Grant Agreement Number & Project Short Name:
23FUN08 - MetSuperQ

Deliverable Reference Number and Title:
D5 - Report including procedure of measurement methodology for qubit parameters and state preparation benchmarking

Organisation name of the lead participant for the deliverable:
NPL Management Limited

Due date of the deliverable:
30 November 2025

Actual submission date of the deliverable:
24 November 2025

Confidentiality Status: PU - Public, fully open (remember to deposit public deliverables in a trusted repository)

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EURAMET. Neither the European Union nor the granting authority can be held responsible for them.

The project has received funding from the European Partnership on Metrology, co-financed from the European Union's Horizon Europe Research and Innovation Programme and by the Participating States.

TABLE OF CONTENTS

1 Summary

This deliverable of the MetSuperQ project aims to establish a procedure for the measurement methodology for superconducting qubit parameters and state preparation benchmarking. The goal is to provide a framework that is applicable to different qubit modalities, while at the same time providing a rigorous framework for executing state-of-the-art measurements on superconducting qubits such that different samples, qubits and laboratory settings can be compared. The measurement procedure is here included as an attachment, in its current form at the date of submitting this deliverable. It should be noted that this document is still a living document and has been designed to evolve as future improvements to qubits demand new or more detailed characterisation. In its current state the measurement protocol will form the basis of the MetSuperQ superconducting qubit Round Robin intercomparison, and it will subsequently be updated based on the findings and challenges encountered during this phase of the project.

2 Introduction

Within the context of the MetSuperQ project, this deliverable is the output of combined tasks 3.1, 3.3 and 3.4, relating to measurement of basic qubit properties, qubit gates benchmarking, and qubit parameter analysis, respectively.

These tasks have been concerned with:

1. Establishing a framework for measurement and analysis of single qubit parameters, and to develop measurement and benchmarking tools that are to be applied to different samples made by different institutes to facilitate comparisons. For example, by carrying out repeated measurements of basic (interleaved) qubit parameters (such as T1, T2E, T2*, f0) and determining the specific conditions under which these measurements should be carried out to ensure robust results.

2. Similarly establishing a framework for single qubit gate and readout fidelity benchmarking.

3. Creating a data analysis framework where measurements carried out in preceeding tasks are analysed using the same tools and code to ensure consistent extraction of qubit parameters from the data. This analysis code will also be applied in later stages of the project to facilitate the Round Robin intercomparison. The details of how the analysis is carried out is outlined together with each measurement step in the attached protocol.

The designed protocol balances state-of-art methods with the possibility to be generic (and applicable to different qubit modalities) and accessibility. The full benchmarking protocol is would not be practical if it takes too much time to execute per qubit, or if it requires qubits or auxiliary devices that are not readily accessible in different laboratories. Hence, the aim is that on a familiar setup, a user should be able to tune up the qubit in about a day and then execute the benchmarking protocol in a couple of days. This latter timescale is required to capture long-term drifts that are present in qubits. The protocol is also designed to be hardware agnostic, and different room temperature hardware control electronics can be used, as long as they adhere to relatively basic requirements that are nowadays fulfilled by most qubit control hardware manufacturers. Similarly, the setup in the dilution refrigerator may vary from user to user, and the protocol aims to provide benchmarking in support of understanding the decoherence induced by different hardware setups. The only strict requirement is that the qubit is connected to a (near-) quantum limited parametric amplifier such as to enable single shot readout of the qubit states. This is a key choice of this implementation which enables a range of state-of-the-art measurement techniques and understanding of decoherence budgets and sets it aside from other efforts so far in comparing qubit performance between laboratories.

Included here are two attachments: the detailed measurement protocol for running up superconducting qubits and performing the relevant benchmark measurements. This consists of descriptions of the measurements, detailed measurement methodology and parameters, what data to collect and how to analyse it. We also attach a “checklist” intended to assist the user to get through the whole benchmarking protocol without omitting any steps and recording the most crucial parameters with enough detail to be able to successfully work through the whole benchmarking protocol.

The analysis code for processing the measured data is under development for release at a later stage in the project. This is managed through a shared GIT repository, including also extensive documentation for each specific analysis and data processing step. The outline framework of this repository and functionality is available at https://opensource.orangeqs.info/metsuperq for project members. We note that this is still a living repository and documentation, which will be made publicly available at a later stage in the project.

3 Attachment 1 – Superconducting qubit measurement & analysis protocol

4 Attachment 2 – Superconducting qubit measurement & analysis checklist

This report and the two attachments complete and fulfil D5 “Report including procedure of measurement methodology for qubit parameters and state preparation benchmarking”.

The MetSuperQ superconducting qubit tuneup and benchmarking protocol

Paul Kugler[1], Nicolas Gosling[1], Masum Uddin[2], Asen Georgiev[3], Li-Wei Chang[3], Shroya Vaidya[2], Hiresh Jadoenathmissier[4], Mahmut Çetin[4], Philipp Lenhard[1], Nicolas Zapata[1], Lukas Scheller[5], Robert Gartmann[5], Wang N. Wong[2], Manognya Acharya[2], Tobias Lindström[2], Sebastian de Graaf[2], Lukas Grünhaupt[3], and Ioan Pop[1,6,7]

[1]IQMT, Karlsruhe Institute of Technology, 76131 Karslruhe, Germany
[2]National Physical Laboratory, Hampton Road, TW11 0LW, United Kingdom
[3]Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany
[4]Orange Quantum Systems, Elektronicaweg 2, 2628 XG Delft, The Netherlands
[5]IPE, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany
[6]PHI, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany
[7]Physics Institute 1, Stuttgart University, 70569 Stuttgart, Germany

November 24, 2025

Abstract

This document outlines a superconducting qubit tuneup and benchmarking protocol developed within the EURAMET EPM project "Metrology for superconducting qubits" (MetSuperQ). It aims to provide a basis for detailed comparison and benchmarking of single qubit performance across different devices, qubit designs, materials and experimental setups.

Contents

List of symbols and constants

f_r	Resonator frequency
$f_{r,g}$	Resonator frequency when qubit is in ground state
$f_{r,e}$	Resonator frequency when qubit is in excited state
χ_{ge}	Resonator dispersive shift
Q_i	Resonator quality factor (internal)
Q_c	Resonator quality factor (coupling)
Q_l	Resonator quality factor (total; loaded)
κ	Resonator linewidth/bandwidth
κ_g	Resonator linewidth/bandwidth when qubit is in ground state
κ_e	Resonator linewidth/bandwidth when qubit is in excited state
f_q	Qubit frequency
ω_q	Qubit angular frequency
$f_{0\to1}$ or f_{ge}	frequency of ground to 1st excited state qubit transition
$f_{1\to2}$ or f_{ef}	frequency of 1st to 2nd excited state qubit transition
$f_{0\to2}$ or f_{gf}	frequency of ground to 2nd excited state qubit transition
α	Qubit anharmonicity
$\|g\rangle$ or $\|0\rangle$	Qubit ground state
$\|e\rangle$ or $\|1\rangle$	Qubit first excited state
$\|f\rangle$ or $\|2\rangle$	Qubit second excited state
E_J	Qubit Josephson energy
E_C	Qubit charging energy
E_L	Qubit inductive energy
d	Qubit Josephson junction asymmetry
Φ_0	Magnetic flux quantum
S_{11}	Reflected scattering parameter
S_{21}	Transmitted scattering parameter
ΔSNR	Parametric amplifier signal to noise improvement
SNR	Signal-to-noise ratio (for IQ clouds)
μ_g	mean center position of the IQ cloud for the qubit ground state
μ_e	mean center position of the IQ cloud for the qubit 1st excited state
σ_g	standard deviation of the IQ cloud for the qubit ground state
σ_e	standard deviation of the IQ cloud for the qubit 1st excited state
T_q	Qubit temperature
T_1	Qubit energy relaxation time
T_2^*	Qubit Ramsey dephasing time
T_2^E	Qubit echo dephasing time
τ	Readout pulse length
τ_r	Recovery delay time
τ_d	Delay time
τ_π	π-pulse length
$\tau_{\pi/2}$	$\pi/2$-pulse length
$\varepsilon_{\mathrm{relaxation}}$	Readout relaxation error
$\varepsilon_{\mathrm{assignment}}$	Readout assignment error
n_{ph}	Intra-cavity photon number
n_m	Readout pulse photon number

1 Introduction

Superconducting qubits presents a promising platform for building a quantum computer due to the relative ease of fabricating large-scale circuits using conventional semiconductor processes and the fast qubit control and manipulation (compared to qubit coherence times). However, even single superconducting qubits constitute very rich physical systems, in particular due to the interaction with their environmental degrees of freedom.

The overall goal for building a fault tolerant quantum computer is to construct a large number of logical qubits (each consisting of several physical qubits) capable of running quantum error correction. In this regard, superconducting qubit platforms have demonstrated remarkable progress in recent years [1–4]. The performance of individual qubits play a very important role, as the more coherent and stable they are, the smaller the hardware and calibration overhead required to implement quantum error correction.

In order to improve the performance of individual qubits we need to make advances in fabrication techniques, materials quality, auxiliary hardware, device packaging and shielding, etc. This is because the qubits are very sensitive to their external environment, with many different mechanisms that can cause decoherence and errors. This results in the degraded performance of qubits, but what presents an even more significant challenge is the instabilities in qubit performance. These instabilities, of many origins, increase as the average qubit coherence increases, and this makes it even more challenging to compare the performance of one qubit to another. Comprehensive statistics is often required to correlate material properties with qubit performance [5–8], and deducing subtle improvements due to e.g. a change in fabrication process can be very challenging. Furthermore, due to non-Markovian effects in the qubit environment [9, 10], the way that the experiment is performed has an impact on the result. At the same time the result and interpretation of many experiments will depend on how well the qubits can be controlled and read out; for example a non-ideal readout can cause transitions of the qubit out of the computational two-state subspace, affecting measured signals and qubit dynamics. Comparison is also made more challenging by the many different performance metrics available, many of which all need to be addressed to engineer large-scale quantum processors, and rarely evaluated together for the same device. To accelerate performance improvements the community need to report how a new device or qubit material performs across all the relevant metrics. In many cases in the past only "hero" devices were reported, in which case very little can be said about the actual average performance. However, in recent years an increasing amount of studies are looking at statistical spread of qubit performance both for individual qubits over time [11], and chip-to-chip variations [12, 13]. Most frequently statistics of T_1 times and resonator quality factors are reported, and sometimes also T_2 times. While this provides improvement in how meaningful comparisons can be carried out, given the above challenges, we still need increased consistency in how measurements are carried out, and more extensive metrics are required to disentangle error sources and decoherence channels to improve overall device performance.

This document aims to provide a framework for such a methodology for calibrating ("tuning up") single superconducting qubits and extracting relevant performance metrics ("quantities of interest", QoI). The purpose is to facilitate better comparison between different qubits, measurement setups and hardware. This protocol is designed to be modular, generalizable to different qubit modalities, and compatible with a range of control stacks and software/firmware commonly used for measurements. While the sequence and structure are unified, each lab can adapt the implementation details to their specific hardware.

We note that this protocol is intended to be a living document: it is expected to evolve over time as improved methods, new analysis tools, or alternative calibration strategies become available. Updates should be discussed and shared within the community to ensure long-term consistency and quality across all measurements compared.

The benchmarking procedure is divided into two parts. First a tuneup protocol to ensure all relevant qubit parameters are extracted and that readout and qubit gates are calibrated and optimised. The second part of the benchmarking relies on using the tuned up qubit parameters to measure specific quantities of interest (QoI's) that form the basis of comparing performance between different qubits. All the tuneup and QoI's come with specific analysis procedure and code (to be made available at a later stage).

It is expected that all the data and metadata mentioned throughout is saved together with metadata describing appropriate qubit and setup identifiers and properties.
The measurement protocol relies the following key assumptions:

- Single-shot readout is possible, with good fidelity. I.e. a parametric amplifier is used in the setup, and the qubit has reasonably good state separation fidelity.
- Qubits are measured at dilution refrigerator base temperatures of ~ 10 mK and are reasonably well thermalised to this temperature.
- The protocol is designed to be generic and can be applied to different qubit types, such as transmons, flux qubits or fluxonium.
- Some parts of the protocol are applicable to frequency tunable qubits, which here assumes flux tuning through an external solenoid or on-chip flux line. However, the same methodology applies to qubits with other means of frequency tuning. For fixed frequency qubits these steps in the protocol can simply be omitted.

The Quantities of Interest can, for the diligent experimentalist, be measured as a function of varying other external parameters, such as cryostat temperature, external flux, applied charge or strain. This can serve to provide more information about how relevant physical properties contribute to decoherence and qubit errors. Description of such measurements goes beyond the scope of this document. We also note that this document does not aim to provide a thorough description of how superconducting qubits work. For that we refer to recent reviews on the subject, for example ref. [4].

2 Qubit Tuneup Protocol

The goal of this protocol is to define a standardised procedure for tuning up superconducting qubits across different platforms and laboratories. It serves as a guide for building a tuneup notebook in which all relevant measurements are executed in a consistent and well-documented manner.

Each step in this protocol specifies the measurement objective and procedure, required settings, assumptions and limitations, expected results, and the information that must be recorded for later use. By unifying both the measurement sequence and its analysis, we ensure comparability and reproducibility across experiments and labs.

In addition to the tuneup notebook, a corresponding analysis notebook should be developed. This notebook loads the acquired data, performs the necessary fits and calibrations, and extracts the quantities of interest.

2.1 Overall Tuneup Procedure

The tuneup procedure is organised in two main stages. First, a series of frequency domain measurements is performed to extract initial estimates of key system properties, such as resonator frequencies, visibility of flux dependence, and qubit-resonator coupling strength. Second, these results are used as the starting point for subsequent time-domain calibration. Figure 1 shows a flowchart for the overall tune-up procedure.

The objective of the full tuneup is to obtain reliable values for all relevant device parameters, including physical circuit parameters (e.g., resonator frequency f_r, quality factor Q, Josephson energy E_J, charging energy E_C), establish a QND (quantum non-demolition) readout, and calibrate high-fidelity π- and $\pi/2$-pulses. The outcome is a qubit device that is fully characterised and ready for further experiments or integration into benchmarking routines.

The tuneup procedure also contains a number of checks that certain criteria are met. If not met this is an indication that there is something wrong with the setup which needs fixing, or the qubit in question does not have good enough performance to successfully carry out the subsequent QoI benchmarking steps.

All relevant settings of the measurement hardware (e.g., frequencies, powers, durations) should be saved automatically as part of the raw data files. Each section of this protocol specifies only the quantities of interest to be extracted and recorded for later use or database integration.

2.2 Frequency domain

2.2.1 Attenuation Check

Comparison between the expected and measured attenuation of the fridge.

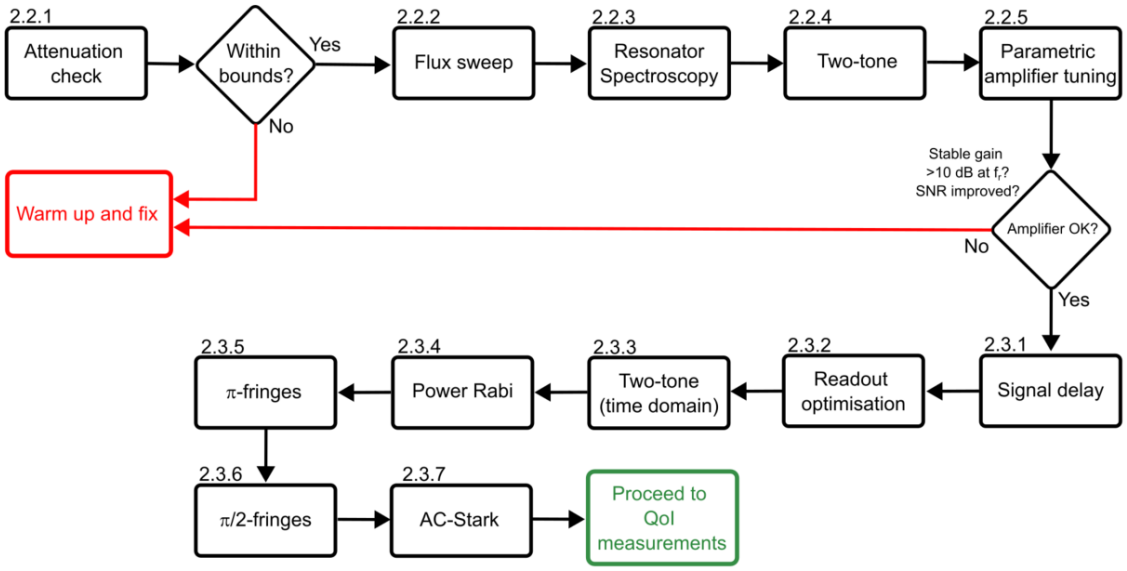


Figure 1: Flowchart for the tune-up procedure. Each step is described in detail in the following sections. A separate flowchart for the readout optimisation (step 2.3.2) is shown in Figure 8.

Description

To verify that the attenuation and gain in the fridge and room temperature setup are within expected limits, we compare an estimated signal budget with a measured transmission trace around the readout frequency. If the difference between the expected and measured values is reasonably small (typically $< 10\,\mathrm{dB}$), we can assume that all components are properly connected and functioning.

This measurement serves as a quick sanity check before proceeding with more detailed calibrations.

Measurement procedure

- Choose a probe frequency close to the resonator readout frequency, but not directly on resonance.
- Calculate the expected attenuation and gain based on:
  - fixed attenuators (datasheet values),
  - estimated cable attenuation at the probe frequency,
  - insertion loss of individual components,
  - gain from cryogenic and room temperature amplifiers.
- Example calculation:

	Attenuation		**Amplification**	
-	60 dB (fixed attenuators)		40 dB (HEMT)	
-	15 dB (cable loss estimate)	+	38 dB (RT amplifier)	
-	3 dB (insertion loss)			
-	78 dB	+	78 dB	= 0 dB total

- Measure the transmission through the cryostat (S_{21}) around the same probe frequency using a VNA or equivalent setup.
- Extract the background signal level away from the resonator dip or peak.
- Compare this measured transmission value with the estimated signal level.

Assumptions and limitations

- All components are assumed to behave linearly at the selected probe power.
- Cable losses and insertion losses must be reasonably estimated or measured beforehand.
- Reflections and impedance mismatches may distort S_{21}, especially in poorly matched systems.
- This is not a precise calibration — deviations within $< 10\,\mathrm{dB}$ are typically acceptable.

Required data

- Estimated total attenuation and gain values (in dB).
- Measured transmission (or reflection) level near the resonator frequency (in dB).
- Final discrepancy between measured and estimated values (in dB).
- Confirmation whether the discrepancy is within acceptable bounds to proceed.

2.2.2 Flux Sweep

Measurement of the resonator response as a function of the applied magnetic flux to the qubit. Depending on the qubit type and coupling strength, characteristic frequency shifts or avoided crossings can be observed as the qubit frequency varies with flux.

Description
For flux-tunable qubits, the transition frequencies between energy levels depend on the externally applied magnetic flux. Due to the qubit-resonator coupling, changes in the qubit frequency affect the resonator frequency. This typically manifests as:

- level anti-crossings (e.g., in flux qubits),
- or a smooth cosine-like modulation of the resonator frequency (e.g., in SQUID transmons).

These features allow for an approximate localization of the qubits sweet spots (typically at $0\,\Phi_0$ and $\pm 0.5\,\Phi_0$) and provide a strong indication of a functional and coupled device. An example is shown in fig. 2.

Measurement procedure

- Find the resonator and set up a reasonable ($\pm 5 - 10$ resonator linewidths) frequency range on the VNA.
- Sweep the current through the flux-bias line or coil over a range wide enough to cover at least one flux period Φ_0.
- At each current point, measure the resonator response:
  - sweep the probe frequency around the expected resonator frequency.
  - fit the resonator frequency using a circle fitting routine as can be found in `https://github.com/qkitgroup/qkit/tree/master/src/qkit/analysis/circle_fit/circle_fit_2019`
- Identify characteristic features in the resonator response:
  - anti-crossings for strongly coupled qubit-resonator systems,
  - periodic shifts for transmons.
- From the periodicity, extract the current range that corresponds to one flux quantum Φ_0 in the qubit loop.
- Rescale the x-axis accordingly and shift it so that sweet spots align with $0\,\Phi_0$ and $\pm 0.5\,\Phi_0$.
- Generate a 2D color plot with:

  - frequency on the y-axis,
  - current (and later calibrated flux) on the x-axis,
  - amplitude or phase of the resonator response as the color scale.
- Save this 2D dataset in a single file (e.g., HDF5) with clear axis labels and units.

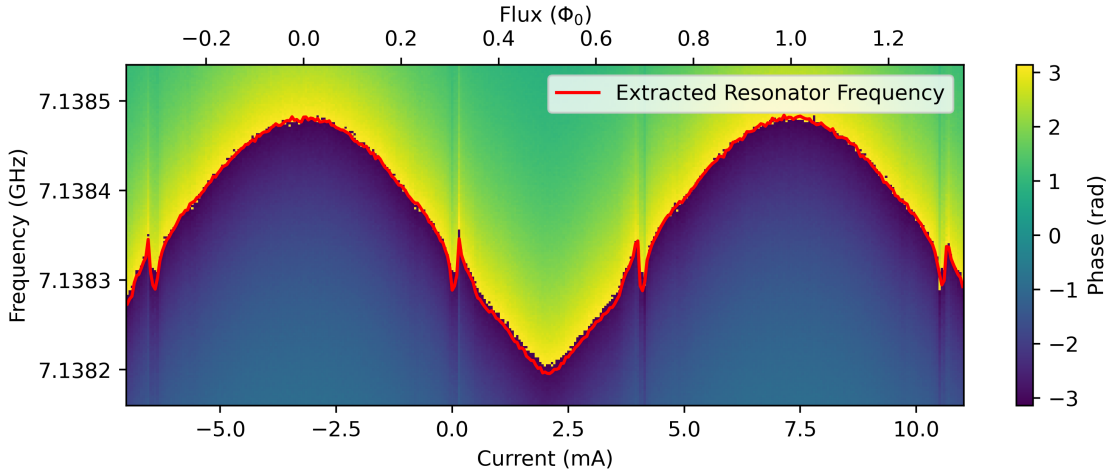


Figure 2: Flux sweep of the resonator for a frequency tunable transmon. The lower x-axis is in current, which is calibrated to flux for the upper x-axis. The resonator frequency is extracted by fitting the spectrum for each current step and is highlighted in red. The periodic appearing feature at 0 mA is due to a $0 \to 3$ qubit transition driven by the readout tone.

Assumptions and limitations

- The resonator is measured with sufficiently low power. For increasing power, the resonance frequency starts to move, we want to be way below this power.
- The qubit is flux-tunable and coupled to the resonator.
- The current must induce a linear and stable magnetic flux through the qubit loop.
- The range of the current sweep should exceed one full flux quantum Φ_0.
- Low-frequency flux noise or hysteresis may blur periodicity or sweet spot identification.

Required data

- Measured 2D dataset (amplitude and phase) as a function of current and frequency.
- Calibrated flux axis (in units of Φ_0), derived from periodicity in current.
- Extracted resonator frequency as a function of current.

2.2.3 Resonator Spectroscopy

Measurement of the resonator for fitting and extracting resonance frequency, bandwidth, and quality factors.

Description

To enable consistent device comparisons and prepare for downstream calibrations, we extract key properties of the readout resonator: the resonance frequency f_r, the total linewidth κ, the loaded quality factor Q_l, and its decomposition into coupling Q_c and internal Q_i components.

The frequency f_r, linewidth κ, and Q_l can be obtained using standard amplitude and phase fitting. Extracting Q_c and Q_i requires fitting the complex resonator response using a circle fit. A typical result is shown in fig. 3.

