

Adaptive Reconstruction for Closed-Loop Optimisation of Bosonic Quantum States



Vasilisa Usova

The cover image is generated with OpenAI's ChatGPT

Adaptive Reconstruction for Closed-Loop Optimisation of Bosonic Quantum States

Vasilisa Vladimirovna Usova, 11940585

Innsbruck, 2025

Dissertation

submitted to the Faculty of Mathematics, Computer Science and Physics of the
University of Innsbruck in partial fulfilment of the requirements for the academic
degree

Doctor of Philosophy

Doctor of Philosophy – Doctoral Program in Physics

Supervisor: Univ. Prof. Dr. Gerhard Kirchmair

Institute of Experimental Physics

Faculty of Mathematics, Computer Science and Physics

Abstract

Continuous-variable qubits offer a promising route toward fault-tolerant quantum computation. However, the high dimensionality of their underlying Hilbert space poses major challenges for efficient state tomography and closed-loop optimisation. Typically, state reconstruction is performed via Wigner [1] or (generalised) Husimi-Q [2] measurements sampled on a regular grid in phase space. More sophisticated approaches employ importance sampling [3] or excitation-number sampling schemes [4]. Nonetheless, the closed-loop optimisation demands reproducible state reconstruction that is robust under phase-space translations and rotations, while keeping the reconstruction time within a few minutes. None of the existing reconstruction techniques can be directly adapted to this setting.

This thesis introduces a novel, efficient reconstruction algorithm that adaptively selects new measurement points in regions of highest model uncertainty and automatically terminates once the target precision is reached. The method requires no preliminary simulations or calibration runs, and it is not restricted to a specific family of bosonic states. Benchmarking on cat states of various sizes demonstrates robust, high-fidelity reconstruction, achieving infidelities on the order of 0.01 across 100 experimental runs within a few minutes. The algorithm is sensitive not only to overall infidelity but also to subtle state imperfections, making it ideally suited for closed-loop applications.

By exploiting the reconstructed state fidelity as an optimisation metric, the direct search Nelder-Mead algorithm enabled closed-loop optimisation of a bosonic (cat) state by applying correction pulses during the generation protocol. These results establish a foundation for future autonomous calibration and control techniques in continuous-variable quantum systems.

Contents

Abstract	v
1 Introduction	9
2 Quantum optimal and adaptive control	11
2.1 Open-loop optimisation techniques	13
2.1.1 Gradient ascent pulse engineering	13
2.1.2 Automatic differentiation	16
2.1.3 Gradient-based optimisation algorithms	18
2.2 Closed-loop optimisation techniques	19
2.2.1 Nelder-Mead algorithm	20
2.2.2 Evolution strategies and CMA-ES algorithm	21
2.2.3 Chopped RAndom Basis algorithm	23
2.2.4 Measurement-based feedback	24
2.2.5 Reinforcement learning	25
2.3 Outlook	27
3 Experimental platform: brief introduction	33
3.1 Harmonic oscillator and the need for nonlinearity	33
3.2 Introducing nonlinearity. Black box quantisation	34
3.2.1 Quantum LC oscillator	35
3.2.2 Transmon	38
3.2.3 Black box quantisation	40
3.3 Experimental characterisation of the Hamiltonian	43
3.4 The necessary set of operations	48
3.5 Tomography and fidelity of bosonic states	49
3.5.1 Wigner and Wigner-based fidelity estimations	52
3.5.2 Husimi Q	53
4 Quantum optimal control of cat states	57
4.1 Bosonic states: challenges and progress outlook	57
4.2 Cat states: creation protocols	58
4.3 Cat states: imperfections	61
4.4 Kerr compensation techniques	65
4.4.1 Geometric phase gates	65
4.4.2 Adiabatic off-resonant drive	67
4.5 Closing the loop: first tries	69
4.5.1 First sampling strategy	71
4.5.2 Second sampling strategy	72
5 Efficient state reconstruction	77
5.1 Model	78
5.2 Bayesian model	81

5.3	Sampling from the distribution	83
5.4	Bootstrapping and exit conditions	84
5.5	Sampling based on the deviation between the bootstrapped models	85
5.6	Pipeline	89
5.7	Experimental runs	91
5.7.1	Parameter configuration: priors and bounds	91
5.7.2	Results & discussion	93
5.8	Outlook: adjusting algorithm for other applications	103
6	Closed-loop optimisation of cat states	105
6.1	Reconstruction parameters and stopping criteria	105
6.2	Closed-loop optimisation configuration	108
6.3	Closed-loop optimisation runs	109
7	Conclusion	113
	Bibliography	117
	Appendix	
	Appendix A Experimental setup and readout	127
A.1	Setup	127
A.2	High-power readout	127
	Appendix B Notes on priors and inverse Wishart	131
	Appendix C Nelder-Mead hyperparameter search	135
	Appendix D Reconstruction convergence for smaller cats	137
	Appendix E GrAI delay line	143
E.1	Motivation	143
E.2	Optical lithography	143
E.3	Electron beam lithography	147
E.4	Transmission and reflection measurements	153
E.5	Kerr measurements	154
E.6	Dispersion relation and kinetic inductance	156
	Acknowledgements	160

Introduction

Quantum information processing has witnessed rapid development across multiple physical platforms, including trapped ions [5], neutral atoms [6], photonic systems [7], and superconducting circuits [8]. Each of these technologies has its own unique set of benefits and challenges in terms of coherence times, controllability, and scalability, while sharing the common goal of realising fault-tolerant quantum computation. Among them, superconducting circuits have become one of the leading competitors due to a combination of reasonably long qubit lifetimes, fast high-fidelity readout and gate operations, and compatibility with modern microfabrication techniques.

The circuit quantum electrodynamics (cQED) platform, traditionally implemented using superconducting circuits, is currently experiencing a burst of milestone experiments and theoretical proposals aimed at implementing error-corrected logical qubits for fault-tolerant quantum computation [9–12]. Quantum error correction relies on encoding information in logical quantum bits (qubits) — quantum systems with redundant degrees of freedom. In cQED, this redundancy can be realised either through a set of discrete-variable qubits, typically approximated as two-level systems, or through multiphoton states in bosonic modes that naturally inhabit large Hilbert spaces. The discrete variable approach was historically the first to enable significant milestones in the field [9, 13]. However, the continuous-variable approach, despite emerging later, has achieved comparable results within a shorter timeframe. In recent years, several experimental demonstrations have successfully realised logical qubits operating beyond the break-even point across different encoding bases, such as GKP [11] and binomial [10].

Scaling up systems of logical qubits requires an increasing number of high-coherence auxiliary qubits and couplers, sophisticated crosstalk mitigation strategies, and precise hardware control over a wide range of DC and microwave signals. At such levels of complexity, analytical models and numerical simulations become insufficient to fully capture the system dynamics, emphasising the need for experimental fine-tuning of control parameters in a so-called closed-loop setting. For continuous-variable qubits, this challenge arises even earlier. Generating a logical state in a bosonic mode requires coupling a nonlinear element to an infinite-dimensional harmonic oscillator, which cannot be simulated precisely without truncating the Hilbert space and neglecting higher-order nonlinearities.

This thesis focuses on the application of closed-loop control techniques to the preparation and characterisation of cat states. Cat states represent one of the most prominent and widely used logical bases for error correction within the cQED community. They can be generated using deterministic protocols that do not require preliminary waveform or

gate sequence simulations [1], making them both experimentally attractive and practically relevant. While simulation-based approaches have become standard for the creation and manipulation of continuous-variable states, the application area of closed-loop methods is strictly limited due to the complexity of tomography in a high-dimensional Hilbert space.

The thesis begins with a literature review of various control and adaptive techniques in Chapter 2, followed by a brief introduction to the experimental platform in Chapter 3. The foundations of cQED have been comprehensively covered across numerous theses and review papers. Including the full background on both optimal control and cQED here would considerably expand the content while reducing clarity. Therefore, it is assumed that the reader is already familiar with the basic concepts of quantum computation and cQED; if not, the relevant introductory material can be found in theses [14, 15] and review [16]. Since optimal control is not a standard topic in cQED theses, Chapter 2 provides a concise and accessible overview of the subject without requiring prior knowledge. Further theoretical background on quantum optimal control can be found in [17], whereas the experimental implementation of open-loop pulses in superconducting circuits is discussed in [18].

Chapter 4 presents the creation protocols for cat states, discusses sources of imperfections, and corresponding mitigation strategies. Chapter 5 introduces an efficient state reconstruction protocol that addresses the challenge of tomography in high-dimensional Hilbert spaces, enabling high-fidelity state reconstruction within a few minutes. Together with the supplementary data in Appendix D, it provides comprehensive benchmarking results for various states and hyperparameter regimes. Finally, Chapter 6 demonstrates the first successful closed-loop optimisation of a bosonic (cat) state. The experimental results from Chapters 5 and 6 establish a solid foundation for future closed-loop optimisation and automated tune-up of system and gate parameters in the continuous variable domain.

Quantum optimal and adaptive control

The first proof-of-principle superconducting circuit experiments demonstrated quantum oscillations between the first two states of the Cooper pair box [19], confirming that coherent quantum control could be achieved in macroscopic electrical systems. This pioneering result established superconducting circuits as a viable platform for quantum computation, enabling precise manipulation of individual qubits using microwave drives. However, these early devices faced severe limitations: short coherence times and imperfect control pulses prevented the implementation of high-fidelity quantum gates. Progress in device design and fabrication, along with improved crosstalk mitigation and advanced pulse-shaping techniques, gradually overcame these limitations. As a result, coherence times now exceed $10^2 \mu\text{s}$, and two-qubit gate fidelities exceeding 0.99 were achieved for gate durations in a range of 10 – 100 ns [20].

Despite these remarkable advances, the next milestone - the transition to logical qubit operations - posed new challenges. Realising a single logical qubit requires either hundreds of discrete-variable qubits [21] or coupling an infinite-dimensional bosonic mode to a nonlinear element [22]. In both cases, eliminating parasitic effects arising from unwanted interactions and experimental imperfections becomes increasingly complex. Techniques that proved effective for calibrations and corrections in few-qubit systems scale poorly with the growing dimensionality of the underlying Hilbert space. Consequently, optimal and adaptive control methods have become indispensable for extending high-fidelity control to larger systems and bringing fault-tolerant quantum computation within reach.

Nowadays, the family of different quantum optimisation algorithms has become enormous. This Chapter covers only the most relevant part of the algorithms used on a superconducting circuits platform. More information about the classification, general idea, and implementation of various optimisation approaches for quantum applications can be found in reviews [23, 24]. The large variety of quantum optimal control (QOC) techniques can be roughly divided into two prominent families of algorithms: open-loop and closed-loop (see Fig. 2.1). In an open-loop, pulse shapes are optimised in a simulation using experimentally characterised Hamiltonian parameters. Closed-loop, on the other hand, relies directly on experimentally measured data to enhance performance. As a general guideline, gradient-based methods, which update model parameters using the gradient of the objective function, are typically employed in open-loop optimisation, whereas gradient-free methods, which do not require gradient calculations, are used in a closed-loop. This rule is very rough and contains a few exceptions when gradient-based methods are applied in a closed-loop as well [11, 25]. For example, reinforcement learning (Section 2.2.5) inherently involves an interaction with a quantum system, which pushes it towards closed-loop

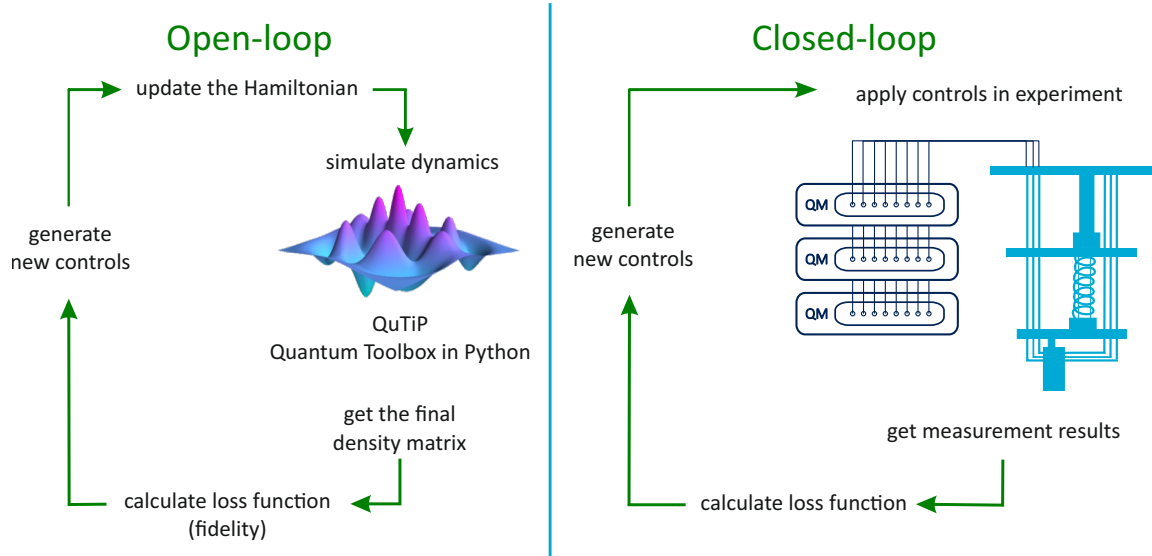


Figure 2.1: Open-loop and closed-loop optimisation organisation workflow. In an open-loop, the optimiser uses the simulated density matrix to calculate a loss function and perform an update. Typically, simulations are done using the **QuTiP** library [26]. In a closed-loop, the optimiser uses measurement outcomes to calculate the loss function and perform an update. QM (Quantum Machines) represents the system control electronics.

techniques. Gradient ascent pulse engineering (Section 2.1.1) has a successful closed-loop implementation, despite the majority of applications referring to an open-loop approach (Section 2.3). The opposite direction, i.e. open-loop applications of gradient-free algorithms, is usually pointless. Gradient-based algorithms are much faster and more reliable in terms of convergence, provided that they have access to the noiseless metric and its derivatives in a simulation.

Although gradient-based methods converge faster and are capable of finding more complex solutions, their closed-loop applications should be designed with great care. During real-time experimental optimisation, the most significant amount of time is spent on measurement, hardware configuration and communication. Therefore, the closed-loop optimisation algorithms should take as few measurements as possible. Gradient-based algorithms, with some exceptions, do not satisfy this criterion because they require several measurements to calculate the gradients numerically.

This Chapter provides a brief overview of the main QOC families used in superconducting quantum circuits. The Chapter first discusses open-loop techniques, then proceeds with closed-loop techniques, and concludes with a general overview of the QOC methods that were successfully experimentally implemented on discrete or continuous variable qubits. In case the method has both open- and closed-loop applications, it is attributed to a more relevant subsection.

2.1 Open-loop optimisation techniques

The first attempt to improve control pulses in superconducting circuits dates back to 2009 when F. Motzoi et al. [27] presented their idea of Derivative Removal (of the leakage to higher transmon levels) by Adiabatic Gate (DRAG). The approach is analytical and implies that the correction can be done by modifying the second $Q(t)$ quadrature (see Section 4.3, Eq. 4.14). The first order quadrature correction $Q(t) = -\dot{I}(t)/\Delta$, $\Delta = \omega_2 - 2\omega_1$ (here $\omega_{1,2}$ are 0 – 1 and 1 – 2 transition frequencies) reduces the leakage out of the computational transmon (see Section 3.2.2) subspace. This leads to a significant improvement: the gate error drops by several orders of magnitude. Although the result was impressive and had an immediate successful implementation [28], the authors mentioned that an open-loop optimisation with the gradient ascent pulse engineering (see Section 2.1.1) outperforms the analytically corrected pulse [27].

Analytical corrections to pulse shapes had further applications, including quantum control of multi-level systems (for example, fluxonium [29]), crosstalk [30], and photon-number-dependent pulses applied to a bosonic mode for Kerr cancellation [31]. Recently, E. Hyppä et al. [32] suggested improving DRAG pulses by adding spectral shaping, which allowed demonstration of a 7.9 ns qubit gate with a low gate error $(1.56 \pm 0.07) \times 10^{-4}$. B. Chiaro et al. [33], alternatively, suggested introducing an additional drive near the unwanted 1 – 2 transition to suppress leakage out of the computational subspace. Such an active leakage cancellation scheme allows reducing the infidelity to 7.5×10^{-5} for a 10 ns long qubit gate.

Even though the analytical corrections provide a significant fidelity improvement, their application area is strictly limited: they are effective on reasonably small systems. For the Hamiltonians of more complicated, higher-dimensional systems, it is difficult to find a good analytical approximation for pulse correction. With the improvements in Hamiltonian characterisation techniques, it became possible to use numerical simulations of the system based on the measured Hamiltonian parameters. Simulation-based optimisations of the control pulses allowed the cQED community to make another crucial step towards fidelity improvements. Recent achievements include the demonstration of a high-fidelity universal gate set on bosonic states [34], which has become a fundamental tool for several experiments, including quantum error correction beyond break-even [11].

Despite requiring accurate measurements of Hamiltonian parameters, open-loop optimisation techniques have a significant advantage: they can explore a huge parameter space by using gradient-based algorithms that the optimal control community has been improving for decades. Open-loop techniques remain the most widely used across the superconducting circuits community.

2.1.1 Gradient ascent pulse engineering

The Gradient Ascent Pulse Engineering (GRAPE) algorithm rapidly became one of the most widely used optimisation techniques, as it drastically reduced the number of required

simulations or experimental runs compared to previous gradient-based methods. The idea was first introduced by Khaneja et al [35] for the optimisation of NMR pulses.

In the optimal control setting, the closed quantum system evolves according to the Schrodinger equation under the Hamiltonian

$$\hat{H} = \hat{H}_0 + \sum_{k=1}^m u_k(t) \hat{H}_k, \quad (2.1)$$

where $\hat{H}_0$ is the system Hamiltonian, and $\hat{H}_k$, $k \in [1, m]$, are drive Hamiltonians with the corresponding time-dependent drive amplitudes $u_k(t)$ of length T . The drive can be decomposed into various kinds of basis functions. The usual choice for the GRAPE algorithm is a piecewise-constant basis, which assumes that the drive takes a constant value within each discretised time step δt :

$$u_k(t) = u_k^j, \quad t \in [j\delta t, (j+1)\delta t], \quad j \in [0, N-1] \quad (2.2)$$

where N is the number of discretisation steps, and the time step is $\delta t = T/N$ (see Fig. 2.2). The time evolution during time step j is defined by the unitary propagator

$$\hat{U}_j = \exp \left[-\frac{i}{\hbar} \delta t \left(\hat{H}_0 + \sum_{k=1}^m u_k^j \hat{H}_k \right) \right], \quad (2.3)$$

and the state of the system at time $t = (j+1)\delta t$ is

$$\psi(t) = \hat{U} |\psi_{\text{initial}}\rangle = \hat{U}_j \hat{U}_{j-1} \dots \hat{U}_0 |\psi_{\text{initial}}\rangle = \mathcal{T} \prod_j \hat{U}_j |\psi_{\text{initial}}\rangle \quad (2.4)$$

with $\mathcal{T}$ denoted as a time-ordering operator.

Since this thesis aims to optimise a bosonic state, this subsection focuses on a state transfer example. A more general description of cost functions for the transfer between Hermitian and non-Hermitian operators for mixed states can be found in the original GRAPE paper [35].

For the quantum state transfer problem, it is straightforward to use the cost function¹ J in the form [36]

$$J = 1 - |\langle \psi_{\text{target}} | \psi(T) \rangle|^2 = 1 - |\langle \psi_{\text{target}} | \mathcal{T} \prod_j \hat{U}_j | \psi_{\text{initial}} \rangle|^2 \quad (2.5)$$

Changing the control pulse amplitude by $\delta u_k = -\epsilon \frac{\partial J}{\partial u_k}$, where $\epsilon > 0$ is a small number, is equivalent to the minimisation of the cost function

$$J(u_k(t) + \delta u_k) = J(u_k(t)) + \frac{\partial J}{\partial u_k} \delta u_k = J(u_k(t)) - \epsilon \left(\frac{\partial J}{\partial u_k} \right)^2 \quad (2.6)$$

¹Cost function, loss function, and objective function are equivalent terms. Figure of merit is a more general term that characterises the performance of the method. In the context of this thesis, all the above-mentioned concepts are equivalent.

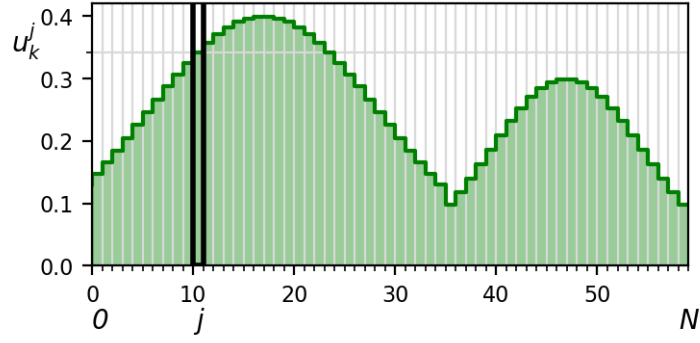


Figure 2.2: A pulse in a piecewise-constant basis. At each iteration of the GRAPE, the algorithm calculates forward- and backwards-propagated states. It allows calculating the derivative $\frac{\partial J}{\partial u_k^j}$ and, therefore, correcting pulse amplitudes for each time step $j\delta t$ by simulating only two evolutions of the system.

As shown in [36], the partial derivative with respect to u_k^j can be expressed as

$$\begin{aligned} \frac{\partial J}{\partial u_k^j} &= -2\Re \left(\langle \psi_{\text{target}} | \hat{U}_{N-1} \dots \hat{U}_{j+1} \frac{\partial \hat{U}_j}{\partial u_k^j} \hat{U}_{j-1} \dots \hat{U}_0 | \psi_{\text{initial}} \rangle \langle \psi_{\text{initial}} | \hat{U}^\dagger | \psi_{\text{target}} \rangle \right) \\ &= -2\Re \left(\langle \xi^j | \frac{\partial \hat{U}_j}{\partial u_k^j} | \phi^j \rangle \langle \psi_{\text{initial}} | \hat{U}^\dagger | \psi_{\text{target}} \rangle \right), \end{aligned} \quad (2.7)$$

where the following abbreviations were used

$$\begin{aligned} |\xi^j\rangle &= \langle \psi_{\text{target}} | \hat{U}_{N-1} \dots \hat{U}_{j+1} \\ |\phi^j\rangle &= \hat{U}_{j-1} \dots \hat{U}_0 | \psi_{\text{initial}} \rangle. \end{aligned}$$

This observation leads to the idea of the GRAPE algorithm:

1. Pick up an initial guess for control pulses $u_k(t)$
2. Repeat the following steps until the optimisation criteria are reached:
 - Calculate the forward propagated states $|\phi^j\rangle = \hat{U}_{j-1} \dots \hat{U}_1 | \psi_{\text{initial}} \rangle$
 - Calculate the backward propagated states $|\xi^j\rangle = \hat{U}_{j+1} \dots \hat{U}_N | \psi_{\text{target}} \rangle$
 - Calculate the gradient of the cost function $\partial J / \partial u_k^j$ for each u_k^j by using the previously calculated states. The control amplitudes u_k^j are re-evaluated depending on the chosen optimisation algorithm (see Section 2.1.3). For the gradient descent, this results in $u_{k,\text{new}}^j = u_k^j - \epsilon \frac{\partial J}{\partial u_k^j}$

The algorithm's structure reveals its big advantage over the widely used Krotov-like gradient-based algorithms [37], [38]. Krotov-like algorithms are sequential and update parameters one after another, which requires $m \times N + 1$ runs in total. In contrast, GRAPE requires only two simulation runs:

- in the "forward" direction for the states $|\phi^j\rangle$, $j \in [0, N - 1]$, starting from the initial state $|\psi_{\text{initial}}\rangle$
- in the "backward" direction for the states $|\xi^j\rangle$, starting from the target state $|\psi_{\text{target}}\rangle$

Such concurrent updates of control parameters have boosted the optimal control community to create new algorithms based on the GRAPE idea (see Fig. 3 in [24] for a tree of QOC methods).

2.1.2 Automatic differentiation

Typically, the cost function includes penalty terms that force the solution to stay within the allowed computational subspace or restrict the waveform shape to match the waveform generator capabilities. GRAPE can deal with most such penalties. However, there are examples of penalty terms that have particularly difficult analytical expressions for gradient calculations [39]. For example,

$$P_1 = \sum_i \sum_{j=1}^N |\langle \psi_i^F | \psi_j \rangle|^2 \quad (2.8)$$

penalises to populate the states within an undesired subspace, specified by a set $\{|\psi_i^F\rangle\}$, at each time step, which becomes particularly important for harmonic oscillators and multi-level systems. For harmonic oscillators, the truncation of the Hilbert space in the simulation introduces an artificial nonlinearity. For transmon [40] or fluxonium [41], it is preferable to take into account higher energy levels while staying within a computational subspace during the evolution. In both cases, it is important to avoid populating higher excited states. Another example is

$$P_2 = 1 - \frac{1}{N} \sum_{j=1}^N |\langle \psi_t | \psi_j \rangle|^2, \quad (2.9)$$

which forces the optimisation algorithm to construct a state close to the target state $|\psi_t\rangle$ in as few iterations as possible.

The summation terms in penalties 2.8 and 2.9 require the wavefunction at every time step j . The original GRAPE algorithm does not account for such complex dependencies in the penalty terms. Introducing them is particularly challenging, and each new penalty term would need explicit programming of the corresponding gradient. Reverse-mode Automatic Differentiation (AD) [42] can efficiently calculate the gradient of any penalty term without requiring any calculations beforehand, making the adjustments in the optimisation process more flexible.

Reverse-mode AD is a generalisation of the backpropagation algorithm, which is used for the gradient calculations in deep neural networks [43]. In simple terms, it sequentially applies the chain rule of differentiation to obtain the gradient of interest. The algorithm begins by constructing an evaluation trace of elementary operations, which is easier to un-

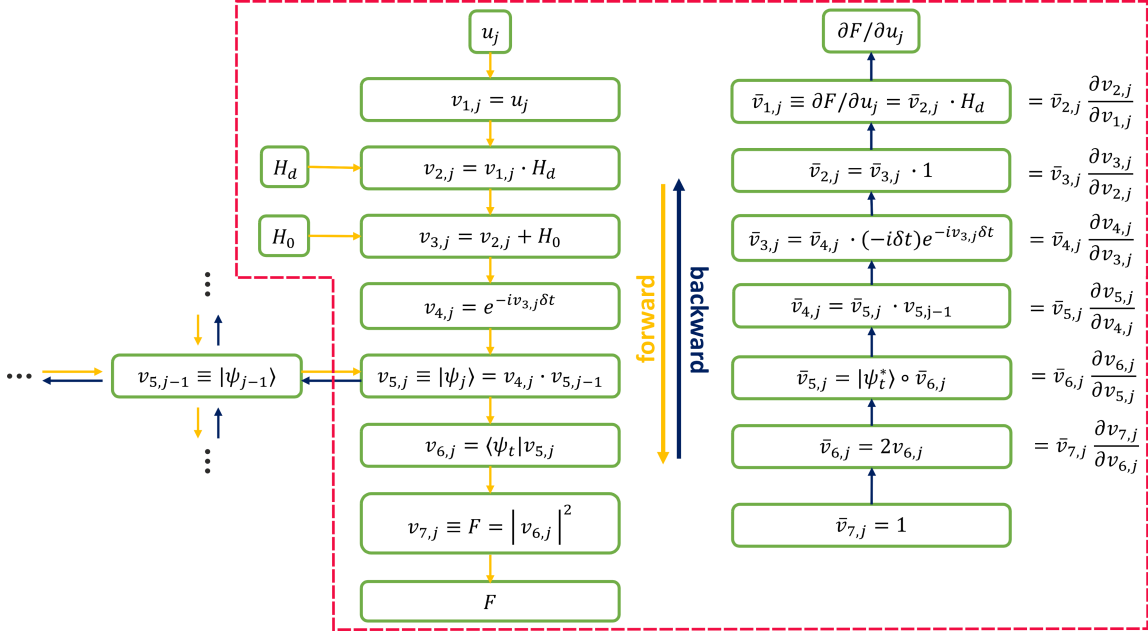


Figure 2.3: Computational graph for a toy optimal control problem. The system with the Hamiltonian H_0 evolves under an applied pulse defined in a piece-wise basis $\{u_j\}$. At each time step t_j , an external drive brings the system to the state $|\psi_j\rangle = \exp(-i(H_0 + u_j H_d)\delta t)|\psi_{j-1}\rangle$, which is calculated in the nodes $v_{2,j} - v_{6,j}$ by using an input variable $v_{1,j} = u_j$, external nodes H_0, H_d , and the quantum state from the previous time step $v_{5,j-1} = |\psi_{j-1}\rangle$. The derivative $\frac{\partial F}{\partial u_j}$ is evaluated in the last adjoint variable $\bar{v}_{1,j}$ as a result of a reverse propagation through the computational graph. The highlighted part is repeated for previous time steps $0, \dots, j-1$. Horizontal lines represent the information flow between consecutive time steps, whereas vertical lines indicate the internal flow of information within each calculation block at time step j . $\circ$ is the Hadamard product.

derstand by examining a specific example. Fig. 2.3 illustrates an example of the evaluation trace for a toy model of an optimal control problem. The model shown here is a simplified version of the one introduced in the previous Section (see Eqns. (2.1) to (2.3)) in the case of $m = 1$. For convenience, the key points are repeated here.

The toy model describes the quantum system defined with the Hamiltonian H_0 that evolves under the drive Hamiltonian H_d , decomposed into piece-wise constant basis 2.2. Assuming at each time step $t_j = t_0 + j \cdot \delta t$ the amplitude of the applied pulse is u_j , the state of the system at the time t_j is

$$|\psi_j\rangle = \hat{U}_j |\psi_{j-1}\rangle = \hat{U}_j \hat{U}_{j-1} \dots \hat{U}_0 |\psi_0\rangle, \quad (2.10)$$

$$\hat{U}_j = \exp\left[-\frac{i}{\hbar} \delta t \hat{H}_j\right] = \exp\left[-\frac{i}{\hbar} \delta t (\hat{H}_0 + u_j \hat{H}_d)\right] \quad (2.11)$$

The computational graph that shows the steps of the evaluation trace starts with declaring the input variables $v_{1,j} = u_j$ for each time step j . The intermediate variables $v_{2,j}, \dots, v_{6,j}$ represent intermediate computation steps corresponding to equations 2.10, 2.11. The output variable $v_{7,j}$ denotes the metric of interest, chosen to be the fidelity $F = |\langle \psi_t | \psi_j \rangle|^2$ for simplicity in this example.

