

Computer vision for automated experiments with superconducting qubits

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: 20 November 2018)

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. The first step towards this goal is to build a noise-insensitive and accurate computer vision tool to process the three-dimensional spectroscopic data that are ubiquitous in the domain. In this work we present and describe two widely-applicable and scalable algorithms that are capable to robustly and efficiently extract all Hamiltonian parameters from the spectroscopic data.

I. MOTIVATION

Experiment automation, automatic data processing and decision making is vital in developing even a small scale quantum processor. For superconducting qubits, large companies such as Google¹, Rigetti² and IBM³ already use to some extent automated procedures for calibration of the qubit parameters and tuning up the single and two-qubit gates. This is not surprising since performing experiments with superconducting qubits is usually quite demanding in terms of human time. For example, to fully characterize a tunable transmon qubit coupled to a resonator (which assumes extracting the Hamiltonian parameters and coherence times) in an ideal scenario one would need to perform at least 10 individual measurements of different kinds^{4,5}. Each of those experiments needs constant monitoring and parameter adjustments which may take days if, for example, several qubit chips are being examined. Moreover, more elaborate experiments such as measuring the coherence times versus some control parameters combinations or versus time are just impossible to do by hand despite having tremendous value to the characterization of the fabrication technology^{6,7}.

Nonetheless, the subject is not represented abundantly in academic literature, and some experiments, for example, spectroscopy, are still being executed by hand⁵. Possibly, this can be attributed to the fact that it is not straightforward to implement a fully autonomous system that would be able to handle the experiment from the very beginning. While in research this state of affairs is acceptable since even a human experimenter does not know beforehand what kind of algorithm he should apply, it is definitely a problem for the development of the quantum computing chips. To develop such system, firstly, it is necessary to create an efficient and robust processing tool that will be able to handle all kinds of experimental data and replace human operator in terms of extracting the

physical qubit parameters.

In this work, we are proposing and implementing several methods of data processing and computer vision that tremendously facilitate the spectroscopy of circuit QED⁸ architectures widely used for the readout the qubit states. Our method is accurate, fast and robust to noise, is applicable to any type of superconducting qubit that has a periodic and parametrizable spectrum, and is fully compatible with the state-of-the-art paradigm of one readout resonator per one qubit^{9,10}. This means that our approach is scalable and can be applied to processors with any number of qubits. To achieve this we have employed several optimization algorithms to find spectral data from the heatmap plots that are ubiquitous in the domain.

II. BACKGROUND

The methods that we use are relying heavily on the theoretical formalism that describes the behaviour of the superconducting qubits and cQED systems¹¹. Thus, below we are introducing briefly the key concepts and equations that will be used in the following sections.

Without significant loss of generality, we will consider a tunable transmon¹² qubit capacitively coupled to a notch-type readout resonator. This architecture is currently widely used⁶ and therefore serves a good example to demonstrate the algorithms performance.

A. Transmon qubit

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. 1(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. 1(b).

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

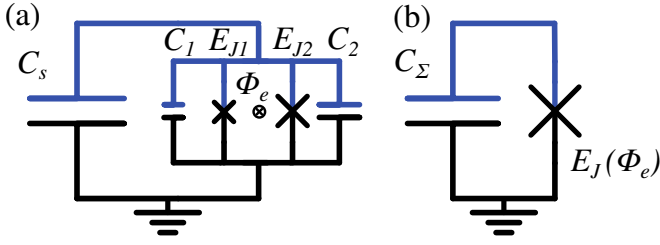


FIG. 1. (a) A tunable transmon circuit with an asymmetric SQUID, $E_{J1} \neq E_{J2}$. (b) Equivalent transmon with tunable energy $E_J(\Phi_e)$ and unified capacitance C_Σ . The qubit island containing its single degree of freedom is in blue.

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

TABLE I. Energies and some transition (single and multi-photon) frequencies for the first 5 levels of the transmon calculated with (3).

The Hamiltonian for such equivalent circuit is as follows:

$$\hat{H}_{tr} = 4E_C \hat{n}^2 - E_J(\Phi_e) \cos \hat{\phi}, \quad (1)$$

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

$$E_J(\Phi_e) = E_{J\Sigma} \cos \phi_e \sqrt{1 + d^2 \tan^2 \phi_e}, \quad (2)$$

where $E_{J\Sigma} = E_{J1} + E_{J2}$, $\phi_e = \pi \Phi_e / \Phi_0$, $d = \frac{E_{J1} - E_{J2}}{E_{J1} + E_{J2}}$ 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 is¹²

$$E_m = m \sqrt{8E_J(\Phi_e) E_C} - \frac{E_C}{12} (6m^2 + 6m), \quad (3)$$

and some of the transition frequencies are presented in Table I. The qubit frequency $f_{ge} = \omega_{ge}/2\pi$ may be approximated as

$$\sqrt{8E_J(\Phi_e) E_C} = f_{ge}^{\max} \sqrt{\cos \phi_e \sqrt{1 + d^2 \tan^2 \phi_e}},$$

$f_{ge}^{\max} = \sqrt{8E_J(0) E_C}$, to make it depending on just two parameters instead of three.