Measurement procedure

- If applicable, tune the qubit to a point where it has low effect on the resonator (e.g. zero-flux).
- Measure the complex response centered on the expected resonator frequency:
  - Use a sufficiently wide frequency span to fully capture the resonance dip, including enough of the shoulders to get a good measure of the background signal required for an accurate fit.
  - Apply low probe power to stay in the linear response regime.
  - Use sufficient averaging and low IF bandwidth to suppress noise.
- Fit the complex response using a circle fitting routine as can be found in `https://github.com/qkitgroup/qkit/tree/master/src/qkit/analysis/circle_fit/circle_fit_2019`
- Extract the following from the fit:
  - Resonator frequency f_r,
  - Loaded quality factor Q_l,
  - Coupling quality factor Q_c,
  - Compute Q_i from: $\frac{1}{Q_l} = \frac{1}{Q_c} + \frac{1}{Q_i}$,
  - Compute the bandwidth from: $\kappa/2\pi = f_r/Q_l$.
- Repeat the measurement across minimum 5 readout powers spanning from a low-power regime (resonator frequency invariant with power) to a high-power regime (resonator frequency shifts with power until the resonator is effectively decoupled from the qubit).

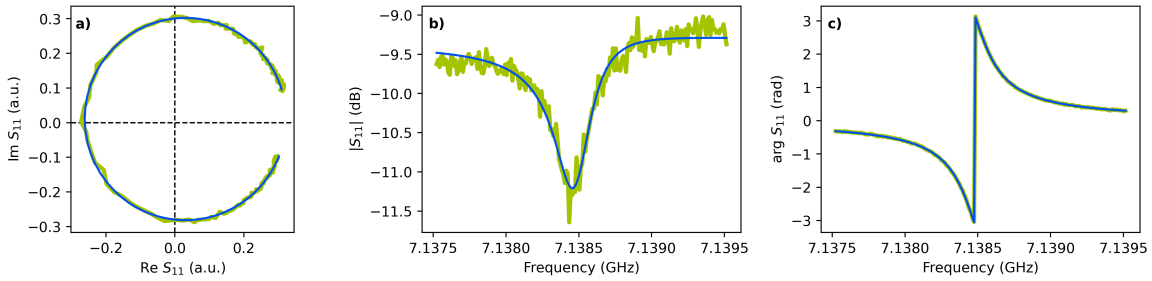


Figure 3: Simultaneous **a)** circle fit to the complex resonator response, **b)** its amplitude, and **c)** phase spectra. The fit here yields: $f_r = 7.1385\,\text{GHz}$, $Q_l = 2.1 \cdot 10^4$, and $Q_c = 2.32 \cdot 10^4$.

In this example, we obtain $\kappa/2\pi = 0.34\,\text{MHz}$ and $Q_i = 2.2 \cdot 10^5$. Note that in the over-coupled regime, Q_i and Q_c become increasingly uncertain due to low sensitivity to internal loss, see [14].

Required data

- Complex measurement data (I/Q or magnitude/phase) vs frequency and power.
- Fitted resonator parameters:
  - Resonance frequency f_r (Hz),
  - Frequency shift from low power to decoupled qubit regime,
  - Loaded quality factor Q_l,
  - Coupling quality factor Q_c,
  - Internal quality factor Q_i,
  - Bandwidth κ (Hz).
- Difference in resonator frequency between low power and punch out regime.

2.2.4 Two-tone

A measurement technique in which the resonator response is monitored while a second pump tone is applied to the qubit. Once the frequency of the pump tone matches a qubit transition, the resonator response changes.

Description
This technique involves applying a secondary microwave tone (the "pump tone") to the qubit while continuously monitoring the complex response of the readout resonator. The pump frequency is swept as a function of applied pump power or magnetic flux. When the pump tone drives a qubit transition (at the right frequency and with sufficient power), the resulting change in the qubit state modifies its dispersive interaction with the resonator, leading to a measurable shift in the resonator signal.

This allows us to map out the qubits energy spectrum, in particular, it allows for:

- Identification of the zero- and half-flux sweet spots with high accuracy,
- Extraction of transition frequencies at those sweet spots (especially $|0\rangle \rightarrow |1\rangle$),
- Observation of the shape and slope of qubit transition lines away from the sweet spots.

This information is valuable as a robust diagnostic of qubit functionality and for fitting the qubit Hamiltonian using numerical models (e.g., `scqubits`). In doing so, we extract circuit parameters such as E_J, E_C, E_L, transition frequencies, anharmonicity, and junction asymmetries. An example of this fit is shown in fig. 4 for a flux tunable transmon. At a later stage in this tuneup protocol, a two-tone experiment is performed in time-domain to extract transition frequencies and sweet spots with higher accuracy.

Measurement procedure The measurement procedure is an iterative process in which the parameters have to be determined step by step. Once all relevant ranges and settings leading to the best contrast are known, a final measurement can be performed.

- Set the readout tone to the current resonator frequency (zero-span (CW) mode, minimal points) and adjust with flux as needed using the previously extracted relation between resonance frequency and flux (current). The readout amplitude should be in the low power regime. IF-bandwidth and averages are set to a value that gives reasonable small fluctuations compared to the expected change in the signal.
- Start by performing a two-tone vs pump power at (if applicable) a fixed flux, ideally close to the sweet spot of the qubit: monitor the resonator response as a function of pump frequency and pump power. Identify the pump power at which the qubit transitions become visible. Usually, the transition appearing at the lowest power belongs to $|0\rangle \rightarrow |1\rangle$, but more transitions become visible at higher powers, as can be seen in fig. 5. If a transition is probed with too much power, the line is getting power broadened. Fix the power at which the transition frequency is visible and sharp. Adjust the pump-frequency step size to be smaller than the transition linewidth.
- (For flux tunable quits) Perform a two-tone vs flux: monitor the resonator response as a function of pump frequency and flux. If possible, cover a flux range that corresponds to a flux quantum and choose a reasonable flux step size of the order of 1-2% of Φ_0.
- Repeat the flux two-tone for at least the $|0\rangle \rightarrow |1\rangle$ and one other qubit transition (with its corresponding ideal pump power).

Fit the measured qubit spectrum using `scqubits` to extract circuit parameters and (if applicable) refine the relationship between current and flux as shown in fig. 4.

Required data

- Measured 3D datasets (amplitude and phase at the cw resonator frequency) of the $|0\rangle \rightarrow |1\rangle$ and one other qubit transition as a function of pump frequency and
  - pump power
  - (For flux tunable qubits) magnetic flux, covering at least a range of one flux quantum Φ_0.

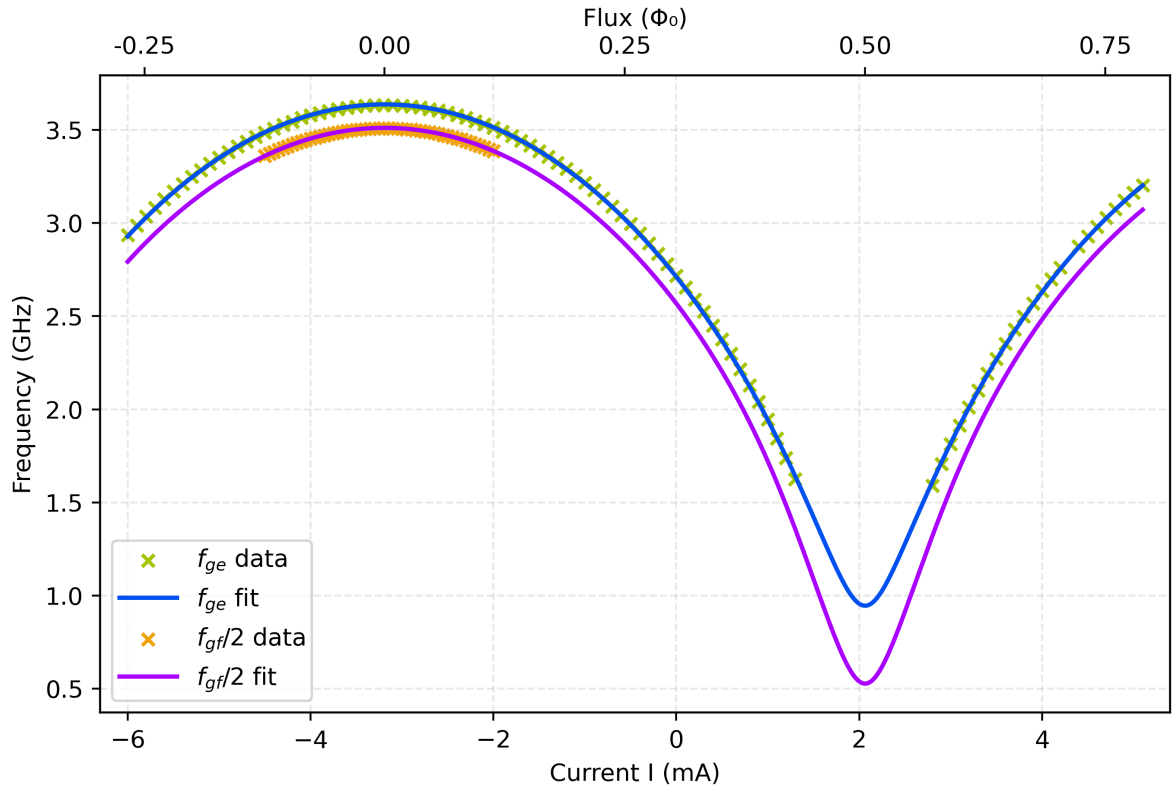


Figure 4: Two-tone spectroscopy data and fit for a flux tunable transmon using the python package scqubits. We extract $E_C = 0.2141\,\text{GHz}$, $E_J = 8.721\,\text{GHz}$, $f_{ge,\Phi=0} = 3.637\,\text{GHz}$, anharmonicity $\alpha = -0.251\,\text{GHz}$ and junction asymmetry $\Delta E_J = 0.456\,\text{GHz}$

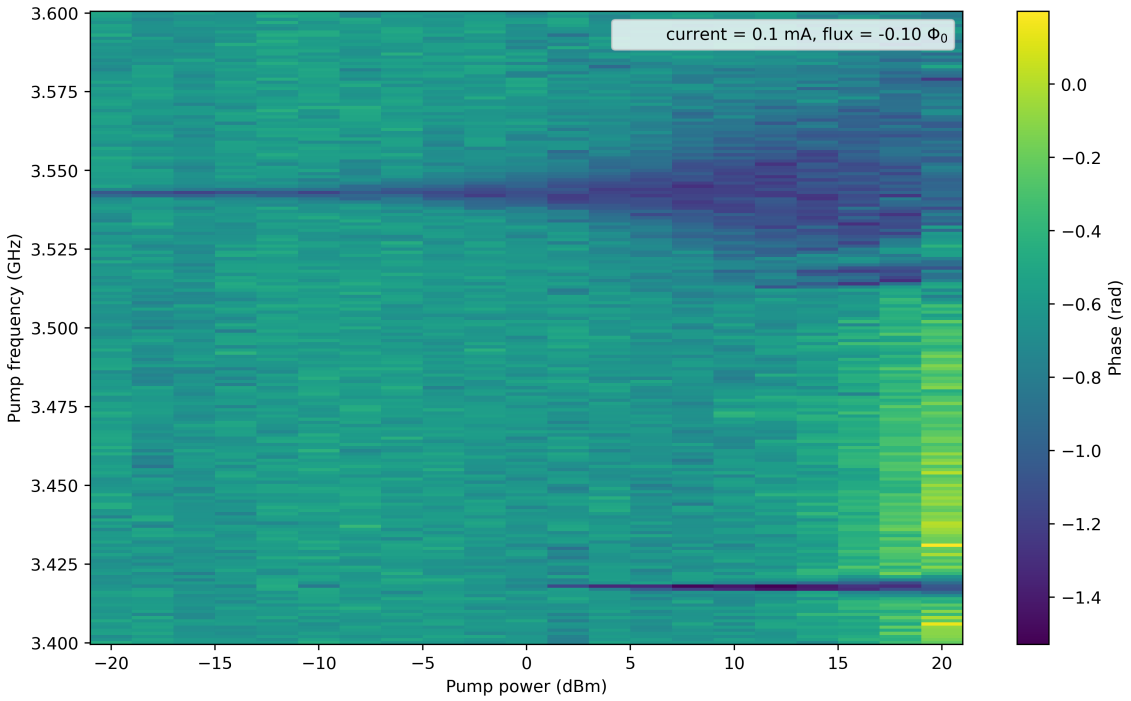


Figure 5: Two-tone spectroscopy of a flux tunable transmon qubit for different pump powers. The $|0\rangle \rightarrow |1\rangle$ transition appears at around 3.54 GHz already for a small pump power. For higher pump powers, the spectral width of the transition becomes visibly wider. Starting from a pump power of 2 dBm, the $|0\rangle \rightarrow |2\rangle$ two photon transition becomes visible at a frequency of $f_{0\rightarrow 2}/2 = 3.42\,\text{GHz}$.

- Circuit parameters of the qubit spectrum fitted with `scqubits`, namely, the transition frequency at the sweet spot $f_{0\to1}$, anharmonicity α, Josephson energy E_J, charging energy E_C and, if applicable, inductive energy E_L and junction asymmetry d.
- Refined current to flux relationship from the fit.

Assumptions and limitations

- The resonator is measured with a sufficiently low power.
- Observed qubit transitions are correctly assigned.
- To reduce measurement time, it is helpful to measure only along the expected transition frequency and not the full x-y range.

2.2.5 Parametric Amplifier Tuning

Setup for the used parametric amplifier so that the signal to noise ratio (SNR) of the output line is maximised.

Description

Tuning a parametric amplifier such as a JPA, JPC, or TWPA is essential to exceed the input noise of the HEMT and achieve high-fidelity amplification of the readout signal. The amplifier must therefore deliver an improvement in the SNR at a typical gain of 10–20 dB near the resonator frequency while remaining stable, low-noise, and free from distortion or saturation. The tuning is an iterative multi-parameter process involving adjustment of:

- the flux bias (for flux-tunable amplifiers),
- the pump tone frequency,
- and the pump tone power.

The parameters are interdependent and must be scanned together to locate an optimal operating point at which the SNR improvement is maximal. The pump tone must be chosen to avoid all qubit transition frequencies and the resonator frequency.

Once an optimal configuration is found, these parameters should be fixed and monitored throughout all subsequent time-domain experiments.

Measurement procedure

In the first part, the amplifier is tuned to give a stable gain in the range of 10-20 dB after which the SNR is optimised.

- Set a weak probe tone at or near the resonator readout frequency and monitor its gain through the full signal chain (ensure that the tone at the amplifier is below its signal saturation power, taking into account attenuation in the input line on the fridge).
- If the amplifier is flux-tunable:
  - Sweep the flux bias and monitor gain to identify bias points with high gain and low insertion loss.
  - Narrow in on promising regions and refine flux bias to maximize stability.
- Iteratively sweep:
  - Pump frequency across the expected operation band.
  - Pump power from below to above threshold, in small steps.
  - For each (flux, frequency, power) combination, measure the probe gain.
- Map out gain vs. all three tuning parameters and select an operating point with:
  - stable gain in the target range ($\sim$10–20 dB),

  - minimal noise or ringing,
  - and no overlap with known qubit or resonator frequencies.

- Connect the output of the fridge to a spectrum analyzer and set it up to measure around the applied probe tone. We now define ΔSNR as the distance between the probe tone and the level of the noise floor within the amplification band in dB. To find the noise floor from the amplification chain, decrease the bandwidth of the spectrum analyzer until the noise does not change anymore.
- Measure this ΔSNR as a function of pump power to get a plot that looks similar to the one shown in fig. 6.
- Ideally repeat this measurement as a function of pump frequency or flux and select an operating point with maximal SNR.
- Confirm that gain and SNR are stable by measuring for 10 minutes. Repeat this check from time to time.
- Lock in the pump and flux settings and use them for the remainder of the tuneup.

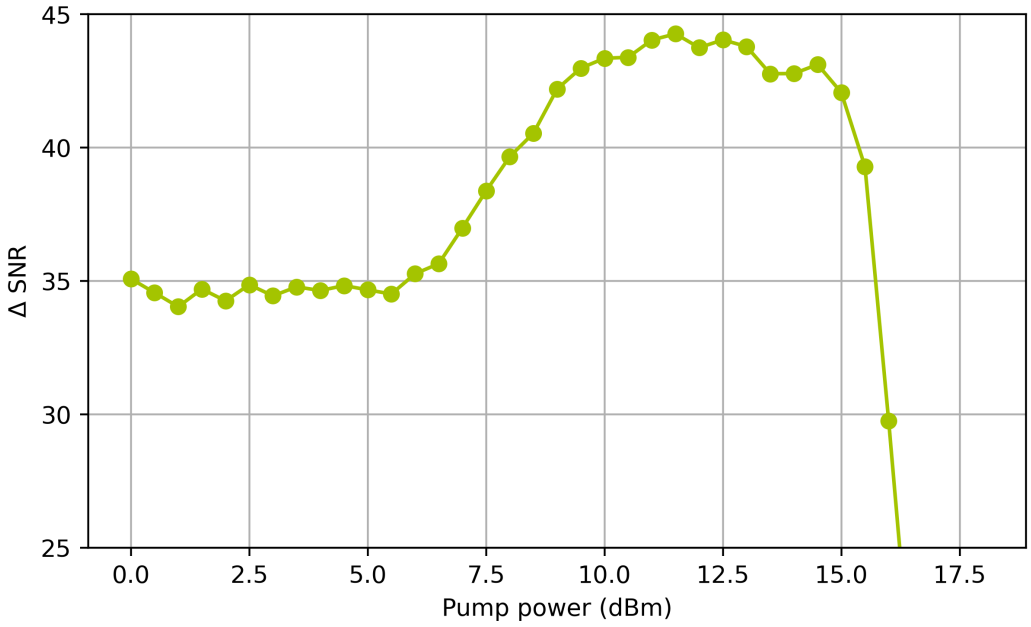


Figure 6: ΔSNR (distance between probe tone at 7.13 GHz and noise floor) as a function of room temperature pump power applied to a TWPA. Starting at 6 dBm, the TWPA produces gain, effectively increasing the SNR by slowly amplifying the signal well above the noise of the HEMT. At 10 dBm, the HEMTs noise is fully saturated and the SNR is limited by the noise of the TWPA, creating a plateau of flat SNR. For pump powers above 14 dBm, the SNR crashes significantly, even below the level before the TWPA increased the gain.

Note: Parametric amplifiers are sensitive to environmental fluctuations (e.g., magnetic fields, temperature drifts), so occasional retuning might be necessary during long measurement sessions.
Note: If $>$ 10 dB gain and clear SNR improvement cannot be obtained, the remaining part of this protocol may be difficult to complete successfully, and it might be advisable to stop at this point and try to correct any hardware issues with the parametric amplifier.

Required data

- Final pump tone frequency and power.
- Final flux bias (if applicable).

- Measured gain profile.
- Final gain at the readout frequency (in dB) and ΔSNR improvement (on/off).
- Confirmation that the pump tone is well-separated from qubit and resonator frequencies.

2.3 Time domain

In the time-domain section, we characterize the dynamic properties of the qubit using pulsed measurements. These measurements rely on heterodyne detection techniques, which allow us to generate and analyze microwave signals with precise frequency and phase control using lower-frequency hardware.

Heterodyne measurement and signal chain

In a typical heterodyne architecture, arbitrary waveform generators (AWGs) or digital-to-analog converters (DACs) produce intermediate frequency (IF) signals, which are then upconverted to microwave frequencies (RF) via IQ mixers. The mixing process uses a continuous-wave local oscillator (LO) signal to shift the IF to the desired RF frequency. This allows for flexible pulse generation at frequencies near the qubit or resonator frequency without requiring direct GHz-range waveform synthesis.

On the measurement side, the reflected or transmitted RF signal is downconverted back to IF using the same LO frequency. The resulting IF signal is digitised and demodulated to extract in-phase (I) and quadrature (Q) components, which fully characterize the amplitude and phase of the signal envelope.

Signal definitions

- **LO (Local Oscillator):** A continuous-wave microwave tone used as a reference for up- and down-conversion. The LO frequency is typically set close to the qubit or resonator frequency.
- **IF (Intermediate Frequency):** A lower-frequency signal (typically tens to hundreds of MHz) generated digitally. It carries the pulse envelope shape and is mixed with the LO to produce the RF signal.
- **RF (Radio Frequency):** The physical microwave signal sent to or received from the device, resulting from the mixing of IF and LO signals.
- **IQ signals:** The complex-valued time-domain signal, represented by its in-phase (I) and quadrature (Q) components. These are used both in waveform synthesis and in demodulation of the readout signal.

This heterodyne architecture allows for high-precision control of qubit operations and flexible measurement strategies with narrowband detection and efficient frequency multiplexing. In the following subsections, we will use this setup to perform relaxation, coherence, and gate calibration measurements in the time domain.

Note: When using this heterodyne architecture, make sure that the LO frequency and mixer harmonics do not coincide with qubit transitions or pump frequencies. Alternatively, make sure that the signal sent in the fridge is well filtered.
Note: It is here assumed that any required mixer calibration for readout at the resonator frequency has been performed, and any sidebands are at least -40 dB below the carrier. Some hardware offer built-in automatic mixer calibration, for which this step is not required.

2.3.1 Signal Delay

Measurement to determine the delay of the readout signal and the response envelope of the resonator in the time domain.

Description

Due to the physical cable lengths and latencies of the control hardware, there is a finite delay between emitting a readout pulse and detecting the returning signal. This includes cable propagation delay as well as internal latencies (e.g., FPGA/ADC delays). Typical total round-trip delays are on the order of hundreds of nanoseconds.

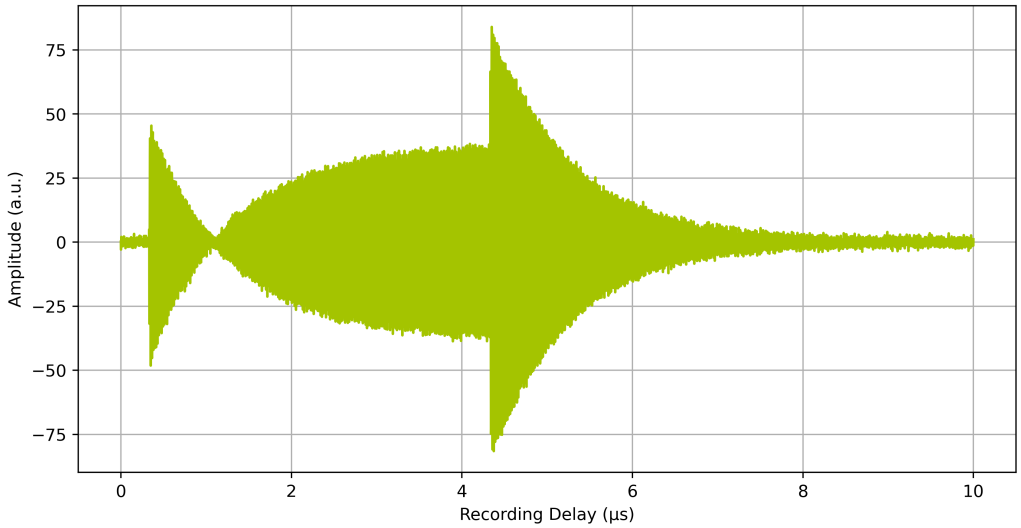


Figure 7: Signal from the resonator–qubit system in response to a $4\,\mu$s readout pulse, measured in reflection. After a delay of approximately $0.35\,\mu$s, the first reflected photons from the resonator reach the acquisition system, causing the recorded amplitude to rise. From this point, the resonator begins to ring up. As the intra-resonator photon population builds up, the emitted photons undergo a π phase shift relative to the incoming pulse. These interfere destructively with the directly reflected drive signal, reducing the measured amplitude until around $1.1\,\mu$s. This marks the point where the recorded signal consists predominantly of photons that have interacted with the resonator, now carrying the full π phase shift. To avoid integrating over uninformative reflections, the data acquisition should begin at this point. After the input pulse ends at $4.35\,\mu$s ($0.35+4\,\mu$s), the last drive photons reach the resonator. The absence of further destructive interference causes a second amplitude jump, after which the resonator rings down until empty.