The reverse-mode AD consists of a forward and backward propagation passes. In the forward propagation step, the variables $v_{i,j}$ at each node of the computational graph are calculated. In the backpropagation step, the gradients $\frac{\partial F}{\partial u_j}$ are computed by evaluating the adjoint variables

$$\bar{v}_{i,j} = \frac{\partial F}{\partial v_{i,j}} \quad (2.12)$$

Here, $i \in [1, 7]$ denotes the calculation step number; see Fig. 2.3 for an explicit example. The adjoint variable $v_{i,j}$ can be rewritten with respect to adjoint variables at the children nodes of the j -th node $ch(j)$ as

$$\bar{v}_{i,j} = \sum_{k \in ch(j)} \frac{\partial F}{\partial v_{i,k}} \frac{\partial v_{i,k}}{\partial v_{i,j}} = \sum_{k \in ch(j)} \bar{v}_{i,k} \frac{\partial v_{i,k}}{\partial v_{i,j}} \quad (2.13)$$

The required derivatives $\frac{\partial F}{\partial u_j} = \bar{v}_{1,j}$ are evaluated by applying 2.13 backwards through the evaluation trace, as shown in Fig. 2.3.

The evaluation trace in Fig. 2.3 contains derivatives of the propagator $U_j = e^{-iH_j\delta t}$ (intermediate variable $v_{4,j}$) at each time step. Reference [39] gives more details about some tricks for calculating the derivative of the matrix exponential, which speed up reverse-mode AD for QOC applications.

2.1.3 Gradient-based optimisation algorithms

GRAPE and reverse-mode AD are the tools for calculating the gradient. The rules for updating parameters, however, depend on the type of optimisation algorithm that is chosen. There are many different types of optimisation algorithms [44], which vary in terms of convergence speed for different types of cost functions and parameter landscapes, optimisation constraints, and memory requirements. The choice of the algorithm is always problem-specific. QOC applications exploit only a small part of the wide variety of optimisation options and can be roughly divided into two big groups: based on first-order and second-order methods.

First-order methods utilise only the gradient of the cost function. The simplest and most widely used algorithm in this family is gradient descent, which optimises the model parameters u by taking steps in the direction opposite to the gradient of the cost function.

$$u \leftarrow u - \epsilon \nabla J, \quad (2.14)$$

and its more advanced version with momentum that "accumulates" the decaying velocity v in parameter space in the direction of past gradients

$$v \leftarrow \alpha v - \epsilon \nabla J, \quad \alpha \in [0, 1) \quad (2.15)$$

$$u \leftarrow u + v \quad (2.16)$$

The further development of first-order methods led to algorithms with adaptive learning rates, of which Adam algorithm [45] typically exhibits faster and more stable convergence, regardless of the initial values of the hyperparameters.²

Second-order methods utilise the second derivatives to find the optimum. The evolution of second-order methods started with Newton's method, which is based on the second-order Taylor expansion. Around some point u_0 , the cost function can be expressed as

$$J(u) = J(u_0) + (u - u_0)^T \nabla J(u) + \frac{1}{2}(u - u_0)^T H(u - u_0) + O(u^3), \quad (2.17)$$

where H is a Hessian matrix. The minimum of $J(u)$ is achieved at

$$u \leftarrow u_0 - H^{-1} \nabla J(u_0) \quad (2.18)$$

The Newton update rule 2.18 finds the minima in one step for convex functions. For non-convex functions, it should be applied iteratively until convergence is achieved. However, calculating the inverse of the Hessian matrix is computationally expensive. Quasi-Newton algorithms overcome this problem by iteratively approximating H^{-1} from gradient evaluations. One of the most popular methods of the quasi-Newton branch is Broyden–Fletcher–Goldfarb–Shanno algorithm (BFGS) [44], which was used for open-loop optimisations of the system in this thesis (see Chapter 6). BFGS exhibits a superlinear convergence rate and typically requires a lot fewer iterations for convergence, compared to gradient descent.

2.2 Closed-loop optimisation techniques

Even for a perfectly characterised system, gate or state fidelity improvement eventually runs into a problem: the simulation can never account for all the experimental imperfections. Leakage, undesired crosstalk, setup imperfections, and unaccounted effects in the Hamiltonian always build a barrier between the perfect theoretical prediction and the actual experiment. To take the next step in improving gate or state fidelity, it is necessary to provide the optimiser with a measured cost function instead of a simulated one.

The first attempts at closed-loop optimisation focused on improving the DRAG pulses, given their widespread use within the superconducting circuits community. J. Kelly et al [46] showed that optimising the randomised benchmarking (RB) sequence fidelity with the gradient-free Nelder-Mead algorithm can increase the CZ gate fidelity from 0.984 to 0.993. The idea of using RB fidelity as a figure of merit has been further exploited by various groups [47, 48] due to its obvious benefits: the sequence is easy to implement, and its fidelity decay does not depend on the state preparation and measurement (SPAM) errors. In comparison to quantum state tomography, it gives an advantage in the number of measurements per optimisation iteration.

²Hyperparameters define the configuration of the optimisation algorithm, thus controlling the convergence. Hyperparameters' values are adjusted before the training.

For the optimisation driven by experimental results, there is always a race between the convergence of the chosen optimisation method and the measurement time of the selected cost function. Sometimes it is possible to achieve high fidelity within a reasonable amount of time by using modified gradient-based methods [25]. This, however, is more an exception than the rule. The most popular experimental applications belong to gradient-free Nelder-Mead, CMA-ES, and real-time feedback (Secs. 2.2.1 to 2.2.4). In recent years, reinforcement learning (Sec. 2.2.5) has become a popular tool for fine-tuning the experimental parameters.

Although gradient-free methods use the measured values directly, it is often not easy to find a set of hyperparameters that ultimately leads to a good result. Depending on the problem, gradient-free algorithms require intensive work in finding an initial guess, designing a cost function and penalties, and defining a proper basis to ensure convergence. Additionally, not every optimisation algorithm exhibits good noise resilience to measurement noise. In general, they tend to find a good solution only for a small number of optimisation parameters. This evolved into the idea of changing the optimisation basis during the optimisation process (see dCRAB algorithm, Sec. 2.2.3), which was successfully experimentally implemented on an NV centre platform [49].

2.2.1 Nelder-Mead algorithm

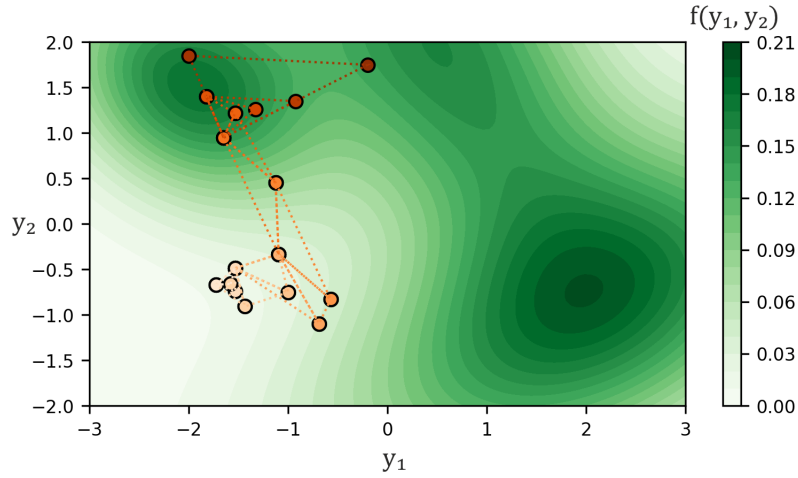


Figure 2.4: First 15 iterations of the Nelder-Mead algorithm on the parameter landscape. The algorithm minimises the cost function $f(y_1, y_2)$ in a two-dimensional space spanned by parameters y_1 and y_2 . In an $n = 2$ parameter space, the algorithm starts with $n + 1 = 3$ randomly chosen points $\bar{y}^{(0)} = (y_1^{(0)}, y_2^{(0)})$, $\bar{y}^{(1)} = (y_1^{(1)}, y_2^{(1)})$, $\bar{y}^{(2)} = (y_1^{(2)}, y_2^{(2)})$, which form an initial simplex. At each iteration, the point with the largest value of the cost function f is replaced by using the simplex update rule 2.19. The vertices of the simplices are plotted as markers of different colours. Here, the first iterations correspond to a more intense orange colour.

The Nelder-Mead algorithm [50] is a gradient-free, direct-search algorithm, which means that it explores the parameter landscape based on the cost function $f(\bar{y})$, $\bar{y} \in \mathbb{R}^n$ values only. The algorithm starts by choosing a simplex of $n + 1$ vertices $\{\bar{y}^{(0)}, \bar{y}^{(1)}, \dots, \bar{y}^{(n-1)}, \bar{y}^{(n)}\}$

arranged in an ascending order of the function values $f(\bar{y}^{(0)}) \leq f(\bar{y}^{(1)}) \leq \dots \leq f(\bar{y}^{(n)})$. At each iteration, it replaces the point $\bar{y}^{(n)}$ with the biggest function value on the simplex, so that the whole simplex slowly "moves" in the direction of the minima (see Fig. 2.4). For that, three main types of operations on the simplex are used: reflection, contraction, and expansion. The operations are specified with respect to the simplex centroid $\bar{y}^c = \sum_{i=0}^{n-1} \bar{y}^{(i)} / n$ and can be described by the general formula

$$\bar{y}^{\text{new}} = \bar{y}^c + \delta(\bar{y}^c - \bar{y}^{(n)}) \quad (2.19)$$

where the usual choice of parameters implies $\delta = 1$ for reflection, $\delta = \pm 0.5$ for contraction and $\delta = 2$ for expansion. After each step, the simplex vertices are rearranged in the ascending order of $f(\bar{y}^{(i)})$ again. The choice of the applied simplex transformations depends on the relations between $f(\bar{y}^{(0)})$, $f(\bar{y}^{(n)})$ and $f(\bar{y}^{(\text{new})})$. The comprehensive algorithm procedure and possible convergence problems are discussed in detail in [51]. The example of the Nelder-Mead simplex descent in the parameter landscape is shown in Fig. 2.4.

2.2.2 Evolution strategies and CMA-ES algorithm

Evolution strategies (ES) [52], as their name suggests, exploit the concept of biological evolution. In a biological population, the new generation is formed as a result of the genetic variation and natural selection. The individuals from the current population produce the new offspring by combining the genetic information from different individuals, which is subjected to mutation. In the ES algorithm, the described above genetic variation is realised through recombination and mutation steps, which are explained in detail below. As a result of the genetic variation, the population contains offspring with different kinds of mutations, but not all of them appear to be viable. Only the mutations that enhance the ability to adjust to the environment survive and form the next population generation, as stated in the natural selection principle. The natural selection laws are introduced via the evaluation step, where the fitness of the offspring is estimated, and the selection step, where the best offspring are retained based on the results from the evaluation step.

The ES algorithm [52] aims to minimise the cost function $f(\bar{x})$ on a d -dimensional parameter space $\bar{x} = (x_1, x_2, \dots, x_d)$. There are different types of ES algorithms. The (μ, λ) -ES algorithm, which is briefly described in this Section, is parametrised with a population size μ and the number of generated offsprings λ . The algorithm starts with generating μ d -dimensional parameter vectors $\bar{x}^{(k)}$, $k \in [1, \mu]$. Continuing the analogy with biological evolution, it is equivalent to the generation of μ individuals, which serve as parents for the next generation. Offspring and the next population generation are formed as a result of the following steps:

- **Recombination.** λ offspring are generated from μ parent individuals. For example, one can calculate the mean over each component for a subset of parent parameter vectors.

- Mutation. λ offspring from the recombination step are subjected to a mutation by applying the following transformation

$$\begin{aligned}\sigma_k^{(g+1)} &= \sigma^{(g)} e^{\xi_k} \\ \bar{x}^{(k,g+1)} &= \bar{x}^{(k,g)} + \sigma_k^{(g+1)} y^{(k)}, \quad y^{(k)} \sim \mathcal{N}(0, I)\end{aligned}\quad (2.20)$$

where $k \in [1, \lambda]$ is an offspring number, g and $g + 1$ are numbers of the current and the next generations, $\mathcal{N}(0, I)$ is a d -dimensional isotropic normal distribution, ξ_k is a random number with zero mean. $\sigma^{(g)}$ is a step size for the generation g , which forms individual step sizes for offspring $\sigma_k^{(g+1)}$.

- Evaluation. The cost function $f(\bar{x})$ is evaluated on λ newly constructed offspring vectors $\bar{x}^{(k,g+1)}$, $k \in [1, \lambda]$. The best offspring (solution vector $\bar{x}$) and the best cost function values are updated.
- Selection. The best μ individuals from the updated population (consisting of parent and new offspring individuals) are selected and will be used for the recombination in the next generation step.

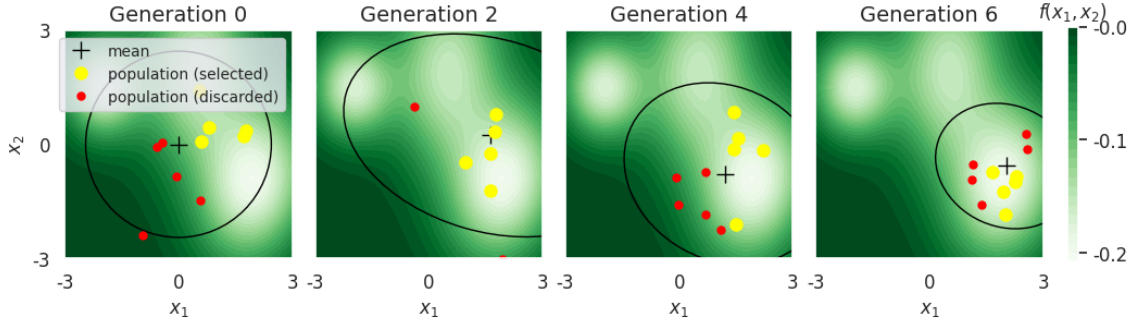


Figure 2.5: CMA-ES algorithm generation steps # 0, 2, 4, and 6 in the parameter space for minimising the cost function $f(x_1, x_2)$ in a two-dimensional parameter space. At each iteration, as a result of recombination and mutation steps, the CMA-ES algorithm forms a new population (red and yellow dots). The cost function is evaluated at those points. During the selection step, only the best representatives of the population form the new generation. The discarded individuals are coloured in red, the selected individuals are coloured in yellow. The $2\text{-}\sigma$ confidence area of CMA-ES, which is defined by the covariance matrix $C^{(g)}$ at the generation step g , is plotted as a black ellipse. The plots are generated using the **cma-es** Python library [53].

Among all the different strategies, the best results were obtained by using a de-randomised mutation that was successfully implemented by Hansen et al [54, 55]. The de-randomised Covariance Matrix Adaptation (CMA) algorithm effectively updates the individual step sizes in the parameter space. This is done by applying a different mutation step

$$\bar{x}^{(k,g+1)} = \langle \bar{x} \rangle_w^{(g)} + \sigma^g y^{(k)}, \quad y^{(k)} \sim \mathcal{N}(0, C^{(g)}). \quad (2.21)$$

Compared to 2.20, mutation rule 2.21 undergoes two important differences:

1. The offspring $\bar{x}^{(k,g)}$ is replaced with a weighted mean of μ best individuals from generation g , $\langle \bar{x} \rangle_w^g = \sum_{i \in [\mu \text{ best}]} w_i \bar{x}^{(i,g)} / \sum_{i \in [\mu \text{ best}]} w_i$;
2. $y^{(k)}$ is drawn from the normal distribution $\mathcal{N}$ with an arbitrary covariance matrix, compared to an isotropic distribution with $C^{(g)} \equiv I$ in 2.20. The covariance matrix $C^{(g)}$ of the normal distribution determines the shape and inclination of the distribution's ellipsoid in the parameter space, which allows for adjusting step sizes for each parameter separately, see Fig. 2.5.

The bottleneck of this approach is the calculation of the covariance matrix eigendecomposition, which is required for sampling $y^{(k)}$ and becomes computationally expensive in higher dimensions. However, CMA-ES is efficient on ill-conditioned problems, and its self-adaptation behaviour typically leads to a much faster convergence.

Typically, it is more challenging to adjust hyperparameters for CMA-ES than for the Nelder-Mead algorithm. Regardless, experimentalists sometimes prefer to use the CMA-ES algorithm. For example, in the case of M. Weringhaus et al. [48], the largest amount of optimisation time is taken up by the instruments' initialisation. Therefore, algorithms that use more points per evaluation before the experiment re-initialisation ensure faster convergence in a closed-loop.

2.2.3 Chopped RAndom Basis algorithm

The pulse optimisation on a piece-wise basis introduces a large number of control parameters, depending on the pulse discretisation step and length. It significantly increases the parameter space, takes longer for convergence, and decreases the probability of reaching the global minimum. That brings the idea of representing the pulse in a different basis, which would require a much smaller number of control parameters.

In the Chopped RAndom Basis (CRAB) ansatz [56], the control pulse $A(t)$ is expanded into a chosen finite orthogonal basis. Possible basis options include Lagrange polynomials, Hermite functions, etc. Currently, the most widely used basis is a Fourier basis

$$A(t) = \sum_{i=1}^{N_c} c_i A_i(t)$$

$$A_i = a_i \sin(\omega_i t) + b_i \cos(\omega_i t)$$

where N_c is the basis dimension. The frequencies ω_i are picked up randomly around the principal harmonics

$$\omega_i = \frac{2\pi i}{t_{\text{pulse}}} (1 + r_i),$$

where t_{pulse} is the length of the pulse, and $r_i \in [-0.5, 0.5]$ is a random number. By changing N_c one can change the dimension of the Fourier space, which leads to changes in pulse bandwidth. The optimisation problem on such a randomly defined basis can be solved with, for example, gradient-free direct search methods.

The basic version of the CRAB ansatz has two disadvantages:

- the convergence depends on the initial choice of the frequencies;
- the "ideal" pulse constructed on a Fourier basis cannot be exactly reproduced by an arbitrary waveform generator due to its limitations.

To overcome the issue of getting trapped in local minima (false traps) caused by an unfortunate choice of basis functions A_i , Rach et al. [57] proposed the dressed CRAB algorithm, which sequentially runs the CRAB optimisation while updating the basis after each iteration. Each run is called a super-iteration and takes an optimised control pulse from the previous iteration as a guess for the next run. After the super-iteration j , the algorithm optimises basis expansion coefficients c_i for a new randomly chosen basis $\{A_i^{j+1}, i \in [0, N_c]\}$

$$A^{j+1}(t) = c_0^j A^j(t) + \sum_{i=1}^{N_c} c_i^{j+1} A_i^{j+1}(t)$$

based on the optimisation result $A^j(t)$ from the previous super-iteration. The basis change after each super-iteration leads to a change in the optimisation landscape, which mitigates the false trapping problem and results in faster convergence of the optimisation routine compared to the standard CRAB algorithm. Another important advantage of dCRAB includes weak dependence of performance on the space dimension N_c and outstanding performance even for the small N_c values, which significantly decreases the computational time [57].

The hardware limitations can be, to a large extent, eliminated by choosing a sigmoid basis [17]. It is worth mentioning here that the idea of the CRAB ansatz is widely used not only in combination with gradient-free methods but also with gradient-based methods. In this thesis, the chopped basis ansatz combined with any gradient-free algorithm will be referred to as the CRAB algorithm, as it was introduced in the original paper.

2.2.4 Measurement-based feedback

The algorithms from the previous sections focus on optimising the control pulses. The goal of measurement-based feedback, on the contrary, is to apply a set of operations to correct errors arising in the quantum system or to get it to a desired target state by using pre-calibrated pulses. The set of correction rules remains constant throughout the experiment, making the technique quick but inflexible.

The concept of quantum measurement-based feedback becomes straightforward when examining one of the experimental realisations. For example, in [58], N. Ofek et al demonstrate quantum error correction of a logical qubit with real-time feedback. The logical qubit is encoded in a cat state basis and stored in a cavity. The cavity is coupled to a transmon, which serves as an ancilla for syndrome measurements and encoding/decoding of the logical qubit state.

The protocol (see Fig. 2 in [58]) begins with the mapping of the ancilla state to the desired even-parity cat state (logical encoding step). Next, during the error correction step, a single-shot measurement is performed on the ancilla approximately every $13 \mu s$. In the cat code, the dominant source of error is a single photon loss, which flips the parity of the encoded state. Therefore, tracking the change in parity during the system's evolution allows for the detection of up to 2 single-photon errors during the error correction step. Finally, depending on the error syndrome, the feedback controller chooses a decoding pulse and calculates an accumulated phase rotation. At the end of the protocol, the state is encoded back in the ancilla.

The successful implementation of measurement-based feedback requires low hardware latency and fast, high-fidelity readout. Although the second one can be easily fixed by adding a parametric amplifier and adjusting the readout parameters, the first one remains challenging for many research groups. The response of the controller should be faster than the characteristic evolution time of quantum systems. It means that the calculations should be performed on a nanosecond scale. This is the reason why real-time feedback protocols are running on Field Programmable Gate Arrays (FPGAs). FPGAs typically operate at a 4 ns clock cycle, which allows it to dramatically reduce the loop time compared to a standard ADC-computer-DAC setup. FPGA programming requires special qualifications from researchers, making it accessible only to a small subset of experimental groups. However, ongoing improvements at the hardware level mitigate the problem. For example, QM OPX+ can be programmed for real-time feedback through the **qm-qua** Python library [59].

More details about the measurement-based feedback implementation challenges can be found in Yehan Liu's thesis [60].

2.2.5 Reinforcement learning

Compared to other optimisation and control methods, the reinforcement learning (RL) approach stands out by considering an actual interaction between the agent and its environment. In the RL paradigm, an RL agent, based on the current state of the environment at each time step, performs an action that potentially maximises its reward.

The interaction between the agent and its environment is usually described in terms of the Markov Decision Process, which is defined by

- set of states S that the environment can take,
- set of actions A that the agent can apply,
- transition probability from state $s_t \in S$ to $s_{t+1} \in S$ after an action $a_t \in A$, $P(s_{t+1}|s_t, a_t)$,
- reward function $R_t = R(s_{t+1}, s_t, a_t)$, which determines appropriateness of the action a_t .

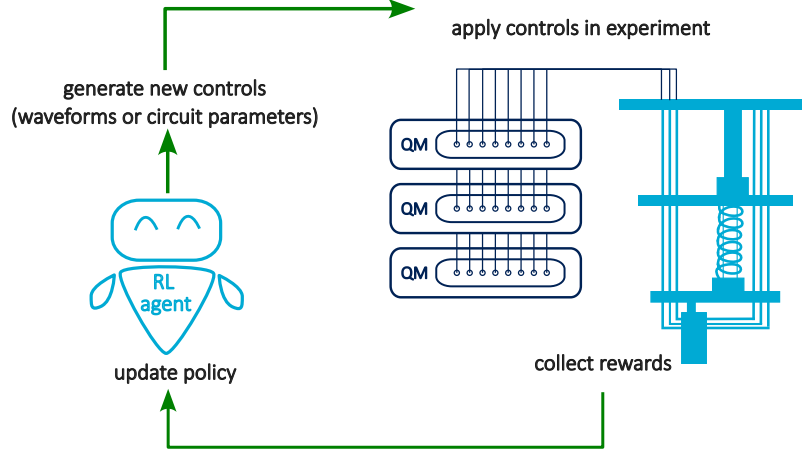


Figure 2.6: Reinforcement learning pipeline. See details in the main text.

The action is selected according to the agent’s policy $\pi(a_t|s_t)$, which is corrected after each episode of agent-environment interaction. The goal is to find a policy that maximises the expected return:

$$R_t = E \left[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \right], \quad (2.22)$$

where $\gamma \in [0, 1]$ determines how much immediate rewards, compared to future rewards, influence the agent’s decision of taking an action.

In the context of cQED experiments, the state space S consists of all the possible states of the observed quantum system. Instead of getting direct information about the state of the system s_t (by accessing the state fidelity), the agent may receive an observation o_t instead, which can be, for example, a result of the projective measurement at the end of the reward circuit or a value of Husimi-Q/Wigner functions. Based on the observation, the reward is calculated. Action space A contains all the available parameters of the control circuit or control pulses. Sivak et al. [61] conducted extensive research on possible RL cQED-oriented applications and corresponding rewards. In their paper, the reward is a projective measurement outcome m at the end of the reward circuit. To keep notation short, in this section, the measurement outcome at the end of the reward sequence composed of the gates $G_1 G_2 \dots G_k$ is denoted as $m[G_1 G_2 \dots G_k]$. Each reward circuit is designed for a specific task. Thus, the task of Fock state $|n\rangle$ preparation may utilise the reward

$$R = -m[|n\rangle\langle n|], \quad (2.23)$$

which is equivalent to a selective conditional π -pulse on the qubit $X^n(\pi)$ followed by a projective qubit measurement (or a generalised Husimi Q measurement, see section 3.5.2). The preparation of the GKP stabiliser states can be learned with

$$R = m[|X\rangle\langle X|] \text{ or } m[|Z\rangle\langle Z|], \quad (2.24)$$

where $|X\rangle, |Z\rangle$ are eigenstates of stabiliser generators $S_{x/p,\Delta}$ ³, and the final projective measurement depends on the first stabiliser displacement in x or p direction, which are chosen with equal probability.

The preparation of arbitrary states can be based on the reward

$$R = m[P_\beta] \text{sign} W_{\text{target}}(\beta), \quad (2.25)$$

where P_β is a parity at the phase space coordinate β , which is multiplied by the sign of the Wigner function at the same coordinate (see Section 3.5.1). Importantly, the reward 2.25 remains stochastic even in the absence of measurement noise and requires averaging over 10 - 100 different phase space points⁴ for convergence in an open-loop.

Rewards 2.23, 2.24, and 2.25 rely on stochastically measured outcomes. Although open-loop runs show good convergence, the authors [61] did not investigate the effect of measurement noise on the convergence behaviour. Later, the authors experimentally showed that the reward 2.24 improves the performance of the quantum error correction code [11]. Nevertheless, the reward based on the stochastically measured outcome provides the most elegant solution by utilising the probabilistic nature of the quantum measurement. The rewards based on the fidelity of the operation or state require much more measurement time, although they may converge much faster.

2.3 Outlook

Although the number of theoretical proposals for different optimal control methods continues to grow from year to year, experimental implementations remain far behind. Table 2.1 gives a brief overview of the superconducting circuits experiments that used open- and closed-loop optimisation methods in recent years. It doesn't include open-loop optimisations from purely theoretical articles. The Table is organised as follows: it lists closed-loop experiments first, and open-loop experiments right after. The articles within open- and closed-loop optimisation sections are grouped by the optimisation method used in the experiment. The system type in the second column is designated as 'Q' for a Josephson junction-based superconducting qubit (transmon, fluxonium, etc) and as 'B' for a bosonic qubit (cat/binomial/GKP state, etc). In case there is no direct information about the figure of merit (FoM) in the paper, the corresponding cell is filled with a "-" sign.

³for the GKP code,

$$S_{x/p,\Delta} = E_\Delta S_{x/p,0} E_\Delta^{-1}, \quad E_\Delta = \exp(-\Delta^2 a^\dagger a), \quad S_{x,0} = D(\sqrt{\pi}), \quad S_{p,0} = D(i\sqrt{\pi})$$

⁴but still with only one measurement per point; the effect of measurement noise is not taken into account

Reference	System type	Optimisation target	Optimisation method	FoM/reward
J. Kelly et al [46] (2014)	Q	DRAG $X_{\pi/2}$, CZ gate	Nelder-Mead (closed-loop)	RB ⁵ fidelity
M. A. Rol et al [47] (2017)	Q	DRAG $X_{\pi, \pm\pi/2}$, $Y_{\pi, \pm\pi/2}$	Nelder-Mead (closed-loop)	restless RB error fraction
M. Werninghaus et al [48] (2021)	Q	DRAG $X_{\pm\pi/2}$, $Y_{\pm\pi/2}$	CMA-ES (closed-loop)	RB fidelity
Z. Zong et al [25] (2021)	Q	CZ gate	data-driven GRAPE based on quantum state/process tomography (closed-loop)	$\ U_d^\dagger U_c - U_{CZ}\ $, $\ \rho(T) - \rho_{ideal}\ $ (see eqs. 8, 9 in the main text for notations)
K. Reuer et al [62] (2023)	Q	initialisation of the transmon in the ground state	RL, real-time feedback (closed-loop)	the reward approximates initialisation fidelity and punishes trajectories with a high number of cycles, see eq. 4 in Supplementary Material
Y. Baum et al [63] (2021)	Q	$X(\pi/2)$, $ZX(-\pi/2)$, SWAP	RL for all the gates; simulated annealing algorithm for ZX gate	reward is a weighted sum of measured fidelities for sequences employing different numbers of control pulse repetitions
N. Ofek et al [58] (2016)	B	monitoring and correcting errors	real-time feedback for syndrome tracking, GRAPE for encoding/decoding (open-loop)	fidelity of the state transfer
P. Campagne-Ibarcq et al [64] (2020)	B	preparation of square and hexagonal GKP code states	real-time feedback	-

⁵randomised benchmarking

Reference	System type	Optimisation target	Optimisation method	FoM/reward
V. V. Sivak et al [11] (2023)	B	quantum error correction circuit parameters (45)	improvement of [64]: RL agent to fine-tune virtual angles, complex-valued amplitudes of displacement gates, etc. (closed-loop) Pauli eigenstates are created with a method from [65] (open-loop) minimising circuit depth with Adam feedback for ancilla reset and cavity cooling	$\frac{1}{2}(E[X_L] + E[Z_L])$, where $E[X_L], E[Z_L]$ - expectations of logical Pauli operators
R. W. Heeres et al [34] ⁶ (2017)	B	encoding, decoding, $X_{\pi, \pm\pi/2}$, $Y_{\pi, \pm\pi/2}$, T , H , I (cat code basis); Fock state	GRAPE (open-loop), feedback (for the preparation of the system in the ground state)	fidelity of the state transfer
L. Hu et al [69] (2019)	B	encoding, decoding, correction pulses (binomial code basis), $X_{\pi, \pm\pi/2}$, $Y_{\pi, \pm\pi/2}$, T	GRAPE (open-loop), real-time feedback	not mentioned, but most probably fidelity of the state transfer
A. Eickbusch et al [65] (2022)	B	Fock states, squeezed states, binomial states, GKP	two-stage (open-loop): gradient descent (Adam) for minimising circuit depth (1), Nelder-Mead for optimising ECD pulses' parameters (2); opt. done for comparison (open-loop): RL for SNAP-gate sequence optimisation (depth, SNAP parameters); repetition of [34] approach;	(1) fidelity of the state transfer, (2) cost function designed explicitly for ECD gate optimisation - see S27 from Supplementary Material

⁶The same approach is used in the following papers: Rosenblum et al [66], Chou et al [67], S.Touzard et al [68]

Reference	System type	Optimisation target	Optimisation method	FoM/reward
M. Kudra et al [70] (2022)	B	sequence of displacements and SNAP gates to generate binomial, cat, cubic phase and GKP states	two-stage: (1) gradient descent for SNAP phases and displacements, (2) unspecified method from Boulder Opal (most probably GRAPE) (open-loop)	(1) see D1 [70], (2) infidelity ⁷
Y. Ma et al [71] (2020)	B	adiabatic QEC pulses, encoding/decoding pulses	GRAPE (open-loop)	fidelity of the state transfer
Ni et al [10] (2023)	B	encoding and decoding pulses, recovery gate pulse shape	real-time feedback, GRAPE (open-loop) for pulse optimisation	fidelity of the state transfer
Hutin et al [3] (2024)	B	control pulses for a family of cat states $ C_\alpha^\phi\rangle \propto \alpha\rangle + e^{-i\phi} -\alpha\rangle$	Adam (open-loop) optimisation of a dense neural network parameters (3 hidden layers); GRAPE (open-loop) for comparison	infidelity of the state transfer with a first-order correction for decoherence (see eq. 4 in the main text)

Table 2.1: Experimental implementations of optimisation techniques on a superconducting circuits platform.

