. Song, P. Bushev, A. Ustinov, and M. Weides, "Efficient and robust analysis of complex scattering data under noise in microwave resonators," *Review of Scientific Instruments* **86**, 024706 (2015).

¹¹J. A. Nelder and R. Mead, "A simplex method for function minimization," *The computer journal* **7**, 308–313 (1965).

¹²E. Jones, T. Oliphant, P. Peterson, *et al.*, "SciPy: Open source scientific tools for Python," (2001–), [Online; accessed <today>].

¹³A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, "Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation," *Physical Review A* **69**, 062320 (2004).

LIST OF CORRECTIONS

- Fatal: Is abstract supposed to be re-written completely? 1
- Note: Δ should not be used to denote a range here. Furthermore in the article Δ is used at least in three different meanings: (1) a range (2) a reminder (3) a standardized sample(i.e. with zero mean - see note below) - this should be fixed! 1
- Note: It should be explain why "scattering parameter" is used (i.e. is it the only measurement available?) and how it is related to Hamiltonian parameters? 2

Error: Change y-axis label from f_p to f_r to on Fig. 2 (c) and x-label from coil current to I . It is enough to have in the figure comments that I denotes coil current 2

Note: At this point of text the possible linkage between the experimental data and the model is not clear enough (apart from some visual similarities) 2

Note: dimensionality reduction typically means either feature selection or feature extraction in machine learning. Neither meaning appears to be applicable here, so would it be possible to chose another name for the procedure? 3

Error: You should probably talk about "likelihood" function rather than "loss" function - see note in "Full model fitting". A more ore less arbitrary loss functions are used for example in decision theory. This is hardly the case here - you should justify the application of a particular loss function 3

Note: Probably close to be true but the reference to the proof should be provided 4

Note: It is a standard procedure called ... "standardizing the data" (make the sample has zero mean and unit variance) - you should refer to it 4

Note: Why would finding of a global maximum of the zero-lag correlation function .. give the required parameter value? It appears to be an interesting idea and should be elaborated further 4

Note: What is square wave signal? 4

Note: Is the above mentioned wave signal complex?Is it S_{12}^{squared} ? 4

Note: It should be epxlained why does minimization of this loss function lead to optimal parameters values. 4

Note: Essentially, you are minimizing log likelihood assuming guaussian noise - it should be explained in this way 5