We perform this measurement to determine a global integration delay — the point in time after which IQ demodulation of the readout signal should begin. Additionally, the measurement reveals the ring-up and ring-down characteristics of the resonator, shaped by both the system transfer function and the applied pulse envelope.

This delay must be fixed and used in all subsequent time-domain experiments to ensure consistent readout timing and fidelity. An example of a measurement of the resonator response is shown in fig. 7.

Measurement procedure

- Apply a readout pulse at the previously extracted resonator frequency.
- Use a flat-top or shaped pulse with duration longer than the resonator lifetime.
- Record the full heterodyne response: complex I and Q signals vs. time, covering:
  - a short baseline before the pulse starts,
  - the duration of the pulse,
  - and several microseconds beyond the end of the pulse.
- Analyze the envelope of the complex signal to extract:
  - approximate delay until signal onset (round-trip time),
  - time to reach steady-state (after ring-up),
  - a suitable integration window start time (global readout delay).

The chosen delay should ensure that the signal is integrated only after it contains photons that left the resonator, but before significant decay has begun. This point will define the global readout delay used in all future qubit tuneup steps.

Required data

- Full time trace of the readout response.
- Pulse shape and duration used for excitation.
- Extracted round-trip delay (from pulse start to signal onset).
- Chosen integration start time (global delay, in ns) and end (usually the integration is as long as the pulse itself).

2.3.2 Readout optimization

Optimizing the readout is a crucial step in the tuneup to ensure that it is QND (quantum non demolition) and gives as small error as possible. The parameters that have to be optimised for this comprise pulse length, recording duration (usually as long as the pulse length), frequency, and amplitude (later calibrated photon number). Optionally, the pulse shape can be optimised to maximize the fraction of time during readout in which the resonator is in its steady state. The resulting faster rise and decay times minimize MIST (measurement induced state transitions) effects but also allow for overall shorter readout pulses.

The idea behind readout optimization is outlined in fig. 8 and breaks down into different iterative loops, where the individual steps are described in the following sections.

1. Find clouds and optimize until SNR = 2.5. SNR is defined as the distance between the ground- ($\boldsymbol{\mu}_g$) and excited-state ($\boldsymbol{\mu}_e$) centers divided by the sum of their standard deviations σ: $\mathrm{SNR} = \dfrac{\boldsymbol{\mu}_g - \boldsymbol{\mu}_e}{\sigma_g + \sigma_E}$.
   - IQ clouds are looked for with a rather strong readout as a starting point.
   - Once clouds are found, perform a frequency optimization as explained in 2.3.2.2 and repeat this step every time the readout pulse length or amplitude is changed.
   - If the SNR is less than 2.5, increase readout pulse length τ. If this does not improve the SNR, increase the readout pulse amplitude (n_{ph}).
   - If the SNR is greater than 2.5, decrease readout pulse amplitude.
2. Match assignment and relaxation/readout error.
   - Measure T_1 via measurement preparation with readout off as explained in 2.3.2.3. (The readout on step in 2.3.2.3 can also be carried out, but at this stage the data obtained from that step is not used, and time can be saved by skipping this step.)
   - Calculate the assignment error $\epsilon_{\mathrm{assignment}}$ (IQ cloud fit output).
   - Calculate the relaxation/readout error $\epsilon_{\mathrm{relaxation}}$ from T_1 and the used readout pulse length τ.
   - If $\epsilon_{\mathrm{assignment}} > \epsilon_{\mathrm{relaxation}}$, increase readout pulse length τ. (See Sec. 2.3.2.3 for definitions.)
   - If $\epsilon_{\mathrm{assignment}} < \epsilon_{\mathrm{relaxation}}$, reduce readout pulse length τ.
3. Evaluate QNDness
   - Remeasure T_1 via measurement preparation but this time alternating with readout off and on.
   - Compare $T_{1,\mathrm{on}}$ with $T_{1,\mathrm{off}}$.
   - Compare the excited state probability for the smallest delay time of the T_1 measurement for both cases ($P(|e\rangle)|_{t=0,\mathrm{on/off}}$) and check if the difference to 1 is explainable by the relaxation error $\epsilon_{\mathrm{relaxation}}$.
   - If there is a substantial mismatch, reduce the readout pulse amplitude and rematch assignment and relaxation error (previous loop).
4. Remeasure IQ clouds.

2.3.2.1 IQ clouds

Repeated stroboscopic measurements of single I/Q points give rise to clouds corresponding to the ground and excited states when visualised as a 2D histogram.

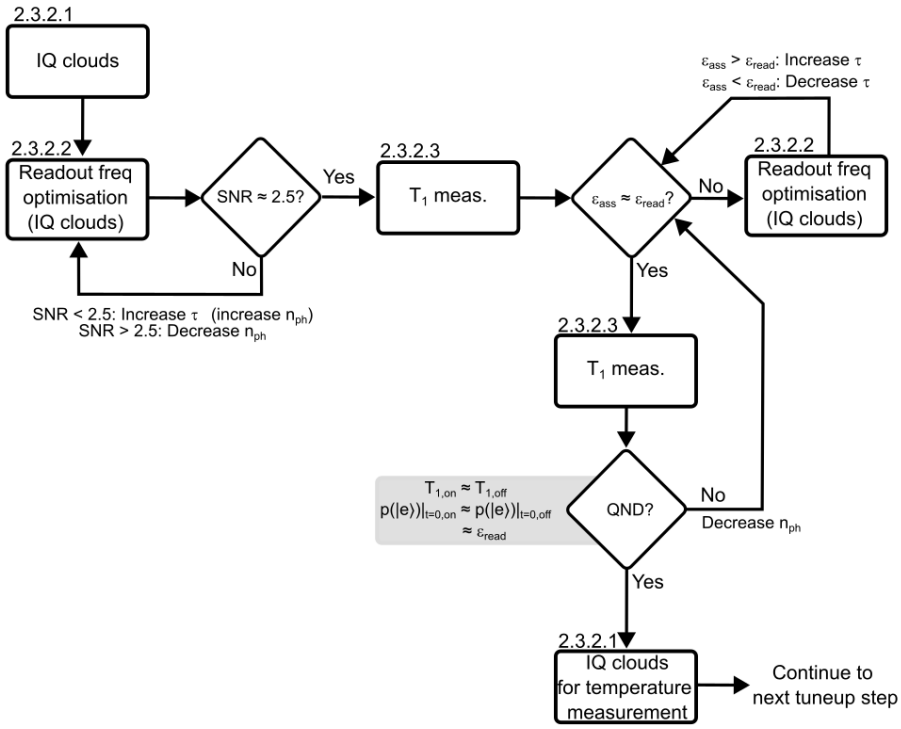


Figure 8: Flowchart for the readout optimisation procedure. Each step is described in detail in the following sections.

Description

This measurement provides a statistical view of the thermal state of the qubit by repeatedly recording the complex heterodyne response (I and Q) to a readout pulse. Performing several hundred thousand shots reveals two (or more) distinct "clouds" in the IQ plane corresponding to different qubit states.

When plotted as a 2D histogram, these clouds typically follow 2D Gaussian distributions. By fitting the histogram with a sum of Gaussians, we can extract:

- the relative population of ground and excited states (thermal equilibrium),
- the signal-to-noise ratio (SNR),
- the readout state assignment error (from SNR),
- the effective qubit temperature T_q, assuming thermal equilibrium and a known $|0\rangle \rightarrow |1\rangle$ transition frequency.

The SNR is defined as the Euclidean distance between the fitted Gaussian centers divided by the sum over cloud variances. This follows the procedure in the `qkit IQCloudAnalysis` implementation, which performs Gaussian mixture modeling using scikit-learn and extracts rotated cloud axes for optimal discrimination.

Note: In practice, it often requires several iterations with varying readout parameters (e.g., frequency, amplitude, readout pulse length, and integration time) before well-separated clouds become visible. For very cold qubits with high transition frequencies, the excited state population is very small. In this case, an additional pulse at the transition frequency must be played to artificially populate the excited state. A typical IQcloud histogram is shown in fig. 9.

Measurement procedure

- For qubits with vanishing excited state probability: play a continues tone at the $|g\rangle \rightarrow |e\rangle$ transition to increase the excited state population.
- Apply a readout pulse at the resonator frequency. (In order to find the the clouds faster, start with a strong readout pulse)

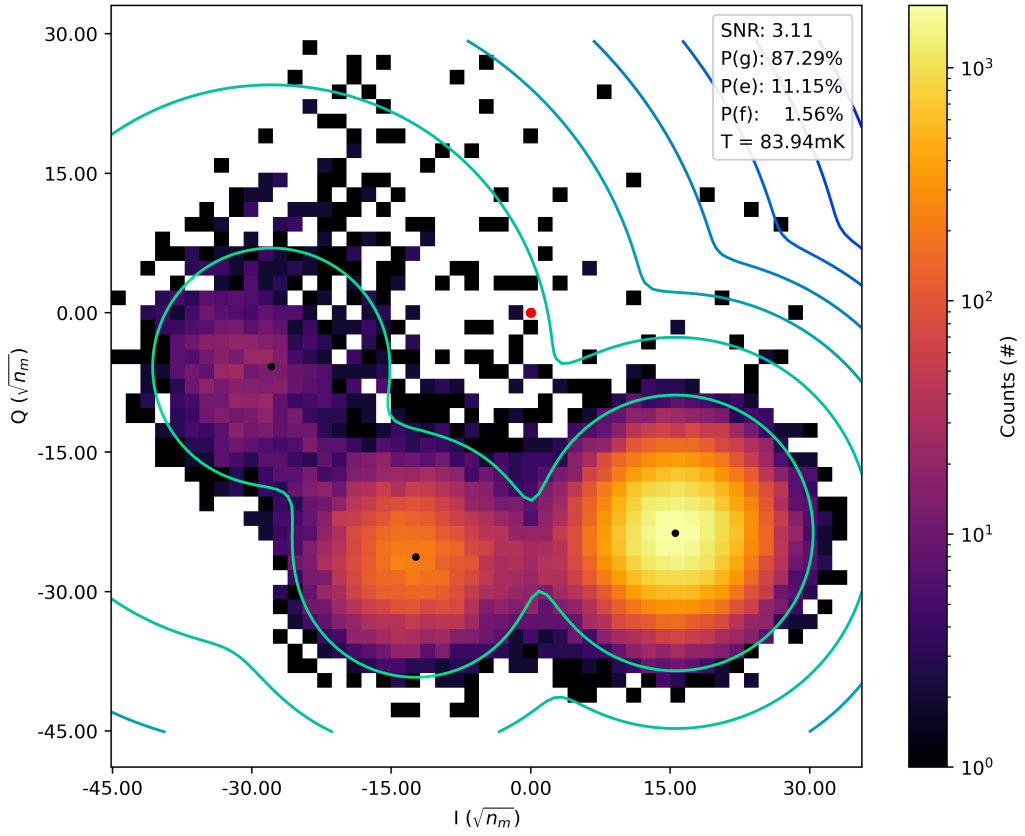


Figure 9: IQ cloud histogram of a flux tunable transmon. From fitting multiple 2D Gaussians to the histogram and under use of Bose-Einstein statistics, we extract populations of ground, excited and f-state that correspond to a 3 level system temperature of 84 mK for transition frequencies of $f_{\mathrm{ge}} = 3.63\,\mathrm{GHz}$ and $f_{\mathrm{ef}} = 3.38\,\mathrm{GHz}$. From the distance between ground state and excited state divided by their variance, we extract a signal-to-noise ratio (SNR) of 3.11. The clouds have been rotated so that they align with a single quadrature; this enables later state discrimination using a vertical threshold line. The axis are in units of the square root of the number of measurement photons, which can be calculated once the power calibration in section 2.3.7 is done.

- Wait for the global integration delay determined previously.
- Record a single IQ sample per shot (i.e., demodulated complex value).
- Wait for resonator ring-down, $^{10}/_{\kappa}$.
- Repeat this cycle $\sim 10^6$ times.
- Visualize the resulting IQ samples as a 2D histogram with 100 bins.
- Fit with 2D Gaussians to extract state populations, SNR and according assignment error. Use manual input of initial parameters (especially cloud center positions) to facilitate the fitting process.
- For not artificially heated qubits, use Bose-Einstein statistics and the known qubit frequency to extract the effective qubit temperature.

The fitting and extraction of the temperature will be handled by a IQ cloud analysis class that will be written and made available at a later stage of the project. Until then, it is recommended to use `qkit IQCloudAnalysis`.

Note: The same dataset can also be interpreted as a quantum jump trace. This enables further analysis such as noise filtering, post-selection, or the study of correlated fluctuations due to two-level systems or quasiparticles.

Once the optimised readout is determined, repeat this measurement with 10^7 shots and save it.

Required data

- Full set of recorded IQ points for each shot.
- Readout pulse parameters: frequency, amplitude, length, integration duration, integration delay, pulse shape and wait time in between readouts.
- 2D histogram or density map of IQ cloud distribution.
- Fit results:
  - Ground and excited state populations,
  - Cloud centers and widths (covariances),
  - Signal-to-noise ratio (SNR),
  - Readout state assignment fidelity,
  - Effective qubit temperature,
  - Indication if an excitation pulse was played
- (Optional) Quantum jump trace for post-selection or time-resolved analysis.

2.3.2.2 Frequency optimization

Measuring IQ clouds as a function of readout frequency to find the optimal condition for maximal signal-to-noise ratio (SNR) and minimal unwanted transitions.

Description

In this step, we optimize the readout frequency to maximize the separation between the ground- and excited state IQ clouds, thereby minimizing the state assignment error. Furthermore, the cloud populations should be constant close to the readout frequency. This optimization is enabled by repeating the IQ-cloud measurement (see previous section) across a sweep of readout frequencies centered around the resonator frequency. At each frequency, the IQ histogram is fitted with a Gaussian mixture model to extract state populations and the signal-to-noise ratio (SNR) as shown in fig. 10.

The IQ cloud fit further allows to split the resonator response in a ground- and excited state contribution. Fitting them individually with a circle fit gives the resonator bandwidth for each state $\kappa_{\mathrm{g,\,e}}$ and resonant frequencies $f_{\mathrm{g,\,e}}$. The dispersive shit χ_{ge} is now defined as the resonance frequency difference $\chi_{\mathrm{ge}}/2\pi = f_{\mathrm{r,e}} - f_{\mathrm{r,g}}$ as shown in fig. 11.

For the AC Stark calibration of the average photon number in the resonator, χ_{ge} and κ are required in dependence of readout power. Accordingly, once section 2.3.7 is reached, the measurements in this section have to repeated for varying readout pulse amplitude.

Measurement procedure

- Define a sweep range centered on the extracted resonator frequency ($\pm 5 - 10$ resonator linewidths).
- For each readout frequency:
  - For qubits with vanishing excited state probability, play a tone at the qubit frequency to equalize state probability. Use a $\pi/2$-pulse once available.
  - Perform an IQ cloud measurement with 10^3–10^4 shots.
  - Fit the resulting histogram with two 2D Gaussians and extract SNR and state populations. Using the assignment to ground and excited state, obtain phase and amplitude for each state.
- Determine the optimal frequency by maximizing:
  - Cloud separation (SNR),
  - Stability and thermal consistency of extracted populations,
  - Absence of cloud merging or fit anomalies.
- Fit the spectrum using a circle fit to obtain bandwidth and resonant frequency for ground and excited state respectively.
- Calculate the dispersive shift χ_{ge} from the resonant frequency difference between excited and ground state.

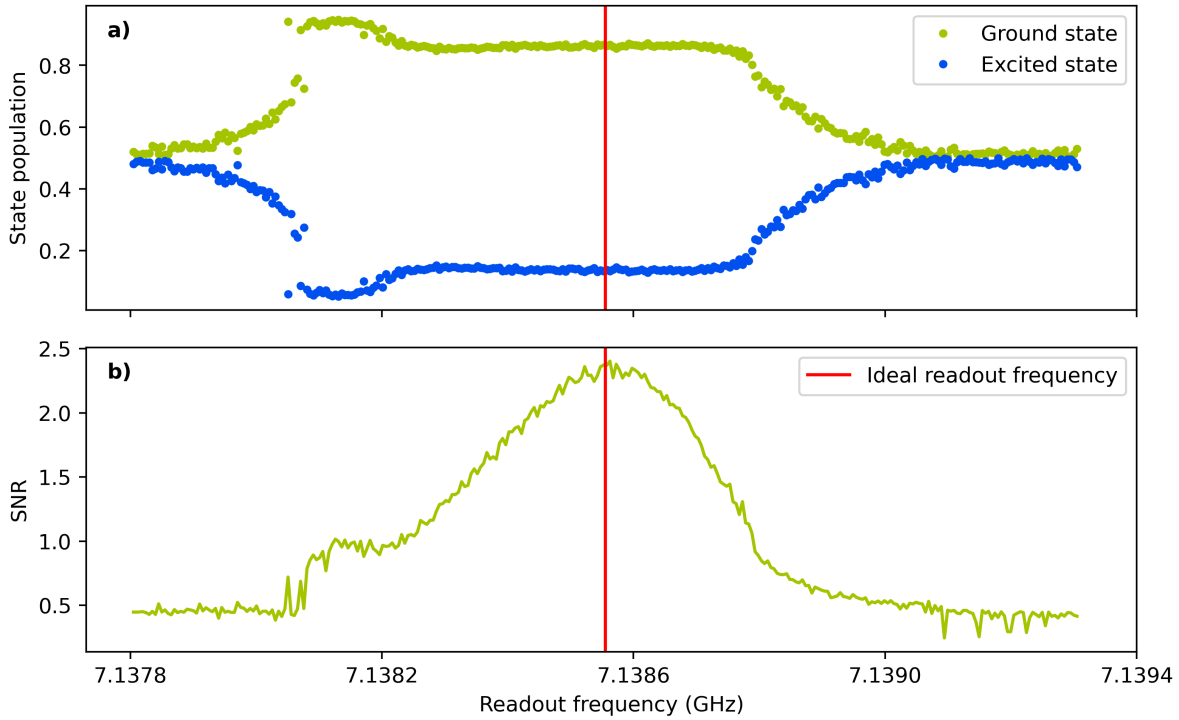


Figure 10: **Readout optimization. a)** Ground and excited state populations versus readout frequency, extracted from individual IQ-cloud fits. At 7.1381 GHz, the state populations behave differently, which can be explained by the not considered f state contribution. **b)** SNR versus readout frequency. The optimal frequency maximizes SNR while maintaining stable populations and is marked by the red line. At the sweep edges, cloud merging leads to poor fits.

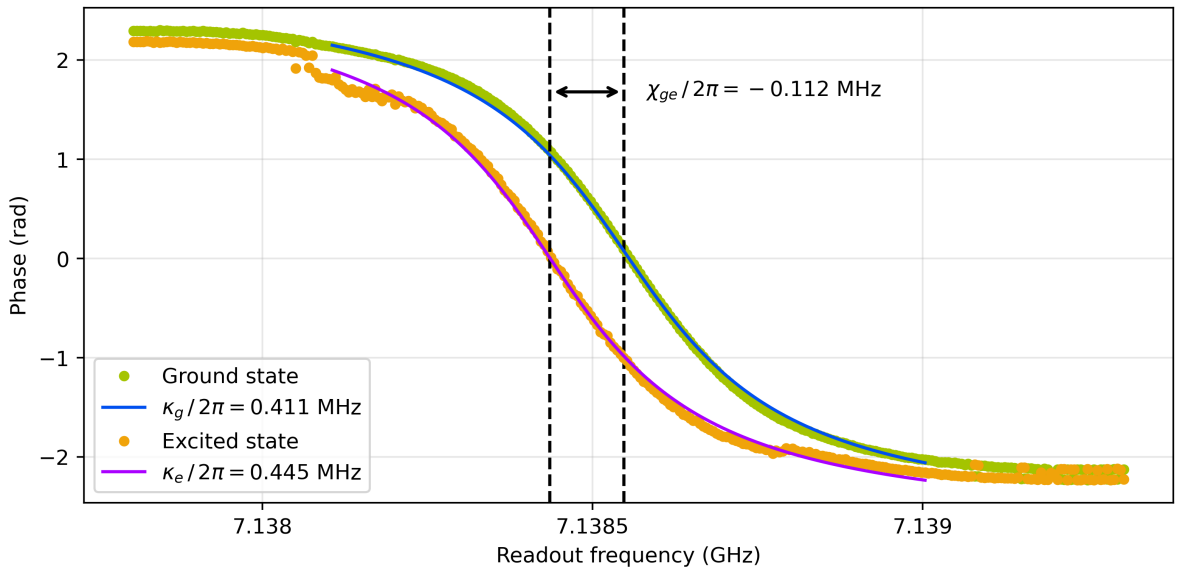


Figure 11: From the frequency optimization measurement extracted phase response of ground and excited state. A circle fit is performed on ground and excited state to extract their linewidth and resonance frequencies (dashed black lines). $\chi_{ge}/2\pi$ is defined as the difference between resonance frequencies. The jumps and deviations in the excited state data are explained by the f-state contribution, increasing uncertainty in χ_{ge}.

Notes on bad fits

- IQ clouds may merge at off-resonant readout frequencies, reducing SNR.
- The fitting routine may incorrectly assign more or fewer states than physically present.
- Cloud centers may fail to track with frequency due to low SNR or distorted distributions.
- Narrow banded parametric amplifiers might distort the readout pulse
- If the resonator frequency is on the slope or peak of a TWPA ripple, the circle fit might have problems capturing the additional phase response. This can be calibrated by moving the resonator away for flux tunable devices with sufficient change in resonator frequency. Alternatively the TWPA can be set up that the resonator frequency is at a position with flat phase response.
- Ideally, once a $\pi/2$- pulse is calibrated, measure the frequency optimization with equal state populations on all visible clouds by playing according pulses. Repeat the measurement in the thermal state and fix cloud positions and variance for the IQ cloud fit from the measurement with equalised populations.

Required data

- Readout frequency sweep (Hz).
- For each readout frequency:
  - IQ histogram data.
  - Fitted cloud positions (2D centers), widths, and weights.
  - Signal-to-noise ratio (SNR).
- Extracted $\chi_{\mathrm{ge}}/2\pi$ (Hz), $\kappa_g/2\pi$ (Hz) and $\kappa_e/2\pi$ (Hz).
- Final chosen readout frequency (Hz).

2.3.2.3 T_1 (measurement preparation)

Evaluate the non-demolishing nature of the readout. Perform relaxation measurements using state preparation by measurement: waiting until the qubit is measured in a desired state.

Description

Some platforms — especially those using FPGA-based control — support real-time feedback or conditional execution based on the outcome of a measurement. This functionality can be used to prepare the qubit in a specific state without applying any pulses to the qubit: the system simply waits until a measurement confirms that the qubit is in the desired state and then proceeds.

This measurement-based initialization enables experiments like T_1 decay from a known state without requiring calibrated gates. On platforms without conditional branching, the same result can be achieved via post-selection in software. This first T_1 measurement allows us to define the error due to decay during readout length τ:

$$\epsilon_{\mathrm{relaxation}} = 1 - e^{-\tau/T_1}.$$

To minimize the overall error, the relaxation error and the state assignment error, defines as

$$\epsilon_{\mathrm{assignment}} = \frac{1}{2}\mathrm{erfc}\left(\frac{\mathrm{SNR}}{\sqrt{2}}\right)$$

should match.

This technique is also a useful tool to evaluate the **QNDness** of the readout — that is, whether the measurement itself influences the quantum state of the qubit. By comparing the T_1 relaxation behavior under different readout conditions, we can determine whether the measurement introduces backaction. An example can be seen in fig. 12.