⁷<https://docs.q-ctrl.com/references/boulder-opal/boulderopal/superconducting/optimize.html>

As shown in Table 2.1, **GRAPE** stands out as one of the most widely used optimisation methods, particularly in experiments involving bosonic qubits. In [34], GRAPE was successfully used to create qubit and cavity control pulses for establishing a mapping between transmon and cavity subspaces and performing the gates on logical cat qubits. The average fidelity of the encoded states was 0.96, and the average infidelity of the gate set was 0.75%. These encouraging results gave a boost to a series of bosonic logical qubits and bosonic states experiments in Schoelkopf's lab, including stabilisation of the cat states manifold [68], demonstration of a CNOT gate between two multiphoton qubits in two microwave cavities [66], teleportation of CNOT gate between logical qubits [67]. Interestingly, despite its gradient-based nature, GRAPE was successfully implemented in a closed-loop for CZ-gate optimisation [25].

Nelder-Mead showed a good performance in both closed- and open-loop tasks from both gate fidelity and time convergence perspectives. Proper choice of cost function and its measurement method improves the Nelder-Mead optimisation time dramatically. Thus, optimisation with the upgraded RB⁵ version - restless RB [47] - allowed to reach T_1 -limited Clifford gate fidelity within a minute, while the classic RB approach requires eleven. The significant speedup comes from the fact that the restless RB concept skips the system initialisation step and estimates the qubit initial state based on the final state at the end of the previous RB sequence.

CMA-ES algorithm applications are not often reported, but demonstrate outstanding results as well. S. Filipp's group [48] achieved single-qubit gate fidelity of 0.9976 for 4.16 ns long pulses and 0.9987 fidelity for pulses longer than five ns. Both gradient-free Nelder-Mead and CMA-ES were used in discrete qubit experiments only.

The real-time feedback is less widely used than other quantum control techniques. However, most of the breakthrough experiments in recent years wouldn't be possible without this adaptive control tool: approaching the break-even point for cat [58], binomial [10], and GKP [11] codes became possible only with real-time correction of the quantum systems' evolution. Another popular application of feedback is preparing the system in the ground state [34].

The experimental use of **reinforcement learning** (RL) agents is just starting to evolve. Amongst all the RL applications, the work from Q-CTRL [63] deserves special attention. The paper demonstrates a $3\times$ reduction in $X(\pi/2)$ gate duration (down to 12.4 ns) and a $2.4\times$ reduction in $ZX(-\pi/2)$ gate infidelity compared to default IBM calibrated pulses. Considering that the RL agent required only 0.5 - 1 h to converge, the future usage of RL for different types of single- and two-qubit gates looks promising. It is worth mentioning that the authors also optimised the ZX gate using a simulated annealing (SA) algorithm. Both SA and RL showed the same reduction in gate error. Since there is no detailed comparison between the two techniques, it is unclear whether, in the chosen application, RL wins over SA. For the bosonic state creation task, theoretical studies [61] show that RL outperforms Nelder-Mead and simulated annealing approaches in fidelity for the same sample size. The other applications of RL include open-loop optimisation of the gate sequence depth and gate parameters [65] and initialisation of the qubit in the ground state through the feedback loop [62]. The most remarkable application was reported in the break-even experiment [11], where an RL agent learned 45 parameters (virtual corrections,

displacement gates amplitudes, etc) in a closed-loop. The authors specifically pointed out that it was not possible to achieve a break-even point using calibrations and educated guesses alone, and RL was crucial for improving overall performance.

Summary

Superconducting circuit experiments employ a wide range of versatile quantum optimal and adaptive control techniques. For discrete variable qubits, both open- and closed-loop approaches have been successfully implemented to improve gate fidelities. For continuous variable (bosonic) qubits, only open-loop optimisation has been widely adopted so far. The closed-loop optimisation for bosonic qubits remains challenging mainly due to two reasons:

- The optimisation algorithm requires a cost function related to a fidelity-like metric, which allows for estimating how close the state or gate is to the target. For continuous spaces, it is tricky to estimate such a metric in a reasonable amount of time. The bosonic state fidelity requires Wigner or Husimi Q measurements (see Section 3.5) in the relevant area of phase space, which typically takes hours. Fidelity-like metrics, as shown later in Section 4.5, do not show good resilience to noise.
- Bosonic qubits, due to the complexity of an underlying Hilbert space, require much more complicated pulses than discrete qubits. It makes the optimisation problem more challenging, especially for gradient-free optimisers, which show good convergence for a low number of parameters.

One of the rare examples of closed-loop optimisation of a continuous system is the break-even experiment [11], where the RL agent has access to the circuit parameters (which significantly reduces the number of optimisation parameters), and the reward circuit applies the projection into stabiliser subspace. While the projection onto the stabiliser subspace significantly reduces the problem complexity, it limits the application area. Closed-loop optimisations of bosonic states without employing projections onto a stabilised subspace have not yet been demonstrated.

Experimental platform: brief introduction

Bosonic states provide the foundation for continuous-variable error correction codes [22] and have been demonstrated across a range of platforms, including optics [7], trapped ions [5], mechanical oscillators [72], and atomic spin ensembles [6]. In this thesis, superconducting circuits serve as a chosen experimental platform. This chapter outlines the essential components for generating and characterising bosonic states in superconducting circuits: theoretical derivation and experimental characterisation of the system Hamiltonian (Secs. 3.1 to 3.3), operations required for state preparation (Section 3.4), and tomography techniques for state reconstruction (Section 3.5).

3.1 Harmonic oscillator and the need for nonlinearity

Bosonic states are created, manipulated and stored in systems that can be considered as harmonic oscillators. In circuit quantum electrodynamics (QED), a superconducting 3D cavity (or a 2D resonator) plays the role of a harmonic oscillator. A superconducting cavity produces a set of bosonic modes with a frequency spacing depending on the boundary conditions. In the experiment, bosonic states are typically created in a particular (typically the fundamental) mode of the cavity; therefore, the Hamiltonians in this Chapter will omit the summation over all the harmonic oscillator modes and include only the relevant mode.

The cavity driven with an external classical drive $\varepsilon(t)$ can be described with the Hamiltonian

$$\hat{H} = \hbar\omega\hat{c}^\dagger\hat{c} + \varepsilon(t)(\hat{c}^\dagger + \hat{c}), \quad (3.1)$$

where ω is the bosonic mode frequency, and c is the mode annihilation operator. In the rotating frame, i.e. under the unitary transformation $\hat{U} = \exp(-i\omega\hat{c}^\dagger\hat{c})$, the Hamiltonian becomes

$$\tilde{H} = \varepsilon(t)(\hat{c}^\dagger e^{-i\omega t} + \hat{c}e^{i\omega t}) \quad (3.2)$$

The corresponding evolution of the system can be simplified by applying the Magnus expansion, which terminates after the second term:

$$\begin{aligned}\tilde{U} &= \mathcal{T}^1 \exp \left(-i \int_0^t d\tau \tilde{H} \right) = \exp \left[\sum_k \Omega_k \right] \\ \Omega_1 &= -i \int_0^t dt_1 \tilde{H}(t_1) = -i \left[\hat{c} \int_0^t dt_1 \varepsilon(t_1) e^{i\omega t_1} + \hat{c}^\dagger \int_0^t dt_1 \varepsilon(t_1) e^{-i\omega t_1} \right] \equiv -i(\hat{c}\alpha^* + \hat{c}^\dagger \alpha) \\ \Omega_2 &= -\frac{1}{2} \int_0^t dt_1 \int_0^{t_1} dt_2 [\tilde{H}(t_1), \tilde{H}(t_2)] = -\int_0^t \varepsilon(t_1) dt_1 \int_0^{t_1} \varepsilon(t_2) \sin(\omega(t_2 - t_1)) dt_2 \equiv -i\phi \\ \Omega_3 &= \frac{i}{6} \int_0^t dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3 \left([\tilde{H}(t_1), [\tilde{H}(t_2), \tilde{H}(t_3)]] + [\tilde{H}(t_3), [\tilde{H}(t_2), \tilde{H}(t_1)]] \right) = 0\end{aligned}$$

It means that applying an external drive is equivalent to displacing the cavity state in the phase space²:

$$\tilde{U} = \exp(-i(\alpha\hat{c}^\dagger + \alpha^*\hat{c}) - i\phi) \propto \hat{D}_\alpha \quad (3.3)$$

Starting from the vacuum state in the cavity, an external resonant drive can create only coherent states, which behave nearly classically. To create an arbitrary quantum state inside the cavity to an arbitrary degree of accuracy, one needs a system with a higher-order Hamiltonian, like a Kerr Hamiltonian [73]. Such a nonlinearity can be embedded by coupling the linear cavity to a system with nonlinearity, such as a transmon or fluxonium qubit. To introduce the idea of a transmon, it is necessary to become familiar with the circuit quantum electrodynamics formalism first.

3.2 Introducing nonlinearity. Black box quantisation

Superconducting circuits are electrical circuits composed of two types of elements: classical-like linear capacitances and inductors, and nonlinear Josephson or phase-slip junctions, which arise from macroscopic quantum phenomena. Operating such a circuit requires low temperatures (below the critical temperature of the superconductor) and low (typically tens of MHz³) driving field strength that can add or remove excitations from the circuit without breaking superconductivity (superconducting gap is typically tens of GHz).

Superconducting circuits are described using a circuit QED formalism that originates from a rich background of cavity QED. The circuit QED (cQED) workflow for an arbitrary circuit can be roughly split into four main blocks, which are explained in more detail in the following sections:

1. definition of the circuit branches and loops;
2. definition of the independent circuit variables;

¹Time-ordering operator

²Redefining the complex displacement amplitude α and omitting the global phase ϕ yields the standard form of the displacement operator, $\hat{D}_\alpha = \exp(\alpha\hat{c}^\dagger - \alpha^*\hat{c})$.

³Equivalently, around -127 dBm for a 20 ns square pulse at 6 GHz

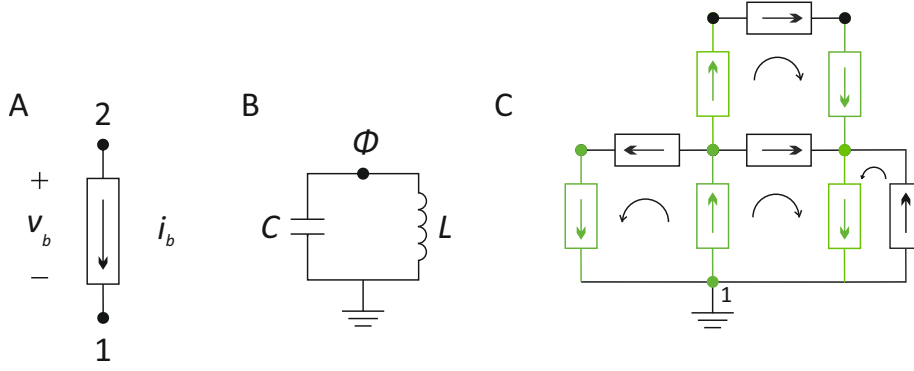


Figure 3.1: cQED conventions and basics. **A** An eclectic circuit can be divided into branches that contain a lumped element (shown as a rectangle) and connect start port 1 and end port 2. According to convention, the end port 2 has a positive voltage. The branch current i_b flows in the opposite direction, from port 2 to port 1. **B** Electric circuit of an LC oscillator. The dynamics can be described by a single independent variable Φ only. **C** A more complicated circuit can be decomposed into a set of branches in a spanning tree (green) and the remaining chords (black). The spanning tree is chosen so that each node is reachable from the reference node (1) without introducing loops. The spanning tree branches together with chords form loops (shown with curved directed arrows). Chord variables can be expressed in terms of independent spanning tree variables by applying Kirchhoff's laws to the loops.

3. writing down the Lagrangian of the circuit, obtaining the formula for conjugate variables;
4. obtaining the Hamiltonian and subjecting it to the quantisation procedure.

This section begins with a definition of the circuit variable, branch flux, and continues with a basic circuit that does not require decomposing the circuit into branches and loops — an LC oscillator. It continues with applying canonical quantisation to an LC oscillator and demonstrating how this approach scales to more complex circuits. Finally, it introduces the transmon and the black box quantisation approach to derive the Hamiltonian of the system used in this thesis. The section contains only the main points and omits some important but not relevant topics to this thesis, such as flux tunability or losses. For an in-depth theoretical description and detailed experiment-oriented overview, see [74] and [75].

3.2.1 Quantum LC oscillator

A (linear) quantum LC oscillator consists of a capacitor and an inductor connected in parallel (see Fig. 3.1 B). Due to Kirchhoff's voltage conservation law, the voltage drop over the capacitor and the inductor is equal, leaving only one independent variable. In cQED, opposed to the classical approach, the dynamics is considered from the perspective of branch fluxes rather than voltages. A branch is a lumped element⁴ that connects two

⁴the element whose physical length L is much smaller than the typical operating wavelength λ , $L \ll \lambda$

nodes (terminals), as shown in Fig. 3.1 A. It can be capacitive or inductive. By definition, the branch flux is

$$\Phi_b(t) = \int_{-\infty}^t v_b(\tau) d\tau, \quad (3.4)$$

or, equivalently, $v_b = \dot{\Phi}_b$. In this definition, branch flux is equivalent to the coordinate x of the classical mechanical harmonic oscillator. It is a convenient choice for a Josephson junction-based circuit, but not the unique one. For example, for a phase-slip junction-based circuit, it is more convenient to use the branch charge as a variable.

The Lagrangian of the circuit has the general form

$$\mathcal{L} = T(\{\dot{\Phi}_b\}) - U(\{\Phi_b\}) \quad (3.5)$$

For the linear capacitive branch, $i_b(t) = C\dot{v}_b(t)$, and the energy stored in the capacitive branch is

$$T_b(t) = \int_{-\infty}^t i_b(\tau) v_b(\tau) d\tau = \int_{-\infty}^t C v_b(\tau) \dot{v}_b(\tau) d\tau = \int_{v_b(-\infty)}^{v_b(t)} C v_b dv_b = \frac{C v_b^2(t)}{2} = \frac{C \dot{\Phi}_b^2(t)}{2} \quad (3.6)$$

Since the energy depends on the derivative of the branch flux variable, the capacitive branch stores kinetic energy in the circuit. For the linear inductive branch, $\Phi_b(t) = L i_b(t)$, and therefore, the energy stored in inductive branch is

$$U_b(t) = \int_{-\infty}^t i_b(\tau) v_b(\tau) d\tau = \int_{-\infty}^t \frac{\Phi_b(\tau)}{L} \dot{\Phi}_b(\tau) d\tau = \int_{\Phi_b(-\infty)}^{\Phi_b(t)} \frac{\Phi_b}{L} d\Phi_b = \frac{\Phi_b^2(t)}{2L} \quad (3.7)$$

and represents the potential energy of the circuit. For both 3.6 and 3.7, it is assumed that there is no electromagnetic field at $t \rightarrow -\infty$, meaning $v_b(-\infty) = \dot{\Phi}_b(-\infty) = 0$. A more general formula can be derived for a non-linear inductive branch, depending on the relation between the branch current and branch flux $i_b(t) = g(\Phi_b(t))$:

$$U_b(t) = \int_0^{\Phi_b(t)} g(\Phi_b) d\Phi_b = G(\Phi_b(t)) \quad (3.8)$$

For an LC oscillator, equation 3.5 depends on one branch variable $\Phi = \Phi_b$:

$$\mathcal{L} = T(\dot{\Phi}) - U(\Phi) = \frac{C \dot{\Phi}^2}{2} - \frac{\Phi^2}{2L} \quad (3.9)$$

Together with the conjugate variable $Q = \partial \mathcal{L} / \partial \dot{\Phi} = C \dot{\Phi}$, it leads to the Hamiltonian

$$H = Q \dot{\Phi} - \mathcal{L} = \frac{Q^2}{2C} + \frac{\Phi^2}{2L} \quad (3.10)$$

which is equivalent to the Hamiltonian of the classical mechanical harmonic oscillator with $x \equiv \Phi$, $p \equiv Q$, $k \equiv L^{-1}$, and $m \equiv C$. Once the conjugate variables are defined, one can perform the circuit quantisation:

$$\begin{aligned} \Phi, Q, H & \qquad \qquad \hat{\Phi}, \hat{Q}, \hat{H} \\ \{\Phi, Q\} = 1 & \qquad \qquad \{\hat{\Phi}, \hat{Q}\} = i\hbar \\ \dot{\Phi} = \{\Phi, H\} & \qquad \qquad \Rightarrow \quad \dot{\hat{\Phi}} = -\frac{i}{\hbar}[\hat{\Phi}, \hat{H}] \\ \dot{Q} = \{Q, H\} & \qquad \qquad \dot{\hat{Q}} = -\frac{i}{\hbar}[\hat{Q}, \hat{H}] \end{aligned}$$

which finally results in the Hamiltonian

$$\hat{H} = \frac{\hat{Q}^2}{2C} + \frac{\hat{\Phi}^2}{2L} = 4E_C \hat{q}^2 + \frac{E_L}{2} \hat{\phi}^2 \quad (3.11)$$

$$E_C = \frac{e^2}{2C}, \quad E_L = \frac{\Phi_0^2}{4\pi^2 L}, \quad (3.12)$$

with capacitive and inductive energies E_C and E_L . The Hamiltonian 3.11 is rewritten in terms of dimensionless flux $\hat{\phi} = 2\pi\hat{\Phi}/\Phi_0$ and charge $\hat{q} = \hat{Q}/2e$ ⁵. From now on, all the variables are assumed to be operators, and the operator hats will be omitted for simplicity. Finally, by introducing the annihilation operator,

$$\begin{aligned} a &= \sqrt{\frac{C}{2\hbar}}\Phi + i\sqrt{\frac{L}{2\hbar}}Q & \Phi &= \Phi_{\text{ZPF}}(a + a^\dagger) = \sqrt{\frac{\hbar Z_0}{2}}(a + a^\dagger), \quad Z_0 = \sqrt{L/C} \\ a^\dagger &= \sqrt{\frac{C}{2\hbar}}\Phi - i\sqrt{\frac{L}{2\hbar}}Q & Q &= iQ_{\text{ZPF}}(a - a^\dagger) = i\sqrt{\frac{\hbar}{2Z_0}}(a^\dagger - a) \\ [a, a^\dagger] &= 1, & [\Phi, Q] &= i\hbar\mathbb{1} \end{aligned} \quad (3.13)$$

and substituting it into 3.11, the second quantisation Hamiltonian can be derived

$$H = \hbar\omega\left(aa^\dagger + \frac{1}{2}\right), \quad (3.14)$$

$$\omega = 1/\sqrt{LC}. \quad (3.15)$$

The quantum LC oscillator is identical to the standard quantum harmonic oscillator and obeys the Heisenberg uncertainty principle $(\Delta\Phi)^2(\Delta Q)^2 \geq \Phi_{\text{ZPF}}^2 Q_{\text{ZPF}}^2 = \hbar^2/4$.

For a more complicated circuit, the Lagrangian 3.5 can be obtained in the same way. The only, but essential, difference lies in the definition of a set of independent variables. As a first step, one node of the circuit is chosen as a reference node. Then, circuit paths from the reference node to every other node are constructed without imposing any loops, see Fig. 3.1 C for an example. The set of branches on the chosen paths is called a spanning tree. Every branch on the spanning tree imposes an independent variable. The branches that are not a part of the spanning tree are called chords and defined in a way that, together with the spanning tree branches, form closed loops. Applying Kirchhoff's current and voltage laws to those loops allows expressing chord variables in terms of spanning tree variables.

⁵ $\Phi_0 = h/2e$ is the superconducting flux quantum, e is electron charge

3.2.2 Transmon

In superconducting circuits, nonlinearity is introduced via a Josephson junction, which consists of two superconducting electrodes connected by a weak link. Typically, Josephson junctions (JJs) are made out of Aluminium, and a weak link is realised by an insulating layer of Aluminium oxide between the two electrodes. The JJ dynamics is described by Josephson relations⁶

$$i_b(t) = I_c \sin \phi, \quad (3.16)$$

$$\frac{d\phi}{dt} = \frac{2\pi}{\Phi_0} v_b. \quad (3.17)$$

The first one emphasises that the supercurrent of Cooper pairs exists even in the absence of applied voltage and is defined by the phase difference ϕ between two superconducting electrodes. The second defines the evolution of the phase difference ϕ when an external potential difference is applied. The equation 3.16 also implies that there exists a maximum (critical) current I_c that can flow through the JJ. The critical current can be estimated from the room temperature resistance measurements by using the Ambegaokar-Baratoff formula [76]

$$I_c = \frac{\pi \Delta}{2e R_n}, \quad (3.18)$$

where R_n is a junction normal state resistance, and Δ is a superconducting gap. By introducing the correspondence between the branch flux and phase difference $\phi = 2\pi \Phi_b / \Phi_0$ (refer to [74]) and using equations 3.8 and 3.16, one can obtain the JJ potential energy

$$U_J = -\frac{\Phi_0 I_c}{2\pi} \cos\left(\frac{2\pi}{\Phi_0} \Phi\right) = -E_J \cos\left(\frac{2\pi}{\Phi_0} \Phi\right). \quad (3.19)$$

This result can be used to obtain the Hamiltonian of the capacitively shunted Cooper pair box, which consists of a superconducting island connected via JJ to a reservoir, see Fig. 3.2 A for the circuit schematics. The Lagrangian of the circuit is

$$\mathcal{L} = \frac{\tilde{C} \dot{\Phi}^2}{2} + \frac{C_g (\dot{\Phi} - v_g)^2}{2} + E_J \cos\left(\frac{2\pi}{\Phi_0} \Phi\right). \quad (3.20)$$

Note that the JJ by construction always has an internal capacitance C_J , which results in the modified capacitance prefactor of the first term $\tilde{C} = C_s + C_J$. Hence the corresponding conjugate variable and Hamiltonian read

$$Q = \frac{\partial \mathcal{L}}{\partial \dot{\Phi}} = C_{\text{tot}} \dot{\Phi} - C_g v_g, \quad C_{\text{tot}} = \tilde{C} + C_g, \quad (3.21)$$

$$H = 4E_C (q + n_g)^2 - E_J \cos \phi, \quad (3.22)$$

where $n_g = \frac{C_g v_g}{2e}$ is an offset charge that defines the number of Cooper pairs on the superconducting island. The Hamiltonian 3.22 eigenvalues can be found in the charge basis, see [74] for more details. Depending on the charging energy $E_C = e^2 / (2C_{\text{tot}})$, the energy levels of 3.22 show different behaviour. For $E_J / E_C \ll 1$, energy levels heavily depend on

⁶Here, Josephson relations are rewritten in terms of branch variables directly.

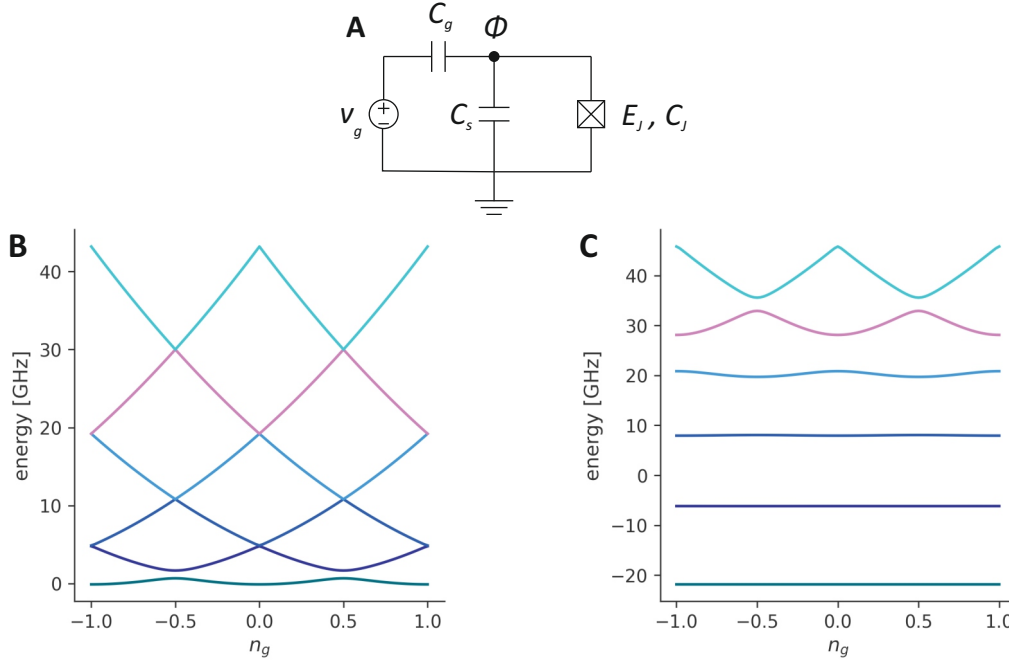


Figure 3.2: **A** Electrical circuit of the Cooper pair box, shunted by capacitance C_s . The Cooper pair box is a superconducting island between a gate capacitance C_g on one side and a JJ with inductive energy E_J and intrinsic capacitance C_J on the other side. The number of Cooper pairs on the island is controlled by the applied gate voltage v_g . **B** Energy levels of the Cooper pair box in the regime $E_J/E_C \ll 1$. The spectrum heavily depends on the offset charge n_g . **C** Energy levels of the Cooper pair box in the transmon regime $E_J/E_C \gg 1$. The lower transitions are independent of the offset charge. The spectrum is calculated with the **scQubits** Python library [77].

the offset charge, which makes them susceptible to charge noise. For $E_J/E_C \geq 20$, the charge dependence (and, therefore, charge noise susceptibility) is lifted, but it does not come without a cost. One can selectively drive transitions only when they are separated in frequency space. Increasing E_J/E_C leads to a decrease in the difference between transition frequencies (e.g. anharmonicity $\alpha = \omega_{01} - \omega_{12}$), which makes the manipulation of different transitions challenging. Typically, for the range $E_J/E_C = 20 - 80$, there is sufficient charge noise suppression, while the anharmonicity remains high enough.

The transmon is a Cooper pair box shunted by a large capacitance, $E_J/E_C \gg 1$ [40]. In this regime, the system behaves as a slightly anharmonic oscillator with the wavefunctions localised around $\phi = 0$. Expanding 3.22 around $\phi = 0$ gives

$$H = 4E_C q^2 + \frac{E_J}{2} \phi^2 - \frac{E_J}{4!} \phi^4 + O(\phi^6) \quad (3.23)$$

Note that here the linear part of the Hamiltonian is $H_{\text{lin}} = 4E_C q^2 + \frac{E_J}{2} \phi^2$ since it is essentially a Hamiltonian of a quantum LC oscillator 3.11, and the non-linear part contains the terms of the forth order and higher, $H_{\text{nl}} = -\frac{E_J}{4!} \phi^4$.

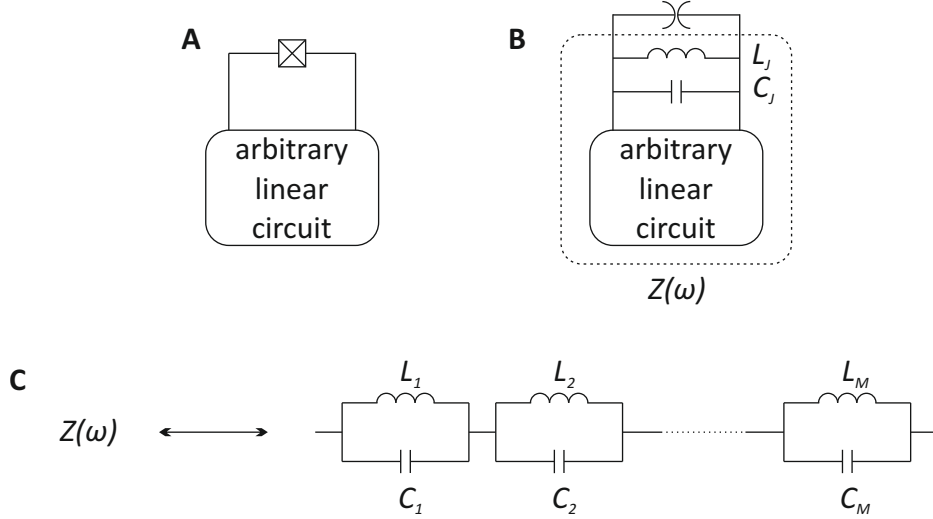


Figure 3.3: The main steps of black box quantisation. **A** In the simplest scenario, the Josephson junction is coupled to an arbitrary linear circuit. **B** The Josephson junction can be divided into linear and non-linear parts. The arbitrary linear circuit is redefined to include linear capacitance and inductance of the JJ. Now, a purely non-linear element (portrayed with a spider symbol) is coupled to the linear circuit with total impedance $Z(\omega)$. **C** According to Foster's theorem, an arbitrary linear circuit can be decomposed into a series of M parallel LC oscillators with frequencies $\omega_m = 1/\sqrt{L_m C_m}$, $m \in [1, M]$.

Introducing the annihilation and creation operators similarly to 3.13 (with replacing Z_0 by $Z_t = \sqrt{L_J/C_{\text{tot}}}$) leads to

$$H = \sqrt{8E_J E_C} b^\dagger b - \frac{E_C}{12} (b + b^\dagger)^4 \quad (3.24)$$

Finally, one can neglect the terms that do not conserve the excitation number (rotating wave approximation, which neglects fast-rotating terms in the interaction picture). The Hamiltonian in this form is called a Duffing oscillator Hamiltonian

$$H = \hbar\omega b^\dagger b - \frac{\hbar\alpha}{2} b^\dagger b^\dagger b b. \quad (3.25)$$

3.2.3 Black box quantisation

All the derivations in the previous sections are done under the assumption that the circuits consist of lumped elements only. This is not always the case: for example, in this thesis, the transmon is coupled to a 3D cavity, which cannot be approximated as a lumped element. Black box quantisation allows circumventing this problem.

From the black box quantisation perspective, the system is a set of Josephson junctions connected through their ports to an arbitrary linear circuit. In the simplest, one-port case, the circuit contains only one Josephson junction coupled to a circuit with impedance $Z(\omega)$ (see Fig. 3.3). Since the Josephson junction can be equivalently represented as a parallel

connection of a capacitor, linear and nonlinear inductors, the linear part of the JJ circuit can be included in the total circuit impedance $Z(\omega)$.