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):

$$\hat{H}/\hbar = \frac{\omega_q}{2} \hat{\sigma}_z + \omega_c \hat{a}^\dagger \hat{a} + g(\hat{\sigma}^- \hat{a}^\dagger + \hat{\sigma}^+ \hat{a}), \quad (4)$$

where ω_q is the qubit frequency, ω_c is the cavity frequency and g is the coupling strength. As long as the RWA is done, this Hamiltonian may be diagonalized analytically¹¹:

$$E_{g,0} = \frac{\omega_c - \omega_q}{2}, \quad (5)$$

$$E_{\pm,n} = (n+1)\omega_c \pm \frac{1}{2} \sqrt{4g^2(n+1) + (\omega_q - \omega_c)^2}. \quad (6)$$

This is very convenient for our purposes. By substituting the dependence of the qubit frequency $\omega_q \equiv \omega_{ge} = 2\pi f_{ge}$ on some control parameter into these equations we can get straightforwardly the full system spectrum in dependence on that parameter. In Fig. 2 we have used the equations (3), (5) and (6) 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 $\omega_c = 2\pi f_c$ on Φ_e ; for example, the well-known avoided crossing pattern can be directly observed in Fig. 2(a), and the other two possible behaviours for the qubit entirely above or below the resonator in Fig. 2(b),(c). To shorten the notation, in the following we will define the corresponding branch frequencies of (6) as $f_{\pm} = (E_{\pm,0} - E_{g,0})/2\pi$.

In Fig. 2(a) it is also possible to see the entire spectrum of a transmon ge transition predicted by equations (2), (3). 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 (5), (6) 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. 2(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. 2(a, bottom) is required, and thus no ambiguity occurs.

III. METHODS

In this section the actual algorithms that perform fitting and data extraction are discussed. Additionally, we describe important peculiarities of the data itself and some essential experimental details.

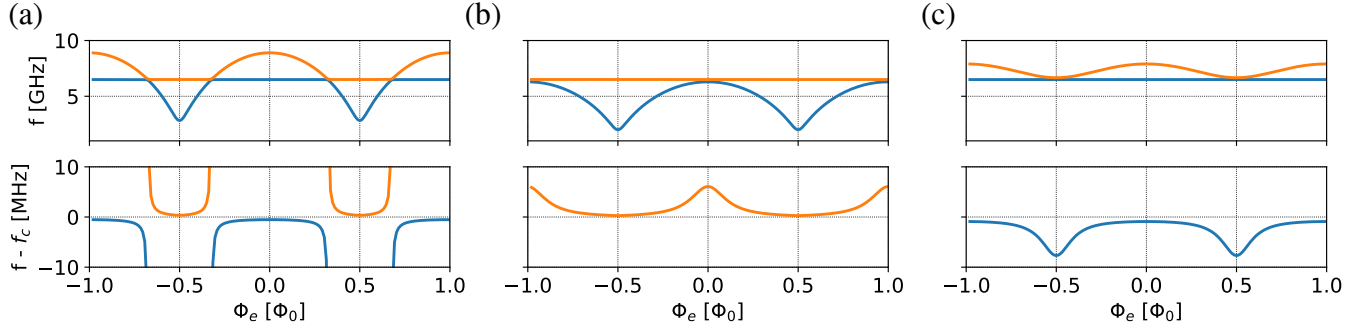


FIG. 2. Frequency spectrum of the transmon-resonator system. Parameters used: $f_{ge}(0) \approx \sqrt{8E_C E_J(0)}/2\pi = 8.5$ GHz, $d = 0.3$, $f_r = 6.4$ GHz, $g = 30$ MHz. For each subplot two transition branches $f_{\pm} = (E_{\pm,0} - E_{g,0})/2\pi$ (orange and blue, respectively) both forming the 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.

A. Analysing resonator spectra

In the real life applications, the spectrum of the resonator is usually recorded with a vector network analyzer which measures the complex scattering parameter S_{21} for each probe frequency f_p in a certain area around the resonance. Hence, after the VNA scan there are two 2D plots for the amplitude and the phase of the transmission versus frequency. When the magnetic flux sweep is added, the data becomes three-dimensional and has to be represented visually using a heatmap plot. One such heatmap for a notch-port resonator coupled to a transmon qubit is presented in Fig. 3(a). Please notice, that in all cases it is not possible to know which flux Φ_e is applied to the SQUID. The experimenter usually knows only the current (or voltage) which he applies to some coil and which is connected linearly to Φ_e : $\Phi_e = MI + \Phi_r$, where M stands for the mutual inductance of the coil and the SQUID, and Φ_r is some residual flux. We will call the currents for the high and low transmon sweet spots I^{max} , I^{min} , respectively.

The idea is to fit the image shown Fig. 3(a) with the model equations (5), (6). This task is not trivial because it belongs to a class of optimization problems where the loss function (i.e. in the least squares algorithm) is ill-defined. This is due to the periodic dependence of the frequencies on Φ_e (and I) and unknown position of the qubit sweet spot due to the local residual flux. Fortunately, it is possible to get a good initial guess for these parameters and then brute force the solution since we have a well defined bounds for each of the variables left.

The outline of the method which uses Fig. 3(a) as an example is presented below. The chosen type of the resonator-qubit arrangement which yields the avoided crossing pattern is not universal: there are two other cases when the qubit spectrum lies entirely below or above the resonator frequency, as in Fig. 2(b),(c). However, they are treated exactly the same way.