Here two methods are compared:

1. The qubit relaxation is measured from the excited state as in a typical T_1 measurement.

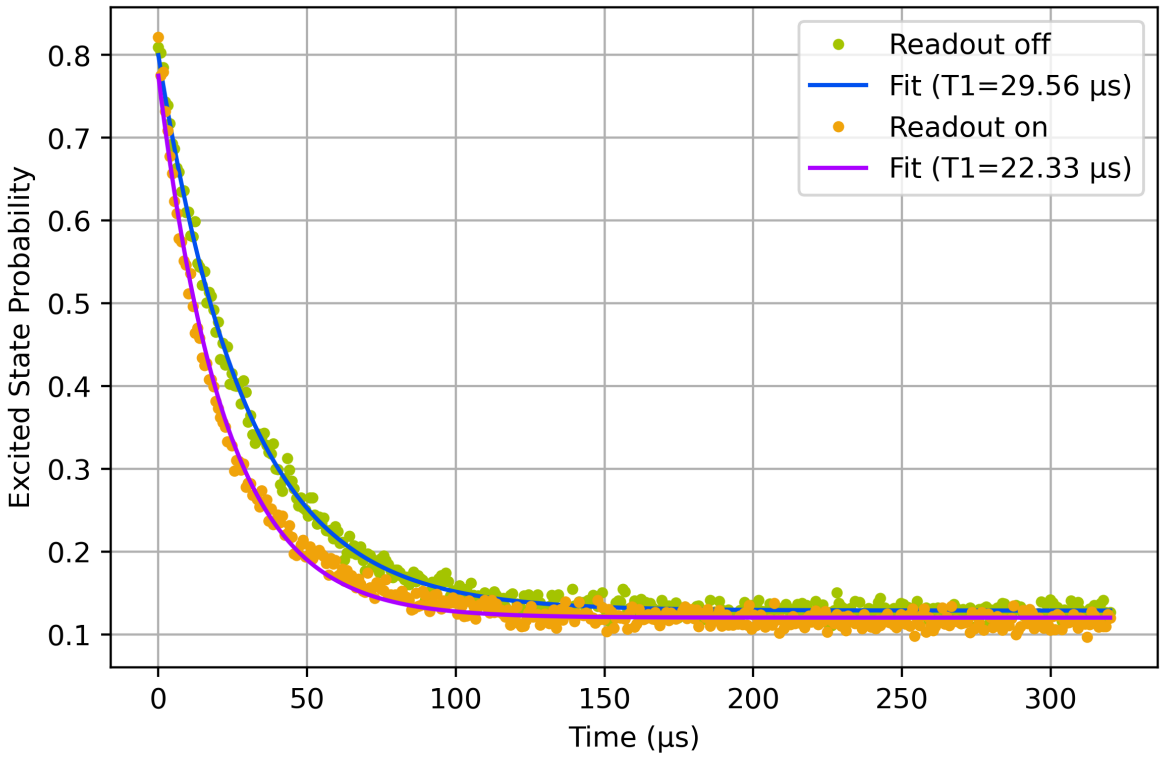


Figure 12: **Measurement preparation.** T_1 measurements performed my measurement preparation. The difference between the curves is, that for the orange one, the readout tone was never turned off. This leads to a degradation of T_1 but also the shape of the decay deviates from the pure exponential shape. In the ideal QND readout case, both decay curves would remain identical.

2. During the T_1 delay time, the readout pulse stays active (cw) and is only read out and turned off once the delay time passed.

The two resulting decays, characterised by timescales $T_{1,off}$ and $T_{1,on}$, should be in good agreement if the measurement is QND. A substantial contribution on how QND the readout can be, is the ring up and ring down time of the resonator. Usually, the readout is more likely QND for short ringup times. This can be realised by pulse shaping the readout pulse. To artificially enhance the influence of the ringup and ring down times, the readout can be turned on and off stroboscopic during the delay times. In addition to T_1, the population of the first point in the decay curve should be compared for the readout on and off case. Ideally, the population for these points is also equal and explainable by the relaxation error. If a good agreement is not found for the specific readout conditions, return to equalize assignment and relaxation error starting from a readout pulse with reduced length.

Measurement procedure
This method relies on state preparation by repeatedly measuring the qubit state until it is found to be excited. If this functionality is not possible to implement, post selection can be used to obtain the same result. In this case, the experiment is performed several times and only the data is kept, where the decay starts from the qubit being in the excited state. This approach is described below. For qubits with vanishing thermal population in the excited state, a drive pulse played at the qubit frequency can be used to equalize populations.

- Use a feedback loop to prepare the qubit in the excited state:
  - Repeatedly measure the qubit state.
  - Once a measurement result indicates the qubit is excited, trigger the T_1 sequence.
- Alternately perform two versions of the T_1 experiment for a few times to compensate for temporal drifts:
  1. **Standard T_1:** The qubit is read out after the delay time passed. The readout is off during this delay time.
  2. **Continuous-readout T_1:** The readout tone remains on throughout the delay time until the final readout.
- Compare the decay curves from both versions:
  - If the readout is QND, both T_1 traces should agree. A faster decay in the continuous-readout case indicates measurement-induced backaction.
  - The excited state population at the first data point of the decay curve should be the same for both cases and explainable by the relaxation error.

  If there is a substantial mismatch between the readout off and readout on case, reduce the readout pulse length and restart the optimization by re-balancing assignment and relaxation error. The goal is to end up with a QND readout that still has sufficient SNR to distinguish the different qubit states well.

Post selection approach
When using post selection, replace the first measurement procedure step from above with the following:

- Perform either step (a) or step (b) depending on qubit type.

a For a high-frequency transmon qubit with small thermal population in the excited state:
  - Wait a time $> 10T_1$ since any previous measurement.
  - Do a measurement (M_0) and save the outcome.
  - Play a pulse to the qubit to excite it.

b For a low frequency fluxonium qubit with significant thermal population in excited state:
  - Do a measurement (M_0) and save the outcome.

c Proceed with the standard or continuous-readout T_1 measurement described above. When analysing the data, keep only the measurements that correspond to $M_0 = |0\rangle$.

Required data

- Time traces from both T_1 experiments:
  - Standard T_1 (no readout during wait),
  - Continuous-readout T_1.
- Extracted T_1 values and fit error for both cases.
- Excited state population at the beginning of the decay curve for both cases and expected excited state population due to relaxation error.
- Readout parameters used:
  - amplitude,
  - frequency,
  - integration time,
  - integration delay.
- Feedback logic used or post-selection criteria.
- Indicator whether feedback was hardware-based or software-based.

2.3.3 Two-tone (time-domain)

Measurement of the qubit transition spectrum by applying a drive tone while monitoring the time-domain resonator response. This technique is similar in concept to section 2.2.4, but implemented with pulsed control and IQ-based detection.

Description

In this time-domain version of two-tone spectroscopy, we apply a microwave drive tone (the "pump") to the qubit while performing a time-domain readout of the resonator. When the pump frequency matches a qubit transition, the excited state population increases and shifts the resonator response due to the dispersive interaction.

Compared to the frequency-domain variant, this approach provides:

- quantitative access to excited state populations via IQ cloud analysis,
- more precise control of pulse timing and amplitude,
- improved SNR through integration-based readout,
- compatibility with state assignment and population fitting.

In general, this measurement should be used to finetune the flux value at the sweet spot positions for flux tunable qubits and to extract the transition frequency at the sweet spot with more precision. This defines the operating point for subsequent timedomain measurements.

A typical example of such a measurement is shown in fig. 13, where the extracted excited state population is plotted as a function of pump frequency and flux.

Measurement procedure

Start from the parameters extracted by the frequency domain two tone measurement in section 2.2.4.

- Apply a strong pump tone at a fixed flux point, long enough to saturate the qubit (approx. 50/50 population).
- Immediately apply a standard readout pulse and measure the IQ response.
- If no qubit transition is found, increase the pulse power and duration. If the response is broad, reduce pulse power and duration.
- Repeat the process for a sweep of pump frequencies around the investigated transition.
- For each point, extract the excited state population from the IQ histogram.

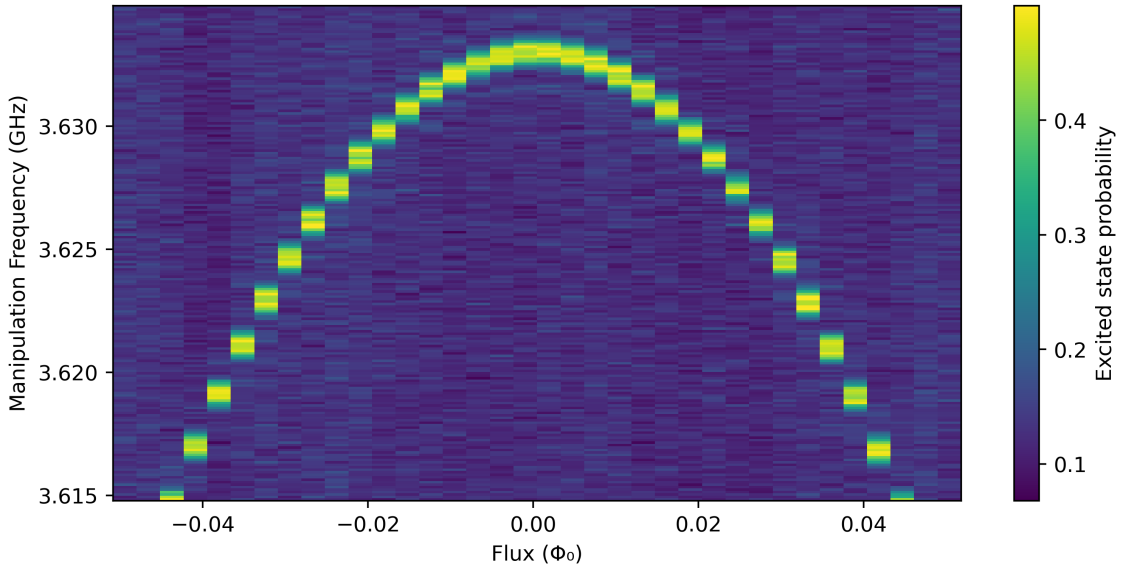


Figure 13: Time-domain two-tone spectroscopy of a flux tunable transmon near its zero-flux sweet spot. The $|0\rangle \rightarrow |1\rangle$ transition at zero-flux is at 3.633 GHz. The color map shows the extracted excited state population from IQ cloud fits.

- Repeat for several pump tone amplitudes or lengths and fix the amplitude and length of the pump tone to give a sharp response.
- (for flux tunable qubits) Repeat for several flux points around the sweet spot of the qubit.
- Update the current-to-flux calibration using the sweet spot positions.
- Update the qubit transition frequencies at the sweet spot.

Required data

- 2D dataset: excited state population (from IQ cloud fit) of the $|0\rangle \rightarrow |1\rangle$ and one other transition as a function of pump frequency and
  - pump power,
  - (for flux tunable qubits) magnetic flux around the sweetspot.
- Identified $|0\rangle \rightarrow |1\rangle$ and other transition frequency at the flux sweet spot.

2.3.4 Power Rabi

Driving the qubit at its transition frequency with a pulse of fixed duration but varying amplitude reveals coherent oscillations in the excited state population. This provides initial calibration estimates for the π and $\pi/2$-pulse amplitudes.

Description

In the Power Rabi experiment, the shape and duration of the qubit drive pulse are held constant, while its amplitude is swept. As the drive strength increases, the qubit undergoes Rabi oscillations between $|0\rangle$ and $|1\rangle$, resulting in a sinusoidal variation of the measured excited state population.

The first maximum of the oscillation corresponds to a π-pulse, ideally inverting the population. The amplitude at half this value corresponds to a $\pi/2$-pulse that prepares an equal superposition on the equator of the Bloch sphere. These values serve as starting points for gate calibration in later sections.

The drive frequency must be aligned with the $|0\rangle \rightarrow |1\rangle$ transition to observe coherent oscillations. The pulse duration should be long enough to resolve at least one full oscillation period. An example is shown in fig. 14.

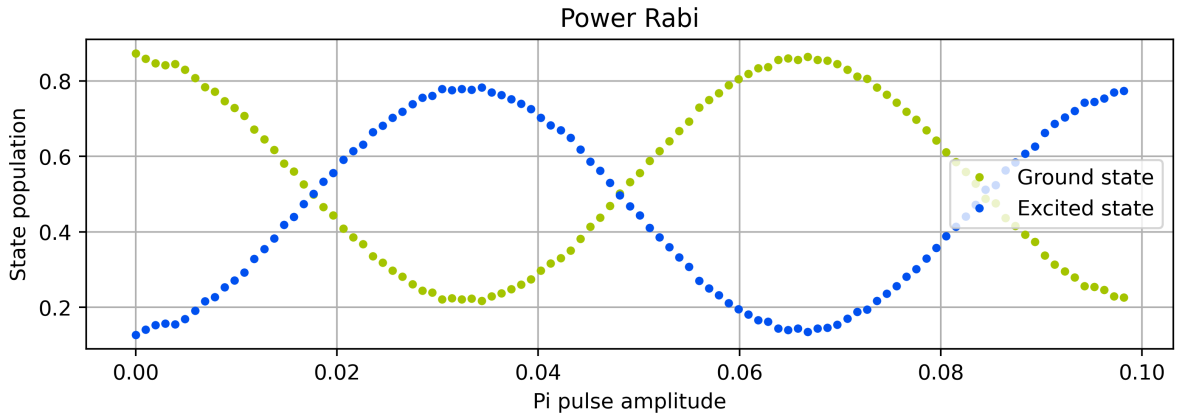


Figure 14: Power Rabi: Fitted excited state population as a function of increasing qubit drive amplitude. The population oscillates due to coherent Rabi dynamics. A π-pulse corresponds to the amplitude that maximally inverts the qubit (here at amplitude 0.033), while a $\pi/2$-pulse corresponds to half that amplitude (here: 0.017).

Measurement procedure

- Fix the duration and shape of the qubit drive pulse (e.g., tens to hundreds of ns, flat-top or Gaussian envelope).
- Sweep the amplitude of the pulse over a range wide enough to observe several Rabi cycles.
- After each pulse, apply the standard readout and extract the qubit state using IQ demodulation and state assignment.
- Repeat each point multiple times to obtain good statistics on the excited state population.
- Plot the population as a function of drive amplitude.
- Fit the oscillation to a sine curve and extract:
  - the amplitude at the first maximum $\rightarrow$ π-pulse amplitude,
  - the midpoint amplitude $\rightarrow$ $\pi/2$-pulse amplitude.

Required data

- Qubit drive amplitudes used in the sweep (arb. units).
- Pulse shape and duration used for the Rabi drive.
- Extracted excited state population for each amplitude.
- Fitted sine curve parameters (frequency, amplitude, phase).
- Identified π and $\pi/2$-pulse amplitudes.

2.3.5 π-fringes

Measurement for fine-tuning the amplitude of the π-pulse.

Description

Starting from the π-pulse amplitude estimated via the Power Rabi experiment, we refine the pulse by measuring how accurately it inverts the qubit state over repeated applications. In this experiment, we start from a thermal state, and a sequence of n back-to-back π-pulses is applied. After each pulse, the qubit state is measured.

Ideally, each π-pulse swaps the population between ground and excited states. If the amplitude is calibrated correctly, this results in a clean oscillation that slowly decays due to energy relaxation (T_1).

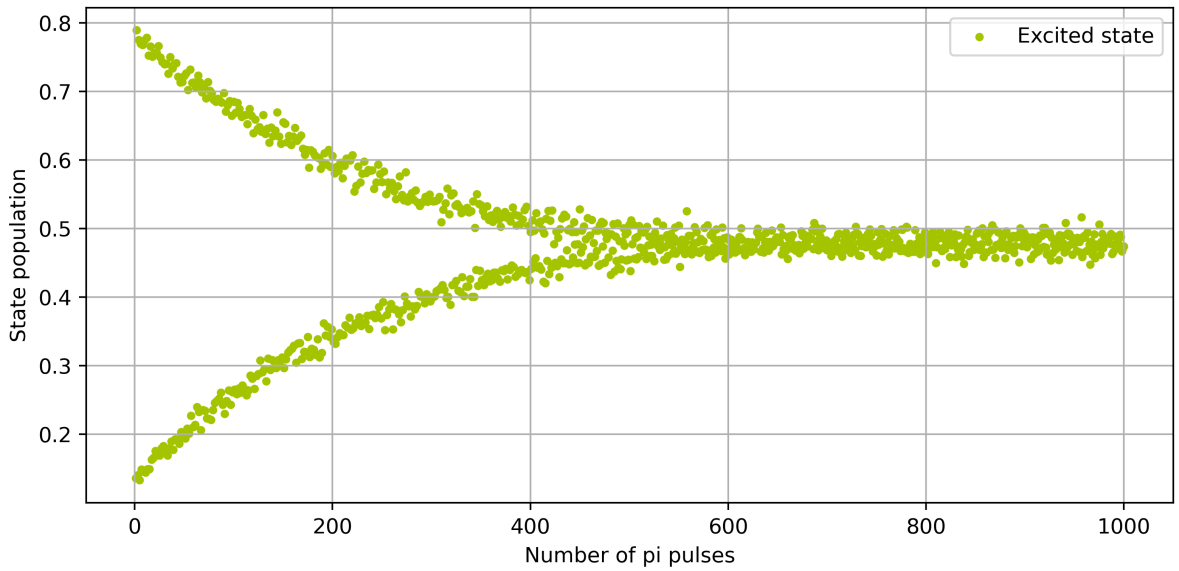


Figure 15: State population of the excited state versus the number of sequential π-pulses. Ideally, the qubit swaps between states with each pulse, decaying slowly towards 50%, with a small offset due to population changes during readout. A bad calibrated π-pulse introduces a beating pattern, faster decay and reduced fringe visibility. The optimal amplitude maximizes the number of visible swaps.

However, if the pulse is slightly under- or over-rotating, a beating pattern emerges due to cumulative phase error on the Bloch sphere. This leads to a faster decay of the oscillation, typically toward 50% occupation.

To find the optimal amplitude, a binary search is performed to maximize the number of visible swaps before the oscillation decays. An example result is shown in fig. 15.

The π-train approach resembles the π-fringe approach, where a measurement is only performed after an even number of π-pulses.

Measurement procedure

- Initialize the qubit in the ground state.
- Apply a sequence of n π-pulses, each followed by a measurement.
- For each n, record and average the IQ values to extract the excited state population.
- Plot the population versus number of π-pulses.
- If the amplitude is miscalibrated, the oscillation will decay faster than expected from T_1.
- Vary the pulse amplitude using a binary search to find the setting that yields:
  - The slowest decay,
  - The largest number of clean state swaps before saturation at 50%.
- Quantify the decay rate by fitting the population of every second point (e.g., all even-numbered pulses) to extract an exponential decay constant.

Required data

- Number of sequential π-pulses applied (n).
- Extracted state populations after each n.
- Final calibrated π-pulse amplitude (arb. units) and duration τ_π.
- Fitted decay constant from the oscillation envelope (quantifying swap visibility lifetime).
- Plot of excited state population versus number of π-pulses.

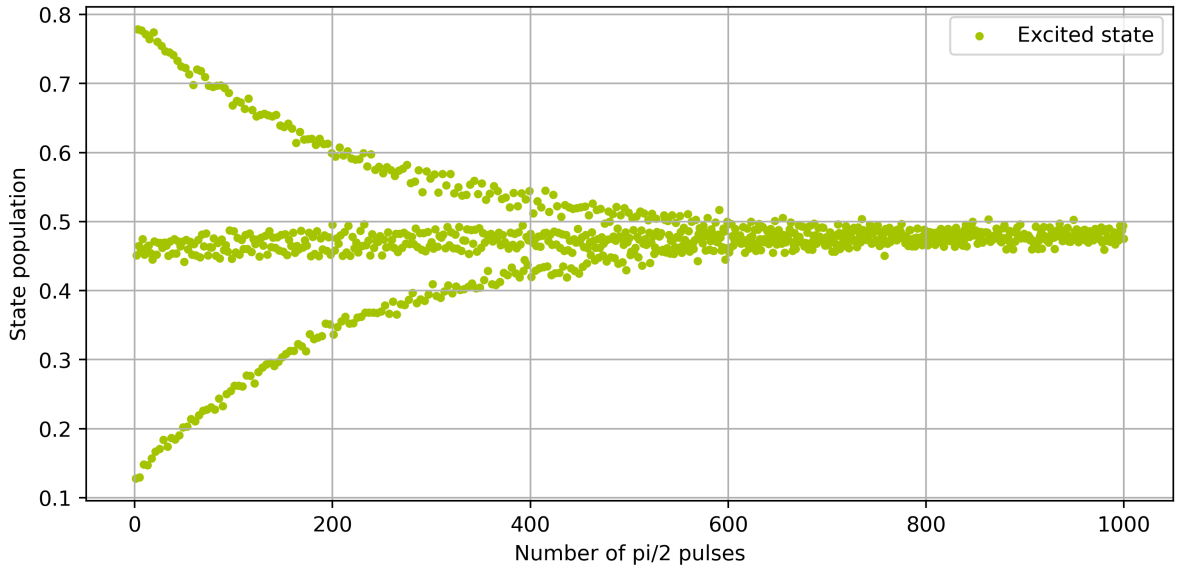


Figure 16: State population of the excited state versus the number of played $\pi/2$-pulses. Ideally, the population oscillates between excited, mixed, and ground states. The $\pi/2$ amplitude is optimised by minimizing the drift in even-numbered points or maximizing oscillation coherence.

2.3.6 $\pi/2$-fringes

Measurement for fine-tuning the amplitude of the $\pi/2$-pulse.

Description

This calibration step refines the $\pi/2$-pulse amplitude, building on the result of the Power Rabi experiment. Unlike a full inversion, a $\pi/2$-pulse rotates the qubit state onto the equator of the Bloch sphere, ideally producing a superposition with equal populations in $|g\rangle$ and $|e\rangle$.

The experiment is conceptually similar to the π-fringe calibration (section 2.3.5): we apply a sequence of nominal $\pi/2$-pulses, each followed by measurement, and observe oscillations in the state population. A well-calibrated $\pi/2$-pulse will produce a symmetric population pattern alternating between ground, mixed, and excited states.

As with other early-time calibration steps, the qubit is not actively reset. Each sequence begins from thermal equilibrium, resulting in a baseline population offset. This shifts the oscillation center away from 50%, but the periodicity and symmetry of the signal remain diagnostic of pulse fidelity.

An example result is shown in fig. 16.

Measurement procedure

- Let the qubit thermalize to its steady state (no initialization pulse).
- Apply a sequence of n nominal $\pi/2$-pulses, each followed by measurement.
- For each n, average the measured IQ data and extract the excited state population.
- Plot the excited state population as a function of the number of $\pi/2$-pulses.
- Identify whether the oscillation:
  - drifts in amplitude or symmetry (sign of over/under rotation),
  - stabilizes around a shifted average due to thermal bias.
- Tune the $\pi/2$ amplitude to:
  - minimize even-numbered point deviations (from a flat mixed state),
  - or maximize the overall fringe coherence and symmetry.

Required data

- Number of sequential $\pi/2$-pulses applied (n).
- Extracted state population after each n.
- Final calibrated $\pi/2$-pulse amplitude (arb. units) and duration $\tau_{\pi/2}$.
- Oscillation symmetry or decay behavior (qualitative or fitted).
- Plot of excited state population versus number of $\pi/2$-pulses.

2.3.7 AC-Stark calibration

Calibrating the number of photons in the readout resonator by measuring the qubit transition frequency shift and dephasing rate change as a function of ac power applied at the resonator frequency.