According to Foster's theorem [78], the linear part of the circuit is equivalent to a system of M LC oscillators connected in series (see Fig. 3.3 C)

$$Z(\omega) = \sum_{m=1}^M Z_m = \sum_{m=1}^M \left(j\omega C_m + \frac{1}{j\omega L_m} + \frac{1}{R_m} \right)^{-1} \quad (3.26)$$

Here, the losses are incorporated via resistances R_m . Using results from quantum LC oscillator quantisation section 3.13 and 3.14, for the linear part of the circuit, one gets

$$H_{\text{lin}} = \sum_{m=1}^M \hbar\omega_m a_m^\dagger a_m \quad (3.27)$$

The linear part of the circuit is therefore decomposed into M normal modes with resonance frequencies $\omega_m = \sqrt{L_m/C_m}$. Each of the modes embeds information about an underlying circuit via its characteristic impedance Z_m , which goes into the operator definition

$$\phi = \sum_{m=1}^M \frac{2\pi}{\Phi_0} \Phi_m = \sum_{m=1}^M \frac{2\pi}{\Phi_0} \sqrt{\frac{\hbar Z_m}{2}} (a_m^\dagger + a_m) = \sum_{m=1}^M \frac{2\pi}{\Phi_0} \varphi_m (a_m^\dagger + a_m) \quad (3.28)$$

For a simple one-port configuration, when one nonlinear element is connected to an arbitrary linear network, the flux across the junction is $\Phi_J = \Phi$. The system Hamiltonian in the normal mode basis is then

$$H = H_{\text{lin}} + H_{\text{nl}} = \sum_m \hbar\omega_m a_m^\dagger a_m + H_{\text{nl}}(\phi), \quad (3.29)$$

with the nonlinear part for the Josephson junction being

$$H_{\text{nl}} = -E_J \left[\cos \phi - \left(1 - \frac{\phi^2}{2} \right) \right] = -E_J \frac{\phi^4}{4!} + E_J \frac{\phi^6}{6!} + O(\phi^8) \quad (3.30)$$

Note that the Taylor expansion of the nonlinear part (Eq. 3.30) may not be possible for some systems [79]. Additionally, the black box approach does not account for charge dispersion, which is fortunately negligible in the transmon regime.

The power of the black box quantisation comes from the fact that it lifts the lumped element assumption and requires only the normal mode resonance frequencies ω_m and impedances (admittances) at those frequencies. Both can be derived from the classical electromagnetic simulations. More details can be found in [80].

In this thesis, the system consists of a transmon coupled capacitively to a 3D cavity and readout resonator. In most cases, the resonator does not affect the dynamics during the creation and manipulation of bosonic states and can therefore be omitted. Following 3.28, the phase operator for the transmon with a mode q coupled to the cavity with the mode c is

$$\varphi = \varphi_q (q^\dagger + q) + \varphi_c (c^\dagger + c). \quad (3.31)$$

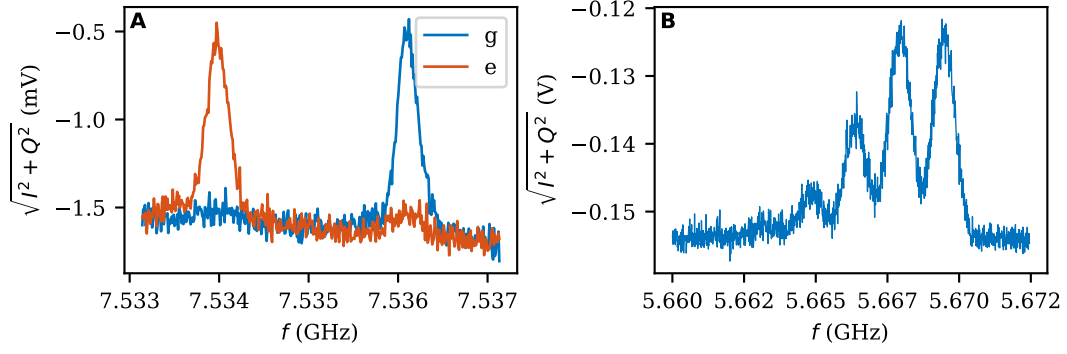


Figure 3.4: Splitting of qubit and cavity/resonator peaks due to dispersive interaction. **A** Dispersive shift of the readout resonator due to qubit excitation. **B** Number splitting of the qubit peak due to photon population in the high Q cavity.

Substituting it into the Hamiltonian 3.29, 3.30 under a rotating wave approximation leads to

$$H = \sum_{q,c} \hbar \omega_m a_m^\dagger a_m - \frac{E_J}{4!} \left(\frac{2\pi}{\Phi_0} \right)^4 (6\varphi_c^4 c^\dagger c^\dagger c c + 6\varphi_q^4 q^\dagger q^\dagger q q + 24\varphi_c^2 \varphi_q^2 c^\dagger c q^\dagger q) + O(\varphi^6) \quad (3.32)$$

Finally, one gets the black box Hamiltonian

$$\begin{aligned} H = & \hbar \omega_q q^\dagger q + \hbar \omega_c c^\dagger c - \\ & - \frac{K_q}{2} q^\dagger q^\dagger q q - \frac{K_c}{2} c^\dagger c^\dagger c c - \chi q^\dagger q c^\dagger c + \\ & + \frac{\chi_{qqc}}{2} q^\dagger q^\dagger q q c^\dagger c + \frac{\chi_{qcc}}{2} q^\dagger q c^\dagger c^\dagger c c \end{aligned} \quad (3.33)$$

where a cavity self-Kerr K_c , a self-Kerr (anharmonicity) of the qubit K_q , and qubit-cavity dispersive shift χ are defined as

$$K_{q,c} = \frac{E_J}{2} \left(\frac{2\pi}{\Phi_0} \right)^4 \varphi_{q,c}^4 \quad (3.34)$$

$$\chi_{qc} = E_J \left(\frac{2\pi}{\Phi_0} \right)^4 \varphi_q^2 \varphi_c^2 \quad (3.35)$$

The second line in 3.33 contains the terms that correspond to the expansion of 3.30 up to φ^4 , and the last line to φ^6 . Usually, 6th-order terms are negligibly small. The interaction term of the Hamiltonian $H_{\text{int}} = -\chi q^\dagger q c^\dagger c$ is a necessary component for qubit-cavity state manipulation: if the qubit (cavity) is excited, the cavity (qubit) frequency experiences a shift depending on the excitation number $q^\dagger q$ ($c^\dagger c$) (see Fig. 3.4), giving rise to a set of conditional and unconditional operations. The qubit-cavity dispersive shift determines the rate at which a large set of operations can be done, along with the bandwidth of control pulses.

3.3 Experimental characterisation of the Hamiltonian

The experimental characterisation of the Hamiltonian [3.33](#) parameters starts with:

- spectroscopy measurements on the qubit, cavity, and readout resonator to determine their frequencies;
- Rabi/Ramsey qubit measurements to calibrate qubit rotations and correct for qubit frequency detuning;
- adjusting the readout parameters to get the best readout contrast. In this setup, a non-QND high-power readout [\[81\]](#) is used; see the Appendix [A](#) for more details.

This is a classic set of techniques for an initial characterisation. To avoid recurrence, those measurements are omitted in this section, but can be found, for example, in [\[82\]](#) or [\[80\]](#).

Higher-order nonlinearities can be obtained from qubit and cavity revival experiments. A stable fit of the data from revival experiments requires a preliminary calibration of the displacement scaling. The pulse sequence of displacement scaling calibration consists of a cavity displacement with varied amplitude, followed by a selective π -pulse at the number split peak $n = 0, 1$, etc. (see Fig. [3.5](#)). In case the thermal population in the cavity is not negligible, the photon number probability distribution satisfies the formula

$$\begin{aligned}
 P_n(\alpha) &= \sum_k^{n+1} \frac{n_{th}^k}{(n_{th} + 1)^{k+1}} P_n^k(\alpha) \\
 P_n^l(\alpha) &= \frac{n!}{l!} \alpha^{2(l-n)} e^{-|\alpha|^2} (L_n^l(|\alpha|^2))^2 \\
 L_n^l(x) &= \begin{cases} \mathcal{L}_n^{(l-n)}(x), & l \geq n \\ \frac{l!}{n!} (-x)^{n-l} \mathcal{L}_n^{(n-l)}(x), & l < n \end{cases}
 \end{aligned} \tag{3.36}$$

with $\mathcal{L}_n^{(\beta)}$ is denoted as generalised Laguerre polynomials, and α - as a cavity displacement value. For negligible cavity thermal population, the formula [3.36](#) simplifies to

$$P_n(\alpha) = e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!} \tag{3.37}$$

Fitting all the curves in Fig. [3.5](#) with the same set of fitting parameters allows the extraction of the conversion factor between AWG amplitude and cavity displacement α .

The pulse sequence of the cavity revival experiment is shown in Fig. [3.6](#). The sequence begins with two displacements by $+\alpha$ and $-\alpha$, separated by a variable time delay t . During the variable delay, the coherent state $|\alpha\rangle$ rotates in phase space due to the cavity nonlinear contributions in the Hamiltonian. Each time the acquired coherent state phase is a multiple of 2π , the final selective X_π^0 pulse fully flips the qubit population, resulting in a "revival" in the measured data. The revival frequency is typically very low, which is why artificial phase oscillations at the frequency ω are added to the second displacement pulse. It allows for

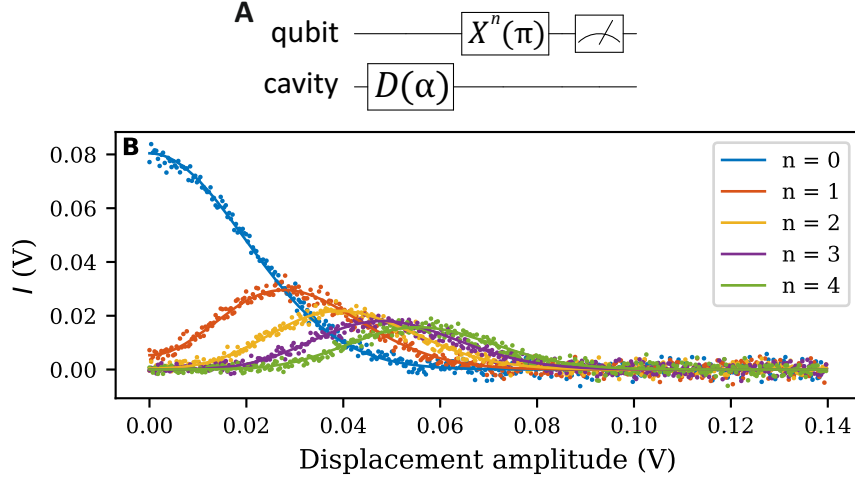


Figure 3.5: Displacement scaling calibration. **A** The pulse sequence consists of a cavity displacement of varying amplitude, followed by a selective qubit pulse $X^n(\pi)$. To get rid of the cavity-resonator cross-Kerr that biases the readout for large α , a background measurement (i.e. with no pulse applied instead of $X^n(\pi)$) is subtracted from the original measurement. **B** The measurements (dots) for different n are fitted together (solid lines) with the formula 3.36 to obtain both the scaling and thermal population in the cavity. The residual thermal cavity population is $n_{th} = 0.071 \pm 0.004$.

observing the revivals when the phase acquired due to nonlinear Hamiltonian contributions is $\varphi = \omega t$, significantly decreasing the required range of time delays for a stable fit.

For the sequence starting with the qubit in $|g\rangle$, the probability of finding the cavity in $|0\rangle$ at the end of the measurement sequence in the interaction picture⁷ is

$$\begin{aligned}
 |\alpha(t)\rangle &= C_t D_\alpha |0\rangle = e^{-i\tilde{H}t} |\alpha\rangle = |\alpha e^{-i\frac{K_c}{2}|\alpha|^2 t}\rangle, \quad C_t = e^{-i\tilde{H}t} \\
 P &= |\langle 0 | D_{-\alpha e^{i\omega t}} C_t D_\alpha | 0 \rangle|^2 = |\langle -\alpha e^{i\omega t} | \alpha(t) \rangle|^2 = |\exp(-|\alpha|^2(1 + e^{-i\omega t - i\frac{K_c}{2}|\alpha|^2 t}))|^2 = \\
 &= \exp\left(-2|\alpha|^2\left(1 + \cos\left(\omega t + \frac{K_c}{2}|\alpha|^2 t\right)\right)\right) = \exp\left(-2|\alpha|^2(1 + \cos \tilde{\omega} t)\right) \quad (3.38)
 \end{aligned}$$

with $\tilde{\omega} = \omega + \frac{K_c}{2}|\alpha|^2$. The frequency of the revival pattern $\tilde{\omega}$ depends on the initial state of the qubit. For the sequence starting with the qubit in $|e\rangle$, the revival frequency is modified to

$$\tilde{\omega} = \omega + \chi + \frac{K_c}{2}|\alpha|^2 + \frac{\chi_{qcc}}{2}|\alpha|^2 \quad (3.39)$$

Even higher nonlinear terms modify $\tilde{\omega}$ in a similar way. In addition, if the cavity drive is detuned by Δ , the revival frequency acquires an additional offset $\tilde{\omega} \rightarrow \tilde{\omega} + \Delta$.

⁷i.e. under the transformation $U = \exp(it(\omega_q q^\dagger q + \omega_c c^\dagger c))$, so that $\tilde{H} = U H U^\dagger + i\hbar \dot{U} U^\dagger = -\frac{K_q}{2} q^\dagger q^\dagger q q - \frac{K_c}{2} c^\dagger c^\dagger c c - \chi q^\dagger q c^\dagger c + O(\phi^6)$

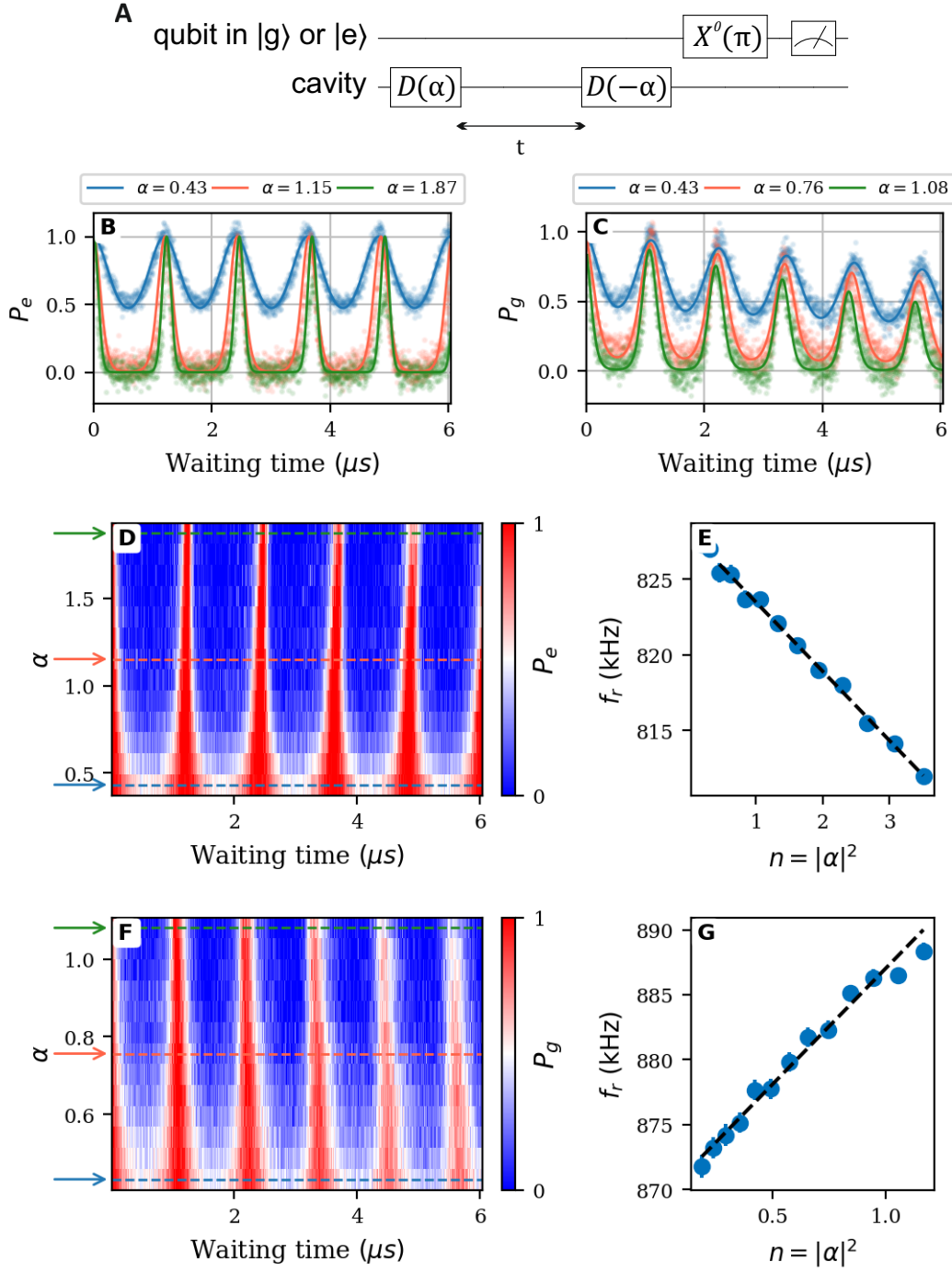


Figure 3.6: Cavity revival experiment. **A** Pulse sequence consists of two displacement pulses separated by a waiting time in between, followed by a selective qubit pulse. To get rid of the cavity-resonator cross-Kerr that biases the readout for large α , a background measurement (i.e. with no pulse applied instead of $X^0(\pi)$) is subtracted from the original measurement. Typically, an artificial phase is applied to the second displacement pulse (e.g. $D(-\alpha e^{i\omega t})$ instead of $D(-\alpha)$ for each delay t) to induce faster cavity revivals. For this set of measurements, it results in 5 artificial oscillations. **B**, **C** Cavity revival measurements starting with qubit in $|g\rangle$ and $|e\rangle$ for selected displacement values (dots) and corresponding fits with Eq. 3.38 (lines). **C**, **D** Cavity revival maps for $|g\rangle$ and $|e\rangle$ initial states of the qubit. Cuts through α -axis are fitted with the formula 3.38 to get the frequency of the revivals. **E**, **F** Fitted frequency of revivals $f_r = \tilde{\omega}/2\pi$ (blue dots) as a function of the number of photons $n = |\alpha|^2$. K_c and χ_{qcc} , as well as the detuning from the true cavity frequency, are calculated from the fit (black line) using Eq. 3.39.

First- and second-order dispersive shifts can be either derived from the number-splitting experiment (see Fig. 3.4) for different coherent states or from the qubit revival experiment. The advantage of the qubit revival experiment is that it gives a perfectly calibrated waiting time for future Wigner tomography measurements by taking into account the effect of non-instantaneous qubit pulses, and, therefore, saves one calibration step.

The qubit revival sequence is shown in Fig. 3.7 A. After the first displacement and $\pi/2$ pulse, the state of the system is

$$|\psi_1\rangle = R_x(\pi/2)D_\alpha|g, 0\rangle = \frac{1}{\sqrt{2}}(|g, \alpha\rangle - i|e, \alpha\rangle)$$

which further evolves in the phase space similar to the cavity revival case,

$$|\psi_2\rangle = e^{-i\tilde{H}t}|\alpha\rangle|\psi_1\rangle = \frac{1}{\sqrt{2}}(|g, \alpha\rangle - i|e, \alpha e^{-i\tilde{\omega}t}\rangle),$$

$$\tilde{\omega} = \chi + n\frac{\chi_{qcc}}{2}.$$

After the final qubit rotation,

$$|\psi_3\rangle = R_x(\pm\pi/2)|g, 0\rangle = \frac{1}{2}(|g\rangle(|\alpha\rangle \mp |\alpha e^{-i\tilde{\omega}t}\rangle) - i|e\rangle(\pm|\alpha\rangle + |\alpha e^{-i\tilde{\omega}t}\rangle)) \quad (3.40)$$

$$P_{g,\pm} = |\langle g, \alpha|\psi_3\rangle|^2 = \frac{1}{4}(2 \mp \langle \alpha|\alpha e^{-i\tilde{\omega}t}\rangle \mp \langle \alpha e^{-i\tilde{\omega}t}|\alpha\rangle)$$

$$= \frac{1}{4}(2 \mp \exp(|\alpha|^2(e^{-i\tilde{\omega}t} - 1)) \mp \exp(|\alpha|^2(e^{i\tilde{\omega}t} - 1)))$$

$$= \frac{1}{4}(2 \mp e^{|\alpha|^2(\cos \tilde{\omega}t - 1)}[\exp(0.5|\alpha|^2(e^{-i\tilde{\omega}t} - e^{i\tilde{\omega}t})) + \exp(0.5|\alpha|^2(e^{i\tilde{\omega}t} - e^{-i\tilde{\omega}t}))])$$

$$= \frac{1}{2}(1 \mp e^{|\alpha|^2(\cos \tilde{\omega}t - 1)} \cos(|\alpha|^2 \sin \tilde{\omega}t)).$$

Subtracting two measurement results with opposite signs of the final qubit rotation improves the contrast by a factor of 2

$$P_{g,-} - P_{g,+} = e^{|\alpha|^2(\cos \tilde{\omega}t - 1)} \cos(|\alpha|^2 \sin \tilde{\omega}t) \quad (3.41)$$

The qubit revival measurement at a specific delay value $t = \pi/\chi$ is also called a parity measurement since it maps the parity of the cavity state on the state of the qubit. It can be seen by rearranging $|\psi_3\rangle$ in the equation 3.40 for $|g\rangle$ - and $|e\rangle$ -branches:

$$|g\rangle : \quad |\alpha\rangle \mp |\alpha e^{-i\tilde{\omega}t}\rangle = e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n (1 \mp e^{-i\tilde{\omega}tn})}{\sqrt{n!}} |n\rangle \quad (3.42)$$

$$|e\rangle : \quad \pm|\alpha\rangle + |\alpha e^{-i\tilde{\omega}t}\rangle = e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n (e^{-i\tilde{\omega}tn} \pm 1)}{\sqrt{n!}} |n\rangle \quad (3.43)$$

It is clear that the prefactor for the $|g\rangle$ -branch at $t = \pi/\tilde{\omega}$ cancels out for even n , and for the $|e\rangle$ -branch - for odd n . The parity sequence in the ideal case has an equivalent operator representation $P = e^{i\pi c^\dagger c} = (-1)^n$. The imperfections in the parity measurement may come from:

- higher-order nonlinearities, that lead to partial cancellation of the even and odd terms 3.42,
- cross-Kerr between the readout resonator and the cavity, which leads to a change in the readout contrast,
- qubit frequency shift due to excitations in the cavity that may lead to off-axis $\pi/2$ -rotations on the Bloch sphere,
- qubit dephasing during the waiting time.

The first three of these affect the contribution of high-photon-number terms to the total parity measurement result, therefore becoming a problem only for bosonic states with high-photon-number populations.

The system parameters are summarised in the Table 3.1. For the experimental techniques used for characterisation but not mentioned in this section, refer to [80] and [82]. The cryogenic and room-temperature microwave setup has a standard configuration for cQED; see Appendix A for schematics.

Parameter	Value	Measurement type
Resonator frequency	7.5284898(4) GHz	spectroscopy
Resonator coupling κ_c	115.2 ± 0.9 kHz	spectroscopy
Resonator linewidth κ_L	148.2 ± 1.6 kHz	spectroscopy
Qubit ge frequency	5.66936(9) GHz	spectroscopy
Qubit ef frequency	5.479845(7) GHz	spectroscopy
Qubit anharmonicity	189.5 ± 0.1 MHz	calculated
Qubit temperature, mK	149.1 ± 1.5 mK	g-e and e-f Rabi
Qubit T_1	31.6 ± 0.2 μs	T_1 measurement
Qubit T_2	12.3 ± 0.2 μs	Ramsey
Qubit T_2 Echo	14.9 ± 0.4 μs	Echo measurement
Cavity frequency	4.5444444(6) GHz	spectroscopy
Cavity T_1^c	59.0 ± 0.7 μs	cavity T_1 measurement
Cavity self-Kerr $K/2$	4.6 ± 0.1 kHz	cavity revivals
Cavity thermal population n_{th}	0.071 ± 0.004	displacement calibration
Cavity displacement scaling	14.41 ± 0.03	displacement calibration
Qubit-cavity dispersive shift χ	1.514 ± 0.001 MHz	qubit revivals
Second order qubit-cavity dispersive shift $\chi_{qcc}/2$	11.58 ± 0.97 kHz 11.18 ± 0.40 kHz	qubit revivals cavity revivals

Table 3.1: Experimentally characterised Hamiltonian parameters of the system.

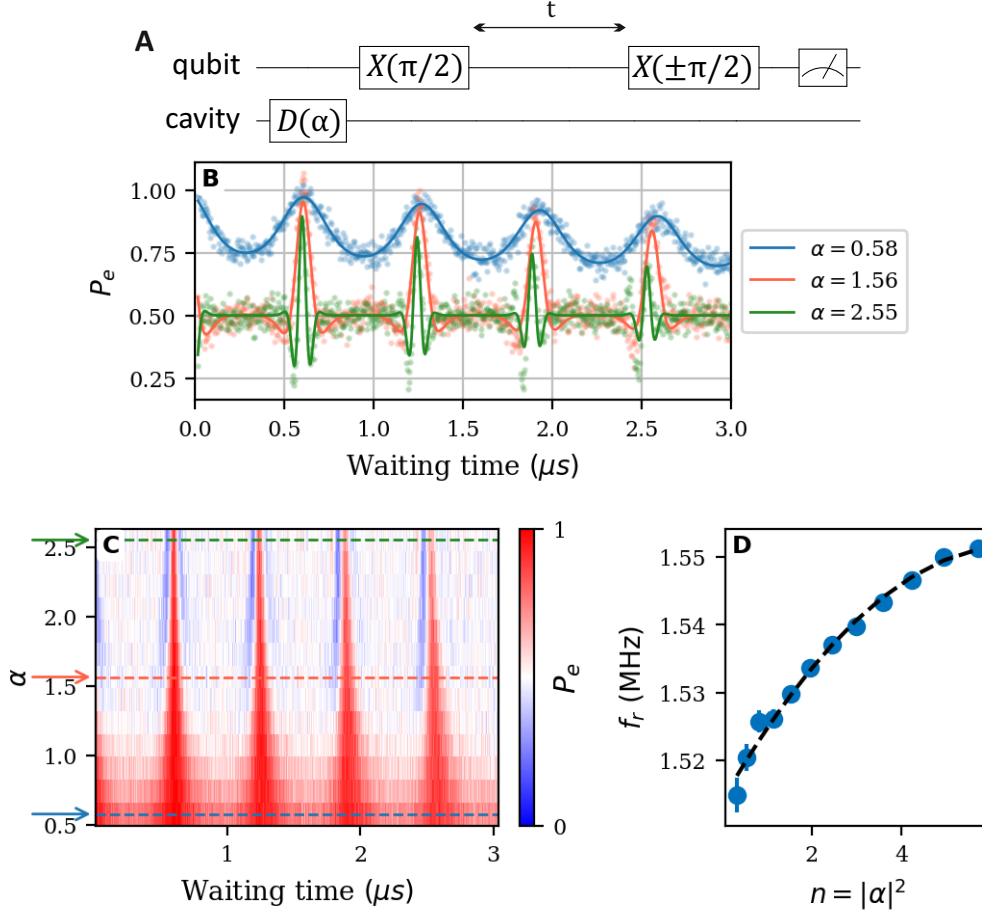


Figure 3.7: Qubit revival experiment. **A** Pulse sequence consists of a displacement pulse, followed by two $\pi/2$ -pulses separated by a waiting time in between. To get rid of the cavity-resonator cross-Kerr that biases the readout for large α and improve the contrast, a background measurement (i.e. with a $\pi/2$ -rotation of the opposite sign) is subtracted from the original measurement. **B** Qubit revival measurements for selected displacement values (dots) and corresponding fits with Eq. 3.41 (lines). **C** Qubit revival map, which shows the qubit excited state population P_e as a function of the waiting time between the two $\pi/2$ -pulses and the first displacement value. Cuts through α -axis are fitted with the formula 3.41. **D** Frequency of revivals $f_r = \tilde{\omega}/2\pi$ (blue dots) as a function of the number of photons $n = |\alpha|^2$. χ and χ_{qcc} are obtained from the fit with the formula $\tilde{\omega} = \chi_{qc} + n\chi_{qcc}/2 + \chi''n^2/6$ (black line).

3.4 The necessary set of operations

There are many approaches to generating bosonic states. However, if the target bosonic state is not created with the waveforms optimised in an open-loop, it boils down to applying a sequence of qubit-cavity gates. The set of possible gates can be divided into three main categories:

1. Unconditional operations on the qubit (cavity). They are realised with short, high amplitude pulses that can cover dispersively-shifted (number-splitted) peaks in frequency space (see Fig. 3.8, **A**, **C**).
2. Conditional operations on the qubit (cavity). They are obtained with long, low-amplitude pulses that are narrow in frequency space. Cavity-selective displacements are usually applied to the vacuum state (Fig. 3.8 **B**). The pulse bandwidth in this case should be smaller than χ . For qubit-selective pulses acting on states of the form $\propto |g, \alpha_1\rangle + |e, \alpha_2\rangle$, the pulse bandwidth must be sufficiently wide to cover the entire frequency range associated with the photon-number distribution of the targeted component (e.g., $|g, \alpha_1\rangle$), while avoiding unintended excitation of the tail of the distribution corresponding to the other component ($|e, \alpha_2\rangle$). Typically, before applying a qubit-selective pulse, the $|e\rangle$ - (or $|g\rangle$ -) branch of the state $|g, \alpha_1\rangle + |e, \alpha_2\rangle$ is displaced to vacuum, resulting in the state $\propto |e, 0\rangle + |g, \alpha_1 - \alpha_2\rangle$ (see Fig. 3.8 **D**). Since $|e, 0\rangle$ corresponds to a single peak in frequency space, the pulse bandwidth can be chosen more freely.
3. Entangling operations. The easiest to calibrate is an entangling conditional phase gate (Fig. 3.8 **E**). Due to the interaction term $-\chi c^\dagger c q^\dagger q$ in the Hamiltonian, it is equivalent to waiting for a time $t = \phi/(\chi + n\chi_{qcc})$ to acquire the desired phase ϕ . More complicated entangling operations, such as ECD [65] and SNAP [31], are presented in Fig. 3.9. The SNAP gate is discussed in more detail in section 4.4.1 as a part of Kerr correction techniques.

This list of operations forms the minimal necessary set for creating arbitrary bosonic states.

3.5 Tomography and fidelity of bosonic states

In quantum optics, there are three canonical ways of describing the bosonic states: with Glauber P-, Husimi Q-, or Wigner quasi-distributions. The prefix "quasi" comes from the fact that although the distributions are normalised to 1 over the complex phase space, they do not necessarily obey the other probability axioms: they can take negative values in the phase space, and do not satisfy the axiom of σ -additivity.