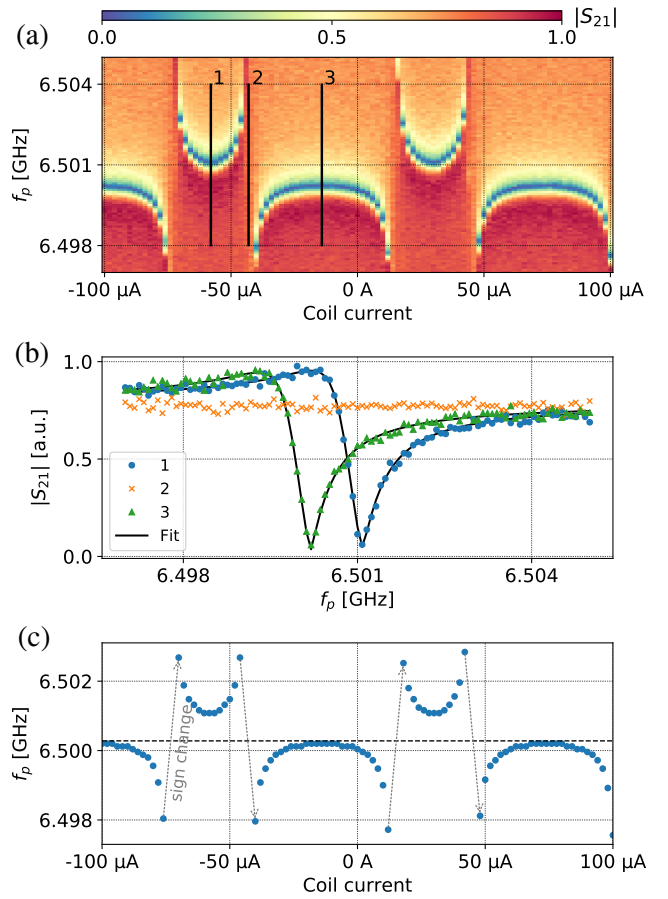


FIG. 3. (a) An experimental spectrum of a resonator strongly coupled to a transmon qubit depending on the coil current generating the magnetic flux Φ_e . (b) Slices of the transmission from (a) showing two slices with fits (1,3) and a plateau with no dip (2) present. (c) Extracted data (blue dots) and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

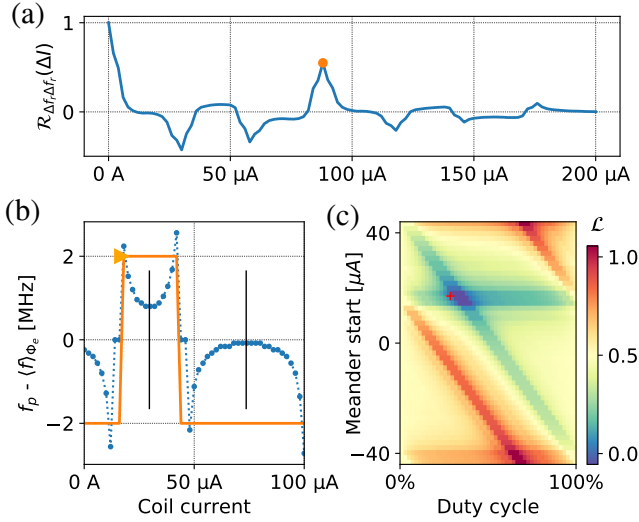


FIG. 4. Period and phase extraction procedure. (a) Autocorrelation function depending on the lag ΔI shows a prominent local maximum at $88 \mu\text{A}$ (orange dot). (b) Construction for phase estimation; $\Delta f_r = f_r - \langle f_r \rangle_{\Phi_e}$ (blue) is fitted with a square wave (orange, start 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.

1. Extracting $f_r(\Phi_e)$ from data

The first thing to do is to reduce the dimensionality of the data, i.e. extract the resonance frequency for each Φ_e and then be left with the desired $f_r(\Phi_e)$.

We do this by employing the *circleft*¹³ library which is capable of fitting various types of microwave resonators. For each Φ_e we fit the complex transmission $S_{21}(f)$ as in Fig. 3(b) (solid black lines) and extract the resonance frequency from the model. The only possible caveat is that for some Φ_e values the resonance dip may disappear (see Fig. 3(b)) and the fit will fail. Such slices are thus excluded in advance via a threshold condition.

The resulting plot of extracted frequency is shown in Fig. 3(c) with blue dots. There are some current values located between the branches for which the points are missing, as expected. Additionally, we plot here another key parameter which is the mean value of the detected frequencies shown as a dashed black line. Firstly, this parameter will serve as an initial guess for the cavity frequency of the model (5), (6): $f_c \approx \langle f_r \rangle_{\Phi_e}$. Secondly, it serves as a baseline used in the following period and phase extraction algorithm which tracks the changes of the sign of the value $\Delta f_r = f_r - \langle f_r \rangle_{\Phi_e}$ (marked as grey dashed lines in Fig. 3(c)). These sign changes are crucial in determining the location of the qubit sweet spot.

2. Extracting period, phase and sweet spot locations

As we have said above, one of the serious obstacles for the fitting is that 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 severely impeded the progress of iterative optimization algorithms. In other words, the unknown coefficients M and Φ_r in the dependence of Φ_e on the coil current I are preventing us to find the global minimum we need. Fortunately, it is possible to find the period and the phase (or, alternatively, Φ_r , or the sweet spot location) of the data without fitting the full model Eqs. (5), (6) which alleviates the stated problem. Of course, in the absence of noise, this would be a trivial task. However, in the real-life conditions the noise to some extent is always present, so below we engineer a solution that is insensitive to local sporadic perturbations of the data.

A very powerful tool for finding the period Π_f 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. However, this is only true when the mean value of the function is zero; otherwise, $\mathcal{R}_{yy}(l)$ would be linear with a very steep slope smearing out all the extrema. Therefore, instead of $f_r(\Phi_e)$ we consider the function $\Delta f_r(\Phi_e)$ introduced above which has zero mean and the same period. For this function, which looks exactly like one from Fig. 3(c) but centered vertically around zero, the autocorrelation function $\mathcal{R}_{\Delta f_r \Delta f_r}(\Delta I)$ is shown in Fig. 4(a). As one can see, it depends on the current offset ΔI which spans $200 \mu\text{A}$ just as the data itself. This means that the data is being zero-padded at all ΔI except for $\Delta I = 0$, and this is why we get a diminishing value of the correlation peaks when the lag reaches Π_f or $2\Pi_f$. 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 $\Delta I = 2\Pi_f$. Note that the autocorrelation function is not smooth. This is because there were some missing points in the data (corresponding to the plateaus of Fig. 3(b)) that were replaced by zeros to ensure correct mapping between current indices and current values. This can be noticed as well in Fig. 4(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 zero-lag correlation function $\mathcal{R}_{\Delta f_r S}(0)$ of Δf_r with a square wave $S(I, \Pi_f, \phi, D)$ signal having the same period but unknown phase ϕ and duty cycle D . An illustration of a meander function satisfying the optimal condition is presented in Fig. 4(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. The high and low levels of the square must be opposite in value, i.e. 1 and -1, and the absolute value does not matter. Generally, the idea behind this is to robustly detect sign changes that were shown back in Fig. 3(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 will fail in the presence of a strong noise. Conversely, the optimization algorithm described above is not sensitive to local short-term sign changes.