Description

When a microwave tone populates the resonator, the qubit transition frequency and dephasing rate change due to the AC-Stark effect. This shift is proportional to the average steady state photon number n_{ph} in the resonator and can therefore be used to calibrate the amplitude-to-photon conversion. In the low ac power regime, the following relations link the change in qubit frequency $\Delta\omega_q$ and dephasing $\Delta\Gamma_\phi$ to the average photon number n_{ph} in the resonator [15, 16]:

$$\Delta\omega_q = \chi_{\mathrm{ge}} n_{\mathrm{ph}}, \qquad \Delta\Gamma_\phi = \frac{2\chi_{\mathrm{ge}}^2}{\kappa} n_{\mathrm{ph}}, \qquad n_{\mathrm{ph}} = c\bar{V}_d^2.$$

The average photon number is proportional to the square of the voltage amplitude V_d at the output of the ac pulse generator and c, a proportionality constant relating n_{ph} and $\bar{V}_d^2$. Dependent on the measurement procedure used, only the frequency shift $\Delta\omega_q = 2\pi(f_q(n_{\mathrm{ph}}) - f_q(0))$ (two-tone based) or also the dephasing change $\Delta\Gamma_\phi = \Gamma_2^*(n_{\mathrm{ph}}) - \Gamma_2^*(0)$ (Ramsey based) can be extracted. From fitting a linear slope ($y = ax$) we obtain the slopes S as shown in fig. 17.

It is required to determine χ_{ge} and κ as explained in section 2.3.2.2 for the same low power regime that matches the ac amplitudes. For the Ramsey-based method, we can compare χ_{ge} extracted from the frequency optimization measurement with

$$\chi_{\mathrm{ge}} = \frac{\kappa S_{\Gamma_\phi}}{2S_{f_q}}.$$

We previously obtained a dispersive shift from the IQ cloud frequency sweep of $\chi_{ge}/2\pi = 112\,\mathrm{kHz}$ (fig. 10). With the fitted slopes, we obtain $\chi_{ge,\mathrm{ac}}/2\pi = 89\,\mathrm{kHz}$ using the average κ of the ground- and excited state. Throughout the rest of this document we will use n_{ph} as the reference quantity to report together with other measurements. Together with τ and κ, n_m can then be evaluated where needed.

The proportionality constant can now be determined using χ_{ge} (here we stick to χ_{ge} extracted from the slopes) and yields $c = 4596\,n_{\mathrm{ph}}/\bar{V}_d^2$. For the calibrated readout used for all previous example measurements shown with fixed readout amplitude, this results in an average photon number of $n_{\mathrm{ph}} = 64.8$. The average measurement photon number in a readout pulse used in the IQ cloud plot in fig. 9 is defined by $\sqrt{n_m} = \sqrt{n_{\mathrm{ph}}\tau\kappa/4}$ where τ is the readout pulse length. This assumes the readout pulse has the same amplitude as the "CW" pulse used during the AC Stark calibration. It defines the amount of photons that were used for the readout and is in the example here $n_m = 173.9$.
We can also calculate the total attenuation of the input line by comparing the output power at the signal generator (measured by a spectrum analyzer or power meter) with the input power at the resonator $P_{\mathrm{in}} = n_{\mathrm{ph}}\hbar\kappa\omega/4$ to refine our knowledge of the input attenuation.

This calibration is essential for consistent interpretation of readout performance and qubit-resonator interaction. Once the relation between readout amplitude and $\bar{n}$ is known, all power-dependent measurements, such as IQ clouds, can be re-expressed in terms of photon number, enabling comparison across devices and setups.

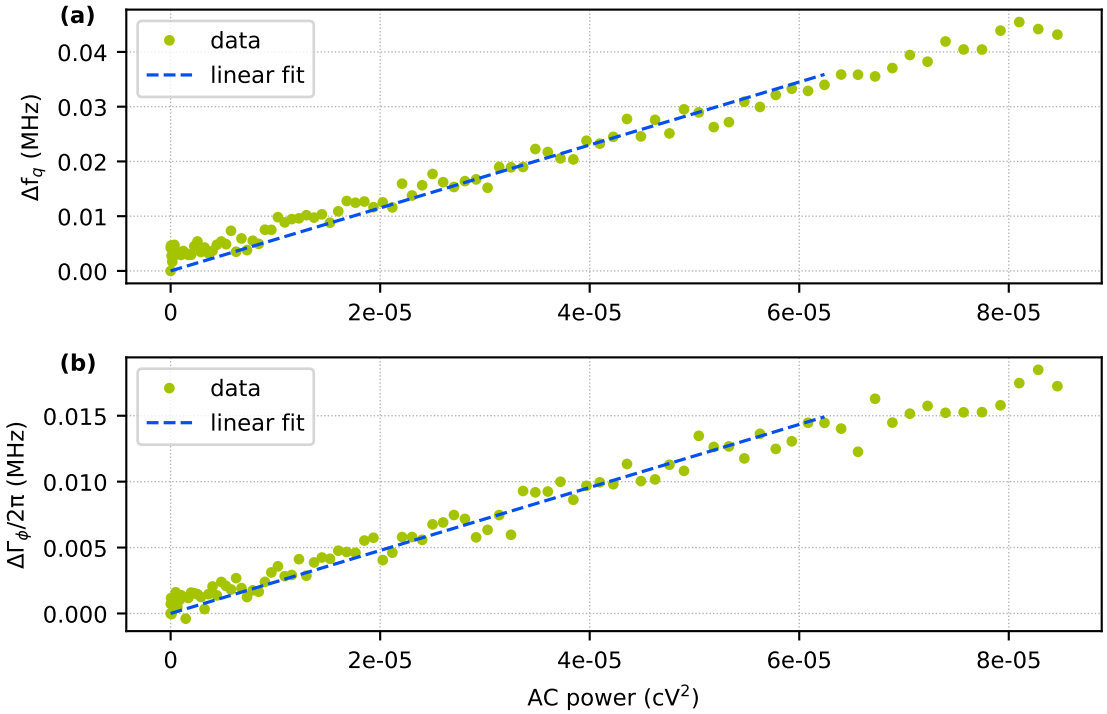


Figure 17: **(a)** AC-Stark shift of the qubit transition frequency f_q extracted from Ramsey measurements as a function of the squared AWG control amplitude V_{AC}^2 ("cV_d^2"). Circles show the measured data, and the dashed line is a linear fit constrained to pass through the origin. From the fit we obtain a Stark-shift slope of $S_{\Delta f_q} = 575.1\ \mathrm{MHz}/(\mathrm{cV}^2)$. **(b)** Increase of the qubit dephasing rate $\Delta\Gamma/2\pi$ obtained under the same drive conditions. The corresponding linear fit yields a dephasing slope of $S_{\Delta\Gamma\phi} = 239.0\ \mathrm{MHz}/(\mathrm{cV}_{\mathrm{d}}^2)$. Only the low-amplitude regime ($V_{\mathrm{AC}}^2 \lesssim 8\times10^{-5}\ \mathrm{cV}_d^2$) is shown, where both frequency shift and dephasing scale linearly with V_{AC}^2. The y-axes are chosen to be in Hz for this plot.

Measurement procedure

Preferred: Ramsey-based AC-Stark
This measurement procedure is similar to the Ramsey sequence presented in section 3.7 just with an additional ac pulse played between π-half pulses.

- Apply a $\pi/2$-pulse to create a superposition state.
- Starting with the Ramsey delay time τ_R, apply a CW tone (the "AC tone") at the readout frequency until $3/\kappa$ before the delay ends such as not to affect the readout by having excess photons in the resonator.
- Apply a second $\pi/2$-pulse with virtual detuning $\Delta\phi = 2\pi f_{\text{detuning}}$ and perform readout. The detuning frequency is chosen to be $f_{\text{detuning}} \approx 3/T_2^*$.
- Wait for $\tau_r = 10\,T_1$ so that the qubit and its environment can relax back to the thermal state.
- Repeat for 200 delay times τ_R, linearly spaced from 0 to $5\,T_1$.
- Measure for 500 acquisitions.
- Repeat the whole experiment with increasing amplitudes of the AC tone, also including 0 (no ac tone played) as a reference. Cover a range of AC amplitudes from the low power regime up to the point where the dephasing rate and frequency shift stop changing linearly when look at as a function of V_d^2.

Fit the IQ clouds for each delay time and ac amplitude to extract the state populations as shown in section 2.3.2.1. Fit the excited state population as a function of delay time for the different ac tone amplitudes using

$$P_e(\tau_R) = A\cos(\omega\tau_R + \phi)e^{-\tau_R/T_2^*} + B$$

where B should be equal to 0.5 minus the relaxation $\epsilon_{\text{relaxation}}$. If a beating pattern is observed (two frequencies) adapt the fitting function to

$$P_e(\tau_R) = A\cos(\omega_0\tau_R + \phi_0)\cos(\Delta\omega\tau_R + \phi_1)e^{-\tau_R/T_2^*} + B \tag{1}$$

with $\omega_1 = \omega_0 + \Delta\omega$ and $\omega_2 = \omega_0 - \Delta\omega$.

Alternative: Two-tone AC-Stark
Perform a two-tone experiment where qubit pump frequency and ac tone amplitude are swept.

- Apply a CW ac tone at the readout frequency with variable amplitude.
- Wait for ring-up of the resonator (e.g. $10/\kappa$).
- Apply a qubit excitation pulse (e.g., π or $\pi/2$)
- Turn off the ac tone and wait for the ring down (e.g. $10/\kappa$).
- Perform a standard readout of the resonator.
- Repeat for several hundred averages.
- Repeat the whole experiment with increasing amplitudes of the AC tone, also including 0 (no ac tone played) as a reference. Cover a range of AC amplitudes from the low power regime up to the point where the dephasing rate and frequency shift stop changing linearly when look at as a function of V_d^2.

Extract the qubit frequencies and continue with the above analysis above.

Additional notes

- This calibration enables reanalysis of previous readout frequency/amplitude sweeps in terms of photon number.
- It also provides a common axis for future readout benchmarking and hardware comparison.
- The square root of measurement photons, $\sqrt{\bar{n}_m}$, can be used as a physically meaningful axis unit in IQ cloud plots.

Required data

- For Ramsey-based approach: 2D array of excited state populations as a function of ac pulse amplitude and delay time.
- Array of applied ac tone amplitudes (or powers), qubit frequencies, and if measured dephasing rates.
- χ_{ge} and κ in the low power regime.
- Fitted slopes and, if applicable, χ_{ge} from the slopes.
- Power calibration c and photon number of a default readout.
- Input attenuation of the fridge.

3 Intercomparison quantities of interest (QoI) measurement procedure

This part describes the set of measurements that will be used to extract qubit performance metrics for comparison between different labs, setups and samples. It is assumed that the qubit has been tuned up and optimised according to the protocol in Section 2.

In this document we briefly outline the key aspects relating to the analysis of the data required to produce each QoI. A more detailed description and open source code will be made available at a later stage, with the intention of providing a common analysis framework for qubit comparison. For this purpose we also outline here all the required auxiliary data and metadata that needs to be captured for each QoI, as required by the data analysis.

3.1 Overall procedure

We start by giving an overview to the measurement protocol for qubit quantities of interest (QoI's). The different measurements outlined in the following sections are as follows:

- Readout performance: QNDness,
- Readout performance: Readout-Induced Leakage benchmarking (RILB),
- *Gate performance: Randomised benchmarking (RB),
- *Gate performance: Quantum state preparation,
- Qubit coherence: Comparison of passive and active qubit energy relaxation measurement (T_1),
- *Qubit coherence: Interleaved coherences over time (T_1, T_2^{echo}, T_2^* (Ramsey) and qubit frequency),
- Qubit coherence & bath: Szilard engine.

The choice of performance metrics aims to give an overarching picture of not only intrinsic qubit coherence but also how well readout and control performs and affect the measurements. The quantities marked with * are repeated over time to capture temporal fluctuations in the measured quantity. It is recommended that the QoI are measured in the order presented here, as some methods require the input from others.

Within the MetSuperQ project a generalised data file format is under development through which all the data generated as part of these QoI's is to be stored. Different laboratories will then need to create their own code to convert data taken by their different hardware setups to this format. This standardised data format will then be used by standardised analysis code that will be made available at a later stage of the project. The aim with this code is that all data analysis can be conducted by anyone independently and consistently to facilitate qubit benchmarking and comparison. As part of the QoI descriptions we give a summary of the outputs from this analysis to clarify the quantities that will be available for comparison.

In the following sections we describe the measurement procedure and analysis for each of these methods. We structure all the measurements using a short description, followed by a detailed description of the measurement procedure, the pulse sequences used, any assumptions or limitations of the method, an overview of the data analysis procedure, and the relevant output quantities.

3.2 QNDness

The quantum non-demolising nature (or "QNDness", denoted $\mathcal{Q}$) of a measurement is a measure of how invasive a qubit readout operation is on the quantum state of the qubit. The ideal case is that the measurement has no effect on the qubit state, i.e. $\mathcal{Q} = 100\%$.

Description

If a quantum measurement is truly non-demolishing this means that the act of measurement of the qubit state does not change the qubit state. In reality there will always be a degree of measurement induced back-action. We start here with he most basic measurement of the QNDness that has been frequently adapted in the past. It should be noted that his method gives an overall readout infidelity which also include factors that are not directly related to the demolishing nature of the measurement, such as T_1 decay during the procedure.

To quantify the QNDness an experiment that make consecutive measurements following the preparation of some state can be used. The method we will use here uses two measurement pulses M_1 and M_2 and we evaluate the the repeatability fidelity as the probability that both measurements give the same outcome.

$$F_r = \frac{P_{gg} + P_{ee}}{2}, \tag{2}$$

where P_{xy} is the probability that measurement 1 and measurement 2 yielded outcome x and y respectively.

The QNDness may depend on:

- Qubit / circuit design parameters,
- The readout frequency, power, duration and readout pulse shape,
- The measurement hardware readout chain (e.g. TWPA settings),
- Assignment & readout fidelity,
- T_1 decay during the QNDness measurement sequence.

The QNDness is reported as a single number in percent (%) where 0% means the measurement is totally destructive and 100% totally invasive. The reported number will have an associated error range, as defined in the data analysis procedure below.

$$\text{Example}: \quad \mathcal{Q} = 95.1 \pm 2.3\%. \tag{3}$$

The number reported should be for the readout conditions used in subsequent QoI measurements. In this protocol we are therefore interested in the resulting QNDness for the optimised readout (as derived in the qubit tuneup) that is subsequently used to measure other QoIs.

Measurement procedure

To measure the QNDness ($\mathcal{Q}$) of the readout operation the qubit is prepared in an initial state (either $|g\rangle$, $|e\rangle$ or $\frac{1}{2}(|g\rangle + |e\rangle)$, interleaved). After a short delay τ_0 a single-shot measurement $M_{1,k}$ (of duration τ) is taken followed by a delay $\tau_d > 1/\kappa$ to allow the intra-cavity field to decay, before another identical measurement $M_{2,k}$ is done. If a JPA is used with a smaller bandwidth than the readout cavity, then the response time of the JPA is instead the relevant parameter determining τ_d. This is followed by a delay $\tau_\mathrm{r} > 10T_1$ to let the qubit relax to its thermal state. This is repeated for all three prepared states $p = |g\rangle$, $|e\rangle$ and $\frac{1}{2}(|g\rangle + |e\rangle)$. Finally, the pulse sequence is repeated $N > 10^6$ times.

The resulting dataset is an array of paired-measurements $M_{1,p,k}$ and $M_{2,p,k}$ (each as an IQ-pair) with $k = 1...N$.

Pulse sequence

Figure 18 shows the pulse sequence to implement.

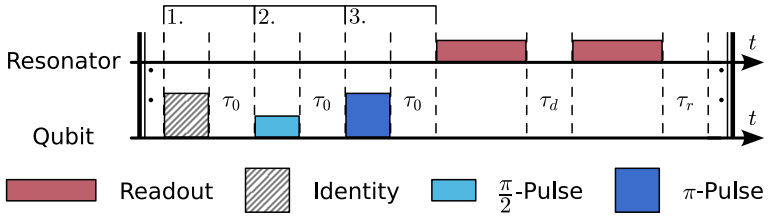


Figure 18: Pulse sequence for QNDness measurement. Numbered brackets show which pulse to use in the respective repetition, i.e. 1st bracket is used in the first iteration of the pulse sequence, 2nd in the second iteration and so on: first the sequence is done with identity operation, the second iteration uses a $\pi/2$-pulse and finally a π-pulse. The prepared state is followed by a short delay τ_0, a measurement, a short delay τ_d and another measurement. Finally wait a reset time τ_r before repeating the whole sequence of three iterations.

Assumptions and limitations

Practical limitations:

1) To obtain the true non-demolishing nature of the readout the readout pulse duration should be significantly smaller than the T_1 relaxation of the qubit, to avoid this mechanism to have a dominant contribution to the measurement. This can be hard to achieve for short-lived qubits.

2) A prerequisite is that the readout state separation fidelity is high, as generally measurements that fall into the overlapping region between g and e will otherwise contribute to a reduced QNDness.

Data analysis procedure

To evaluate the QNDness we start from the measured array of paired-measurements $M_{1,p,k}$ and $M_{2,p,k}$. From this dataset apply the standard state assignment algorithm to classify each measurement result as either $|g\rangle$ or $|e\rangle$. We then compute the conditional probabilities $P_{xy} = P(M_2 = y|M_1 = x)$, i.e. that measurement 1 and measurement 2 both yielded outcome $x, y = g$ or e. From this we calculate

$$\mathcal{Q} = \frac{P_{gg} + P_{ee}}{2}, \tag{4}$$

which is taken across all the different state preparations (p).

This definition of QNDness includes various other sources of error such as T_1-decay during the pulse sequence, state preparation errors, assignment errors and measurement induced state transitions (MIST) beyond the computational subspace. Therefore, it can overestimate the QNDness. We can eliminate some of these errors by discarding measurements where the state assignment is not certain (i.e. in the overlapping region of the assignment histograms).

Estimate the contribution from T_1 relaxation: To estimate the amount of infidelity that comes from T_1 relaxation during the measurement. We take the probability that the qubit relaxes during the wait between measurements and the duration of the second measurement.

$P_r = 1 - \exp(-\tau_{\text{tot}}/T_1)$, where $\tau_{\text{tot}} = \tau_d + \tau$. When propagating the error in T_1 this gives an error in P_r. This number should be reported together with Q, and is expected to be a small fraction of $1-Q$.

Required metadata for analysis:

- Readout pulse duration τ (in μs)
- Readout pulse shape
- Readout pulse rise/fall time (in μs)
- Delay τ_0 between control and first measurement pulse
- Delay between measurement pulses (τ_d) (in μs)
- Resonator Q's and frequency f_r (in GHz)
- Average qubit T_1 time (single number, μs) and error in determination. This should be the spread of the T_1-histogram and not only the error from a single fit to a T_1 measurement. This is the error on T_1 to use for the bounds on the relaxation contribution to the QNDness. This data is obtained in the interleaved coherences QoI in section 3.7.

- Readout state assignment fidelity (from qubit tuneup) (%)
- Readout pulse photon number n_m (from qubit tuneup)

Sanity checks on metadata to be performed by analysis routine:

- Check that $\tau_d > 1/\kappa$
- Check that $\tau_{\mathrm{r}} > 10T_1$

Quantity to report for comparison

- QNDness $\mathcal{Q}$ with error bar, with at least 2 significant digits beyond the last 9.
- The calculated contribution from relaxation, P_r, including an error bar.
- readout pulse photon number n_{ph} and length τ.
- Resonator decay time κ and qubit relaxation time T_1.

QoI references

[17] Gaurav Bothara, Srijita Das, Kishor V. Salunkhe, Madhavi Chand, Jay Deshmukh, Meghan P. Patankar, and R. Vijay. High-fidelity qnd readout and measurement back-action in a tantalum-based high-coherence fluxonium qubit. *APL Quantum*, 2(2):026103, April 2025. ISSN 2835-0103. doi: 10.1063/5.0255892

[18] R. Dassonneville, T. Ramos, V. Milchakov, L. Planat, É. Dumur, F. Foroughi, J. Puertas, S. Leger, K. Bharadwaj, J. Delaforce, C. Naud, W. Hasch-Guichard, J.J. García-Ripoll, N. Roch, and O. Buisson. Fast high-fidelity quantum nondemolition qubit readout via a nonperturbative cross-kerr coupling. *Physical Review X*, 10(1):011045, February 2020. ISSN 2160-3308. doi: 10.1103/PhysRevX.10.011045

[19] Bryan T. Gard, Zachary Parrott, Kurt Jacobs, José Aumentado, and Raymond W. Simmonds. Fast high-fidelity quantum nondemolition readout of a superconducting qubit with tunable transverse couplings. *Physical Review Applied*, 21(2):024008, February 2024. ISSN 2331-7019. doi: 10.1103/PhysRevApplied.21.024008

[20] Y. Sunada, S. Kono, J. Ilves, S. Tamate, T. Sugiyama, Y. Tabuchi, and Y. Nakamura. Fast readout and reset of a superconducting qubit coupled to a resonator with an intrinsic purcell filter. *Physical Review Applied*, 17(4):044016, April 2022. ISSN 2331-7019. doi: 10.1103/PhysRevApplied.17.044016

3.3 Readout-Induced Leakage Benchmarking (RILB)

Quantifies the probability of readout-induced leakage out of the qubit's computational subspace, based on a randomised gate sequence followed by repeated measurements. With a measurement of the QNDness (above) this allows to separate out additional sources of infidelity from the non-demolishing nature of the readout.

Description

Readout-Induced Leakage Benchmarking (RILB) quantifies how likely the readout process is to cause leakage out of the qubit computational subspace ($|g\rangle$, $|e\rangle$) into higher excited or non-computational states such as $|f\rangle$ or $|l_i\rangle$. Leakage is especially problematic in repeated-measurement protocols like quantum error correction, where leaked states can persist and propagate errors.

The core idea is to probe the integrity of repeated readout and control cycles. A sequence of randomly chosen identity (I) and π (X)-pulses is applied, with each gate immediately followed by a readout. In the absence of leakage, the qubit population should alternate predictably with each π gate. If the qubit leaks, subsequent π gates become ineffective, and the output becomes either static or random, breaking this alternation. This method is based on Hazra et al. [21].

To assess the readout-induced backaction more quantitatively, this procedure should be repeated for a range of **readout amplitudes**, expressed in terms of **calibrated photon number** n_{ph} in the resonator (see section 2.3.7) and readout pulse lengths determined to maintain the same SNR. This reveals how leakage probability scales with measurement strength and helps identify a compromise between readout fidelity and invasive-ness. An example, see figure 4g in Ref. [21].

Measurement procedure

- Sweep the readout amplitude (or photon number n_{ph}) over a meaningful range, typically covering the operating point used during readout optimization and adjust the readout pulse length to maintain the same SNR.
- For each amplitude:
  - Choose a fixed number of random cycles N per sequence (e.g. 40).
  - Repeat by generating K different random gate sequences (e.g. 100):
    - Generate a random string $\boldsymbol{i}$ of N gates, where each gate is I or X with 50% probability.
    - Run this gate sequence a large number of times ($M > 1000$):
      1. Measure the qubit by applying the readout pulse at the given amplitude and pulse length and save the observed state (ground state or excited state side of a linear threshold) $r_0 \in \{0, 1\}$ and assignment confidence f_0 (how certain the IQ point can be assigned to $|0\rangle$ or $|1\rangle$).
      2. For each round $n = 1 \dots N$:
         - Apply gate $i_n \in \{I, X\}$.
         - Immediately measure the qubit state and save the observed state r_n and assignment confidence f_n.
      3. Let the qubit and environment relax to its thermal state by waiting for $\tau_r = 10\,T_1$.
      4. Build the observed output bit string $\boldsymbol{r}$ and assignment confidence string $\boldsymbol{f}$ for all measurements $n \in [0, N]$.
    - Save the input bitstring of played gates $\boldsymbol{i}$ (length N) and resulting M output bitstrings of recorded outcome states $\boldsymbol{r}$ and assignment confidences $\boldsymbol{f}$ (each of shape $N + 1 \times M$).
    - Repeat again for another random gate sequence
- The final result should be a $K \times N$ matrix $\boldsymbol{i}$ of played gates and two $K \times N + 1 \times M$ matrices of measurement outcomes $\boldsymbol{r} \in \{0, 1\}$ and assignment confidences $\boldsymbol{f}$. K is the number of random sequences, N the number of gates within a sequence and M the number of repetitions (for averaging) for each sequence.
- repeat for a different readout amplitude.