From a theoretical perspective, quasi-probability distributions originate from the following characteristic functions

$$\chi_P(\lambda) = \text{tr}(\rho e^{i\lambda^* c^\dagger} e^{i\lambda c}) \quad (3.44)$$

$$\chi_Q(\lambda) = \text{tr}(\rho e^{i\lambda c} e^{i\lambda^* c^\dagger}) \quad (3.45)$$

$$\chi_W(\lambda) = \text{tr}(\rho e^{i\lambda^* c^\dagger + i\lambda c}) \quad (3.46)$$

that correspond to normal, anti-normal and symmetric (Weyl) operator ordering. Displacement operator corresponds to the symmetric operator ordering 3.46, and, therefore, its expectation value is a value of the characteristic function χ_W itself.

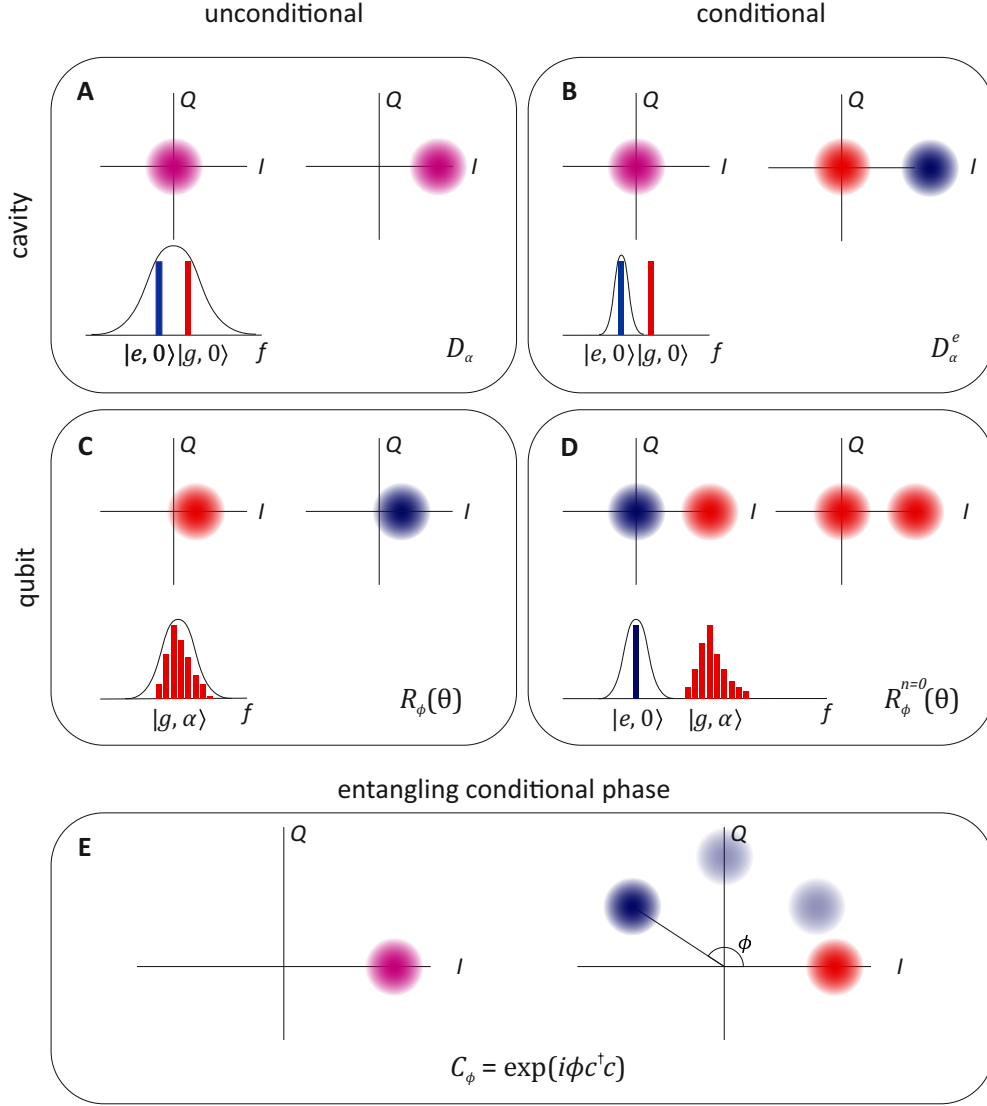


Figure 3.8: The minimal set of cavity-qubit operations for bosonic state creation protocols. Red colour corresponds to the qubit in $|g\rangle$, blue to the qubit in $|e\rangle$, and purple to the superposition of $|g\rangle$ and $|e\rangle$. Each operation is described within a single rectangle and is structured as follows: the upper-left picture shows the state in phase space before the operation, the upper-right picture shows the state after the operation is applied, and the bottom-left picture shows the population in frequency space. The black Gaussian-shaped line illustrates the Fourier transform of the applied pulse. All the pictures are not to scale and are intended to illustrate the idea only. **A** Unconditional (unselective) cavity displacement D_α . **B** Conditional (selective) cavity displacement $D_\alpha^e = D_\alpha|e\rangle\langle e| + I_c|g\rangle\langle g|$. D_α^e can be used at the beginning of any sequence to get rid of a parasitic excited state population of the qubit. **C** Unconditional (unselective) qubit rotation $R_\phi(\theta)$ around axis ϕ for an angle θ . **D** Qubit rotation conditioned on $n=0$ photons in the cavity $R_\phi^0 = |0\rangle\langle 0|R_\phi(\theta) + (I_c - |0\rangle\langle 0|)I_q$. Note that the rotation can be conditioned on any photon number, $R_\phi^n = |n\rangle\langle n|R_\phi(\theta) + (I_c - |n\rangle\langle n|)I_q$, but it would require an even stronger restriction on the pulse width. **E** Entangling conditional phase gate C_ϕ . Due to qubit-cavity dispersive interaction, $|e, \alpha\rangle$ acquires a phase ϕ in the phase space, while $|g, \alpha\rangle$ stays on axis.

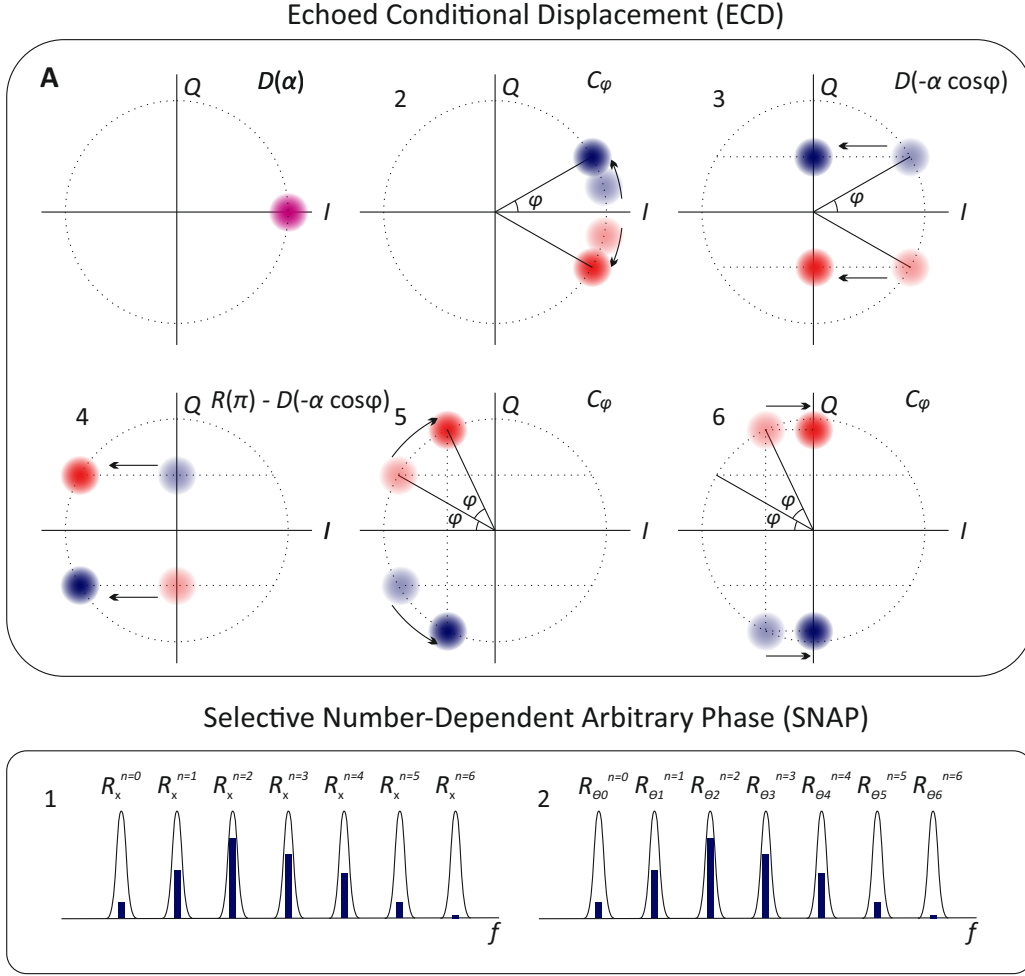


Figure 3.9: Entangling operations that require more extensive calibration procedures. Red colour corresponds to the qubit in $|g\rangle$, blue to the qubit in $|e\rangle$, and purple to the superposition of $|g\rangle$ and $|e\rangle$. **A** ECD gate steps in phase space. The gate sequence starts with displacing the state $(|g\rangle + |e\rangle)/\sqrt{2}$ by α at frequency $(\omega_c^g + \omega_c^e)/2$ (1). Since the displacement drive is detuned from both g and e , the states rotate in opposite directions in phase space during the idle time t (2). The sequence proceeds with two displacements by $-\alpha \cos \varphi$, $\varphi = \chi t/2$, with an unselective π -pulse in between (3-4). Waiting for the system to evolve over time t increases the separation of branches in phase space (5). A final displacement returns $|g\rangle$ and $|e\rangle$ to the Q -axis (6). At the end of the gate, the states are separated by $\beta = 2i\alpha \cos \varphi$. **B** SNAP gate action in frequency space. The gate consists of first applying selective qubit π -rotations around the x -axis at the frequencies corresponding to number-split peaks $\omega - \chi n - \chi_{qcc} n^2$ (1), and then applying selective qubit π -rotations around θ_n -axes calibrated for each peak n (2), see Section 4.4.1 for more details.

The quasi-probability distributions are defined as the Fourier transform of the corresponding characteristic functions:

$$P(\beta) = \frac{1}{\pi^2} \int \chi_P(\lambda) e^{\beta\lambda^* - \beta^*\lambda} d^2\lambda \quad (3.47)$$

$$Q(\beta) = \frac{1}{\pi^2} \int \chi_Q(\lambda) e^{\beta\lambda^* - \beta^*\lambda} d^2\lambda \quad (3.48)$$

$$W(\beta) = \frac{1}{\pi^2} \int \chi_W(\lambda) e^{\beta\lambda^* - \beta^*\lambda} d^2\lambda \quad (3.49)$$

In cQED, bosonic states are usually characterised with Wigner and, less frequently, Husimi Q distributions. The P-distribution requires the measurement of the normal order moments, which is tricky for the cavity field. The situation is opposite for propagating microwave fields. Normally ordered moments $\langle (a^\dagger)^m a^n \rangle$ can be directly obtained from the cross-correlation measurements between output complex amplitudes in the two-channel microwave detection scheme [83].

Reconstruction of the bosonic states with characteristic functions instead of quasi-probability distributions is typically more intricate. However, it is the only reconstruction option for the number-unresolved systems [65].

3.5.1 Wigner and Wigner-based fidelity estimations

The Wigner function 3.49 has a well-known expression in cQED [84]. Switching to a continuous eigenbasis $\{|x\rangle\}$ of the quadrature operator $X = (c + c^\dagger)/2$ after some mathematical transformations [85] leads to

$$W(\beta) = \frac{2}{\pi} \text{tr}(D_\beta^\dagger \rho D_\beta P) = \frac{2}{\pi} \langle P_\beta \rangle, \quad \beta = x + ip \quad (3.50)$$

Therefore, the Wigner function value at each point β of phase space can be obtained by first displacing the state by $-\beta$ and then measuring the parity P_β , see Fig. 3.10 A for the pulse sequence.

It follows from Eq. 3.50 that the density matrix of the bosonic state is

$$\rho = 2 \int W(\beta) P_\beta d^2\beta \quad (3.51)$$

Therefore, ρ can be obtained by density matrix reconstruction techniques. However, in most cases, the value of interest is not the density function itself, but the fidelity of the state, which is much easier to calculate. The fidelity between the target bosonic state ψ_T with the Wigner function W_T and the measured bosonic state with the measured Wigner function W_M can be calculated as

$$F = \langle \psi_T | \rho_M | \psi_T \rangle = \frac{1}{\pi} \int W_T(\beta) W_M(\beta) d^2\beta. \quad (3.52)$$

The fidelity 3.52 is defined for the pure target state ψ_T , while the measured state can be either pure or mixed. In the experiments, strictly speaking, the state cannot be considered pure due to noise and coupling to an environment. For example, bosonic states in this thesis are not pure as a result of a significant thermal population in the cavity. Due to experimental limitations, it sometimes makes sense to calculate the fidelity with respect to the best achievable quantum state in the current setup, which can be mixed as well. In such a case, the fidelity definition 3.52 should be generalised. There is no strict rule on how to characterise a mixed state; for the full review on this topic, please refer to [86]. However, the most popular fidelity estimates are F_1 fidelity that corresponds to a maximum transition probability between the purifications of density matrices

$$F_1 = \left(\text{tr} \sqrt{\sqrt{\rho_T} \rho_M \sqrt{\rho_T}} \right)^2, \quad (3.53)$$

and F_2 fidelity that normalises the usual fidelity definition with respect to the lowest purity of participating states

$$F_2 = \frac{\text{tr}(\rho_T \rho_M)}{\max[\text{tr}(\rho_T^2), \text{tr}(\rho_M^2)]}. \quad (3.54)$$

F_1 requires density matrix reconstruction, while F_2 needs only Wigner functions of target and measured states, which makes it easier to use in most of the experiments. One has to take into account that the F_2 fidelity always gives a lower value: $\forall \rho_T, \rho_M : F_2(\rho_T, \rho_M) \leq F_1(\rho_T, \rho_M)$.

The fidelity estimate depends on the measured "window" in phase space and the phase space grid size. In case the experimental quantum state is known and close to a theoretically expected one, it may be beneficial to sample phase space measurement points from the distribution [3, 61]

$$p(\beta) \propto |W_{|\psi_T\rangle\langle\psi_T|}(\beta)| \quad (3.55)$$

and calculate the fidelity estimate as

$$F = \frac{\pi}{N} \sum_i \frac{W_{|\psi_T\rangle\langle\psi_T|}(\beta_i)}{p(\beta_i)} W_M(\beta_i). \quad (3.56)$$

Although this approach requires fewer phase space samples, it avoids sampling in the supposedly "blank" areas of phase space. In an experiment, such areas can show some parasitic features, leading to higher fidelity estimates compared to the actual state fidelity.

3.5.2 Husimi Q

By substituting Eq. 3.45 into Eq. 3.48 and using a closure relation on a coherent states basis $\int d^2\beta |\beta\rangle\langle\beta| = \pi$, it can be shown that the Husimi Q function gives an overlap with the coherent state β :

$$Q(\beta) = \frac{1}{\pi} \langle\beta|\rho|\beta\rangle \quad (3.57)$$

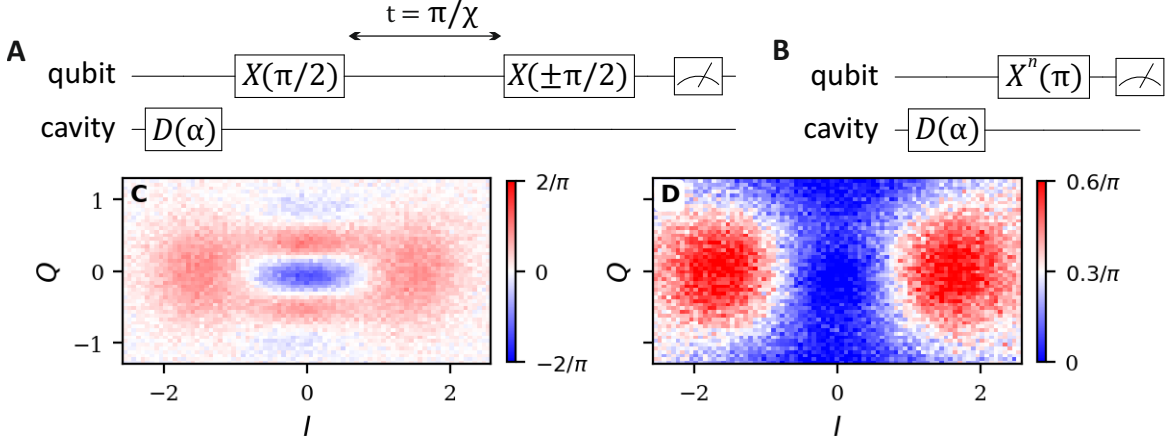


Figure 3.10: Tomography of an $\alpha = 1.5$ cat state (see Section 4.2). **A** Pulse sequence for Wigner measurement. **B** Pulse sequence for generalised Husimi Q measurement. **C** Wigner tomography of the cat state. **D** Husimi Q tomography of the cat state. Compared to the Wigner tomography, the quantum features of the state are not pronounced.

or, equivalently, an overlap between the displaced bosonic state and the vacuum state $|0\rangle$, since $Q = \frac{1}{\pi} \langle 0 | D_\beta^\dagger \rho D_\beta | 0 \rangle$. Compared to the Wigner function, Husimi Q gives less information about the quantum features of the state (see Fig. 3.10 C and D). For example, a cat state and a mixture of coherent states look the same in the Husimi Q representation. It has an intuitive explanation [85]. The anti-normal-ordered characteristic function is exponentially suppressed compared to the symmetrically-ordered one: $\chi_Q = e^{-|\lambda|^2/2} \chi_W$. Because of this, the Fourier transform of the χ_Q function (Husimi Q distribution) is less sensitive, especially around 0, to quantum features than the Fourier transform of the χ_W function (Wigner distribution).

Non-sensitivity to quantum features can be solved by measuring generalised Husimi Q functions, which give an overlap with higher Fock states $|n\rangle$ [2]

$$Q_n(\beta) = \frac{1}{\pi} \langle n | D_\beta^\dagger \rho D_\beta | n \rangle. \quad (3.58)$$

Generalised Husimi Q functions allow the reconstruction of the Wigner function [2]:

$$W(\beta) = \frac{2}{\pi} \sum_n (-1)^n \langle n | D_\beta^\dagger \rho D_\beta | n \rangle = 2 \sum_n (-1)^n Q_n(\beta) \quad (3.59)$$

as well as the density matrix

$$Q_n(\beta) = \frac{1}{\pi} \sum_{i,j} M_{nmij} \langle i | \rho | j \rangle \quad (3.60)$$

$$M_{nmij} = \langle n | D_\beta^\dagger | i \rangle \langle j | D_\beta | m \rangle \quad (3.61)$$

by solving a set of linear equations with respect to M_{nmij} . This makes the estimation of fidelity more experimentally straightforward for the F_1 fidelity (Eq. 3.53) in case a set of measured generalised Husimi Q functions is available. In most cases, the Wigner

function measurement is preferable since it requires less measurement time and gives more information about the quantum state. However, in some cases, measuring the Husimi Q function provides an advantage. For example, T. Krisnanda et al [4] suggest optimising the excitation number $|n\rangle$ to measure a set of tomography displacements $|\alpha_k\rangle$, $k \in [0, D^2 - 1]$, where D is a predefined truncated dimension. Then any bosonic state within the truncated Hilbert space can be efficiently reconstructed with $D^2 - 1$ measurements of displaced generalised Husimi Q functions Q_n by acquiring the least-squares estimator and applying a Bayesian inference approach. The reconstruction technique was named ORENS (optimised reconstruction with excitation number sampling). The authors show that the reconstruction fidelity for Fock states up to $D = 6$ exceeds 0.95. The method was benchmarked for Fock states, their superpositions, and small $\alpha = 1$ cat states. The reconstruction fidelity for bigger and distorted cat states was not investigated. By examining examples of suggested ORENS sampling points, one might infer that capturing distortions and differentiating more complex, entangled states in high-dimensional Hilbert spaces would result in a significant decrease in reconstruction fidelity. In addition, even though ORENS is capable of reconstructing any state within the truncated Hilbert space, it does not guarantee that the choice of measurements is optimal for a specific state. Replacing Husimi Q with Wigner measurements yields a comparable reconstruction fidelity; however, the Wigner measurement requires twice as many measurements as Husimi Q (see Fig. 3.10 A, B).

Summary

Bosonic states are generated in the modes of high-Q superconducting cavities. For the creation and manipulation of bosonic states, a nonlinear element, like a transmon, should be coupled to the cavity. By applying a sequence of conditional, unconditional and entangling operations on the qubit and the cavity, an arbitrary state can be created in the bosonic mode. The state fidelity or density matrix can be calculated from the Wigner or generalised Husimi Q distributions measurements.

Quantum optimal control of cat states

Over the past years, significant progress has been made in the development of bosonic quantum error correction schemes [22]. Different codes rely on distinct families of bosonic logical states, and achieving high-fidelity preparation of these states is essential for fault-tolerant quantum computation. This Chapter provides an overview of recent advances in bosonic state generation (Section 4.1), with particular emphasis on one of the most extensively used bosonic encodings - the cat states - discussed in Secs. 4.2 to 4.3. Among the primary challenges in cat state preparation are distortions arising from the Kerr effect. Techniques for Kerr compensation are reviewed in Section 4.4, and the first steps towards the demonstration of closed-loop Kerr correction are presented in Section 4.5.

4.1 Bosonic states: challenges and progress outlook

A wide variety of high-fidelity bosonic states are currently created through open-loop optimisation. The approaches can be roughly divided into three main branches¹:

- waveform-based [34], where the pulse waveform is optimised directly on a piece-wise basis with the gradient-based algorithm;
- SNAP-based [70], where the state creation sequence consists of N SNAP gates $S(\theta)$ interleaved with $N + 1$ displacements D_α . The optimisation is carried out in two stages: first, the gate parameters (θ_i, α_i) are optimised using gradient descent, and second, the pulse envelopes are corrected based on the Hamiltonian parameters. In the first stage, the sequence length is fixed for each simulation and decreased step by step for each new round of simulations until the target state fidelity becomes unsatisfactory. In the original theoretical proposal, there is an additional step for optimising the circuit depth.
- ECD-based [65], where the state creation sequence consists of N blocks of ECD(β)-gates and unselective qubit rotations $R_\phi(\theta)$. It is a two-step optimisation as well: first, the gate parameters $(\beta_i, \phi_i, \theta_i)$ and the sequence depth are optimised with a gradient descent algorithm, and, second, the ECD-gate parameters are optimised to

¹According to Yehan Liu [60] and Zhang et al. [87], creating the bosonic state through dissipative engineering can be categorised as driven-dissipative or Markovian feedback. That would add a feedback-based approach to the list. However, dissipative engineering is usually not considered as a type of feedback (and, therefore, adaptive control) by the superconducting circuits community.

account for the photon loss, Kerr and second-order dispersive shift. A. Eickbusch et al [65] point out that the ECD-based approach is about a magnitude faster than the SNAP- and waveform-based ones.

Since all approaches work in an open-loop, they highly rely on the experimentally characterised Hamiltonian. This leads to an unanticipated reduction in state fidelity if the exact Hamiltonian is unknown or if the experimental setup introduces unexpected perturbations to the control pulses. The most straightforward way to mitigate the problem is to use closed-loop optimisation techniques, but applying them to a continuous-variable state is challenging. In discrete closed-loop applications, the figure of merit is usually based on a (randomised benchmarking) fidelity. Due to the long measurement time required for continuous variable states, it is not suitable to perform full quantum state reconstruction and calculate the state fidelity. This raises a question: is it possible to calculate a fidelity-like metric that is good enough for the optimiser based on a limited number of measurements?

Among all the possible bosonic computation bases, the cat state basis is one of the easiest to create and manipulate experimentally. Compared to, for example, binomial or GKP states, cat states have well-known deterministic creation protocols and do not require any preliminary open-loop optimisation of the gate sequence. It makes them perfect candidates for the first proof-of-principle closed-loop optimisation experiments on bosonic states.

4.2 Cat states: creation protocols

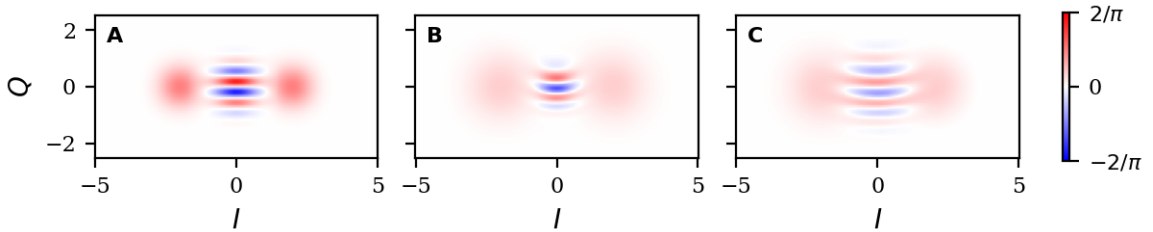


Figure 4.1: Impact of thermal cavity population on the $\alpha = 2$ cat state created with qcMAP and ECD protocols. **A** Cat state formed from a vacuum state. The state does not depend on the creation protocol. **B** Cat state formed with the qcMAP protocol from a thermal state with $n_{\text{th}} = 0.8$. **C** Cat state formed with ECD protocol from a thermal state with $n_{\text{th}} = 0.8$.

A cat state in a bosonic mode is defined as a superposition of two coherent states, $|\psi_{\text{cat}}\rangle = (|\alpha\rangle + |-\alpha\rangle)/\sqrt{2}$. There are two deterministic protocols for creating a cat state with an arbitrary phase - qcMAP and ECD - that result in two different types of cat states. Namely, the action of the qcMAP protocol is equivalent to the operator [88]

$$S_{\text{qcMAP}} = \frac{1}{\sqrt{2}}(D_{\alpha} + e^{i\phi} D_{-\alpha}), \quad (4.1)$$

while the ECD protocol is equivalent to

$$S_{\text{ECD}} = \frac{1}{\sqrt{2}}(1 - e^{i\phi}\Pi)D_{\alpha}. \quad (4.2)$$

Accordingly, the Wigner function of the cat created with qcMAP protocol is

$$W_{\text{qcMAP}}(\beta) = \frac{1}{\pi} \left[\mathcal{P} \left(e^{-2\mathcal{P}|\beta-\alpha|^2} + e^{-2\mathcal{P}|\beta+\alpha|^2} \right) - 2 \cos(4\Im\{\alpha^*\beta\} + \phi) e^{-2|\beta|^2/\mathcal{P}} \right], \quad (4.3)$$

and of the cat created with ECD protocol is

$$W_{\text{ECD}}(\beta) = \frac{\mathcal{P}}{\pi} \left[e^{-2\mathcal{P}|\beta-\alpha|^2} + e^{-2\mathcal{P}|\beta+\alpha|^2} - 2 \cos(4\Im\{\alpha^*\beta\} + \phi) e^{-2|\beta|^2/\mathcal{P}} \right], \quad (4.4)$$

where $\mathcal{P} = (2n_{\text{th}} + 1)^{-1}$ is the purity of the initial cavity state.

The difference between these two types of cats becomes more visible with an increase in the thermal population of the cavity (see Fig. 4.1). Although the protocols result in identical states at the low thermal population of the cavity $n_{\text{th}} \rightarrow 0$, it is good to keep in mind that the ECD protocol "washes out" the fringes and can be one of the reasons for the low fringe contrast of an experimentally created cat state. This Chapter assumes that the cavity is initially in the vacuum state $|0\rangle$ and the qubit is initially in the ground state $|g\rangle$. For more information regarding the properties of the qcMAP and ECD cat states in the case of the significant thermal population in the cavity, see [88].

The **qcMAP protocol** for a cat of size α consists of the following steps (see Fig. 4.2 for the illustration in phase space):

	Gate	Description	State at the end of the step
1	$X_{\pi/2}$	Bring the qubit state to the Bloch sphere equator by applying a $\pi/2$ -pulse. The rotation axis determines the phase of the cat fringes	$\frac{1}{\sqrt{2}}(g\rangle - i e\rangle) \otimes 0\rangle$
2	D_{α}	Displace the cavity by α	$\frac{1}{\sqrt{2}}(g, \alpha\rangle - i e, \alpha\rangle)$
3	$T_{\pi/\chi}$	Wait for a time π/χ , so that $ e, \alpha\rangle$ undergoes a π -rotation in phase space due to qubit-cavity dispersive interaction	$\frac{1}{\sqrt{2}}(g, \alpha\rangle - i e, -\alpha\rangle)$
4	D_{α}	Displace the cavity by α to bring the state $ e, \alpha\rangle$ to the centre of phase space	$\frac{1}{\sqrt{2}}(g, 2\alpha\rangle - i e, 0\rangle)$
5	X_{π}^0	Apply a selective (disentanglement) π -pulse on the qubit to flip it from $ e, 0\rangle$ state to $ g, 0\rangle$. It disentangles the cavity from the qubit and forms a cat state centred at $(\alpha, 0)$	$\frac{1}{\sqrt{2}}(g, \alpha\rangle - g, 0\rangle)$
6	$D_{-\alpha}$	Bring the cat state back to the centre by displacing it by $-\alpha$	$\frac{1}{\sqrt{2}}(\alpha\rangle - -\alpha\rangle) \otimes 0\rangle$

A crucial part of the protocol is correctly choosing the length of the disentanglement pulse in step # 5. In the frequency domain, the pulse width should be much smaller than the qubit-cavity dispersive shift χ (see Fig. 3.8 D). Otherwise, it can partially flip $|g, 2\alpha\rangle$. Flipping the population in step # 5 at $n \leq N$ is equivalent to applying the operator

$$X_\pi^N = X_\pi \otimes \sum_{i=0}^N |n\rangle\langle n| + I \otimes \sum_{i=N+1}^{\infty} |n\rangle\langle n|,$$

which would result in only a partial qubit-cavity disentanglement. Surely, the longer the pulse, the better, but experimental limitations put an upper bound on the pulse length. In the end, the length of the pulse is a trade-off between the pulse spectral width, qubit and cavity coherence times, and parasitic distortions due to higher-order nonlinearities (see Section 4.3).