The global optimal ϕ and D are found using a brute force algorithm. It calculates the loss function $\mathcal{L} = -\mathcal{R}_{\Delta f_r S}(0)$ on a

50×50 grid of (ϕ, D) and takes the minimal value of all. It is feasible due to the evident boundaries on $\phi \in [-\Pi_{f_r}/2, \Pi_{f_r}/2]$ 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. 4(c). One can see that instead of a single minimum it has an area of the same minimal value. Again, the roots of this effect lie in the missing f_r points. However, any value from this valley suits well enough for our purposes, and the algorithm finds no difficulty in locating it.

With values for $\Pi_{f_r} = \Phi_0/M$, ϕ and D we now can calculate the currents of the two transmon sweet spots in the case of the avoided crossings pattern:

$$I^{max} = \phi + \Pi_{f_r} \left(\frac{1+D}{2} \right) = \Phi_r/M, \quad (7)$$

$$I^{min} = \phi + \Pi_{f_r} \left(\frac{D}{2} \right). \quad (8)$$

For the smooth dependences Fig. 2(b)-(c) the situation is opposite: $I^{max(min)} \rightarrow I^{min(max)}$. To distinguish the two cases when the noise is not large, it is enough to calculate the maximal differential of the frequency data and compare it to its peak-to-peak amplitude. For the avoided crossings these values are close and for the smooth curves they are not. However, in the presence of noise this indicator will fail, and we will have to resort to checking both current values for being I_{max} based on the best fit among the two possibilities.

3. Full model fitting and testing

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 Nelder-Mead¹⁴ search. Firstly, we define the loss function for the problem as follows. For the known probe frequency span of the data Δf_p (Fig. 3(a), whole y-axis) and the set of N extracted points $\{p_i\} = \{(I_i, f_{r,i})\}$, we define the loss function per one point as

$$\mathcal{L} = \frac{1}{N} \sum_{i=0}^N [f_i - \mathcal{M}(MI_i + \Phi_r, f_c, g, f_{ge}^{max}, d)]^2, \quad (9)$$

$$\mathcal{M} = \begin{cases} f_+, & |f_+ - f_c| < \Delta f_p/2 \\ f_-, & \text{otherwise,} \end{cases} \quad (10)$$

where, as in the previous section,

$$f_{\pm} = f_{\pm}(MI_i + \Phi_r, f_c, g, f_{ge}^{max}, d) = E_{\pm,0} - E_{g,0}.$$

The condition of Eq. (10) means that we choose only the frequencies that lie within a window Δf_p around the model f_r 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. 5 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. From the plots

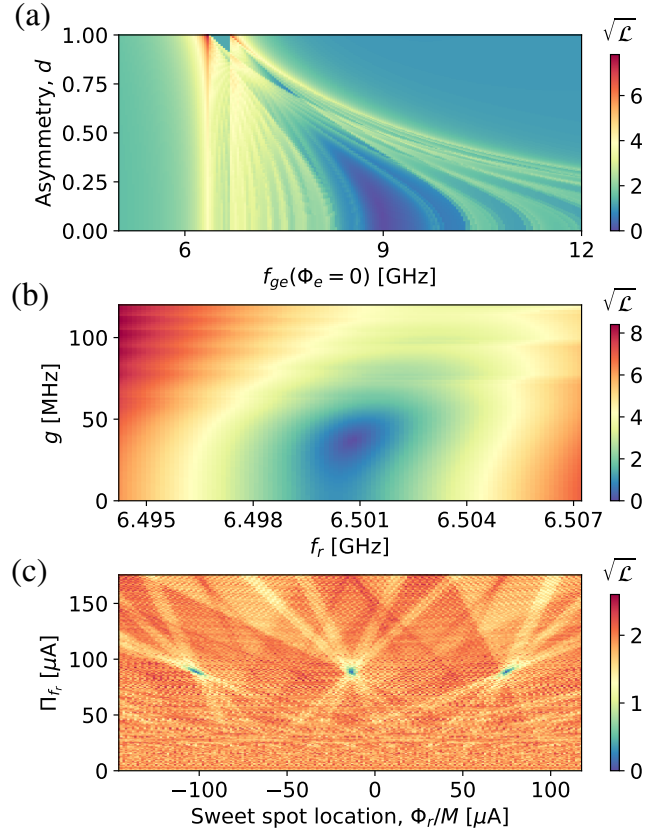


FIG. 5. Slices of the loss function (9) for the full model and experimental data from the example made around the optimal point. Root-mean-square per-point loss $\sqrt{\mathcal{L}}$ is in MHz. (a) f_{ge}^{max} and d are varied, other parameters optimal. A large valley is located near 9 GHz, and some 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) Period $\Pi_f = \Phi_0/M$ and Φ_r/M varied, others optimal. This subplot illustrates a complex structure of local minima around the true one which we find analytically.

it is obvious that it is ill-defined and has a lot of local minima which would impede robustness in real life applications. The f_{ge}^{max} and d parameters present the most significant difficulty in terms of false minima as can be seen from Fig. 5(a). In contrast, f_r presents the least difficulty, as can be seen from Fig. 5(b). The last plot Fig. 5(c) shows a very complex structure of the loss function and a very small-sized optimal valley and serves as an illustration of why the period and phase extraction is important.