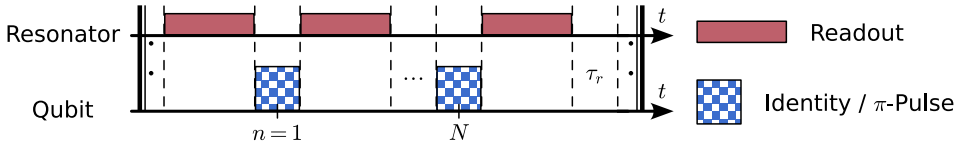


Figure 19: Pulse sequence for the RILB measurement. Generate a randomised bitstring $\boldsymbol{i}$ of length N containing identity or π-pulse at each position with a 50% propability. Alternatingly measure and play i_n, starting and ending with a readout. Wait for a delay of $\tau_r = 10\,T_1$ to relax back to the thermal state. Repeat M times with the same randomised bitstring. Repeat everything for K different randomised bitstrings.

Pulse sequence

Figure 19 shows the pulse sequence to use for the RILB experiment.

Required data

- Calibrated readout photon number n_{ph} sweep and according readout pulse lengths τ that give the same SNR.
- Number of random sequences (K), number of repetitions per sequence (M) and total number of random gates in each sequence (N).
- All generated random gate strings $\boldsymbol{i}$ (I, X) used in the experiment ($K \times N$ matrix).
- Measured output bit strings $\boldsymbol{r}$ and their assignment confidence $\boldsymbol{f}$ ($K \times N + 1 \times M$ matrices).
- Readout pulse parameters: amplitude, frequency, duration and envelope.
- π-pulse parameters: amplitude, duration and envelope.
- Inter-round timing: delay between gate and readout.

Assumptions and limitations

For this QoI to be successful it is required that the readout duration is much shorter than T_1 ($\tau_{\text{Readout}}/T_1 \ll 1$). Otherwise, the population outside of the computational subspace might relax back in the computational subspace during readout. The length of the π-pulses is ideally chosen to have a spectral content smaller than the anharmonicity of the qubit, allowing to neglect the leakage originating from the bit-flip operation. It is also assumed, that the qubit lives inside a Markovian environment.

Data analysis procedure

For each measured readout amplitude:

1. for each random gate sequence $k \in (1, K)$:
   (a) For each experimental realisation M compute the flipped-or-not bitstring $\boldsymbol{o}_n = \boldsymbol{r}_{n-1} \oplus \boldsymbol{r}_n$.
   (b) Calculate the bitwise correlation with the played pulse sequence $\boldsymbol{C}_n = 1 - 2(\boldsymbol{i}_n \oplus \boldsymbol{o}_n)$. $\boldsymbol{C}$ is a $N \times M$ matrix.
   (c) Calculate the average $\bar{\boldsymbol{C}}$ over all experimental realisations remaining in M. $\bar{\boldsymbol{C}}$ is a vector with N elements.
2. Calculate the average correlation across all random sequences (across K) to get $\langle \bar{\boldsymbol{C}}_n \rangle$, a vector with N elements.
3. Fit $\langle \bar{\boldsymbol{C}}_n \rangle$ to $\langle \bar{\boldsymbol{C}}_n \rangle = (1/2)(A + B(1-L)^n)$

Plot $\langle \bar{\boldsymbol{C}}_n \rangle$ vs circuit depth n for each readout amplitude n_{ph}.

Required metadata for analysis:

- Readout pulse duration τ (in μs)

- Readout pulse amplitude calibration against photon number.
- Readout pulse shape and rise/fall time (in μs)
- Delay τ_0 between control and first measurement pulse
- Delay between measurement pulses (τ_d) (in μs)
- Resonator Q and frequency f_r (in GHz)

Quantity to report for comparison

- Plot $\langle \bar{\boldsymbol{C}}_n \rangle$ vs circuit depth n for each readout amplitude n_{ph}.
- Calculated readout fidelities, repeatability and leakage rates per readout operation for each photon number.

QoI references

[21] S. Hazra, W. Dai, T. Connolly, P.D. Kurilovich, Z. Wang, L. Frunzio, and M.H. Devoret. Benchmarking the readout of a superconducting qubit for repeated measurements. *Physical Review Letters*, 134(10):100601, March 2025. doi: 10.1103/PhysRevLett.134.100601

3.4 Randomised Benchmarking

Quantifies the accuracy of qubit control pulses by randomising a sequence of Clifford gates to prepare a qubit state.

Description

Clifford gate randomised benchmarking can indicate average gate error through single or multi-qubit Clifford gates. This method of characterising a qubit is robust against state preparation and measurement error (SPAM).

The RB protocol involves running a random sequence of Clifford gates at increasing depths and then running the inverse of those same gates. There are 24 Clifford gates in the single-qubit Clifford group. The randomised gate sequence can be expressed as:

$$G_1 G_2 G_3 ... G_{N-1} G_N G_{inverse}$$

Measurement procedure

1. Determine a list of sequence depths so that the qubit survival probability decays significantly, with at least 10 different circuit depths (N) in the list. An exponential decay curve should fit this decay. At this point the desired circuit outcome can be chosen to be either $|1\rangle$ or $|0\rangle$, the survival probability later will be a measure of how often this was achieved.

2. Determine how many random circuits m_{circ} should be performed for each sequence length N. The number of m_{circ} circuits can be chosen depending on desired accuracy, 10 circuits per gate depth can be used initially.

3. For the list of sequence lengths, generate a sequence of random Clifford gates, randomly selected from the Clifford group and then the last set of gates should be the inverse of all previous gates. These Clifford gates should be broken down into native superconducting qubit pulses.

4. For each circuit at each gate sequence depth, the number of shots (averaging) for each circuit should be based on desired benchmarking precision, 1000 repeats should be sufficient.

5. Once circuit results are acquired, calculate the survival probability p_{survival} for each circuit at each sequence length, the circuit results should be averaged for each sequence length.

After a single dataset is acquired also investigate the temporal stability of the RB fidelity by picking a Clifford gate depth where $p_{\text{survival}} \approx 75\%$, and repeat random circuits at this gate depth and measure how p_{survival} fluctuates over time. Measure for at least one hour.

Pulse sequence

Once Clifford gates are randomly drawn from the Clifford gate set, they must be broken down into native superconducting qubit gates, in simple terms it will be a series of $\pi/2$-pulses with varying phase (ϕ). It is necessary to develop software capabilities that can generate the random sequence of native gates that will be then be appropriately scheduled to be performed on the qubit. This should then be repeated for every circuit for a particular gate depth, of which the gates used should each time be newly generated.

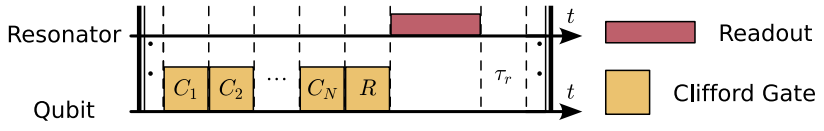


Figure 20: Pulse sequence for a randomised benchmarking circuit at a gate depth N, each random Clifford gate C_i, $i \in (1, N)$, is played in close succession, followed by the inverse sequence of gates (here labeled R). Not shown here is the different (random) phases of the Clifford gates. Each random circuit is repeated several times for averaging, before a new random circuit is generated. In total m_{circ} number of unique circuits are generated at each chosen gate depth.

Assumptions and limitations

RB only works if $\pi/2$-pulses are properly calibrated and do not cause any excitations beyond the computational subspace when many pulses are used. Also, a requirement is the ability to randomly generate a series of native control pulses according to the Clifford gate set.

Data analysis procedure

The resulting data should be characterised into $|0\rangle$ or $|1\rangle$ with p_{survival} indicating the resulting circuit outcome being in the predetermined desired state. The exponential decay of the survival probability can be written as:

$$p_{\text{survival}} = A_0 \alpha^m + B_0, \tag{5}$$

where A_0, α and B_0 are fitting parameters corresponding to amplitude, decay and the baseline respectively. Since α represents the error per Clifford gate, the average Clifford RB gate error, r, can be computed from:

$$r = 1 - \alpha - \frac{(1-\alpha)}{d}, \tag{6}$$

where $d = 2^{N_q}$ is the dimension of Clifford gates and N_q is the number of qubits (here $N_q = 1$, so d=2).

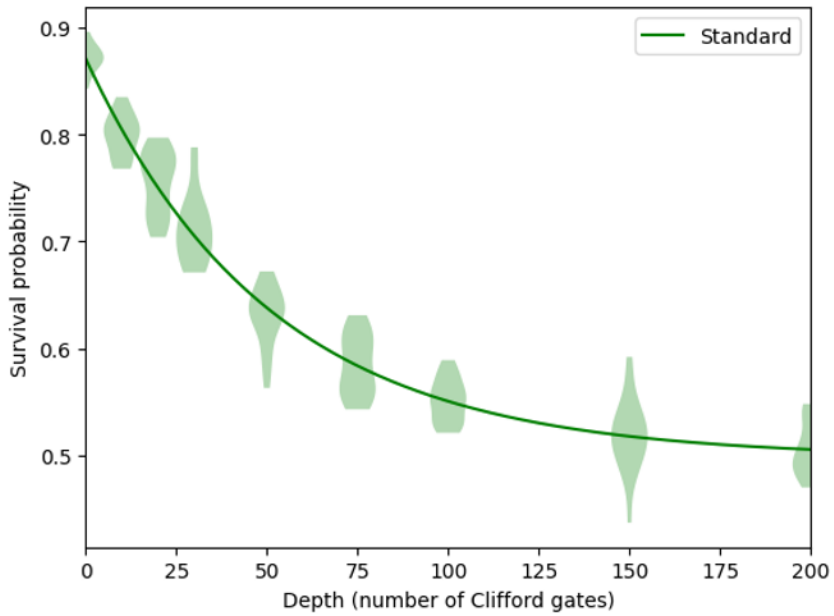


Figure 21: This plot shows qubit gate fidelity characterisation by Randomised Benchmarking using the Clifford gate set, up to a depth of $N = 200$ Clifford gates with 10 circuits executed m_{circ} per gate depth. The gate fidelity from this data set is 99.03%. The green shaded areas depict the range and density of circuit results at that particular depth.

Quantity to report for comparison

- Single gate fidelity α as a percentage (%) with at least 2 significant digits beyond the last 9. Associated error bar from fit.
- Time series data of RB survival probability p_{survival} at the chosen gate depth.
- Calculated mean and standard deviation of the temporal spread of p_{survival}.
- Total time of temporal measurement and number of counts.
- $\pi/2$-pulse parameters ($\tau_{\pi/2}$, shape, frequency).
- Readout parameters: SNR, τ, n_{ph}.

QoI references

[22] Deep Lall, Abhishek Agarwal, Weixi Zhang, Lachlan Lindoy, Tobias Lindström, Stephanie Webster, Simon Hall, Nicholas Chancellor, Petros Wallden, Raul Garcia-Patron, Elham Kashefi, Viv Kendon, Jonathan Pritchard, Alessandro Rossi, Animesh Datta, Theodoros Kapourniotis, Konstantinos Georgopoulos, and Ivan Rungger. A review and collection of metrics and benchmarks for quantum computers: definitions, methodologies and software. doi: 10.48550/arXiv.2502.06717. URL `http://arxiv.org/abs/2502.06717`. arXiv:2502.06717 [quant-ph]

3.5 Quantum state preparation

The target with quantum state preparation benchmarking is to determine the fidelity of achieving the desired (prepared) qubit state.

Description

Measurement procedure

The measurement consists of preparing the qubit in either $|0\rangle$ or $|1\rangle$ states and then measuring if the outcome matches the expected result.

- (if no significant thermal population, e.g. for Transmons) Randomise the state of the qubit by playing a $\pi/2$-pulse. (optional).
- Measure and record the state of the qubit (single shot measurement, IQ value). This is M_1.
- Play a π-pulse (prepare $|0\rangle$ or $|1\rangle$).
- Measure and record the state of the qubit (single shot measurement, IQ value). This is M_2.
- Wait $\tau_r \approx 10\,T_1$ before repeating to relax back to the ground state.

Repeat the above sequence at least 10^6 times to get enough statistics over time (total measurement will take about one hour). For qubits with poor state separation, $< 99\%$, more measurements may be needed.

Pulse sequence

The pulse sequence is shown in figure 22.

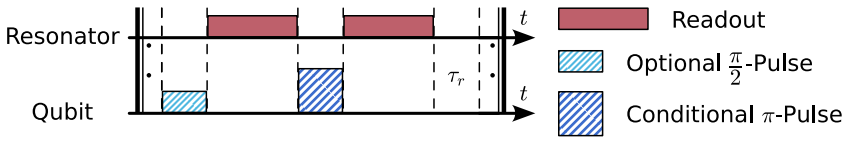


Figure 22: Pulse sequence for quantum state preparation fidelity measurement.

Assumptions and limitations

It is assumed that prior to the measurement, T_1 value is roughly known and the π-pulse is optimised (see tuneup protocol). It is also assumed the readout fidelity and state separation has been optimised, by maximizing the separation between the two states in the IQ plane. Optimizing the readout signal involves optimizing frequency and signal power as a function of pulse length and/or attenuation without driving the readout close or into excited states. A noisy measurement could result in overlapping IQ clouds.

Data analysis procedure

1. Convert the measured IQ values to qubit states using the standard Gaussian mixture model classifier.
2. For qubits with poor readout fidelity (overlapping IQ clouds) there will be many measurements that cannot be classified as 0 or 1 with high certainty. Discard all measurements that have a state classification uncertainty in either M_1 or M_2 greater than η.
3. Calculate the conditional probabilities $P_{xy} = P(M_2 = x|M_1 = y)$ with $x, y = |0\rangle$ or $|1\rangle$. I.e. find the probability that given we were initially in state y, applying the π-pulse brought us to state x. Calculate the state preparation infidelity

$$1 - F = \frac{P_{11} + P_{00}}{2}. \tag{7}$$

4. Repeat analysis for different state classification uncertainties η and plot against $1-F$. Extract the constant level $F_{\max}$ for small $\eta < \eta_0$. This will be the regime where the state separation uncertainty no longer contributes to the state preparation fidelity. Report this $F_{\max}$ as the state preparation fidelity.

5. Segment the data in smaller time-bins and perform the same analysis steps 2-3 using the found $\eta < \eta_0$. Plot $F_{\max}$ vs time.

Quantity to report for comparison

- Calculated $F_{\max}$ with associated error bar. Use the average value across all times measured. Report at least two significant digits beyond the last 9.
- Calculated $F_{\max}$ with associated error bar as a function of time.
- Calculated mean and standard deviation of the variation in $F_{\max}$ over time.
- Calculated $F_{\max}$ with associated error bar as a histogram from the temporal measurement.
- π-pulse parameters (τ_π, shape, frequency).
- $\pi/2$-pulse parameters ($\tau_{\pi/2}$, shape, frequency).
- Readout parameters: SNR, τ, n_{ph}.

QoI references

[4] Philip Krantz, Morten Kjaergaard, Fei Yan, Terry P. Orlando, Simon Gustavsson, and William D. Oliver. A quantum engineer's guide to superconducting qubits. *Applied Physics Reviews*, 6(22):021318, June 2019. ISSN 1931-9401. doi: 10.1063/1.5089550

[22] Deep Lall, Abhishek Agarwal, Weixi Zhang, Lachlan Lindoy, Tobias Lindström, Stephanie Webster, Simon Hall, Nicholas Chancellor, Petros Wallden, Raul Garcia-Patron, Elham Kashefi, Viv Kendon, Jonathan Pritchard, Alessandro Rossi, Animesh Datta, Theodoros Kapourniotis, Konstantinos Georgopoulos, and Ivan Rungger. A review and collection of metrics and benchmarks for quantum computers: definitions, methodologies and software. doi: 10.48550/arXiv.2502.06717. URL `http://arxiv.org/abs/2502.06717`. arXiv:2502.06717 [quant-ph]

3.6 Energy relaxation time T_1: passive vs active

We define T_1 as the characteristic time scale governing the energy relaxation of a qubit's ground and excited state back to its (thermal) equilibrium. To understand if the T_1 time is affected by the active state preparation of a typical T_1 measurement (playing π-pulses) this QoI aims to compare this active method with a passive T_1 measurement where only readout pulses are used. Each relaxation experiment is to be repeated multiple times to minimize the chance of T_1 fluctuations impacting the comparison.

Description

We assume a qubit with ground state $|0\rangle$ and first excited state $|1\rangle$. The transition rate from state $|0\rangle$ to state $|1\rangle$ we call Γ_{up}, while we call the transition rate for the opposite process Γ_{down}. To define the characteristic timescale T_1, we define the relaxation rate $\Gamma_1 = \Gamma_{\text{up}} + \Gamma_{\text{down}}$. The probability to find the qubit in the same state we started from is governed by the equations:

$$P(|1\rangle, t) = [P(|1\rangle, 0) - P(|1\rangle)_{\text{equilibrium}}] \exp(-\Gamma_{\text{down}} t) + P(|1\rangle)_{\text{equilibrium}}, \tag{8}$$

$$P(|0\rangle, t) = [P(|0\rangle, 0) - P(|0\rangle)_{\text{equilibrium}}] \exp(-\Gamma_{\text{up}} t) + P(|0\rangle)_{\text{equilibrium}}. \tag{9}$$

Finally, we define $T_1 = \dfrac{1}{\Gamma_1} = \dfrac{1}{\Gamma_{\text{down}} + \Gamma_{\text{up}}}$. It is known that any interaction with a superconducting qubit might influence its dynamics, either by directly interacting with the qubit or by influencing its environment [10]. The aim here is to compare two relaxation measurements: active (π-pulse is used to flip the qubit state) vs passive (no drive pulse is used).

Measurement procedure

Two measurement procedures employed to quantify T_1. Both are outlined in the corresponding subsections below. Both should be measured and the results compared against each other. To understand if variations in T_1 occurred, these two methods shall be measured multiple times. We recommend at least 5 times each (active, passive, active, passive, ...).

1. **Passive relaxation measurement using state preparation by measurement**: Using repeated measurements we start the T_1 measurement sequence from a known qubit state without applying a gate. Assuming a real-time capable control hardware, once the desired state is found, wait a time τ and measure again the qubit state. For the state detection threshold, the center of the IQ-cloud indicating the respective state should be used to reduce the number of state preparation errors. Then vary the wait time τ following the initial measurement. The longest wait time τ_{max} chosen should be $10 \times T_1$ to adequately capture the decay back to equilibrium. A logarithmic distribution of values $\tau \in (0, \tau_{\text{max}})$ should be used. For each value τ at least 100 state measurements in each state ($|0\rangle$ and $|1\rangle$) must be recorded to calculate the probability of measuring the same state as at the start of the sequence $P(\tau)$ with good accuracy. After each sequence wait a time $\tau_{\text{r}} = 10T_1$ to allow for the qubit state to be randomised again before next measurement.

   For measurement systems without real-time capability, a post-selection strategy can be employed instead. Similarly to the above, measure the qubit state, wait a time $\tau \in (0, \tau_{\text{max}})$ and measure again the qubit state. For each τ take a larger number of measurements to ensure that for each starting state ($|0\rangle$ and $|1\rangle$) at least 100 measurement sequences are available. For very cold qubits this may require a very large number of measurements. Complete the data analysis before proceeding, or repeat the measurement if more statistics are needed.

2. **Active relaxation measurement using state preparation by measurement**: The measurement procedure is similar to the passive measurement above, apart from the use of a conditional π-pulse to set the initial qubit state after the initial measurement. I.e. a measurement is taken showing the desired state, and then a π-pulse is played to flip the qubit, followed by the usual varied wait time τ. The measurement after the wait time τ is to be followed immediately by the next state preparation sequence (reset time $\tau_{\text{r}} = 0$) to speed up the measurement.

   When using the post-processing approach it is also necessary to use the same $\tau_{\text{r}} = 0$ to facilitate comparison between labs. This is because the rate at which the π-pulses are applied is a parameter which affects the measurement outcome.

3. Repeat each technique at least 5 times.

Pulse sequence

Figures 23 and 24 shows the pulse sequences to be implemented.

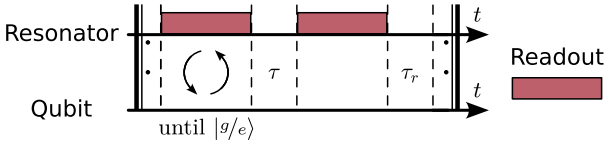


Figure 23: Pulse sequence for T_1 measurement using state preparation by measurement.

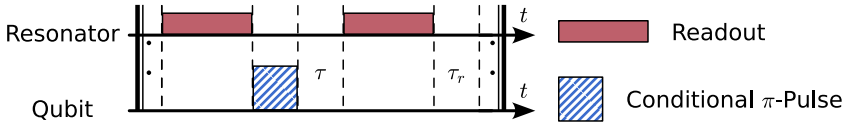


Figure 24: Pulse sequence for T_1 measurement using active state preparation.

Assumptions and limitations

The measurement procedure assumes that the readout is tuned up to be sufficiently QND such that the measurement itself does not lead to state transitions.

Data analysis procedure

For each starting state $|0\rangle$, $|1\rangle$ the time dependent probability $P(t)$ to encounter the state after a given time is calculated from the at least 100 measurements for every time τ between state initialization and measurement. Fitting Eqs.(8) and (9) $\Gamma_{\mathrm{up/down}}$ are extracted and T_1 calculated according to $T_1 = \dfrac{1}{\Gamma_1} = \dfrac{1}{\Gamma_{\mathrm{down}} + \Gamma_{\mathrm{up}}}$. Also calculate the resulting quality factor of the qubit $Q = 2\pi T_1 f_q$.

Required metadata for analysis:

- Qubit frequency f_q,
- Readout pulse photon number n_{ph},
- Readout pulse duration τ,
- Reset time τ_{r},
- Threshold used for state preparation by measurement (if applicable),
- Whether active state preparation was used or post-processing is required.

Quantity to report for comparison

- Γ_{down}, Γ_{up}, T_1 and Q for each of the two methods (active and passive) with at least 3 significant digits and associated error bars from fits.

QoI references

[23] W. Dai, S. Hazra, D. K. Weiss, P. D. Kurilovich, T. Connolly, H. K. Babla, S. Singh, V. R. Joshi, A. Z. Ding, P. D. Parakh, J. Venkatraman, X. Xiao, L. Frunzio, and M. H. Devoret. Spectroscopy of drive-induced unwanted state transitions in superconducting circuits. doi: 10.48550/arXiv.2506.24070. URL `http://arxiv.org/abs/2506.24070`. arXiv:2506.24070 [quant-ph]

[24] Marie Frédérique Dumas, Benjamin Groleau-Paré, Alexander McDonald, Manuel H. Muñoz-Arias, Cristóbal Lledó, Benjamin D'Anjou, and Alexandre Blais. Measurement-induced transmon ionization. *Physical Review X*, 14(4):041023, October 2024. ISSN 2160-3308. doi: 10.1103/PhysRevX.14.041023

[25] Mostafa Khezri, Alex Opremcak, Zijun Chen, Kevin C. Miao, Matt McEwen, Andreas Bengtsson, Theodore White, Ofer Naaman, Daniel Sank, Alexander N. Korotkov, Yu Chen, and Vadim Smelyanskiy. Measurement-induced state transitions in a superconducting qubit: Within the rotating-wave approximation. *Physical Review Applied*, 20(5):054008, November 2023. ISSN 2331-7019. doi: 10.1103/PhysRevApplied.20.054008

[26] Daniel Sank, Zijun Chen, Mostafa Khezri, J. Kelly, R. Barends, B. Campbell, Y. Chen, B. Chiaro, A. Dunsworth, A. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Mutus, M. Neeley, C. Neill, P.J.J. O'Malley, C. Quintana, P. Roushan, A. Vainsencher, T. White, J. Wenner, Alexander N. Korotkov, and John M. Martinis. Measurement-induced state transitions in a superconducting qubit: Beyond the rotating wave approximation. *Physical Review Letters*, 117(19):190503, November 2016. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.117.190503

[27] Konstantin N. Nesterov and Ivan V. Pechenezhskiy. Measurement-induced state transitions in dispersive qubit-readout schemes. *Physical Review Applied*, 22(6):064038, December 2024. ISSN 2331-7019. doi: 10.1103/PhysRevApplied.22.064038

3.7 Interleaved T_1, T_2^* (Ramsey), and $T_2^{\mathbf{Echo}}$

Long monitoring of the three fundamental coherence times, using an interleaved acquisition strategy.