The ECD protocol reduces the negative effects of qubit dephasing during the idling time by replacing step #3 with the following steps (see Fig. 4.3):

	Gate	Description	State in the end of the step
3.1	$T_{\frac{\pi}{2\chi}}$	Wait for a time $\frac{\pi}{2\chi}$, so that $ e, \alpha\rangle$ undergoes a $\pi/2$ -rotation in phase space due to qubit-cavity dispersive interaction	$\frac{1}{\sqrt{2}}(g, \alpha\rangle - i e, i\alpha\rangle)$
3.2	$D_{-\frac{\alpha}{2}, -\frac{\alpha}{2}}$	Displace the cavity by $(-\alpha/2, -\alpha/2)$. The states corresponding to the ground and excited states of the qubit are now positioned symmetrically with respect to the phase space origin	$\frac{1}{\sqrt{2}}(g, \frac{\alpha}{2} - i\frac{\alpha}{2}\rangle - i e, -\frac{\alpha}{2} + i\frac{\alpha}{2}\rangle)$
3.3	Y_π	Flip the qubit around the y -axis. This operation swaps the states corresponding to the ground and excited qubit states in the phase space	$\frac{1}{\sqrt{2}}(i g, -\frac{\alpha}{2} + i\frac{\alpha}{2}\rangle + e, \frac{\alpha}{2} - i\frac{\alpha}{2}\rangle)$
3.4	$D_{-\frac{\alpha}{2}, -\frac{\alpha}{2}}$	Displace the cavity by $(-\alpha/2, -\alpha/2)$. The $ g, -\alpha\rangle$ state is back on the I -axis	$\frac{1}{\sqrt{2}}(i g, -\alpha\rangle + e, -i\alpha\rangle)$
3.5	$T_{\frac{\pi}{2\chi}}$	Wait for the time $\frac{\pi}{2\chi}$, so that $ e, \alpha\rangle$ finishes its evolution at the point $(\alpha, 0)$ in phase space.	$\frac{1}{\sqrt{2}}(g, -\alpha\rangle + i e, \alpha\rangle)$

Such a sequence reduces qubit dephasing during step #3 of the qcMAP protocol in a manner similar to the widely known echo sequences. The pulse Y_π in the sequence should be as short as possible to make it unselective: the situation is opposite to that with the X_π^0 pulse. Being too selective can lead to parasitic fringes on the bottom of the Wigner map.

Operations 3.1 - 3.5 flip excited and ground qubit states with respect to coherent cavity states $|\alpha\rangle$ and $|\alpha\rangle$ (see Fig. 4.3). This is the reason why displacements in steps #4 and

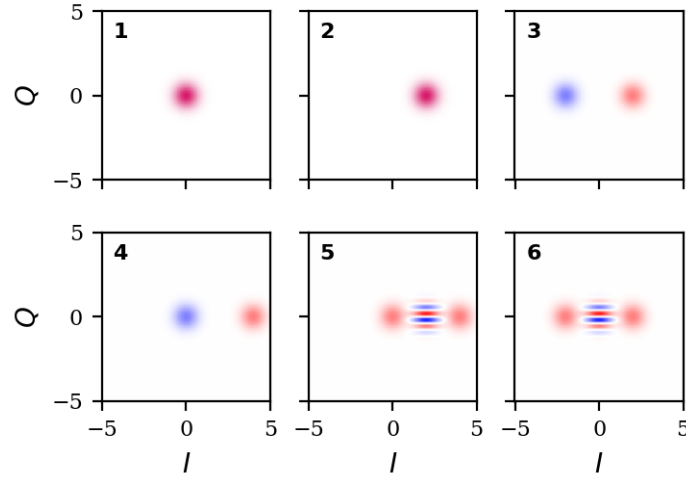


Figure 4.2: qcMAP protocol steps in phase space. The numbers of the plots correspond to the step numbers in the table 4.1. Red, blue, pink colours correspond to qubit in $|g\rangle$, $|e\rangle$, $(|g\rangle + |e\rangle)/\sqrt{2}$ states respectively. Note that only figures 3 - 6 correspond to actual Wigner maps of the states; figures 1 - 2 would appear blank.

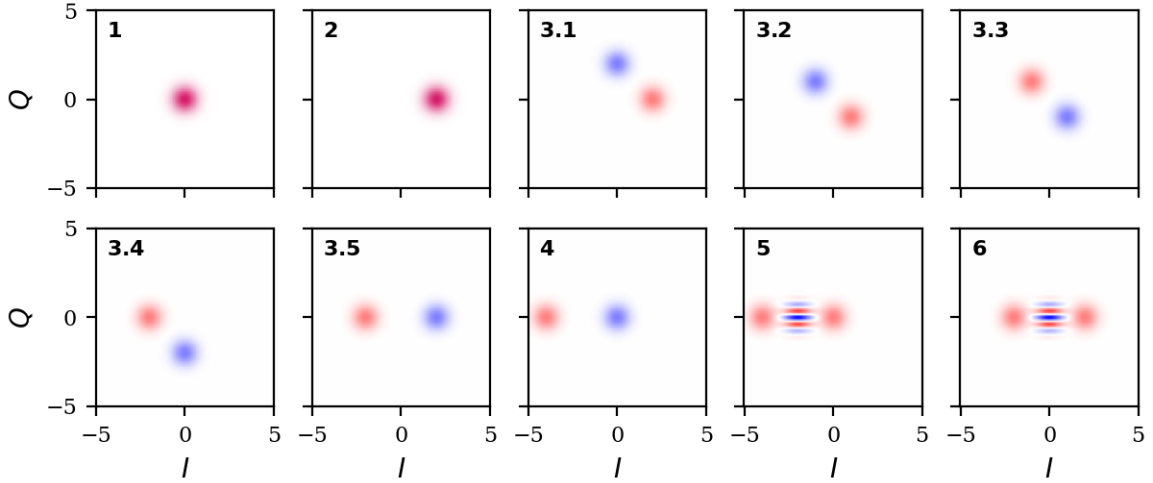


Figure 4.3: ECD protocol steps in phase space. The numbers of the plots correspond to the step numbers in the table 4.2. Red, blue, pink colours correspond to qubit in $|g\rangle$, $|e\rangle$, $(|g\rangle + |e\rangle)/\sqrt{2}$ states respectively. Note that only figures 3 - 6 correspond to actual Wigner maps of the states; figures 1 - 2 would appear blank.

#6 have the opposite sign compared to the same steps in the qcMAP, and the final cat state has a different global phase with respect to the qcMAP protocol.

4.3 Cat states: imperfections

Experimental setup limitations and specifics of the quantum system lead to imperfections in the cat states created with deterministic protocols. The list may include:

1. Too short disentanglement pulse

Although it is desirable to keep the disentanglement pulse as short as possible to reduce distortions arising due to higher-order nonlinearities, a too-short pulse can become too unselective and flip not only the state $|e, 0\rangle$, but also $|g, 2\alpha\rangle$. Assuming that the cavity thermal population is negligible, 99% selectivity corresponds to a Gaussian with the spectral width [15, p. 84],

$$\sigma_\omega = \frac{4|\alpha|^2\chi}{5}, \quad (4.5)$$

The smaller the cat, the bigger the limitation on the spectral width of the disentanglement pulse.

2. Significant contribution of higher order perturbation terms: K and χ_{qcc}

The higher-order nonlinearities in the Hamiltonian

$$H^{(4)} = -\frac{1}{2}(K + \chi_{qcc}q^\dagger q)c^{\dagger 2}c^2 \quad (4.6)$$

result in parasitic distortions of the cat fringes. The distortions occur during the evolution step #3 and the disentanglement step #5. During the evolution step, $|g, \alpha\rangle$ experiences a photon number-dependent distortion $|n(t)\rangle = e^{i\bar{n}^2 K t}|n\rangle$, while $|e, \alpha(t)\rangle$ experiences an even bigger distortion $|n(t)\rangle = e^{i\bar{n}^2(K + \chi_{qcc})t}|n\rangle$, where $\bar{n} = |\alpha|^2$, and $t = \pi/\chi$ for the qcMAP protocol. During the disentanglement step, the distortion affects only the displaced state $|g, 2\alpha\rangle$. Due to a larger displacement value, the prefactor in the exponent is significantly larger. However, the disentanglement pulse duration is shorter

$$t = 4\sigma_t = \frac{4}{2\pi\sigma_w} = \frac{10}{\bar{n}\pi\chi}, \quad \bar{n} = |2\alpha|^2. \quad (4.7)$$

The Kerr distortions in this case can be approximated as $|n(t)\rangle = e^{i\bar{n}^2 K t}|n\rangle = e^{i10\bar{n}K/(\pi\chi)}|n\rangle = e^{i40|\alpha|^2 K/(\pi\chi)}|n\rangle$. The relative strength of distortions during evolution and disentanglement steps is defined by the fraction of their subexponential expressions:

$$\frac{\pi|\alpha|^4(K + \chi_{qcc})/\chi}{40|\alpha|^2 K/(\pi\chi)} = \frac{\pi^2|\alpha|^2}{40} \left(1 + \frac{\chi_{qcc}}{K}\right),$$

which depends on the cat's size and the values of the nonlinearities. For large cat states, distortions during the evolution step dominate.

3. Non-instantaneous cavity and qubit pulses compared to the qubit-cavity dispersive shift χ

A large qubit-cavity dispersive shift χ can lead to additional undesired system evolution of the states $|g, \alpha(t)\rangle$ and $|e, \alpha(t)\rangle$ during the displacement pulses. This can be seen, for example, on the Wigner map after the first displacement in the qcMAP sequence (step #2). If χ is big enough, the blue Gaussian, which corresponds to

the $|e, \alpha\rangle$ state, starts to rotate in phase space during the disentanglement pulse and acquires an additional phase, while the red Gaussian, which corresponds to the $|g, \alpha\rangle$ state, remains on the axis. An additional recalibration of the evolution time is required in such a case. Disentanglement pulses should be as fast as possible to bring them closer to the ideal theoretical instantaneous displacements. The same problem appears during the Wigner measurement: the $|e, \alpha\rangle$ state can start evolving in phase space during the qubit $\pi/2$ -pulse.

In most experiments, the "out-of-the-box" displacement pulses can be considered as instantaneous. The system used in this thesis required additional calibrations.

4. High initial excitation population of the qubit

A high initial qubit population leads to a systematic error in cat creation. To avoid that, one can use a qubit reset before the creation sequence or postselection techniques. Another option is to push the state $|e, 0\rangle$ out of measurable Hilbert space by applying large conditional displacements at the beginning of the creation protocol [15] (see Fig. 3.8 B).

5. Low qubit/cavity coherence time

Cavity coherence plays a big role during the cat formation sequence. The coherent state decays in the cavity as $e^{-t/T_{\text{cav}}}$, which means that the system during the evolution and disentanglement steps #3 and #5 would be affected. In this case, an additional adjustment of the displacements' amplitudes is required. In addition, the cat fringes decay as $T_{\text{cav}}/(2\alpha^2)$ after the cat formation and can lead to a lower fringe contrast at the end of the Wigner or Husimi Q measurement.

Qubit decay affects the cat state during all stages of state generation and tomography. The relaxation during the creation sequence can lead to a significant residual population of the qubit's excited state at the end of the protocol, reducing the measurement contrast. This unwanted population can be removed by applying a conditional reset before the tomography sequence. Qubit dephasing mostly scrambles the phase between $|g, \alpha(t)\rangle$ and $|e, \alpha(t)\rangle$ during the evolution step #3 and can be reduced by using the ECD sequence instead of the qcMAP.

To show the influence of different imperfection mechanisms, the cat state creation sequence is simulated with the **QuTip** library [26]. The system's evolution during step #3 is described with a Lindblad equation, which can be numerically solved with the **QuTip** `mesolve` function:

$$\frac{\partial \rho}{\partial t} = -\frac{i}{\hbar} [H_{\text{tot}}, \rho] \quad (4.8)$$

$$H_{\text{tot}} = H + H_{\text{env}} \quad (4.9)$$

$$H_{\text{env}} = \sum_n \frac{1}{2} \left(2C_n \rho C_n^\dagger - \rho C_n^\dagger C_n - C_n^\dagger C_n \rho \right) \quad (4.10)$$

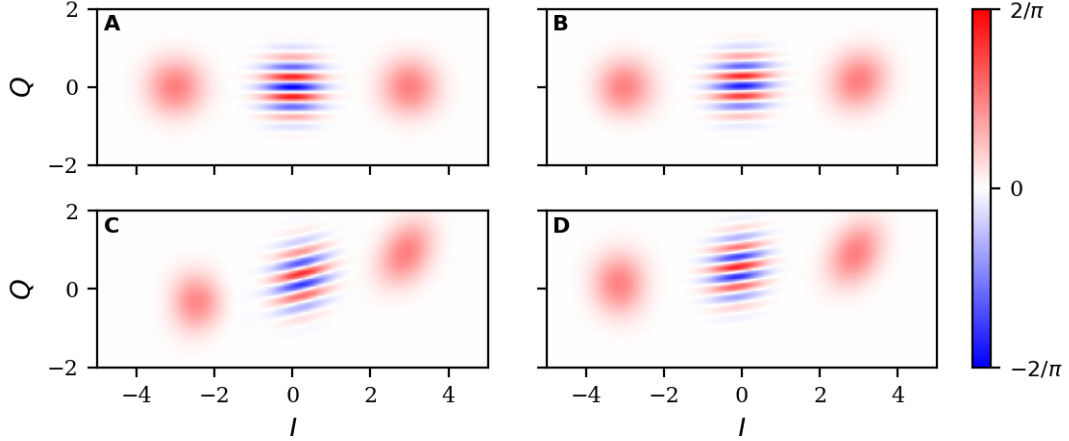


Figure 4.4: Cat state imperfections. The Hamiltonian parameters and decoherence rates take experimentally measured values. All the qubit pulses are 4σ -Gaussians with $\sigma_t = 9$ ns. The fidelities of the simulated states are: $F_A = 1.00$, $F_B = 0.96$, $F_C = 0.62$, $F_D = 0.33$ (the drop in fidelity for the last two is due to a wrong fringe phase, which is altered by the Kerr effect). **A** Simulation of the $\alpha = 3$ cat state creation sequence. The pulses are assumed to be perfect, and higher-order nonlinearities and decoherence are neglected. **B** Simulation taking into account higher-order nonlinearities and decoherence. Fidelity decrease due to decoherence is about 1.5 %. The corresponding parameters are taking experimental values. **C** Same as **B**, but the exact evolution during displacements and qubit pulses is taken into account. The evolution time is $t = \pi/(\chi + n\chi_{qcc}) = 325$ ns. **D** Same as **C**, but the evolution time takes an experimentally calibrated value.

Here, $C_1 = \sqrt{\Gamma_1}q$, $C_2 = \sqrt{\Gamma_2/2}\sigma_z$, and $C_3 = \sqrt{\Gamma_{\text{cav}}}c$ - qubit and cavity collapse operators, and H is the system Hamiltonian 3.33. The evolution during the cavity and qubit pulses is simulated with the **mesolve** function as well by taking into account the applied drive

$$\frac{\partial \rho}{\partial t} = -\frac{i}{\hbar} [H_{\text{tot}} + H_{\text{dr}}, \rho] \quad (4.11)$$

In the case of a Gaussian qubit pulse,

$$H_{\text{dr}} = \frac{\Omega(t)}{2} (e^{i\phi}q + e^{-i\phi}q^\dagger) \quad (4.12)$$

$$\Omega(t) = \frac{\theta e^{-(t-T/2)^2/2\sigma_t^2}}{\sqrt{2\pi\sigma_t^2} \text{erf}(T/2^{3/2}\sigma_t)} \quad (4.13)$$

where for $X_{\pi/2}$ and X_π pulses θ is defined as $\pi/2$ and π accordingly, the phase $\phi = 0$, and the total pulse time $T = 4\sigma_t$ for a 4σ -Gaussian pulse [88]. In the case of an arbitrary qubit (cavity) drive defined with I and Q envelopes,

$$H_{\text{dr}} = I(t) \frac{a + a^\dagger}{2} + iQ(t) \frac{a - a^\dagger}{2} \quad (4.14)$$

with $a = c$ for cavity drive and $a = q$ for qubit drive. Note that in this convention the cavity quadratures are flipped, e. g. displacement by $\alpha = \alpha_1 + i\alpha_2$ would move the state by $-\alpha_2 - i\alpha_1$ in the simulation. Simulation results are presented in Fig. 4.4.

Once the cryogenic part of the experimental setup is fixed, there are not so many options for improving qubit and cavity coherence times (rather than incorporating echo sequences). The length and shape of the control pulses can be adjusted to best suit hardware and creation protocol limitations. However, the distortions due to higher-order nonlinearities can still be reduced by applying correction pulses during or after the creation protocol.

4.4 Kerr compensation techniques

Higher-order nonlinearities

$$H^{(4)} = -\frac{1}{2}(K_c + \chi_{qcc}q^\dagger q)c^{\dagger 2}c^2 = \tilde{K}c^{\dagger 2}c^2$$

induce parasitic rotation and squeezing of the coherent state in phase space for $\tilde{K}t \ll 1$ [89], resulting in an undesired photon-number-dependent phase evolution (phase scrambling) of a bosonic state when the nonlinear coefficients are large and the evolution times are long. Currently, there are two schemes for compensating Kerr and higher-order nonlinearities: using geometric phase gates and tuning the nonlinearity of the cavity-qubit system with an off-resonant drive.

4.4.1 Geometric phase gates

Due to higher-order nonlinearities, different Fock states $|n\rangle$ gain different phases during evolution. The most straightforward way of compensating for it is to apply a phase correction to each state $|n\rangle$ that has a significant contribution in the Fock state decomposition of the target cavity state $|\psi_{\text{cav}}\rangle = \sum_{n=0}^{n_{\text{max}}} c_n |n\rangle$. Consider a qubit-cavity system described by the Hamiltonian 3.33. In a number-split regime, it is possible to drive each photon number peak selectively at $\omega_{n,q} = \omega_q - n\chi$. One can apply two subsequent π rotations on the qubit at frequency $\omega_{n,q}$: the first one around the Y axis, and the second one around a tilted axis (see Fig. 3.9 B). As a result, the qubit ends up back in the $|0\rangle$ state but acquires an additional geometric phase θ , depending on the qubit trajectory on the Bloch sphere (see Fig. 4.5 A). This idea is utilised by a Selective Number-Dependent Arbitrary Phase (SNAP) gate [31] and can be equivalently described by the operator

$$S_n(\theta) = e^{i\theta|n\rangle\langle n|}$$

By applying several SNAP gates at $n = 0, \dots, n_{\text{max}}$ Fock states

$$S(\theta) = \prod_{n=0}^{n_{\text{max}}} S(\theta_n),$$

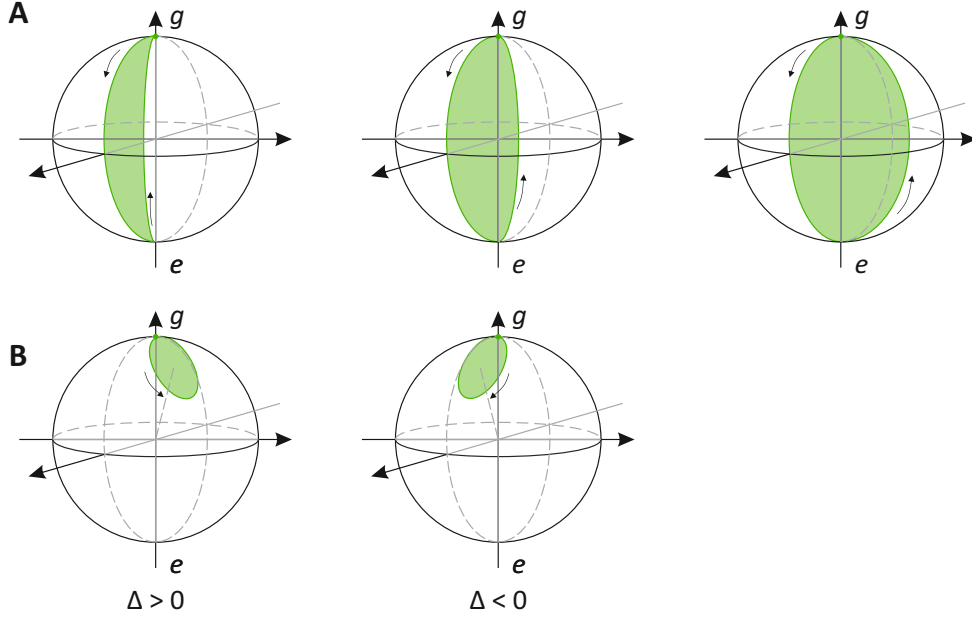


Figure 4.5: Actions of the geometric phase gates on the Bloch sphere. **A** SNAP gate. Each Bloch sphere corresponds to a different number-split peak. Depending on the rotation axis of the second selective π -pulse θ , the qubit imprints a different geometric phase (green area). **B** PASS gate. Depending on the detuning between the drive and the photon number peak, the qubit imprints the phase in a clockwise or counterclockwise direction.

it is possible to cancel out the phase space scrambling: if the phases of different Fock states $\theta_n = \arg(c_n)$ constructively interfere, undesired effects cancel out. By calibrating out the relative Fock states' phases, Heeres et. al. [31] show that it is possible to achieve 0.97 fidelity of the coherent state after $1 \mu s$ evolution time. The set of displacements and SNAP gates allows the creation of arbitrary states [90], making it a perfect basis for open-loop optimisation.

Another way to exploit a geometric phase is to apply a continuous off-resonant drive on a qubit-cavity system in a number-split regime. If the drive is off-resonant with any of the photon number qubit peaks, the trajectory on the Bloch sphere forms a small loop around the pole (see Fig. 4.5 B). Depending on the sign of the detuning between $|g, n\rangle$ and the drive frequency, the rotation on the Bloch sphere occurs in a clockwise or counterclockwise direction. The solid angle for the trajectory on the Bloch sphere for the state $|g, n\rangle$ depends on the drive detuning Δ_d and the drive amplitude Ω :

$$\theta = 2\pi \left(1 - \frac{\Delta_d}{\sqrt{4\Omega^2 + \Delta_d^2}} \right) \quad (4.15)$$

It is equivalent to inducing a photon-number-resolved ac Stark shifts (PASS) δ_n of energy levels, resulting in the effective Hamiltonian

$$H_{\text{PASS}} = \sum_n \delta_n |n\rangle\langle n| = \sum_n \frac{\Omega^2}{\Delta_d - n\chi} |n\rangle\langle n| \quad (4.16)$$

By carefully choosing the amplitudes and drive detunings, it is possible to accumulate the desired geometric phase or cancel out the cavity Kerr, similarly to the SNAP gate. Y. Ma et al. [71] showed a successful implementation of PASS for error-transparent² Kerr cancellation for the binomial logical qubit using two PASS drives: between the Fock states $|2\rangle - |3\rangle$ at $\Delta_d = 3.37\chi$, and $|3\rangle - |4\rangle$ at $\Delta_d = 2.27\chi$. The acquired phase shift of the logical state after 250 μs remained at the level of $2\pi \times 0.02(3)$ kHz.

4.4.2 Adiabatic off-resonant drive

Kerr and higher-order cavity nonlinearities can be tuned by applying an off-resonant adiabatic drive [91] on the ancilla (transmon). Depending on the transmon anharmonicity, cavity-transmon coupling and drive detuning, the nonlinearities show different behaviour as a function of $|\Omega_d/\delta_d|^2$, where Ω_d is proportional to the drive amplitude and $\delta_d = \omega_d - \omega_{10}$ is a drive detuning from the transmon 0 – 1 transition frequency. Of particular interest is the situation when

$$|\delta_a| \gg \alpha, \quad \delta_a = \omega_a - \omega_{10}, \\ g_a/(\omega_a - \omega_c) \ll 1$$

which is within the usual experimental range, and

$$|\delta_a| \gg |\delta_d|, \quad \delta_d = \omega_d - \omega_{10},$$

which allows for avoiding drive-induced resonances (see Section V of [91]). Zhang et. al. [91] in Sections IV and VI show that within the rotating wave approximation, after several unitary transformations, the Hamiltonian in the presence of an off-resonant drive can be rewritten as follows:

$$\begin{aligned} H_{\text{RWA}} &= H_C + H_{AB} \\ H_C &= -\delta_{dC} N_C - \frac{\alpha}{2} N_C(N_C + 1) + \Omega_d C^\dagger + \Omega^* C \\ H_{AB} &= -\delta_{dA} - \frac{1}{2} \chi_{AA} N_A N_A \\ \delta_{dC} &= \delta_{dC} + \chi_{AC} N_A, \text{ where} \end{aligned} \quad (4.17)$$

A - annihilation operator of the cavity-dressed mode

C - annihilation operator of the transmon-like dressed mode

²Error-transparent (ET) gates have a deterministic trajectory in code and error spaces, regardless of when the error occurs during the gate operation. Since the final state in the error space is known, it can always be mapped back to the code space by applying a corresponding correction pulse after the gate. ET gates do not require the measurements of error syndromes. The correction pulse leaves the code state unchanged and maps the corrupted state from the error space back to the code space.

$\delta_{dA} = \omega_d - \omega_A$ - cavity-drive detuning

$\delta_{dC} = \omega_d - \omega_C = \omega_d - (\omega_{10} + \alpha)$ - transmon-drive detuning

The operators have the following connection to the bare transmon annihilation operator:

$$c = \xi_A A + \xi_C C$$

with $\xi_{A(C)}$ being a participation ratio of the cavity (transmon) mode that one can easily get from HFSS simulations. In the dispersive regime, the mode is mostly contained in the cavity, $|\xi_A| \ll 1$, $|\xi_C| \approx 1$. The self- and cross-Kerr nonlinearities are defined here as

$$\chi_{XX} = \alpha |\xi_X|^4, \quad \chi_{XX'} = 2\alpha |\xi_X \xi_{X'}|^2$$

which makes it clear that self-Kerr is the dominant nonlinear term in the Hamiltonian.

To cancel the Kerr effect and suppress the higher order nonlinearities, it is preferable to use a large detuning $\delta_{dC} \gg \alpha$ (see Fig. 4 in [91]). Besides, the larger the detuning, the easier to hit an optimal drive power range. An example of Kerr suppression for the $\alpha = 6$ coherent state after a $2 \mu s$ evolution time can be seen in Fig. 4.6.

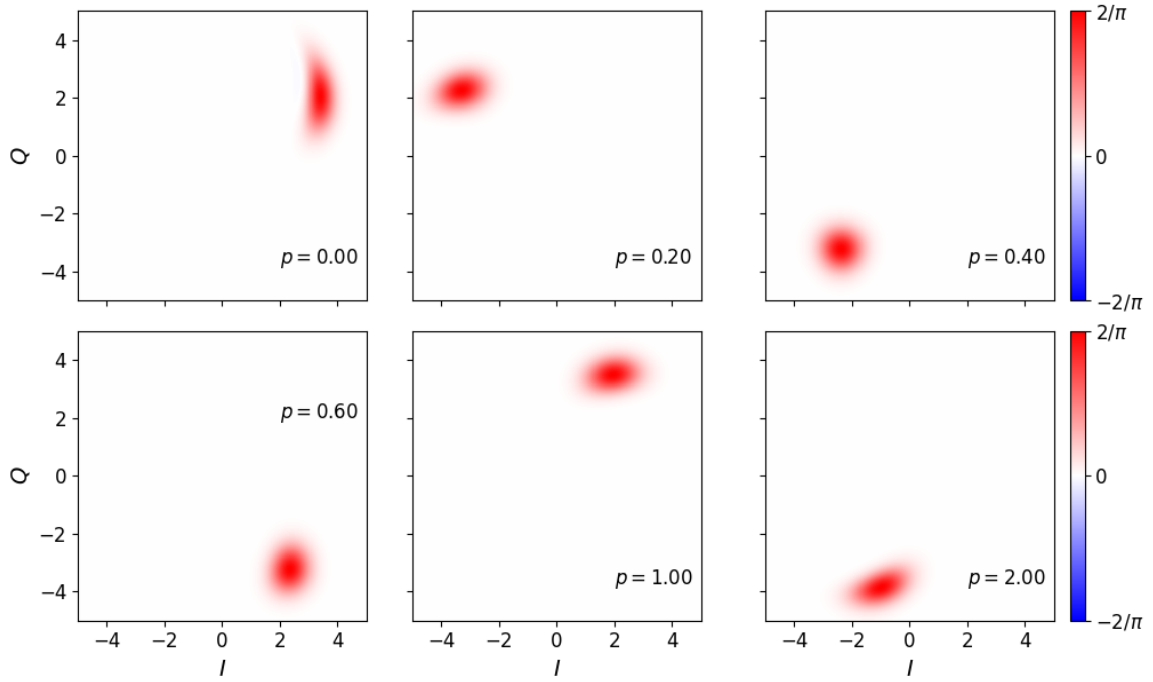


Figure 4.6: QuTiP simulations of the effect of adiabatic off-resonant drive for the coherent state $\alpha = 4$ after $2 \mu s$ evolution. Different plots correspond to different values of the parameter $p = |\Omega_d / \delta_d|^2$. The Hamiltonian parameters are: $\omega_{01}/2\pi = 5$ GHz, $\chi_{qc}/2\pi = 1.25$ MHz, $\chi_{cc}/2\pi = 4$ kHz, $\delta_d/2\pi = \alpha/2\pi = 180$ MHz. Increasing the drive amplitude leads to decreased phase scrambling. Once the drive amplitude passes the optimal point ($p \approx 0.4$), the Kerr effect starts to corrupt the state again. Note that the state acquires an additional rotation in phase space.

The off-resonant drive cancellation scheme allows for the cancellation or reversal of the Kerr effect and is simple to use: the only parameter that needs to be tuned is the drive amplitude. However, the scheme has some disadvantages. First of all, cavity self-Kerr depends on the transmon state. It means that one can turn on the Kerr-compensation scheme only when the qubit and the cavity are disentangled, i.e. after the creation of the bosonic state is complete. In the case of the qcMAP protocol, it means that the method would mitigate the Kerr errors on coherent states $|\alpha\rangle$ and $-|\alpha\rangle$ equally (even though they were affected by Kerr differently during the protocol), unless one of the coherent states is shifted to the centre of phase space. Second, applying a largely detuned drive may require an additional analog output of the AWG in the experiment. Nevertheless, the scheme provides a perfect solution for mitigating uncorrectable errors arising from higher-order nonlinearities in many bosonic quantum error correction codes.

4.5 Closing the loop: first tries

Before running closed-loop optimisation, three main questions should be addressed: which type of optimiser to use, which pulses in the sequence to correct, and what to use as a figure of merit. The answers to these questions are determined by the type of optimisation problem and its constraints, as well as hardware limitations.

In experimental applications, gradient-free algorithms often outperform gradient-based ones, as they require fewer measurements per optimisation iteration. Following the most widespread optimisation approaches in superconducting circuits (see Chapter 2.3), there are two options: Nelder-Mead and CMA-ES. CMA-ES is preferable in setups where the bottleneck arises from the reconfiguration of experimental hardware³, since after each step it returns a set of best offspring from the generated population. In this thesis, initialisation or reconfiguration of the experimental setup is not a limiting factor from a time perspective.