The brute force algorithm acts on the grid specified in Table II. The ranges in the grid are based on the usual design parameters of the modern qubit samples and the number of steps is chosen to be able to catch the optimal valley in the loss function.

After the coarse brute force optimization is done and the optimal valley is located, we apply the Nelder-Mead simplex downhill algorithm to polish the brute force result. This algorithm works very well in our case, as one can see from the

Parameter	Value range	Steps #	Brute est.	Final
M	Φ_0/Π_{f_r}	1	$11.4 \Phi_0/\text{mA}$	same
Φ_r	$M I^{max}$	1	$-0.13 \Phi_0$	same
f_c	$\langle f_r \rangle_{\Phi_e} \pm 1 \text{ MHz}$	3	6.5003 GHz	6.5007 GHz
g	20 - 40 MHz	5	24 MHz	36.1 MHz
f_{ge}^{max}	4 - 12 GHz	80	8 GHz	8.9 GHz
d	0 - 0.9	9	0.5	0.08

TABLE II. Grid specifications for the brute force algorithm and the resulting optimal points. The number of steps is taken with some margin to ensure correct location of the optimal valley even for difficult cases. The optimal parameters estimated by brute force and the final ones have the RMS loss values of 0.258 MHz and 0.024 MHz, respectively.

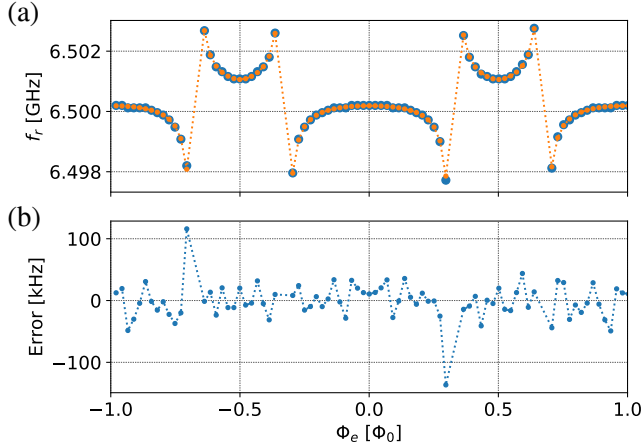


FIG. 6. The result of the algorithm execution. (a) The data (blue dots) and the model (orange connected dots). (b) Residual error which is mostly random and is of order of 25 kHz per point.

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 .

Finally, the resulting fit is presented in Fig. 6 along with the plot of the residuals. The match is almost perfect, and the noise in the residuals is random and thus shows no systematic errors. In Fig. 7 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 tests¹³; we take distance $2r$ between two maximally remote points of $S_{21}(f_p, \Phi_e)$ among any pair on the complex plane (i.e., the diameter of the resonance circle) as the signal amplitude, and the noise of the form $\frac{\xi_1 + i\xi_2}{\sqrt{2}}$, where ξ_1, ξ_2 distributed normally with zero mean and variance (amplitude) σ . The SNR then 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

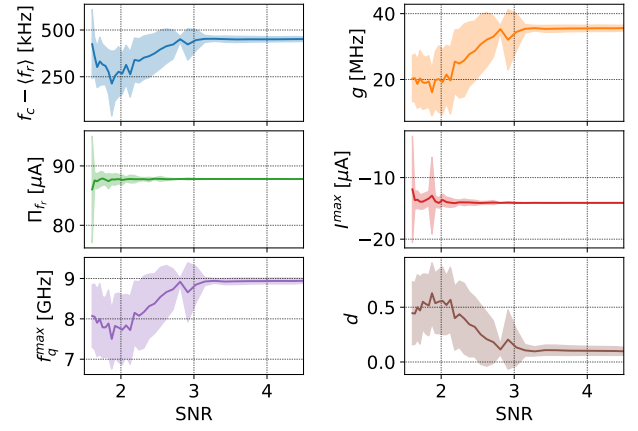


FIG. 7. Behaviour of the algorithm on the data with the gaussian noise of varying noise power added. The clouds show standard deviations of optimal parameters after 50 tries, solid lines show mean values. The algorithm is very stable above SNR=8 but is still robust at detecting avoided crossings down to SNR=3-4 but with reduced accuracy in f_{ge}^{max} , g and d . Other three parameters are still very accurate.

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 algorithm was implemented in Python using brute and Nelder-Mead routines from Scipy. 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. (3) and (6) and because that the brute routine is not parallelized. However, this is still fast, firstly, because recording 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. 2. Example fitting results for the variants in Fig. 2(b,c) are provided in Fig. 8. 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. 9(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 $\langle f_c \rangle_{\Phi_e}$. 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

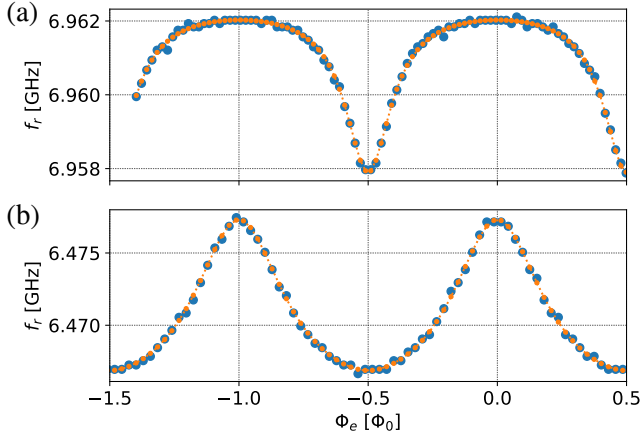


FIG. 8. Fitting two other samples of our dataset presenting other two possible cases (blue dots are data, orange connected dots are the model). (a) The qubit is entirely above the resonator. Per-point RMS error is 60 kHz. (b) The qubit is below the resonator. Error is 150 kHz (larger due to the proportionally larger amplitude of the frequency variation).