Description

This protocol characterizes the qubit's T_1, T_2^* (Ramsey), and T_2^{Echo} coherence times in an interleaved fashion to monitor their joint evolution under constant environmental conditions. In each repetition cycle, data points for one delay point[1] from each of the three experiments are acquired sequentially. This is repeated until each time delay has been measured and then the total sequence is repeated for averaging. This removes deviations between the extracted values of the different coherence times due to slow drifts and ensures fair comparison. The qubit is passively initialised from thermal equilibrium without active reset. A fixed delay of $\tau_{\mathrm{r}} = 10\,T_1$ ensures full relaxation before the T_2 measurements. For the T_1 measurements, an even longer delay of $\tau_{\mathrm{TLS}} = 1\,\mathrm{ms}$ guarantees that even long-lived two-level systems (TLS) have relaxed back to their thermal state.

Measurement procedure

Each repetition cycle consists of the following sub-sequences.

1. T_1**:** A blind π-pulse prepares the excited state, followed by a variable wait time τ_1 and measurement.
2. T_2^* **(Ramsey):** A $\pi/2$-pulse brings the qubit to the equator of the Bloch sphere. After a variable delay τ_R, a second $\pi/2$-pulse with a virtual Z rotation is applied: a phase shift $\Delta\phi = 2\pi f_{\mathrm{detuning}} \cdot \tau_R$ mimics a detuned drive, enabling Ramsey fringes without frequency detuning. The detuning frequency should be chosen to be $f_{\mathrm{detuning}} \approx 3/T_2^*$.
3. $T_2^{\mathbf{Echo}}$**:** A $\pi/2$–wait–π–wait–$\pi/2$ sequence with total separation τ_E is used, with the final $\pi/2$-pulse phase-shifted by π to rotate the qubit into the excited state for maximal contrast.

The full repetition cycle consists one T_1, one T_2^{Echo}, and four T_2^* measurements and is executed for all delays $\tau \in \{\tau_1, \tau_R, \tau_E\}$, forming a complete interleaved acquisition. The delay times τ are chosen to cover $5\,T_X$ for each experiment. For T_1 and T_2^{echo}, delays are logarithmically spaced with a total of 50 points. For T_2^*, a linear spacing with 200 delays is used to allow a good fit to the oscillating Ramsay fringe. This is why the Ramsey sequence is performed four times within each cycle, using different delay values for τ_R. The measurement of a single interleaved IQ acquisition corresponds to $\approx 3000\,T_1 + 50\,\mathrm{ms}$. The full experiment consists of 500 such acquisitions and is run continuously for at least 12 hours to capture slow fluctuations. For a representative $T_1 = 10\,\mu\mathrm{s}$, a single experiment takes $\approx 40\,\mathrm{seconds}$, yielding 1080 experiments over the 12 hour period.
As a last step, the measured IQ data are sorted by experiment time and delay step. For each delay, the IQ clouds are fitted using a Gaussian mixture model (as introduced in section 2.3.2.1) to extract the state populations $P(\tau)$.

Pulse sequence

Figure 25 shows the pulse sequence to implement.

Assumptions and limitations

Readout and gate calibrations are assumed to be stable and accurate, as established during tuneup. The protocol assumes that T_1, T_2^*, and T_2^{Echo} are approximately constant over the timescale of each acquisition block. Long-term drift is expected and is the primary quantity under investigation. The thermal state is assumed to approximate the ground state sufficiently well; this may not hold for hot or low-frequency qubits, which could compromise contrast and fitting quality.

[1] There are 4 delay points measured for the Ramsey- and one for the T_1- and Echo-experiments within a sequence.

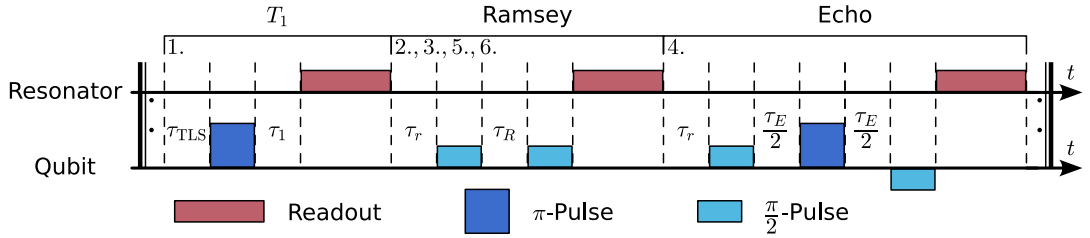


Figure 25: **Pulse sequence for interleaved T_1, T_2^* (Ramsey), and T_2^{Echo} measurements from the thermal state.** Each cycle consists of one T_1, one T_2^{Echo}, and four T_2^* measurements, repeated 500 times for varying delays $\tau \in \{\tau_1, \tau_R, \tau_E\}$. Horizontal (volta) brackets indicate the order in which sections are played within each cycle (T_1, Ramsey, Ramsey, Echo, Ramsey Ramsey), to accommodate 4x more points in τ_R. In the T_1 section, a waiting time of $\tau_{\mathrm{TLS}} = 1\,\mathrm{ms}$ ensures that the qubit environment is well thermalised. Afterwards, the qubit is initialised in the excited state via a blind π-pulse and measured after a variable delay τ_1. For the Ramsey sequence, first wait for $\tau_r = 10\,T_1$. Then, two $\pi/2$-pulses separated by τ_R are applied, with the second pulse carrying a virtual phase shift to emulate detuning. For T_2^{Echo}, after waiting for $\tau_r = 10\,T_1$, a refocusing π-pulse is inserted midway between two $\pi/2$-pulses, each separated by $\tau_E/2$. The final $\pi/2$-pulse is phase-shifted by π to rotate the qubit to $|e\rangle$, increasing readout contrast.

Data analysis procedure

The excited state populations as a function of delay time are fitted as follows:

$$
\begin{aligned}
T_1: &\quad P_e(\tau_1) = A\, e^{-\tau_1/T_1} + p_{th} \\
T_2^*: &\quad P_e(\tau_R) = A\cos(\omega\tau_R + \phi)e^{-\tau_R/T_2^*} + B \\
T_2^{\mathrm{Echo}}: &\quad P_e(\tau_E) = A\, e^{-\tau_E/T_2^{\mathrm{Echo}}} + B
\end{aligned}
$$

From these fits, the decay constants T_1, T_2^*, and T_2^{Echo} are extracted, along with parameters such as the virtual detuning frequency ω, phase ϕ, and baseline offset B. The baseline offset B should correspond to $0.5 - e^{\tau/T_1}$, where τ is the readout pulse length. The output includes coherence time traces with fit residuals, histograms, PSD and ADEV plots, and optionally all intermediate population and IQ data for quality checks. Figure 26 shows an example of temporal analysis over a series of T_1 measurements.

Note: In some cases the Ramsey oscillations may include additional beating frequency components. In such situations a fit to the above equation will not result in the correct T_2^*. It is therefore required to have an analysis procedure that checks against additional frequency components and determines the right model to use (c.f. eq. (1)).

Quantity to report for comparison

- Time series of extracted T_1, T_2^{Echo}, T_2^* and f_q, with associated error bars from fits.
- Calculated means and standard deviations of all four quantities taken over the entire time period. Report these with at least 3 significant digits.
- Histogram over time series data
- Calculated PSDs and ADEVs of the time series data.
- Total time of temporal measurement and number of counts.
- Readout parameters: SNR, τ, n_{ph}.

QoI references

[11] Jonathan J. Burnett, Andreas Bengtsson, Marco Scigliuzzo, David Niepce, Marina Kudra, Per Delsing, and Jonas Bylander. Decoherence benchmarking of superconducting qubits. *npj Quantum Information*, 5(1):1–8, June 2019. ISSN 2056-6387. doi: 10.1038/s41534-019-0168-5

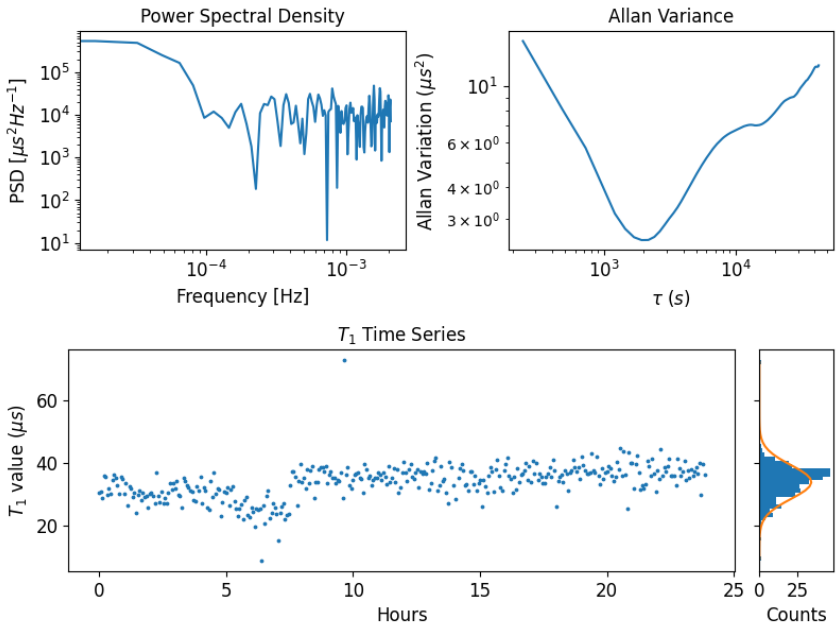


Figure 26: Example visualisation of the temporal variations of T_1 from an interleaved coherence measurement.

3.8 Szilard engine

Detection of long lived TLS's though hyper-polarization of environment by active control of the qubit.

Description

The Szilard engine is a method by which the long lived two-level system (TLS) environment of the qubit can be measured indirectly. This method consists of two distinct sequences. First, the polarization sequence is run in which the qubit is measured and a conditional π-pulse is played to excite (de-excite) the qubit. This introduces heating (cooling) into the TLS environment through the qubit. After the polarization sequence, the qubit is prepared in a target state. It is important to run the experiment here from both excited- and ground-state to ensure that the extraction of the dynamics is complete. Finally, a quantum jump trace is measured, which maps the qubit-TLS dynamics.

Measurement procedure

To be able to extract the coupling rates and detunings of the TLS's to the qubit, multiple traces need to be measured at different thermal populations of the TLS environment. Also, the qubit-TLS dynamic needs to be measured starting from both the ground and excited state. In the following the terms polarization, initialization, and quantum jump trace will be explained:

- Polarization is the active control of the qubit to a state. For the heating (cooling) experiment, the qubit must be polarised to the excited (ground) state. A single polarization consists of a measurement of the qubit, with a conditional π-pulse played afterwards. The condition on the π-pulse is chosen such that after the polarization procedure the qubit is in the excited (ground) state. This polarization is then repeated N times to complete the sequence for a single trace. The qubit and the TLS's interact in between the polarisations, to facitlitate this interaction a wait time of $\tau_p = 10/T_1$ is added before the next polarisation. The whole polarization sequence is repeated for $N \in [1, 10, 100, 1000, 10000]$.
- Similarly to a polarization, the initialization actively controls the qubit to the wanted state. In both the heating and cooling protocols, the qubit is initialised once in the excited state and once in the ground state. This results in 4 different traces with 5 different polarization sequence lengths.
- Finally, the stroboscopic quantum jump trace is measured for at least 100 ms. The points in the quantum jump trace must be dense with respect to the T_1 of the system. Each measurement is separated by a idle time $\tau_M \approx$ readout pulse length. These traces must then be averaged 400 times for all different types of polarizations and initializations to capture the dynamics of the system. This gives a total of 8000 traces per Szilard experiment.
- Each measurement in the qubit monitoring sequence produces one IQ pair which is stored for the final comparison including their timestamp.

Pulse sequence

Figure 27 shows the pulse sequence to implement.

Assumptions and limitations

It is assumed that the TLS's are weakly coupled and long lived. Deviations from this will result in limited polarizability of the TLS environment.

Data analysis procedure

The measured stroboscopic Quantum Jump traces are sorted by polarization number N, heating or cooling, and ground or excited state initialization. The individual traces for each setting are averaged giving the dynamics of the system for different preparations. The heated (cooled) dynamics are then bundled together for the different polarization sequences and simultaneously fit the to the Solomon rate equations following Ref. [10, 28]:

$$\dot{\mathbf{p}} = -\mathbf{M}\mathbf{p} + \begin{pmatrix} \Gamma_q \\ \Gamma_t \end{pmatrix} \mathrm{p_{th}} \tag{10}$$

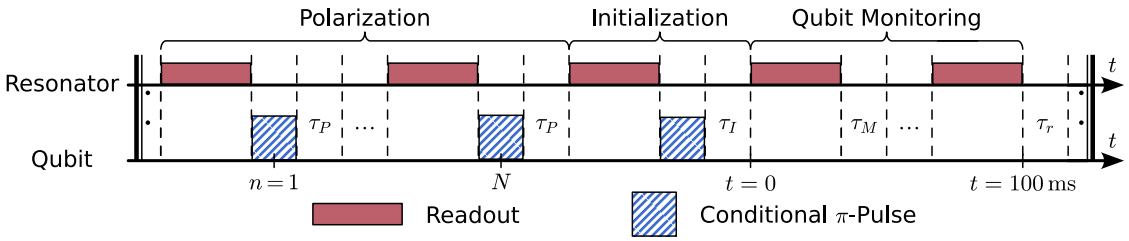


Figure 27: **Pulse sequence for the Szilard engine.** We begin by polarizing the qubit's environment by stabilizing the qubit in either $|g\rangle$ or $|e\rangle$ with active feedback preparation N times, each separated by a delay τ_P. After active initialization to either $|g\rangle$ or $|e\rangle$ and a delay of τ_I, the qubit state is monitored stroboscopically for duration of 100 ms, with individual readouts separated by a delay τ_M. At the end, a delay of $\tau_r = 10\,\mathrm{ms}$ makes sure, that the qubit and environment relaxed back to their thermal state. The sequence is then repeated for all combinations of polarization (heating/cooling, different N) and initialization ($|g\rangle$, $|e\rangle$) for 400 times.

$$\mathbf{M} = \begin{pmatrix} \Gamma_q + \sum_i \Gamma^i_{qt} & -\mathbf{\Gamma_{qt}} \\ -\mathbf{\Gamma_{qt}} & \mathbf{\Gamma_{tls}} \end{pmatrix}$$

$$\mathbf{\Gamma_{TLS}} = \begin{pmatrix} \Gamma^1_{qt} + \Gamma_t & 0 & ... & 0 \\ 0 & \Gamma^2_{qt} + \Gamma_t & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & ... & 0 & \Gamma^n_{qt} + \Gamma_t \end{pmatrix}$$

Here $\mathbf{p}$ is a vector containing the populations of the qubit and the TLS's, this is also the only time dependent variable, where $\mathrm{p_{th}}$ is the steady state population. Γ_q (Γ_t) is the decay of the qubit (TLS's) in the absence of any TLS's (qubit). $\mathbf{\Gamma}_{qt}$ is a vector containing all individual coupling rates between the qubit and each TLS ($\Gamma^\mathrm{i}_\mathrm{qt}$). Each trace in the bundled qubit dynamics start with different initial populations for all considered TLS's. This population can be calculated by:

$$p^i_\mathrm{off} = \frac{\Gamma_t p_\mathrm{th} + \Gamma^i_\mathrm{qt}}{\Gamma_t + \Gamma^i_\mathrm{qt}}$$

$$p^i(0) = (p_\mathrm{th} - p^i_\mathrm{off}) e^{-(\Gamma_t + \Gamma^i_\mathrm{qt}) t_\mathrm{pol}} + p^i_\mathrm{off} \tag{11}$$

Here t_pol corresponds to the polarization time of the TLS which is N times $\tau_p + \tau_r + \tau_\pi$. This initial condition follows from the Solomon model and assumes that each polarization step is fast enough to keep the qubit in the excited state.

To reduce the complexity of the equations and not overfit the data, we assume that the TLS's are spaced equally from each other in a ladder with the qubit centered in the middle of the ladder (also in the middle between two TLS's) and each TLS has the same coupling g to the qubit. This approximation is valid, as we are interested in long time scale dynamics and any changes in the exact positions will only lead to changes in the short time scale. This gives a constraint on each $\Gamma^\mathrm{i}_\mathrm{qt}$. This constraint reduces our parameters significantly going from n (number of TLS's) parameters to two (density of TLS's, coupling g both in units of Γ_2).

Quantity to report for comparison

- TLS ladder density
- Coupling rate
- Number of coupled TLS's
- T_1 of the qubit with and without the TLS's

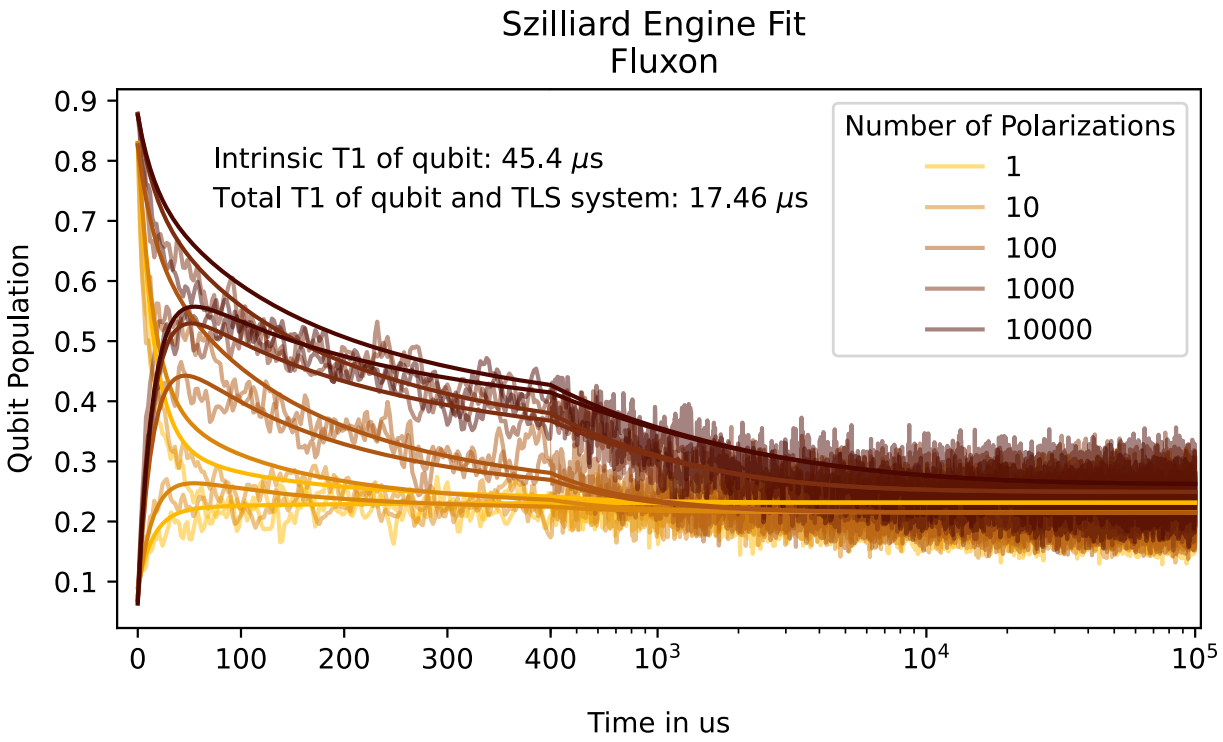


Figure 28: Example of a Szilard hyperpolarization experiment including fit to Solomon equations.

QoI references

[10] Martin Spiecker, Patrick Paluch, Nicolas Gosling, Niv Drucker, Shlomi Matityahu, Daria Gusenkova, Simon Günzler, Dennis Rieger, Ivan Takmakov, Francesco Valenti, Patrick Winkel, Richard Gebauer, Oliver Sander, Gianluigi Catelani, Alexander Shnirman, Alexey V. Ustinov, Wolfgang Wernsdorfer, Yonatan Cohen, and Ioan M. Pop. Two-level system hyperpolarization using a quantum szilard engine. *Nature Physics*, 19(9):1320–1325, September 2023. ISSN 1745-2473, 1745-2481. doi: 10.1038/s41567-023-02082-8

[28] Martin Spiecker, Andrei I. Pavlov, Alexander Shnirman, and Ioan M. Pop. Solomon equations for qubit and two-level systems: Insights into non-poissonian quantum jumps. *Physical Review A*, 109(5): 052218, May 2024. ISSN 2469-9926, 2469-9934. doi: 10.1103/PhysRevA.109.052218

References

[1] Rajeev Acharya et al. Quantum error correction below the surface code threshold. *Nature*, 638 (8052):920–926, February 2025. ISSN 1476-4687. doi: 10.1038/s41586-024-08449-y.

[2] Rajeev Acharya et al. Suppressing quantum errors by scaling a surface code logical qubit. *Nature*, 614(7949):676–681, February 2023. ISSN 1476-4687. doi: 10.1038/s41586-022-05434-1.

[3] Sebastian Krinner, Nathan Lacroix, Ants Remm, Agustin Di Paolo, Elie Genois, Catherine Leroux, Christoph Hellings, Stefania Lazar, Francois Swiadek, Johannes Herrmann, Graham J. Norris, Christian Kraglund Andersen, Markus Müller, Alexandre Blais, Christopher Eichler, and Andreas Wallraff. Realizing repeated quantum error correction in a distance-three surface code. *Nature*, 605 (7911):669–674, May 2022. ISSN 1476-4687. doi: 10.1038/s41586-022-04566-8.

[4] Philip Krantz, Morten Kjaergaard, Fei Yan, Terry P. Orlando, Simon Gustavsson, and William D. Oliver. A quantum engineer's guide to superconducting qubits. *Applied Physics Reviews*, 6(22): 021318, June 2019. ISSN 1931-9401. doi: 10.1063/1.5089550.

[5] Mustafa Bal et al. Systematic improvements in transmon qubit coherence enabled by niobium surface encapsulation. *npj Quantum Information*, 10(1):1–8, April 2024. ISSN 2056-6387. doi: 10.1038/s41534-024-00840-x.

[6] Akshay A. Murthy and et al. Identifying materials-level sources of performance variation in superconducting transmon qubits. doi: 10.48550/arXiv.2503.14424. URL `http://arxiv.org/abs/2503.14424`. arXiv:2503.14424 [quant-ph].

[7] Anjali Premkumar, Conan Weiland, Sooyeon Hwang, Berthold Jäck, Alexander P. M. Place, Iradwikanari Waluyo, Adrian Hunt, Valentina Bisogni, Jonathan Pelliciari, Andi Barbour, Mike S. Miller, Paola Russo, Fernando Camino, Kim Kisslinger, Xiao Tong, Mark S. Hybertsen, Andrew A. Houck, and Ignace Jarrige. Microscopic relaxation channels in materials for superconducting qubits. *Communications Materials*, 2(1):72, July 2021. ISSN 2662-4443. doi: 10.1038/s43246-021-00174-7.

[8] M. Virginia P. Altoé, Archan Banerjee, Cassidy Berk, Ahmed Hajr, Adam Schwartzberg, Chengyu Song, Mohammed Alghadeer, Shaul Aloni, Michael J. Elowson, John Mark Kreikebaum, Ed K. Wong, Sinéad M. Griffin, Saleem Rao, Alexander Weber-Bargioni, Andrew M. Minor, David I. Santiago, Stefano Cabrini, Irfan Siddiqi, and D. Frank Ogletree. Localization and mitigation of loss in niobium superconducting circuits. *PRX Quantum*, 3(22):020312, April 2022. ISSN 2691-3399. doi: 10.1103/PRXQuantum.3.020312.