The choice of the optimiser fell on the direct search algorithm **dCRAB** from the QuOCS library [92]. dCRAB (see Section 2.2.3 for more details), due to a random choice of the basis vectors after each superiteration, escapes traps in the optimisation landscape and avoids the usage of piece-wise constant basis. The optimisation of long pulses on a piece-wise basis introduces numerous optimisation parameters, which complicate the convergence of gradient-free algorithms and significantly increase the optimisation time. In this Chapter, the pulses were constructed in a Fourier basis, which usually serves as a good starting point for the optimisations.

dCRAB from QuOCS can use two different types of direct search algorithms: Nelder-Mead and CMA-ES. In the case of Nelder-Mead, after each superiteration (or after some time) it is possible to account for the drifts of the figure of merit. The drift compensation feature is not applicable for CMA-ES due to its operating principles. This peculiarity, together with the absence of time constraints from the hardware side, determined the choice of **Nelder-Mead algorithm** as a direct search method for closed-loop optimisations.

³For example, this is the case for Zurich Instruments AWG, see [48]

The cat states in this thesis are formed with the **qcMAP protocol**. The protocol contains the idling step #3, during which the state $|e, \alpha\rangle$ evolves in a phase space. Instead of waiting for π/χ , until $|e, \alpha(t)\rangle$ makes a π -rotation in phase space, it is reasonable to apply correction pulses during that time. Therefore, the optimiser is expected to compensate for the errors raised during the step and pre-compensate for the errors that arise after. Compared to the ECD protocol, qcMAP is preferable for closed-loop applications because the optimiser finds better solutions for longer pulses.

Since it is impossible to measure the entire Wigner map after each optimisation step due to time constraints, it makes sense to select a set of measurement points and construct a fidelity-like figure of merit defined on this set of points. Although this solution lacks flexibility and heavily relies on the choice of sampling points, it is very intuitive and simple, and can serve as a good starting point for developing a closed-loop optimisation strategy. The figure of merit based on N measured points can be defined as

$$\text{FoM} = \frac{\sum_{i=1}^N w_i \cdot W_i^{\text{id}} W_i^{\text{meas}}}{\sum_{i=1}^N (W_i^{\text{id}})^2} - \sum_j P_j, \text{ where} \quad (4.18)$$

w_i - weighting coefficient for measurement point i ,

W_i^{id} - Wigner value for the ideal state for measurement point i ,

W_i^{meas} - measured Wigner value for measurement point i ,

P_j - penalty term.

To sum up, the chosen optimisation configuration is the following:

- Optimisation algorithm: Nelder-Mead in dCRAB formalism
- Optimisation basis: Fourier
- Optimised pulses: I and Q quadratures of qubit and cavity pulses during the evolution step of the qcMAP protocol
- Figure of merit: (see Eq. 4.18), defined on a specifically designed set of measurement points.

Prior to implementation in the experiment, various optimisation configurations and sampling strategies were evaluated and adjusted in an open-loop setting. The evolution of the system during the protocol was simulated in the **QuTiP** library [26]. The result of the simulation is a density matrix of the final state, providing direct access to the state fidelity and enabling the calculation of Wigner values. The only important difference from a closed-loop optimisation lies in the absence of measurement noise.

Since the optimisation approach in the future has been significantly changed, the following sections provide only a brief guideline through closed-loop optimisations on sparse sampling and do not include statistical investigations or proof of resilience to different initial cat phases. The proper characterisation of the final version of the optimiser can be found in Chapter 6.

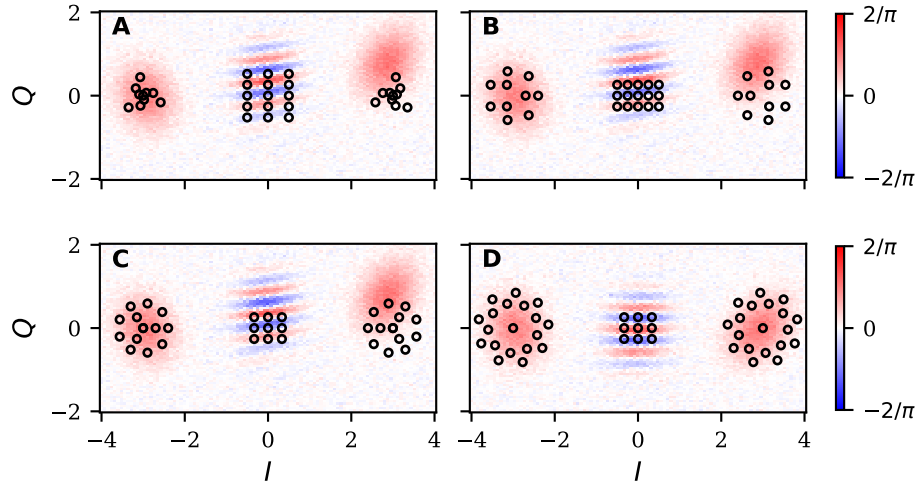


Figure 4.7: **A - C** The sampling strategies used in the first closed-loop optimisation experiments. **D** The sampling strategy that was designed to correct the Kerr distortions based on open-loop simulations. All the optimisation runs for the second sampling strategy started from a manually rotated cat state, as shown in **D**. The states in **A - C** and **D** differ only by rotation and displacement corrections, although they may seem different due to the sampling pattern on top.

4.5.1 First sampling strategy

Fig. 4.7 **A** presents the first sparse sampling strategy. 35 sampling points are located next to the most distinct areas in the phase space: interference fringes and Gaussians. The sampling shape in the Gaussians' area is designed to correct not only the position of the Gaussians but also the shape. By using sparse sampling, the time required to get the figure of merit reduces from several hours to several minutes, as shown in Fig. 4.8 **B**. In the first experiments, the figure of merit was simplified to

$$\text{FoM} = \frac{\sum_{i=1}^N W_i^{\text{id}} W_i^{\text{meas}}}{\sum_{i=1}^N (W_i^{\text{id}})^2} \quad (4.19)$$

$$\text{FoM}_{\text{tot}} = w_{\text{fr}} \cdot \text{FoM}_{\text{fr}} + w_G \cdot \text{FoM}_G = \frac{1}{2} \cdot \text{FoM}_{\text{fr}} + \frac{1}{2} \cdot \text{FoM}_G,$$

meaning the weights of the fringes (fr) and Gaussians (G) were chosen to be equal⁴.

To run an optimisation efficiently, one needs to find a trade-off between the number of averages, FoM deviation and time per optimisation step. The corresponding graphs are plotted in Fig. 4.8. The higher number of averages reduces the noise in FoM, but dramatically increases the measurement time, which has the dominant contribution to the total optimisation step time. Fig. 4.8 **B** shows that 600 - 1000 averages give an optimal result from the perspective of FoM standard deviation vs time constraints.

⁴Changes in the weights did not lead to significant changes in the distorted Gaussian shape or the fringe amplitude.

An example of an optimisation run is shown in Fig. 4.9. The optimiser corrects for rotation and displacement in phase space and the wrong cat state phase. Still, it is incapable of remedying Kerr distortions, even for a cat state with manual rotation and displacement corrections. The open-loop simulations showed that for Kerr compensation, it is required not only to change the sampling strategy around the Gaussians but also to add the penalty terms to the FoM: increasing the number of points in the Gaussians' area and introducing the random rotation of the sampling pattern during the optimisation was not enough for the Kerr correction.

4.5.2 Second sampling strategy

The revised sampling strategy should be more sensitive to the asymmetries in the Gaussian shape, which can be achieved by sampling more points equidistantly at a larger distance from the centre. Fig. 4.7 D shows an example of one of the suggested sampling approaches that gives reproducible results and corrects for Kerr distortions in a numerical simulation. The sampling strategy was tested first on the coherent state after 300 ns of free evolution. It is important to note that the optimisation converges only if most of the sampling points are not in a "zero" Wigner map area. It means that the Kerr correction in an open-loop works reliably only on the states that were rotated (aligned to the I axis) before the start of the optimisation. The penalty terms that force the optimiser to correct for the coherent state asymmetry are:

$$\text{FoM} = F - C \cdot (P^2|_{\text{inner}} + P^2|_{\text{outer}}) \quad (4.20)$$

$$P = \sum_{i=1}^n (x_i - \bar{x})^2, \text{ where}$$

F is defined as in 4.19 and calculated on all the sampling points,

$P|_{\text{inner}}$ is defined on the points of the inner circle (see Fig. 4.7 D),

$P|_{\text{outer}}$ is defined on the points of the outer circle (see Fig. 4.7 D),

C is a penalty prefactor.

For a sampling strategy similar to the one for the Gaussian in Fig. 4.7 D and $C = 20$, the final optimised coherent state reaches 0.99 fidelity in simulation.

For the cat state the FoM is set to

$$\begin{aligned} \text{FoM} = & F_{g,l} - C_{g,l}(P_{g,l}^2|_{\text{inner}} + P_{g,l}^2|_{\text{outer}}) + \\ & + F_{g,r} - C_{g,r}(P_{g,r}^2|_{\text{inner}} + P_{g,r}^2|_{\text{outer}}) + \\ & + 2F_f \text{sign}(F_f), \text{ where} \end{aligned} \quad (4.21)$$

$F_{g,l}(F_{g,r}, F_f)$ is defined as in 4.19 and calculated on all the sampling points for the left Gaussian (right Gaussian, cat fringes),

$P_{g,l}|_{\text{inner}} (P_{g,r}|_{\text{inner}})$ is defined on the points of the inner circle for the left (right) Gaussian,

$P_{g,l}|_{\text{outer}} (P_{g,r}|_{\text{outer}})$ is defined on the points of the outer circle for the left (right) Gaussian,

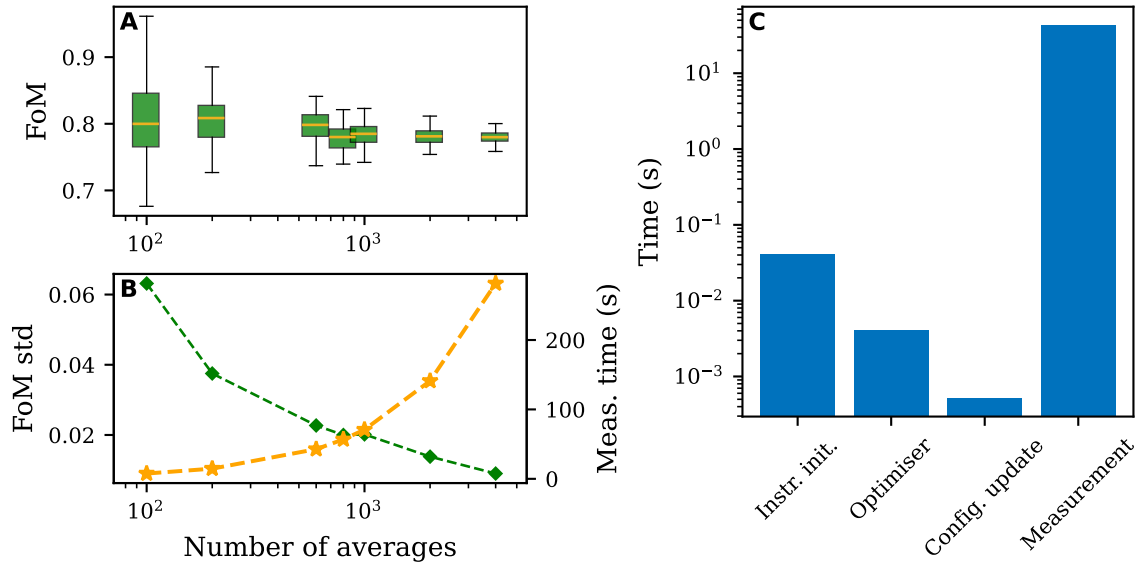


Figure 4.8: Adjusting the number of averages with respect to time constraints and figure of merit deviation. All the measurements are done on the initial cat state (before the correction pulses). **A** Figure of merit vs the number of averages. The mean is shown with the yellow line. The green box contains the data from the first to the third quartile. **B** Standard deviation of the figure of merit (green) and measurement time (yellow) vs number of averages. **C** The time spent on sub-steps during one optimisation step for 600 averages.

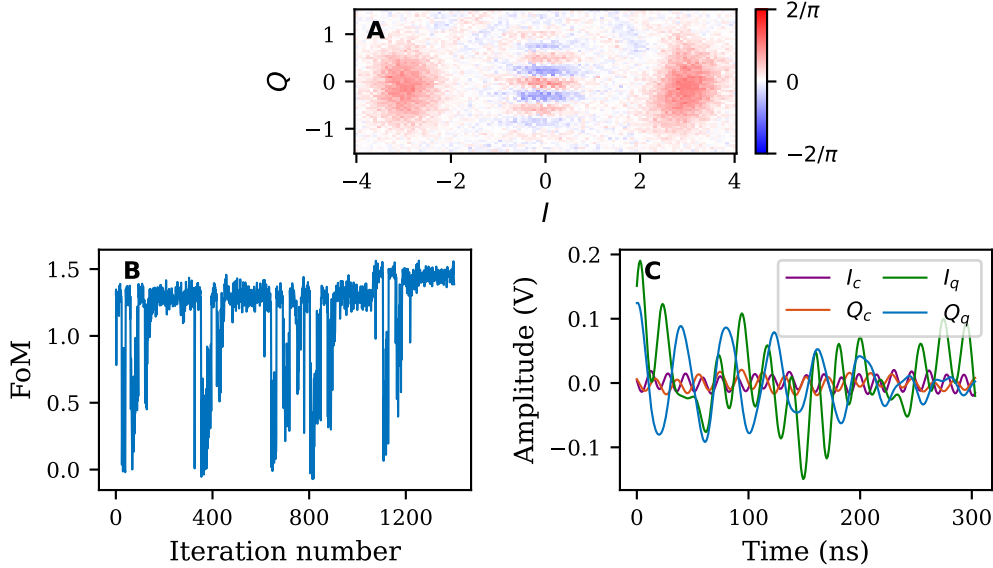


Figure 4.9: Example of a closed-loop optimisation run with predefined sparse sampling. This run started with the optimised pulses from a different optimisation run, which already aligned the cat on the axis, as an initial guess. **A** Wigner map of the optimised cat state. **B** Figure of merit as a function of iteration number. **C** The best set of cavity and qubit optimised pulse quadratures applied for the cat correction during the evolution time. These pulses were applied to measure the Wigner map from Figure A.

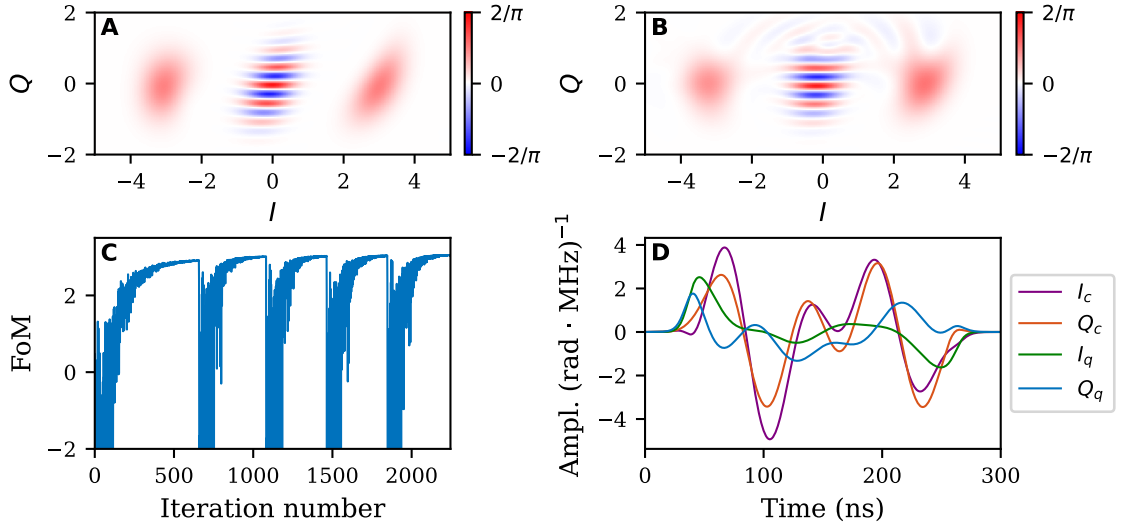


Figure 4.10: Open-loop optimisation run of dCRAB based on the Nelder-Mead algorithm. The optimisation showed a fidelity improvement from 0.81 to 0.96 and used sampling **C** in Fig. 4.7. **A** Wigner map of the cat state before the optimisation run with preliminary rotation correction. **B** Wigner map of the optimised cat state. **C** Learning curve of the optimisation algorithm. **D** Quadratures of qubit and cavity optimised pulses.

$C_{g,l}$ ($C_{g,r}$) is a penalty prefactor for the left (right) Gaussian.

An example of an open-loop optimisation run with $C_{g,l} = 1000$ and $C_{g,r} = 2000$ is presented in Fig. 4.10. Although the new figure of merit and sampling strategies worked well in an open-loop, the same approach has never been successful in a closed-loop. The variation of the figure of merit induced by the measurement noise is high enough to prevent Nelder-Mead from converging to the optimal solution for both coherent and cat states.

Summary

The closed-loop optimisation of bosonic states, and, in particular, cat states, remains a challenging task. The imperfections that can be corrected with optimisation include distortions due to nonlinearities. Although there are some possible solutions for the Kerr compensation, they are not optimal or applicable to all types of bosonic states. A more general approach would require generating an arbitrary correcting waveform in a closed-loop. The first steps towards it are presented in the last section of this Chapter.

Naive closed-loop optimisation on predefined sparse sampling with Nelder-Mead can correct for effects that dramatically decrease state fidelity, such as rotation, displacement, and incorrect fringe phase. However, these cases can be corrected manually. Rotation and displacement compensations can be calculated from the fit of the Wigner map of the state, and the fringe phase can be controlled by the phase of the first pulse in the creation sequence. The changes in the figure of merit for weaker effects, such as Kerr distortion, are washed out due to measurement noise. This means that either the figure of merit or the optimisation algorithm should be improved in terms of noise resilience to make closed-loop optimisation more beneficial.

One possibility is to use a stochastic cost function as a figure of merit (see Section 2.2.5). It yields good results for several problems in an RL setting, as first shown in a theoretical proposal and later confirmed in a quantum error correction experiment [11]. However, widely used model-free algorithms (such as Nelder-Mead or simulated annealing) exhibit poor behaviour on stochastic cost functions and require access to the state fidelity to achieve high performance [61]. In addition, even with an RL algorithm, such a cost function demonstrates stochastic behaviour for a cat generation task. The most stable results, of course, are obtained in an ideal situation where an infinite number of points in phase space are measured and averaged, which is equivalent to having access to fidelity as a reward.

Since a noise-resilient fidelity-like figure of merit would benefit any of the possible optimisation algorithms, the next step in this thesis was made in the direction of attempting to design such a figure of merit in the first place.

Efficient state reconstruction

The first experiments from Section 4.5 showed that the closed-loop optimisation requires a more noise-resilient, fidelity-like figure of merit. Measuring a Wigner map for extracting fidelity with a standard square sampling takes hours. Even the measurement of the full Wigner map by using the fidelity estimation 3.56 and optimal importance sampling

$$P(\beta) = \frac{|W_{|\psi_{\text{cat}}\rangle\langle\psi_{\text{cat}}|}(\beta)|}{\|W_{|\psi_{\text{cat}}\rangle\langle\psi_{\text{cat}}|}\|_1} \quad (5.1)$$

takes around 10^6 samples for cats with $\alpha \leq 2$ [3], which in the current measurement setup would take about 17 minutes. Ideally, the fidelity estimation should be obtained much faster. An overnight closed-loop optimisation puts a rough upper bound on the reconstruction time: 8 hours / 500 iterations $\approx$ 58 seconds per iteration.

This Chapter introduces a new reconstruction algorithm that enables high-precision fidelity estimation in the shortest possible time. This is achieved by introducing the model that parametrises the Wigner map, and active sampling, which draws the samples from the areas of the highest uncertainty (variance). The model parameters' update and sampling are interleaved until the exit conditions are satisfied, as shown in the following pseudocode:

Algorithm 1 Reconstruction algorithm

Require: initialised model M , exit conditions

```

 $B \leftarrow []$  ▷ array of complex coordinates in phase space
 $W \leftarrow []$  ▷ array of Wigner values measured at points  $B$ 
 $\beta_{\text{new}} \leftarrow \text{sample\_new\_points}(M)$  ▷ new sampled complex coordinates in phase space
while exit conditions not satisfied do:
     $W_{\text{new}} \leftarrow \text{measure\_in\_experiment}(\beta_{\text{new}})$ 
     $W \leftarrow \text{concatenate}(W, W_{\text{new}})$ 
     $B \leftarrow \text{concatenate}(B, \beta_{\text{new}})$ 
     $M \leftarrow \text{optimise}(B, W, M)$  ▷ update the model parameters
     $\beta_{\text{new}} \leftarrow \text{sample\_new\_points}(M)$ 
end while

```

The following sections provide a detailed explanation of the implementation of each part of the reconstruction algorithm. First, the model is introduced in Sections 5.1 and 5.2. The Chapter then continues by describing a simple sampling procedure in Section 5.3. The exit conditions and the more advanced sampling technique are discussed in Sections 5.4 and 5.5. Finally, the Chapter moves on to algorithm implementation details (Section

5.6) and concludes with an experimental benchmarking of the algorithm. The Chapter covers most of the machine learning background required for understanding the idea of the reconstruction algorithm. If some terms or concepts are unknown to the reader, please refer to Goodfellow [43], which gives an intuitive but exhaustive explanation of machine learning concepts.

For convenience, in this Chapter, the phase space coordinates use an equivalent notation $\beta = \text{Re}(\beta) + i\text{Im}(\beta) = I + iQ$.

The work was carried out in collaboration with Asst. Prof. Samuele Tosatto¹.

5.1 Model

The model-based approach involves designing an efficient parametric model that accurately reflects the main features of the Wigner map by using a limited number of measurements. The first insights may be inspired by analysing the analytical expression of the cat state Wigner function. As mentioned in Chapter 4, the Wigner function of a cat state formed with a qcMAP protocol from a thermal state of purity $\mathcal{P}$ is [88]

$$W_{\text{qcMAP}}(\beta) = \frac{1}{\pi} \left[\mathcal{P} \left(e^{-2\mathcal{P}|\beta-\alpha|^2} + e^{-2\mathcal{P}|\beta+\alpha|^2} \right) - 2 \cos(4\text{Im}\{\alpha^*\beta\} + \phi) e^{-2|\beta|^2/\mathcal{P}} \right], \quad (5.2)$$

which is essentially a sum of three Gaussians, one of which is modulated with a cosine. Experimentally created states can, however, have squeezed left and right Gaussians and be displaced or rotated in phase space (see Section 4.3), as shown in Fig. 5.1. To account for that, the Wigner map of the reconstructed state is parametrised in the following way:

$$W_{\text{model}}(\beta) = A_L \mathcal{N}(\mu_L, \Sigma_L) - 2A_C \mathcal{N}(\mu_C, \Sigma_C) \cos(4\alpha\beta \cdot R(\phi) + \varphi) + A_R \mathcal{N}(\mu_R, \Sigma_R), \quad (5.3)$$

where

$\mathcal{N}(\mu_L, \Sigma_L)$ is the left Gaussian with the mean $\mu_L = (I_L, Q_L)$ and covariance matrix Σ_L ,
 $\mathcal{N}(\mu_R, \Sigma_R)$ is the right Gaussian with the mean $\mu_R = (I_R, Q_R)$ and covariance matrix Σ_R ,
 $\mathcal{N}(\mu_C, \Sigma_C)$ is the centre Gaussian with the mean $\mu_C = (I_C, Q_C)$ and covariance matrix Σ_C ,

A_L, A_C, A_R are the amplitudes of the left, centre and right Gaussians,

$R(\phi)$ is a rotation matrix that accounts for a correct fringe angle ϕ with respect to the Gaussians,

φ is the fringe phase.

¹University of Innsbruck, Department of Computer Science, Technikerstraße 21a, 6020 Innsbruck, Austria
 University of Innsbruck, Digital Science Center, Innrain 15, 6020 Innsbruck, Austria

In addition, every covariance matrix is parametrised as

$$\Sigma_i = R(\theta_i)^T \Sigma_{d,i} R(\theta_i), \quad \Sigma_{d,i} = \begin{pmatrix} \sigma_{x,i}^2 & 0 \\ 0 & \sigma_{y,i}^2 \end{pmatrix},$$

where $i \in [L, C, R]$ represents left, centre, and right Gaussian, $\sigma_{x,i}$ and $\sigma_{y,i}$ are the standard deviations of the corresponding Gaussian in the coordinate system spanned by the principal axes x and y , and θ_i is an angle between the Cartesian coordinate system and the principal axes' coordinate system.

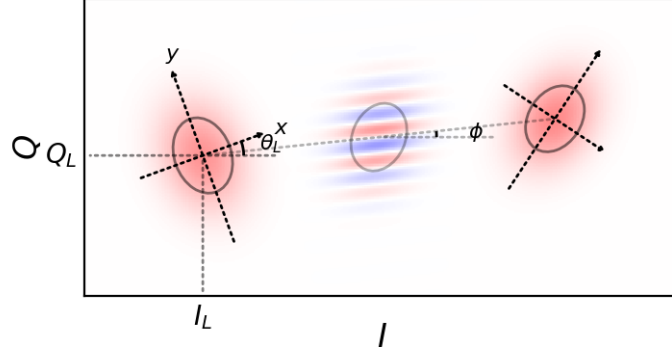


Figure 5.1: Illustration of the model parameters on the Wigner map. For each covariance matrix, eigenvalues and eigenvectors are calculated. Based on that, the minimum and maximum standard deviations σ_x and σ_y , and the angle between the principal axis and the Cartesian x-axis are evaluated.

Some of the introduced parameters can be expressed in terms of others, significantly reducing the number of independent parameters during the optimisation procedure. Fringe angle in the phase space, as well as the position of the fringes' centre, depends on the positions of the left and right Gaussians:

$$\phi = \arctan \frac{Q_R - Q_L}{I_R - I_L}, \quad (5.4)$$

$$\mu_C = 0.5(\mu_L + \mu_R). \quad (5.5)$$

The requirement on the trace of the cat state density matrix $\text{tr} \rho = 1$ allows the elimination of one of the Gaussians' amplitudes

$$A_R = 1 - A_L. \quad (5.6)$$

For a small thermal population of the cavity, one of the standard deviations in each covariance matrix can be eliminated as well. Since small Kerr perturbations are equivalent to the squeezing operation,

$$\sigma_y = 0.25/\sigma_x. \quad (5.7)$$

In the current setup, due to a sufficiently high thermal population, the Wigner map cannot be fitted with a model that satisfies condition 5.7. Removing σ_y from a set of optimisation parameters improves convergence and reproducibility and can be very beneficial for experiments with lower thermal population.

Therefore, it is enough to define the left and right Gaussians with five parameters each, and the centre Gaussian with three parameters only. Together with the left and centre amplitudes and the fringe phase, the model has a total of 16 free parameters:

$$\begin{aligned}\theta = \{ & I_L, Q_L, \sigma_x^{(L)}, \sigma_y^{(L)}, \theta_L, \\ & I_R, Q_R, \sigma_x^{(R)}, \sigma_y^{(R)}, \theta_R, \\ & \sigma_x^{(C)}, \sigma_y^{(C)}, \theta_C, \\ & A_L, A_C, \varphi \}\end{aligned}\tag{5.8}$$

Each measurement step during the reconstruction protocol returns a set of b^2 measured Wigner values $\{W_{\text{meas}}^{(i)}\}_{i=1}^b$ for points in phase space $\{\beta^{(i)}\}_{i=1}^b$, which together with the measurements from the previous iterations form a dataset of N pairs $\{\beta^{(i)}, W_{\text{meas}}^{(i)}\}_{i=1}^N$. The actual Wigner function of the state is unknown; thus, the observed data $\{\beta^{(i)}, W_{\text{meas}}^{(i)}\}_{i=1}^N$ is approximated with a parametric family of distributions $W_{\text{model}}(\beta, \theta)$ from Eq. 5.3. The goal is to fit the observed data with the chosen model as close as possible.

Experimentally measured Wigner values are noisy. Assuming that the measured Wigner values $W_{\text{meas}}^{(i)}$ at each phase-space coordinate $\beta^{(i)}$ follow a normal distribution $\mathcal{N}$ with variance σ^2 due to measurement noise, a likelihood function can be constructed to quantify the probability that the model predicts the observed value $W_{\text{meas}}^{(i)}$ at the given phase space coordinate $\beta^{(i)}$

$$\mathcal{L}^{(i)}(\theta) = p(W_{\text{meas}}^{(i)} | \beta^{(i)}, \theta) = \mathcal{N}(W_{\text{meas}}^{(i)} | \mu = W_{\text{model}}(\beta^{(i)}, \theta), \sigma^2).\tag{5.9}$$

For N independent and identically distributed data points, the likelihood function generalises to

$$\mathcal{L}(\theta) = \prod_{i=1}^N \mathcal{N}(W_{\text{meas}}^{(i)} | \mu = W_{\text{model}}(\beta^{(i)}, \theta), \sigma^2).\tag{5.10}$$

The likelihood function $\mathcal{L}(\theta)$ quantifies the agreement between the model predictions and the observed data. Maximising likelihood $\mathcal{L}(\theta)$ with respect to the model parameters θ therefore yields the parameter set that produces the closest match to the measured values³

$$\theta_{\text{ML}} = \arg \max_{\theta} \mathcal{L}(\theta).\tag{5.11}$$

An estimator θ_{ML} from Eq. 5.11 is called a maximum likelihood estimator. Typically, it is more convenient to maximise log likelihood instead, since it avoids numerical underflow by replacing a multiplication of probabilities with a sum of log probabilities

$$\theta_{\text{ML}} = \arg \max_{\theta} \log \mathcal{L}(\theta) = \arg \max_{\theta} \sum_{i=1}^N \log \mathcal{N}(W_{\text{meas}}^{(i)} | \mu = W_{\text{model}}(\beta^{(i)}, \theta), \sigma^2).\tag{5.12}$$

² b stands for batch size, which is a number of measured/sampled points per reconstruction iteration.

³Typically, regression problems in physics are solved by minimising the sum of squares. Minimising the sum of squares is equivalent to maximising likelihood as described in this Section; see [43] or [93] for a proof.

Once the parameter estimation θ_{ML} is obtained from the maximum likelihood, it is fixed and used for the model predictions. This approach is known as statistical (frequentist) inference.

5.2 Bayesian model

Maximum likelihood is a powerful tool widely used for various tasks in physics and machine learning. However, it tends to overfit the data when the number of available samples is low. This problem is clearly evident during the first reconstruction iterations, when the dataset is limited to a few tens of points. The issue can be addressed by adopting a Bayesian inference paradigm, which incorporates prior knowledge on model parameters into the model.