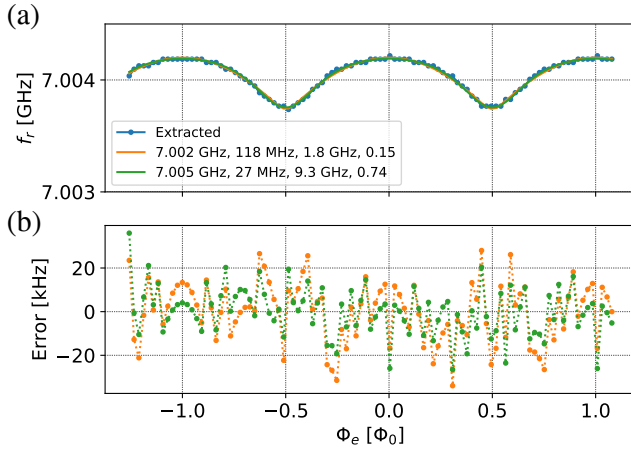


FIG. 9. (a) Two alternative fits (orange and green lines) for the same data (blue dots). The corresponding parameters f_c , g , f_{ge}^{max} , d for the two alternative fits are shown in the legend. (b) Residuals for the two fits, in orange (qubit below) and green (qubit above). RMS errors are 14 kHz and 11 kHz, respectively.

two-tone spectroscopy.

B. Analysing qubit spectra

Unfortunately, the resonator fitting procedure accuracy in terms of the qubit frequency (>100 MHz) is not enough in practical applications where it is necessary to obtain exact value of the qubit frequency down to 1-10 MHz. For the cQED systems, it may be done, in general, by trying to excite the qubit at different frequencies and then read it out expecting the e -state to be measured. The task is to find the exact location and shape of all the spectral lines depending on the

magnetic flux that will be detected using such a method.

Particularly, in many groups, the qubit spectrum is searched for using the so-called two-tone spectroscopy¹⁵. It consists of sending two microwave signals at the cQED system from which one, the probe tone, has a fixed frequency near f_c , and the frequency f_{exc} of the other, excitation tone, is varied. When the second signal becomes resonant with some transition (i.e., qubit ge transition), a shift of the cavity frequency occurs, and thus the transmission at the frequency of the first tone changes. This way, a heatmap showing the dependence of the frequency of various system transitions on the magnetic flux can be obtained.

An example from our database for a tunable transmon coupled to a notch-port resonator is displayed in Fig. 10. To obtain the two-tone spectrum in Fig. 10(b), for each current value the frequency of the first tone f_p is chosen to be equal to the frequency of the minimal transmission from the single-tone spectroscopy (STS), see Fig. 10(a). At the avoided crossing, this is not possible, and thus the frequency f_p may be chosen to be arbitrary within the scan area of the STS. This does not pose a problem since the two-tone spectroscopy is still only effective when the qubit and the resonator are detuned far from each other, i.e. in the dispersive regime¹¹, and would not work well near the resonant regime.

The resulting plot Fig. 10(b) shows the dependence of S_{21} on the magnetic flux and the excitation frequency. Most of the scan is in blue meaning that the resonator is not shifted and the probe tone gets almost entirely reflected (see colorbar). However, there are some different colour areas. Firstly, a bright and narrow horizontal line is observed near 6.5 GHz where the resonator is located. Likewise, there are two sharp hori-

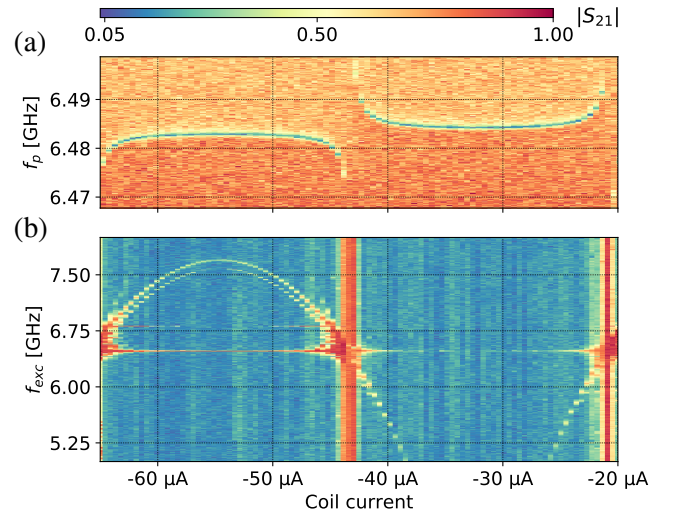


FIG. 10. Two types of spectroscopy results side by side; the colorbar is common, coil current scan range is the same. (a) Single-tone spectroscopy (one period on Φ_e). (b) Two-tone spectroscopy result. Clear flux-dependent transmon transitions are visible (ge and $gf/2$), and some fixed-frequency resonator excitations are present as well. Red vertical stripes are observed when the qubit and the cavity are close to resonance.

zontal lines just above 6.75 GHz which are from the crosstalk coupling to another on-chip resonator. Secondly, bright vertical stripes are seen at the current values corresponding to the avoided crossing areas. Finally, two transmon transitions of varying frequency are visible. The upper one is the ge transition, and the lower one is the two-photon $gf/2$ one; at each coil current, they are separated by $E_C/2h$, or half the anharmonicity of the transmon (see Table I).