[9] Ze-Tong Zhuang, Dario Rosenstock, Bao-Jie Liu, Aaron Somoroff, Vladimir E. Manucharyan, and Chen Wang. Non-markovian relaxation spectroscopy of fluxonium qubits. doi: 10.48550/arXiv.2503.16381. URL `http://arxiv.org/abs/2503.16381`. arXiv:2503.16381.

[10] Martin Spiecker, Patrick Paluch, Nicolas Gosling, Niv Drucker, Shlomi Matityahu, Daria Gusenkova, Simon Günzler, Dennis Rieger, Ivan Takmakov, Francesco Valenti, Patrick Winkel, Richard Gebauer, Oliver Sander, Gianluigi Catelani, Alexander Shnirman, Alexey V. Ustinov, Wolfgang Wernsdorfer, Yonatan Cohen, and Ioan M. Pop. Two-level system hyperpolarization using a quantum szilard engine. *Nature Physics*, 19(9):1320–1325, September 2023. ISSN 1745-2473, 1745-2481. doi: 10.1038/s41567-023-02082-8.

[11] Jonathan J. Burnett, Andreas Bengtsson, Marco Scigliuzzo, David Niepce, Marina Kudra, Per Delsing, and Jonas Bylander. Decoherence benchmarking of superconducting qubits. *npj Quantum Information*, 5(1):1–8, June 2019. ISSN 2056-6387. doi: 10.1038/s41534-019-0168-5.

[12] Janka Biznárová, Amr Osman, Emil Rehnman, Lert Chayanun, Christian Križan, Per Malmberg, Marcus Rommel, Christopher Warren, Per Delsing, August Yurgens, Jonas Bylander, and Anita Fadavi Roudsari. Mitigation of interfacial dielectric loss in aluminum-on-silicon superconducting qubits. *npj Quantum Information*, 10(1):78, August 2024. ISSN 2056-6387. doi: 10.1038/s41534-024-00868-z.

[13] J. Van Damme, S. Massar, R. Acharya, Ts Ivanov, D. Perez Lozano, Y. Canvel, M. Demarets, D. Vangoidsenhoven, Y. Hermans, J. G. Lai, A. M. Vadiraj, M. Mongillo, D. Wan, J. De Boeck, A. Potočnik, and K. De Greve. Advanced cmos manufacturing of superconducting qubits on 300 mm wafers. *Nature*, 634(8032):74–79, October 2024. ISSN 1476-4687. doi: 10.1038/s41586-024-07941-9.

[14] D. Rieger, S. Günzler, M. Spiecker, A. Nambisan, W. Wernsdorfer, and I.M. Pop. Fano interference in microwave resonator measurements. *Physical Review Applied*, 20(1):014059, July 2023. ISSN 2331-7019. doi: 10.1103/PhysRevApplied.20.014059.

[15] Jay Gambetta, Alexandre Blais, D. I. Schuster, A. Wallraff, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Qubit-photon interactions in a cavity: Measurement-induced dephasing and number splitting. *Phys. Rev. A*, 74:042318, Oct 2006. doi: 10.1103/PhysRevA.74.042318. URL `https://link.aps.org/doi/10.1103/PhysRevA.74.042318`.

[16] Nicolas Zapata, Ivan Takmakov, Simon Günzler, Simon Geisert, Soeren Ihssen, Mitchell Field, Ameya Nambisan, Dennis Rieger, Thomas Reisinger, Wolfgang Wernsdorfer, and Ioan M. Pop. Granular aluminum parametric amplifier for low-noise measurements in tesla fields. *Physical Review Letters*, 133(26), December 2024. ISSN 1079-7114. doi: 10.1103/physrevlett.133.260604. URL `http://dx.doi.org/10.1103/PhysRevLett.133.260604`.

[17] Gaurav Bothara, Srijita Das, Kishor V. Salunkhe, Madhavi Chand, Jay Deshmukh, Meghan P. Patankar, and R. Vijay. High-fidelity qnd readout and measurement back-action in a tantalum-based high-coherence fluxonium qubit. *APL Quantum*, 2(2):026103, April 2025. ISSN 2835-0103. doi: 10.1063/5.0255892.

[18] R. Dassonneville, T. Ramos, V. Milchakov, L. Planat, É. Dumur, F. Foroughi, J. Puertas, S. Leger, K. Bharadwaj, J. Delaforce, C. Naud, W. Hasch-Guichard, J.J. García-Ripoll, N. Roch, and O. Buisson. Fast high-fidelity quantum nondemolition qubit readout via a nonperturbative cross-kerr coupling. *Physical Review X*, 10(1):011045, February 2020. ISSN 2160-3308. doi: 10.1103/PhysRevX.10.011045.

[19] Bryan T. Gard, Zachary Parrott, Kurt Jacobs, José Aumentado, and Raymond W. Simmonds. Fast high-fidelity quantum nondemolition readout of a superconducting qubit with tunable transverse couplings. *Physical Review Applied*, 21(2):024008, February 2024. ISSN 2331-7019. doi: 10.1103/PhysRevApplied.21.024008.

[20] Y. Sunada, S. Kono, J. Ilves, S. Tamate, T. Sugiyama, Y. Tabuchi, and Y. Nakamura. Fast readout and reset of a superconducting qubit coupled to a resonator with an intrinsic purcell filter. *Physical Review Applied*, 17(4):044016, April 2022. ISSN 2331-7019. doi: 10.1103/PhysRevApplied.17.044016.

[21] S. Hazra, W. Dai, T. Connolly, P.D. Kurilovich, Z. Wang, L. Frunzio, and M.H. Devoret. Benchmarking the readout of a superconducting qubit for repeated measurements. *Physical Review Letters*, 134(10):100601, March 2025. doi: 10.1103/PhysRevLett.134.100601.

[22] Deep Lall, Abhishek Agarwal, Weixi Zhang, Lachlan Lindoy, Tobias Lindström, Stephanie Webster, Simon Hall, Nicholas Chancellor, Petros Wallden, Raul Garcia-Patron, Elham Kashefi, Viv Kendon, Jonathan Pritchard, Alessandro Rossi, Animesh Datta, Theodoros Kapourniotis, Konstantinos Georgopoulos, and Ivan Rungger. A review and collection of metrics and benchmarks for quantum computers: definitions, methodologies and software. doi: 10.48550/arXiv.2502.06717. URL `http://arxiv.org/abs/2502.06717`. arXiv:2502.06717 [quant-ph].

[23] W. Dai, S. Hazra, D. K. Weiss, P. D. Kurilovich, T. Connolly, H. K. Babla, S. Singh, V. R. Joshi, A. Z. Ding, P. D. Parakh, J. Venkatraman, X. Xiao, L. Frunzio, and M. H. Devoret. Spectroscopy of drive-induced unwanted state transitions in superconducting circuits. doi: 10.48550/arXiv.2506.24070. URL `http://arxiv.org/abs/2506.24070`. arXiv:2506.24070 [quant-ph].

[24] Marie Frédérique Dumas, Benjamin Groleau-Paré, Alexander McDonald, Manuel H. Muñoz-Arias, Cristóbal Lledó, Benjamin D'Anjou, and Alexandre Blais. Measurement-induced transmon ionization. *Physical Review X*, 14(4):041023, October 2024. ISSN 2160-3308. doi: 10.1103/PhysRevX.14.041023.

[25] Mostafa Khezri, Alex Opremcak, Zijun Chen, Kevin C. Miao, Matt McEwen, Andreas Bengtsson, Theodore White, Ofer Naaman, Daniel Sank, Alexander N. Korotkov, Yu Chen, and Vadim Smelyanskiy. Measurement-induced state transitions in a superconducting qubit: Within the rotating-wave approximation. *Physical Review Applied*, 20(5):054008, November 2023. ISSN 2331-7019. doi: 10.1103/PhysRevApplied.20.054008.

[26] Daniel Sank, Zijun Chen, Mostafa Khezri, J. Kelly, R. Barends, B. Campbell, Y. Chen, B. Chiaro, A. Dunsworth, A. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Mutus, M. Neeley, C. Neill, P.J.J. O'Malley, C. Quintana, P. Roushan, A. Vainsencher, T. White, J. Wenner, Alexander N. Korotkov, and John M. Martinis. Measurement-induced state transitions in a superconducting qubit: Beyond the rotating wave approximation. *Physical Review Letters*, 117(19):190503, November 2016. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.117.190503.

[27] Konstantin N. Nesterov and Ivan V. Pechenezhskiy. Measurement-induced state transitions in dispersive qubit-readout schemes. *Physical Review Applied*, 22(6):064038, December 2024. ISSN 2331-7019. doi: 10.1103/PhysRevApplied.22.064038.

[28] Martin Spiecker, Andrei I. Pavlov, Alexander Shnirman, and Ioan M. Pop. Solomon equations for qubit and two-level systems: Insights into non-poissonian quantum jumps. *Physical Review A*, 109 (5):052218, May 2024. ISSN 2469-9926, 2469-9934. doi: 10.1103/PhysRevA.109.052218.

MetSuperQ measurement reporting template and checklist

This checklist follows the “MetSuperQ superconducting qubit tuneup and benchmarking protocol” version: 24 November 2025

Different numbers obtained throughout the procedure that require use in a later stage are denoted a unique number (in a yellow box) which is used to reference this value later.

Part 1 – Sample information & installation

Sample identifier		
Qubit identifier		
Qubit type		
Fridge used		
Fridge base temperature		
Control electronics used		
Parametric amplifier model used		
Complete drawing of setup exists, including all components, filters, attenuators, etc, inside the fridge		
Sample installed by:		
Installation date:		
Cooldown start date:		
Sample measured by:		Date(s):
This document was filled in by:		Date(s):

Section 2 - Tuneup

2.2.1 – Attenuation check

Input line number for qubit readout	
Input line fixed attenuators (dB)	
Estimated input line cable loss (dB)	

Sample insertion loss (dB)	
Total loss expected (dB)	
Output line number for qubit readout	
HEMT gain (dB)	
RT amplifier gain (dB)	
Total Gain expected (dB)	
Expected VNA transmission (total loss – total gain) (dB)	
Measured VNA transmission near resonator (dB)	
Attenuation level within acceptable bounds to proceed?	

2.2.2 – Flux sweep

Found resonator frequency (GHz)		V1
Qubit is flux tunable? (yes/no) if no – skip the rest.		
Current range that corresponds to 1 flux quantum (mA)		V2
Current that biases the qubit at it's half flux quanta sweet-spot (mA, at least 3 significant digits)		V3
Measurement completed: Resonator vs flux bias		
Resulting 2D-plot:		
[paste figure here]		

2.2.3 – Resonator spectroscopy

Verify flux bias is set to the qubit sweet-spot [V3]		
Fitted resonance frequency (GHz, at least 5 significant digits)		V4
Fitted loaded Ql (at least 3 significant digits)		V5
Fitted coupling Qc (at least 3 significant digits)		V6
Calculated internal Qi (at least 3 significant digits)		V7
Calculated bandwidth kappa/2pi = f_r/Ql (MHz, at least 3 significant digits)		V8
Power sweep completed, punchout visible		

2.2.4 – Two-tone

Two-tone measurement completed, qubit found		
Qubit 0->1 frequency (GHz, at least 5 significant digits)		V9
Qubit 1->2 frequency (GHz, at least 5 significant digits)		V10

Calculated anharmonicity (0->1) – (1->2) (MHz, at least 3 significant digits)	

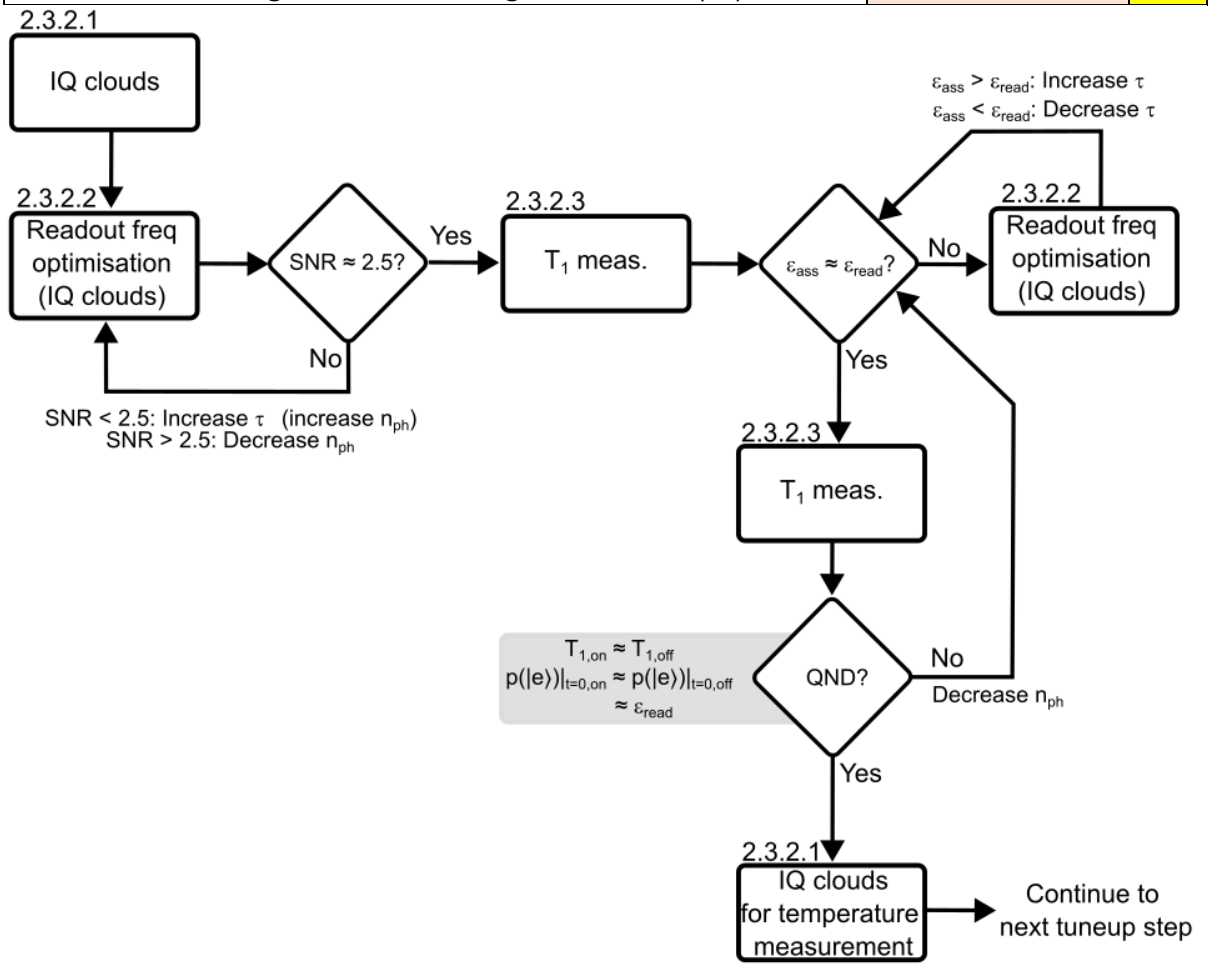
2.2.5 – Parametric amplifier tuning

Chosen operating point	
Pump frequency (GHz, at least 7 significant digits)	
Pump power on source (dBm, at least 3 significant digits)	
DC current bias (mA, at least 4 significant digits) (if applicable)	
Measured gain profile for the chosen operating point:	
[paste figure here]	
Extracted gain at the readout frequency (dB)	
Measured ΔSNR using a spectrum analyser	
Noise floor confirmed limited by amplifier chain	
[paste figure here]	
Extracted optimal ΔSNR at the readout frequency (linear units)	
Check:	
Re-type the resonator frequency from 2.2.3 (GHz) [V4]	
Re-type the qubit frequency from 2.2.4 (GHz) [V9]	
Confirm that the pump tone is well separated from the qubit and resonator frequencies (and multiples/combinations thereof)	
Confirm signal power used was below PA saturation power	
Confirm PA works and performance at the operating frequency meets requirements to proceed	

2.3.1 – Signal delay

Mixer calibration completed near the resonator frequency (if applicable)		
Delay measurement completed		
Measurement setting: Applied pulse duration (ns)		
Measured delay until signal onset (ns)		
Measured time to reach steady state (ns)		
Suitable integration window start time (ns)		V11

Recommended integration window length for readout (ns)		V12



2.3.2 Readout frequency optimisation

Initial settings check.

Confirm optimal parametric amplifier settings are applied	
Verified to implement the readout delay and within range of the recommended integration window in [2.3.1 – Signal delay] [V11, V12]	
Confirm qubit at right operating frequency [V3]	

Run through the steps 2.3.2.1-3 following the flowchart. When an optimal is found, note the final outcomes below.

Final readout settings

Obtained SNR at optimal readout frequency (at least 3 significant digits)	
Obtained $\varepsilon_{assignment}$	
Obtained $\varepsilon_{readout}$	
Obtained $T_{1,on}/T_{1,off}$ ratio	
Readout pulse settings:	V13
Readout pulse frequency (GHz, at least 6 significant digits)	
Readout pulse amplitude (au)	
Readout pulse length (τ, ns, at least 3 significant digits)	
Note on readout pulse shape, rise/fall times, etc.	

2.3.2.1 – IQ clouds

Measurement completed after optimisation completed	
Fitted ground and excited state populations	
|0>	
|1>	
Effective qubit temperature (Kelvin)	
Obtained SNR from fit	
(optional) quantum jump trace dataset taken	

2.3.2.2 – Readout frequency optimisation

Measurement completed after optimisation completed	
Resulting figure of state population and SNR vs readout frequency	
[paste figure here]	

Calculated dispersive shift $\chi/2\pi = \Delta f$ (MHz, at least 4 significant digits)		V14
Repeated measurement vs readout pulse amplitude completed		
Plot of χ vs readout amplitude		
[paste figure here]		

2.3.2.3 Measurement preparation

Resulting figure containing both T_1 measurements		
[Paste figure here]		
Each measurement repeated at least 5 times?		
Measured T1 with readout off, note average fitted T_{1off} value (us, at least 3 significant digits)		V15
Measured T1 with readout on during delay, note average fitted T_{1on} value (us, at least 3 significant digits)		

2.3.6 Two-tone (time domain)

Confirm using optimal readout parameters [V13]		
Updated value for the current applied to reach sweet-spot (mA, at least 4 significant digits)		V16
Qubit frequency at half-flux bias (GHz, at least 5 significant digits)		V17
Parameters from fitted spectrum (if applicable)		
EJ (GHz, at least 3 significant digits)		
EC (GHz, at least 3 significant digits)		
EL (GHz, at least 3 significant digits)		
Resulting 2D map		
[paste figure here]		

2.3.7 Power Rabi

Verify that optimal readout settings are used [V13]		
Verify at sweet-spot [V16]		
Verify pi-pulse frequency is at qubit frequency [V17]		
Note fixed duration of pi-pulse (ns, at least 3 significant digits)		
Measurement sweeping pulse amplitude completed		
Fitted amplitude for pi-pulse (a.u., at least 2 significant digits)		V18
Fitted amplitude for pi/2 pulse (a.u., at least 2 significant digits)		V19
Resulting figure with Rabi oscillations		

[paste figure here]

2.3.8 Pi-fringes

Verify readout settings [V13], qubit frequency [V16, V17]		
Note again measured T_1 value from previously (us) [V15]		
Note repetition rate for application of repeated pi-pulses (i.e. including delay between pulses) (ns)		
Optimised pi-pulse amplitude (a.u., at least 3 significant digits)		V20
Confirm fringes are free of beatings	Yes/No	
Fitted decay time of the pi-fringes (us)		
Confirm the decay time for the fringes compares with the T_1 value	Yes/No	
Resulting figure with fringe decay		
[paste figure here]		

2.3.9 Pi-half fringes

Verify readout settings [V13], qubit frequency [V16, V17]		
Optimised pi-half pulse amplitude (a.u., at least 3 significant digits)		V21
Confirm fringes are free of beatings	Yes/No	
Confirm the mixed state readings are free of drift	Yes/No	
Confirm the fringe coherence and symmetry is maximised	Yes/No	
Resulting figure with pi-half fringes		

[paste figure here]

2.3.10 AC-Stark calibration

Note method of calibration carried out:	Ramsey/Two-tone
Note again from above fitted χ_{ge} (MHz) [V14]	
Note again from above fitted $\kappa/2\pi$ (kHz) [V8]	
Frequency used for CW tone at readout frequency (GHz, at least 6 significant digits)	
Confirm same frequency as above optimised readout pulse frequency	Yes/No
Resulting figure of qubit frequency shift vs CW amplitude:	
[Paste figure here]	
Resulting figure of readout amplitude vs resonator photon number	
[Paste figure here]	
Note the optimised readout amplitude from previous step (a.u.) [V13]	
Calculated number of photons in readout pulse (at least 2 significant digits)	

Final tune-up notes

Confirmed all steps in protocol carried out	
Date completed (dd/mm/yyyy)	
Any additional notes or comments (on e.g. deviations from protocol)	

Section 3 – QoI's

3.2 QNDness: Check. [ ] Readout settings [V13], [ ] PA operating at optimal point.	
3.2 QNDness: Measurement completed (date: dd/mm/yyyy)	
3.2 QNDness: Obtained Q (at least 3 significant digits)	
3.2 QNDness: Estimated Q from T_1 relaxation (at least 3 significant digits)	
3.3 RLIB: Check. [ ] Readout settings [V13], [ ] PA operating at optimal point. [ ] Qubit at sweet-spot [V16]. [ ] Optimised pi-pulse settings used [V17, V20].	
3.3 RLIB: Measurement completed (date: dd/mm/yyyy)	
3.4 RB: Check. [ ] Readout settings [V13], [ ] PA operating at optimal point. [ ] Qubit at sweet-spot [V16]. [ ] Optimised pi-pulse settings used [V17, V20].	
3.4 RB: Measurement vs depth completed (date: dd/mm/yyyy)	
3.4 RB: Chosen depth to carry out temporal measurement at (where survival probability ~ 75%)	
3.4 RB: Measurement vs time at fixed depth completed (date: dd/mm/yyyy)	
3.5 Quantum state preparation: Check. [ ] Readout settings [V13], [ ] PA operating at optimal point. [ ] Qubit at sweet-spot [V16]. [ ] Optimised pi-pulse settings used [V17, V20].	

3.5 Quantum state preparation: Measurement completed over at least 1 hour (date: dd/mm/yyyy)	
3.6 Energy relaxation: Check. [ ] Readout settings [V13], [ ] PA operating at optimal point. [ ] Qubit at sweet-spot [V16]. [ ] Optimised pi-pulse settings used [V17, V20].	
3.6 Energy relaxation: active measurement completed several times, obtained T_1 (us, at least 3 significant digits)	
3.6 Energy relaxation: passive measurement completed several times, obtained T_1 (us, at least 3 significant digits)	
3.7 Interleaved: Check. [ ] Readout settings [V13], [ ] PA operating at optimal point. [ ] Qubit at sweet-spot. [ ] Optimised pi-pulse settings used [V17, V20]. [ ] Optimised pi/2-pulse settings used [V17, V21].	
3.7: Interleaved: Measurement completed (date: dd/mm/yyyy)	
3.7: Interleaved: Note cycle repetition rate (seconds)	
3.7: Interleaved: Note total time of measurement (hours)	
3.8 Szilard: Check. [ ] Readout settings [V13], [ ] PA operating at optimal point. [ ] Qubit at sweet-spot. [ ] Optimised pi-pulse settings used [V17, V20]. [ ] Optimised pi/2-pulse settings used [V17, V21].	
3.8 Szilard: Measurement completed (date: dd/mm/yyyy)	

Final QoI notes

Confirmed all QoI's measured and data properly curated	
Date completed (dd/mm/yyyy)	
Any additional notes or comments (on e.g. deviations from protocol)	