The presence of several bright lines along with noise makes it in general impossible to accurately fit the data with the model (3) using the standard curve fitting procedures. Additionally, when several multi-photon transitions are visible, it is hard for an algorithm to distinguish between them to find only the ge transition. This means that for robust qubit detection we need to fit all possible single- and multi-photon spectral lines simultaneously. Below, we present our approach to solve this problem.

1. Thresholding the data

Similar to the reasoning in section III A 1, the first thing to do is to reduce the three-dimensional picture Fig. 10 to two dimensions. This is done in several steps.

First, we calculate the mean transmission $\langle S_{21} \rangle_{f_{exc}}(\Phi_e)$ for each current value. This operation gives us the background dependent only on Φ_e and not on f_{exc} . This background can be then subtracted from the data, and the absolute value of the difference $Z = |S_{21} - \langle S_{21} \rangle_{f_{exc}}|$ may be then plotted. For the example, the resulting image can be seen in Fig. 11. As one can notice, the background subtraction removes most of the vertical stripes that could be seen in Fig. 10(b).

Next, the data points may be extracted to form a 2D point array where the qubit lines may be searched for. This procedure is done in several steps. First of all, it is necessary to do thresholding of the data to obtain a binary image. We have tested various thresholding methods, i.e. Otsu's method¹⁶; however, the best performance in distinguishing local peaks from background was obtained using another method that is based on estimating the noise in the data. We estimate noise level σ by calculating the numerical difference between adjacent points in a row for each flux value, and then calculate the median absolute value:

$$\sigma = \text{median}_{i,j} (|Z_{i,j+1} - Z_{i,j}|),$$

where i numerates magnetic flux values, and j numerates excitation frequency values. Using median instead mean reduces the impact of singular bright extrema that may be present in the row.

Next, for each row we use the *find_peaks* routine of the Scipy library to detect all local extrema there, and then exclude those whose height is less than σ as probably originating from the noise.

After finishing the procedures above, we get an array of points that correspond to the most prominent peaks in the data. However, already at this point it is possible to remove all points that come from the fixed-frequency resonator spectral lines. This is done by excluding horizontal rows of points that

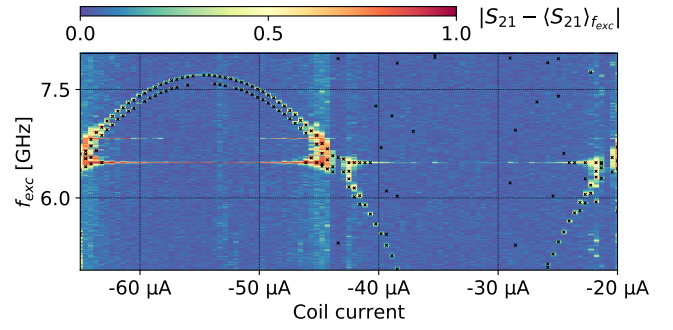


FIG. 11. Processed data and extracted and filtered datapoints (black crosses) that will be subject to fitting.

have exactly the same frequency value within the resolution of the data.

The final extracted points are shown in Fig. 11 in black crosses. As one can see, the horizontal lines are correctly not included, and two qubit lines are almost fully represented. Background noise is mostly ignored, too.

2. Fitting binary data

a. General outline of the method. In most cases, even the filtered binary data contains noise and, importantly, is not single-valued for each flux value. This means that standard curve fitting methods are not applicable, and thus we resort to a hybrid algorithm that combines features of the Hough transform¹⁷ for global optimization and Nelder-Mead search local optimization. The algorithm bases on three main parts: the point selection procedure, solution ranking and parameter sampling. The selection procedure is universal, and the latter two parts come in two variations depending on which type of optimization will be used.

The general idea for the first part is illustrated in Fig. 12. Here, three narrow frequency bands of Δf are chosen around the model spectral lines calculated for tested parameters by (3). Points outside the bands are considered far away and are ignored. All the points inside a band are attributed to its line and included in its loss function. If at a certain flux there are several points inside a band, only the closest one is chosen and the rest (contentious) are ignored. Let $P_{1,2,3}$ be the sets of points selected for corresponding transition lines $ge, gf/2, gd/3$ (in blue, orange and green in Fig. 12); the number of points in the sets are $N_{1,2,3}$, $N_Z = \sum_i N_i$, and M is the total number of measured magnetic flux values; let $D_{1,2,3}$ be the corresponding sets of vertical distances between the selected points and the corresponding spectral lines. Finally, let D be the vertical distances between all the extracted points P (both black and coloured in Fig. 12, total number N) and the upper spectral line.

The second part of the algorithm now has to evaluate the quality of the fit, i.e. be able to rank supplied sets of optimization parameters. The ranking differs for the brute search and for the local optimization.

In case of the brute search, the evaluation is based both on the total number of selected points N_Σ and the normalized squared distances $\langle D_\Sigma^2 \rangle = \sum_i \sum_{P_i} D_i^2 / N_1$ (note that we normalize based just on N_1 to respect the majority voting in case of low N_2, N_3). Maximizing the total number of points included ought to have a higher priority than minimizing the average distances, since otherwise a single selected point lying exactly on the spectral line would win. However, in the presence of spurious spectral lines this prioritization may lead to wrong results since an incorrect solution with a slightly larger N_Σ but high normalized distance measure would overshadow the correct one with a low distance measure. To battle this effect and increase the influence of the distance loss, we quantize N_Σ into bins of size N_b equal to 25% of the total number of fluxes M , and then choose from the best bin the solution having the lowest distance measure.

Alternatively, for the Nelder-Mead algorithm one should build a loss function that returns just a single value instead of a double-valued rank as before. Since we still want to both maximize the number of points and minimize the mean distances, we take again same binned value $[N_\Sigma/N_b]N_b$ and $\langle D_\Sigma^2 \rangle$ but now calculate the function

$$\mathcal{L}_{N-M} = \frac{1}{[N_\Sigma/N_b]N_b} + \langle D_\Sigma^2 \rangle. \quad (11)$$

This loss function has a convex shape near the sought extremum; this makes it possible to run the Nelder-Mead algorithm seamlessly.

Finally, the third part of the algorithm is either the brute force search to find the global extremum or the Nelder-Mead descent to polish the found parameters.

Note the similarity between this construction and the Hough transform. The transform is implemented by quantizing the Hough parameter space into finite intervals or accumulator cells that each point votes for. Likewise, we use the frequency band (frequency quantization) to find the points voting for a certain parameter set; other parameters are as well quantized on a brute search grid. However, our approach is extended by a subsequent local optimization necessary since we want to improve the accuracy beyond brute grid resolution.

To sum up, the optimization is done on points belonging to a narrow frequency band around the spectral line when this

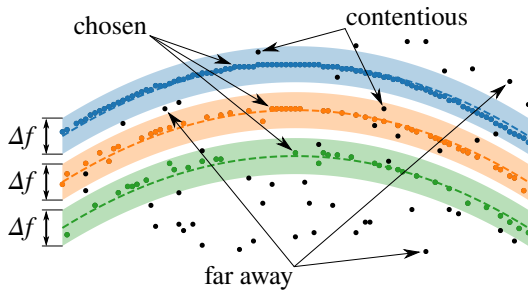


FIG. 12. An illustration for the loss function definition. In different colours are the frequency bands and the corresponding selected points, black points are ignored.

Optimization		Φ_r	f_{ge}^{max}	d
Brute 1	range	$\Phi_r^{(0)} \pm .05 \Pi_{\Phi_e}^{(0)}$	$f_{ge}^{max,(0)} \pm 30\%$	0.1 - 0.9
	steps	10	50	8
Brute 2	range	$\Phi_r^{(1)} \pm .02 \Pi_{\Phi_e}^{(0)}$	$f_{ge}^{max,(1)} \pm 100 \text{ MHz}$	$d^{(1)} \pm 0.1$
	steps	10	20	10

TABLE III. Grid specifications for the brute force algorithm and the resulting optimal points. The number of steps is taken with some margin to ensure correct location of the optimal valley even for difficult cases. The optimal parameters estimated by brute force and the final ones have the RMS loss values of 0.258 MHz and 0.024 MHz, respectively.

band contains as many points as possible. This logic is applicable both to single-line and multi-line fitting. In the following, we will use single-line fitting to find at least one spectral line in the point cloud, and then full multi-line fitting to detect all lines that may be present.

b. Single-line optimization First of all, we refine some given coarse initial values for the qubit parameters (period $\Pi_{\Phi_r}^{(0)}$, sweet spot location $\Phi_r^{(0)}$, $f_{ge}^{max,(0)}$ and $d^{(0)}$). This is necessary because usually the initial guess comes from the results of the single-tone spectroscopy fitting which may be inaccurate due to noise. The refinement is performed using a single spectral line by two brute force searches and then a Nelder-Mead optimization on the chosen points.

First brute force is done with a wide frequency band $\Delta f = 100 \text{ MHz}$ around it and a grid specified in .

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).

²B. J. Bloom, S. A. Caldwell, M. Reagor, and C. Rigetti, U.S. Patent No. US 20180260732A (13 Sep. 2018).

³<https://quantumexperience.ng.blumex.net>.

⁴Resonator characterization, single-tone spectroscopy, two-tone spectroscopy, excitation and readout power calibration, IQ mixer calibration, Rabi oscillations, Ramsey oscillations, fine excitation frequency adjustment, fine π -pulse adjustment, T_1 , T_2 , T_{2E} measurements.

⁵S. M. Smith, *Metrology of Quantum Control and Measurement in Superconducting Qubits*, Ph.D. thesis, University of California Santa Barbara (2018).

⁶R. Barends, J. Kelly, A. Megrant, D. Sank, E. Jeffrey, Y. Chen, Y. Yin, B. Chiaro, J. Mutus, C. Neill, *et al.*, "Coherent josephson qubit suitable for scalable quantum integrated circuits," *Physical review letters* **111**, 080502 (2013).

⁷P. Klimov, J. Kelly, Z. Chen, M. Neeley, A. Megrant, B. Burkett, R. Barends, K. Arya, B. Chiaro, Y. Chen, *et al.*, "Fluctuations of energy-relaxation times in superconducting qubits," *Physical review letters* **121**, 090502 (2018).

⁸A. 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).

⁹R. 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).

¹⁰J. 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).

¹¹A. 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).
- ¹²J. 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).
- ¹³S. 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).
- ¹⁴J. A. Nelder and R. Mead, “A simplex method for function minimization,” *The computer journal* **7**, 308–313 (1965).
- ¹⁵A. Wallraff, D. Schuster, A. Blais, J. Gambetta, J. Schreier, L. Frunzio, M. Devoret, S. Girvin, and R. Schoelkopf, “Sideband transitions and two-tone spectroscopy of a superconducting qubit strongly coupled to an on-chip cavity,” *Physical Review Letters* **99**, 050501 (2007).
- ¹⁶N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE transactions on systems, man, and cybernetics* **9**, 62–66 (1979).
- ¹⁷P. V. Hough, “Method and means for recognizing complex patterns,” (1962), uS Patent 3,069,654.