Contrary to the frequentist perspective, in Bayesian inference, the model parameters are not fixed when making a prediction; instead, they follow a prior probability distribution. The model prediction follows the Bayes' rule

$$p(\theta|W_{\text{meas}}^{(i)}, \beta^{(i)}) \propto p(W_{\text{meas}}^{(i)}|\beta^{(i)}, \theta) \times p(\theta), \quad (5.13)$$

$$\text{posterior} \propto \text{likelihood} \times \text{prior}, \quad (5.14)$$

which claims that the posterior probability $p(\theta|W_{\text{meas}}^{(i)}, \beta^{(i)})$ (probability of having a set of parameters θ given observed data $W_{\text{meas}}^{(i)}$) is a result of a likelihood $p(W_{\text{meas}}^{(i)}|\beta^{(i)}, \theta)$ (probability of getting an observed data from the model given the set of parameters θ) and prior $p(\theta)$ that represents some knowledge on θ before observing the data. The prediction of the Bayesian model, therefore, gives the distribution for all θ and puts more weight on the parameter values that had a higher prior probability $p(\theta)$. The prior $p(\theta)$ is defined by the user before the inference and is usually based on the problem knowledge. The best estimate of the model parameters can be found with a maximum a posteriori (MAP) estimate, which returns θ of maximum posterior probability $p(\theta|W_{\text{meas}}^{(i)}, \beta^{(i)})$:

$$\begin{aligned} \theta_{\text{MAP}} &= \arg \max_{\theta} \log \left(p(\theta) \prod_{i=1}^N p(W_{\text{meas}}^{(i)}|\beta^{(i)}, \theta) \right) \\ &= \arg \max_{\theta} \log \mathcal{L}(\theta) + \log p(\theta). \end{aligned} \quad (5.15)$$

In the Bayesian model used for the reconstruction, the first (likelihood) term in 5.15 is defined according to the Eq. 5.10:

$$P_{\text{likelihood}} = \sum_{i=1}^N \log \mathcal{N}(W_{\text{meas}}^{(i)} | \mu = W_{\text{model}}(\beta^{(i)}, \theta), \sigma = 0.01), \quad (5.16)$$

where $\mathcal{N}$ is a normal distribution, $W_{\text{meas}}(\beta^{(i)})$ - measured Wigner values at phase space points $\beta^{(i)}$, $W_{\text{model}}(\beta^{(i)})$ - Wigner values at phase space points $\beta^{(i)}$ predicted by the model, and σ is fixed to 0.01 before the start of the reconstruction.

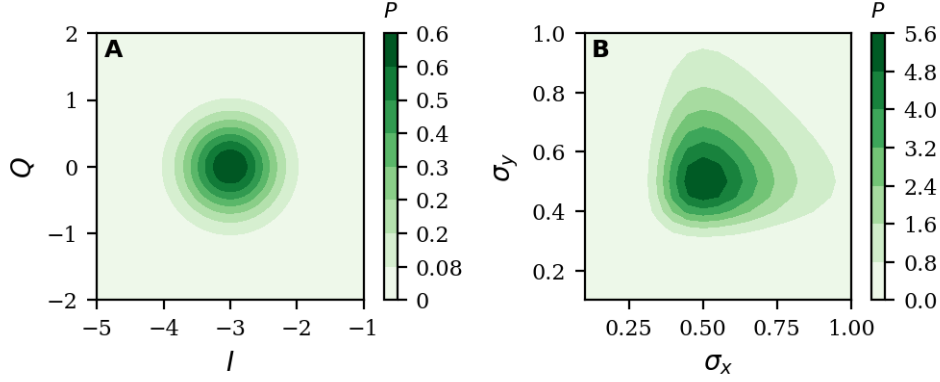


Figure 5.2: Prior distributions for the left Gaussian, which push the model to find a solution close to an ideal $\alpha = 3$ cat state during the first reconstruction steps. **A** The probability to find the left Gaussian centre μ_L follows a normal distribution $\mathcal{N}((-3, 0), \text{diag}(0.25, 0.25))$. The most probable position for the left Gaussian corresponds to coordinate $(-3, 0)$ in phase space, as expected for an ideal cat state. **B** Prior probability distribution for σ_x and σ_y of the left Gaussian, which is defined by an inverse Wishart distribution $\mathcal{W}^{-1}(\psi = \text{diag}(0.25, 0.25), \nu = 1)$. Note that $\mathcal{W}^{-1}$ is defined on 2×2 positive definite covariance matrix, which is parametrised by three parameters: σ_x , σ_y , ϕ . For this plot, ϕ is fixed to 0. Again, the most probable Gaussian is isotropic, as it should be for an ideal cat state.

The prior (second term in 5.15) is defined as

$$\begin{aligned}
 P_{\text{prior}} = & \log \mathcal{N}(\mu_L; \mu_L^{\text{prior}}, \Sigma_L^{\text{prior}}) + \log \mathcal{N}(\mu_R; \mu_R^{\text{prior}}, \Sigma_R^{\text{prior}}) + \\
 & \log \mathcal{W}^{-1}(\Sigma_L; \psi_L^{\text{prior}}, \nu_L^{\text{prior}}) + \\
 & \log \mathcal{W}^{-1}(\Sigma_R; \psi_R^{\text{prior}}, \nu_R^{\text{prior}}) + \\
 & \log \mathcal{W}^{-1}(\Sigma_C; \psi_C^{\text{prior}}, \nu_C^{\text{prior}})
 \end{aligned} \tag{5.17}$$

The first two terms in Eq. 5.17 encode prior knowledge about the positions of the left and right Gaussian centres, μ_L and μ_R , while the centre of the middle Gaussian, μ_C , is determined according to Eq. 5.5. Eq. 5.17 assumes that the probability of finding the left (right) Gaussian centre at a phase space coordinate $\mu_{L(R)}$ is modelled by a prior normal distribution, $\mathcal{N}(\mu_{L(R)}; \mu_{L(R)}^{\text{prior}}, \Sigma_{L(R)}^{\text{prior}})$. Here, $\mu_{L(R)}^{\text{prior}}$ and $\Sigma_{L(R)}^{\text{prior}}$ denote the mean and covariance of the prior distribution, which are predefined based on prior knowledge of the state. For a cat state of size α , for instance, the prior means are set to $\mu_{L(R)}^{\text{prior}} = (\mp\alpha, 0)$, which pushes the model to find the solution close to an ideal cat state case (see Fig. 5.2 A for an example).

The subsequent terms in Eq. 5.17 incorporate prior knowledge about the covariance matrices Σ_L , Σ_R , and Σ_C , which describe the left, right, and central Gaussians, respectively. The distribution for a covariance matrix Σ of a multivariate normal distribution is an inverse Wishart distribution, $\mathcal{W}^{-1}(\Sigma; \psi, \nu)$, parametrised by the scale matrix ψ and the

degrees of freedom ν (see Fig. 5.2 B). The scale matrix ψ mostly affects the mode⁴ of the distribution, while the degrees of freedom ν have a significant influence on its variance: larger ν values yield a sharper distribution, increasing the probability of finding Σ close to the mode (see Fig. B.1). Both ψ and ν for each Gaussian are specified beforehand based on prior problem knowledge. Further details on the inverse Wishart distribution are provided in Appendix B, and all the parameters of the prior distributions used for the reconstruction are presented in Section 5.7.1.

To conclude, the model parameters are updated based on a MAP estimate⁵ (Eq. 5.15), with the likelihood and prior terms defined in equations 5.16 and 5.17:

$$\theta_{\text{MAP}} = -\arg \min_{\theta} (P_{\text{likelihood}}/b) - r \cdot P_{\text{prior}}, \quad (5.18)$$

where r is a coefficient that controls the impact of the prior, and b is the number of measured/sampled points per iteration (batch size). The normalisation of the likelihood term with respect to the batch size ensures similar convergence for any batch size, provided a proper prior is used.

5.3 Sampling from the distribution

The easiest active sampling strategy is to draw the samples directly from the three Gaussians describing the cat Wigner distribution 5.3 with equal probability:

$$\begin{aligned} \beta &\sim \mathcal{N}(\mu_i, \Sigma_i) \\ i &\in \{L, R, C\}, \quad p_L = p_R = p_C = \frac{1}{3} \end{aligned} \quad (5.19)$$

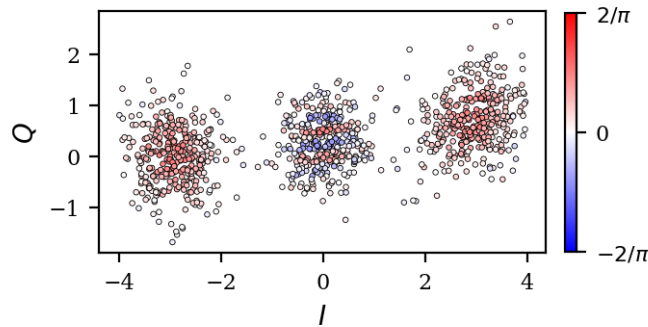


Figure 5.3: Example of sampling from the Gaussians defining the model 5.3. Each new point is sampled from the left, centre, or right Gaussian with an equal probability of 1/3. After sampling and measuring the batch of points, the model parameters are updated before starting the new sampling iteration.

⁴The mode of the probability density function $p(x)$ is the value x , where the density reaches its maximum, i.e. $\arg \max p(x)$

⁵Maximising likelihood is equivalent to minimising negative likelihood

Sampling from the model distribution is very fast, but not always optimal: some points are sampled from the areas that do not improve the model, i.e. "blank" areas in the Wigner map, as shown in Fig. 5.3. Despite its simplicity, direct sampling from the distribution shows decent results in a real experiment, as shown later in Section 5.7.

5.4 Bootstrapping and exit conditions

Although the sampling strategy in Eq. 5.19 is not optimal, it offers a practical starting point for constructing the reconstruction algorithm. At each iteration, the points sampled using 5.19, along with the measurements from previous iterations, are passed to the model. As a result of Bayesian inference (Section 5.2), a new set of model parameters is obtained and used for sampling in the next iteration. The final missing component of Algorithm 1 is the exit condition, which can be determined based on the model's uncertainty. One possible approach to quantify this uncertainty is bootstrapping [94]. Bootstrapping is an effective method for estimating the statistical properties of complex estimators⁶ or those derived from unknown probability distributions. In the context of the state reconstruction problem, bootstrapping addresses the question of how confident the model's predictions are at a given iteration, given the data acquired up to that point.

Consider an original dataset $D = \{(\beta^{(i)}, W_{\text{meas}}^{(i)})\}_{i=1}^N$. From this dataset, one can generate M new datasets D_k , where $k \in \{1, \dots, M\}$, by randomly sampling points from D with replacement. Each resampled dataset D_k has the same size as the original dataset D . The central idea of the bootstrapping method is that the statistical properties of an estimand⁶ (such as its mean or variance) can be inferred by computing the estimator on each resampled dataset D_k and analysing the resulting distribution of estimates.

The concept becomes clear by considering a specific example. Fig. 5.4 shows a schematic representation of the bootstrapping procedure for $M = 10$ models. Each model is initialised with the same set of model parameters, but takes a different input dataset D_k that is sampled from the measured data D with replacement. Due to 10 different input datasets, the MAP estimation returns 10 different sets of model parameters. Each model generates its own Wigner map, which is used to calculate, for example, fidelity. The mean and variance of the fidelity are obtained from the 10 output fidelities that 10 bootstrapped models estimated.

The role of the estimand is not limited to fidelity. For instance, it can also be represented by the mean or integrated variance between M Wigner maps. Algorithm 1 is terminated once the chosen estimand goes below a chosen threshold. To gain time, the MAP estimation for bootstrapped models is calculated in parallel by using the **multiprocessing** Python package.

⁶Estimand is a statistical quantity one wants to estimate. Estimator is a function of the input data, that produces an estimation of an estimand. The most prominent estimators are maximum likelihood (ML) and maximum a posteriori (MAP), introduced in Sections 5.1 and 5.2, respectively. Further reading on this topic is recommended in [43].

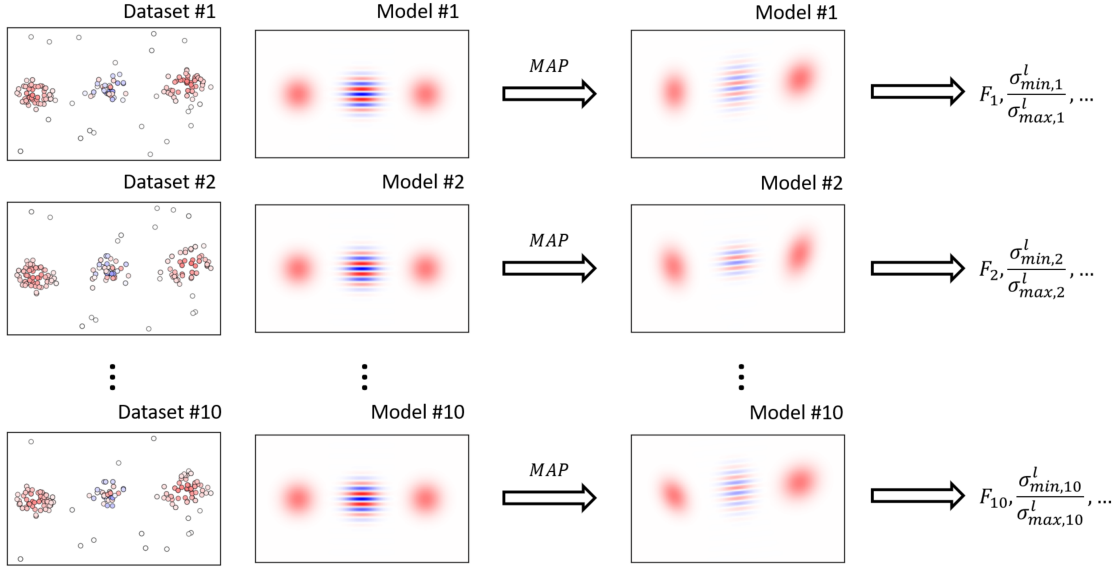


Figure 5.4: Schematic representation of the bootstrapping procedure. 10 identical models are optimised on 10 different datasets $\{D_k\}_{k=1}^{10}$, sampled from an original dataset (D) with replacement. Each generated dataset D_k contains the same number of points as the original dataset D . As a result of maximum a posteriori (MAP) estimation (see Section 5.2), each model converges to a distinct set of parameters and produces a unique Wigner map. Generated Wigner maps can be used to calculate the mean and variance of the estimands of interest (e.g., fidelity or ratio of Gaussians' deviations).

Different bootstrapping initialisation strategies lead to distinct convergence behaviours. In most experimental runs, the models are initialised as follows. Alongside the bootstrapped models, an additional reference model performs a MAP estimation using all the measured data D . The parameters obtained from this reference model serve as initial guesses for the bootstrapped models in the subsequent iteration. In general, it is not necessary to initialise all bootstrapped models identically. Alternatively, the initial parameters can be drawn from the prior distributions. This approach can reduce potential model bias at the cost of increased reconstruction time. In practice, the first strategy proved more advantageous, showing no significant bias while maintaining fast and stable convergence.

5.5 Sampling based on the deviation between the bootstrapped models

The bootstrapping approach not only establishes a termination condition but also provides an alternative active sampling strategy. New measurement points can be selected from regions of highest uncertainty (variance) between the bootstrapped models. Figure 5.5 illustrates the standard deviation between Wigner maps generated by 10 bootstrapped models trained on resampled datasets $\{D_k\}_{k=1}^{10}$. This deviation, once normalised, can serve as a probability distribution for selecting new sampling points. To allocate more

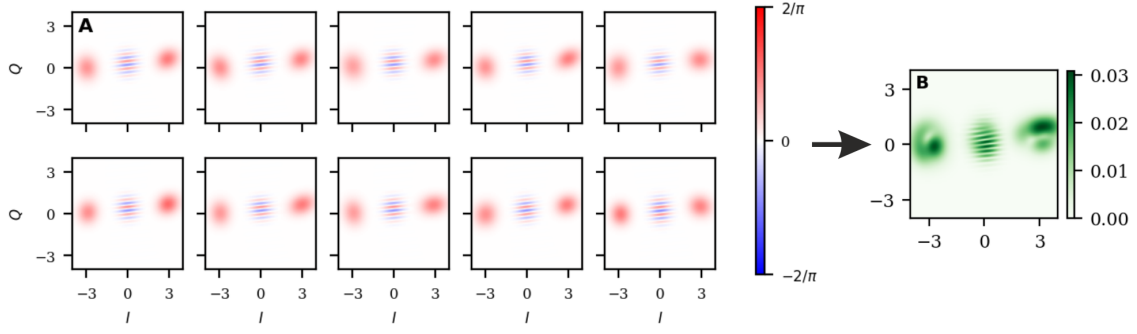


Figure 5.5: Deviation map for Boltzmann distribution 5.20. **A** Wigner maps generated by 10 bootstrapped models. **B** Unnormalised standard deviation $\sigma(I, Q)$ between the Wigner maps from **A**. $\sigma(I, Q)$ values are later used in Eq. 5.20.

probability to high variance regions and ensure a proper normalisation, the samples are drawn from the Boltzmann distribution defined over the deviation values

$$p_B(I, Q) = \frac{e^{\beta_B \sigma(I, Q)}}{\sum_{I, Q} e^{\beta_B \sigma(I, Q)} \delta I \delta Q}, \quad (5.20)$$

where $\sigma(I, Q)$ is a standard deviation between the Wigner function values of bootstrapped models at the point (I, Q) in phase space. δI and δQ are the phase space grid sizes, and $\beta_B = 1/T$ is the inverse temperature of the Boltzmann distribution. The lower the temperature of the Boltzmann distribution, the sharper the resulting sampling distribution becomes.

Eq. 5.20 has no closed-form expression, making the sampling procedure challenging. Fortunately, the problem of drawing samples from complex probability distributions is common in the machine learning community and can be addressed in various ways. One possible solution is to use a rejection sampling method [93].

For simplicity, the method's idea is explained here for a target distribution $f(x)$ defined over a one-dimensional random variable x (see Fig. 5.6). Since it is not possible to sample from the target distribution $f(x)$ directly, the rejection sampling proposes to get a sample x from some other known distribution - proposal distribution $g(x)$ - and accept the point sampled from the proposal distribution if it fulfils the rejection criteria:

$$\begin{aligned} u &\sim C \cdot U(0, 1) \\ u &\leq f(x), \end{aligned} \quad (5.21)$$

where $U(0, 1)$ is the uniform distribution between 0 and 1, and C is a scaling factor which guarantees that the condition $\forall x : f(x) \leq C \cdot g(x)$ is satisfied. It assures that the proposal distribution confines the target distribution; otherwise, some values of x will be over- or undersampled. In this Chapter, the samples that are generated from the proposal distribution are referred to as proposed samples.

The larger the rejection region, the greater the number of proposed samples required to obtain a valid sample x from the target distribution $f(x)$. If the number of rejected

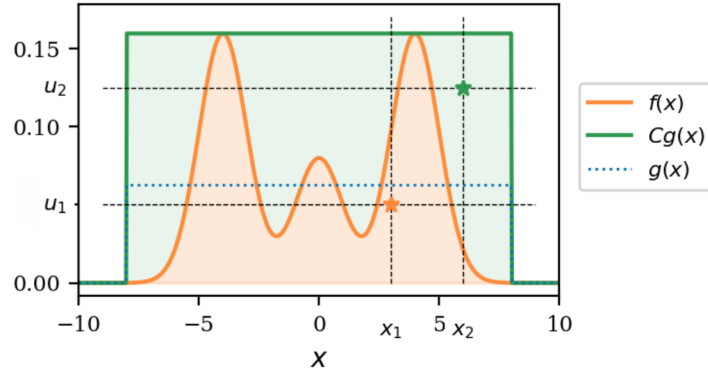


Figure 5.6: Rejection sampling method for a one-dimensional case. $f(x)$ - target distribution, $g(x)$ - proposal (uniform) distribution, from which the samples x_1 and x_2 are generated. u_1 and u_2 - two randomly generated values between 0 and 1 for points x_1 and x_2 . Since $u_1 \leq f(x_1)$, the point x_1 is accepted (orange star). x_2 is rejected, since $u_2 > f(x_2)$ (green star). The rejection area is coloured green, and the acceptance area is coloured orange.

proposed samples is too high, it is more efficient to generate a specified number of points (e.g. with **NumPy**⁷ functions) and then check whether they satisfy the rejection criteria instead of getting proposed sampling points one by one and checking the rejection criteria for each of them, as done in the algorithm below

Algorithm 2 Rejection sampling algorithm

Require: $f(x)$ - target distribution, $g(x)$ - proposal distribution, $U(0, 1)$ - uniform distribution between 0 and 1, N - number of sampling points drawn from the proposal distribution, n - required number of samples

Ensure: $N \gg n$

$C \leftarrow \max \frac{f(x)}{g(x)}$

$x[1, ..N] \sim g$

▷ sample x from proposal distribution $g(x)$

$u[1, ..N] \sim C \cdot U(0, 1)$

▷ sample u from uniform distribution $U(0, 1)$

$n_{\text{sampl}} \leftarrow \text{argwhere } u[1, ..N] \leq C \cdot g(x[1, ..N])$

▷ check the rejection criteria 5.21

$x_{\text{sampl}} \leftarrow x[n_{\text{sampl}}]$

▷ return samples that satisfy the rejection criteria

return $x_{\text{sampl}}[1:n]$

Since the position of the cat state in phase space during the future closed-loop optimisation runs will be unknown, a proposal distribution for the reconstruction is chosen as a two-dimensional uniform distribution:

$$p = \begin{cases} 1/((I_{\max} - I_{\min})(Q_{\max} - Q_{\min})) & \text{for } I \in [-4, 4], Q \in [-4, 4], \\ 0, & \text{otherwise.} \end{cases} \quad (5.22)$$

⁷NumPy is based on pre-compiled, optimised C-code and outperforms any for-loop-based implementation of the same operation in Python. It becomes crucial once such a large number of rejection sampling points is generated.

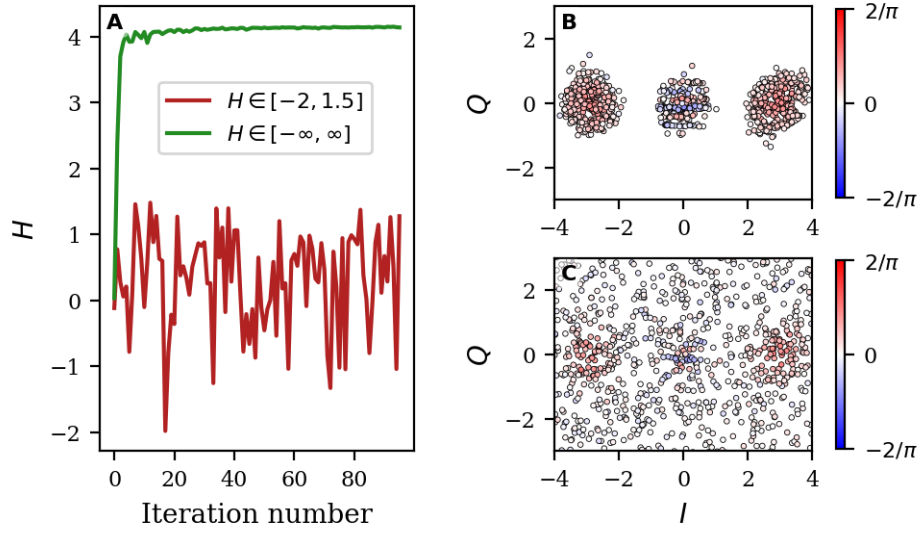


Figure 5.7: Entropy evolution during the reconstruction process for an adaptive strategy, where the algorithm constrains the entropy H within the bounds $[-2, 1.5]$, and a non-adaptive strategy, where the entropy is not adjusted, and their respective influence on the sampling process. **A** Entropy as a function of the iteration number. **B** Sampled points during the reconstruction process when the entropy is kept within the bounds ($H \in [-2, 1.5]$). **C** Sampled points during the reconstruction process when the entropy is not adjusted ($H \in [-\infty, \infty]$).

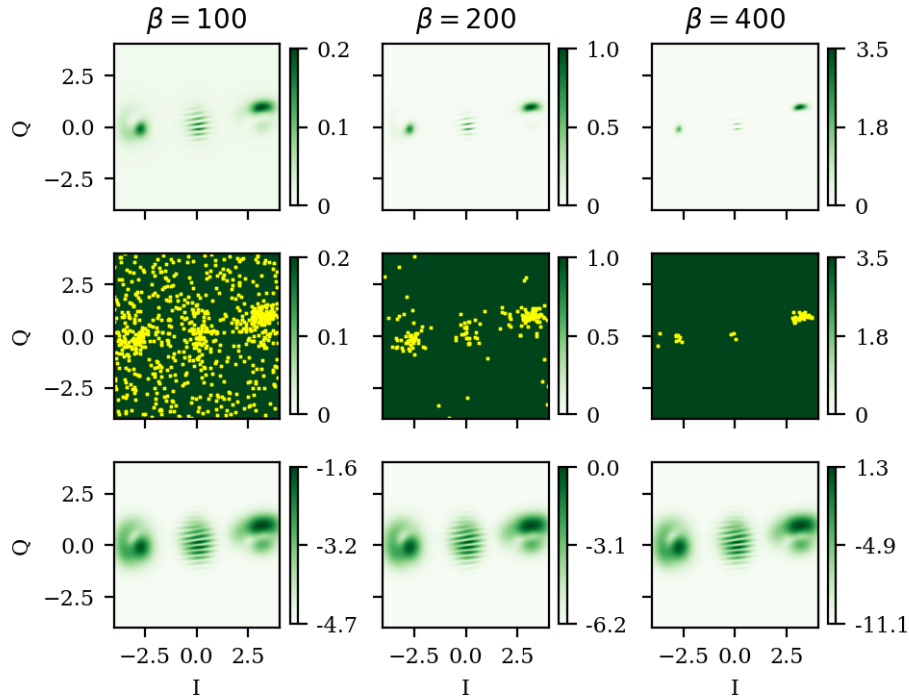


Figure 5.8: Example of the target distribution for rejection sampling at different inverse temperatures $\beta_B = 1/T$. The first row shows the target distribution, calculated with equation 5.20 from the deviation between the Wigner maps $\sigma(I, Q)$, generated by 10 bootstrapped models after the optimisation. The second row shows the scaled proposal probability distribution $Cg(x)$ (here, $g(x)$ is a uniform distribution), and samples generated with rejection sampling (yellow markers). The last row shows the same plots as the first row, but on a logarithmic scale.

Although the uniform proposal distribution ensures that the target distribution is fully encapsulated into the proposal distribution at any time step, it introduces a huge rejection area. Depending on the temperature of the Boltzmann distribution 5.20, Algorithm 2 typically needs $N = 10^4 - 10^6$ proposed samples to get dozens of sampled points (depending on the inverse temperature β_B). If the number of generated rejection sampling points is less than the requested number of samples in the batch, the rest is sampled from the model distribution 5.19.

As the reconstruction proceeds, the bootstrapped models yield increasingly similar predictions. Therefore, the Boltzmann distribution 5.20 becomes smoother, and the new points start to be sampled more frequently from low-informative ("blank") regions of the Wigner map (see Fig. 5.7). To counteract this effect, the distribution temperature is adjusted based on its differential entropy

$$H = - \sum_{I,Q} p(I, Q) \log p(I, Q) \delta I \delta Q = - \sum_{I,Q} p(I, Q) \cdot (\beta_B \sigma(I, Q) - \log N) \delta I \delta Q \quad (5.23)$$

$$N = \sum_{I,Q} e^{\sigma(I,Q)} \delta I \delta Q \quad (5.24)$$

Note that the differential entropy 5.23 can take negative values, which is counterintuitive from a physics perspective. Once the distribution entropy reaches a specified upper limit, the inverse temperature is decreased by 50%. This way, the spread of sampling points in phase space remains confined, and the sampled points are located next to the areas where the predictions of bootstrapped models diverge the most. On the other hand, a too-sharp distribution may result in a large rejection area (see Fig. 5.8). This, in turn, requires many proposed samples and significantly increases sampling time. To avoid this, the inverse temperature is increased if the entropy reaches a lower threshold. The behaviour of the sampling technique for different cat states and entropy range $[-2, 1.5]$ is presented in Fig. 5.9.

Most of the time spent on the sampling procedure is devoted to rejection sampling. Although the idea and implementation of rejection sampling are simple compared to other methods, it suffers from high computational costs if the proposal distribution is far from the target one. Adjusting the entropy only partially fixes this problem. In the current implementation, this limitation does not result in an excessive increase in reconstruction time.

5.6 Pipeline

At this point, all the steps of Algorithm 1 have been covered in the previous sections. The actual implementation of the algorithm has undergone modifications to ensure convergence and reduce reconstruction time. Since the measurement is costly, the hardware measurement and calculations on the CPU for selecting the next measurement points are performed in parallel in two threads (see Fig. 5.10). In this configuration, the time spent on one iteration is limited by the time required for the longer operation (measurement or CPU calculations). In the optimal case, measurement and calculation times are approxi-

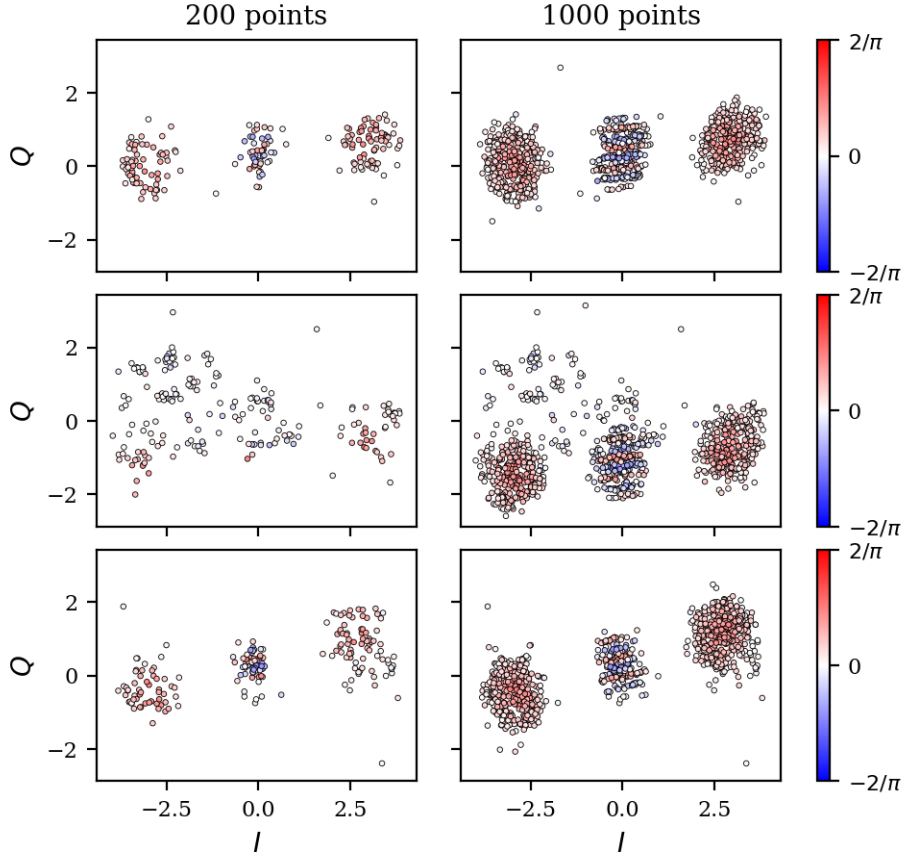


Figure 5.9: Example of sampling based on the deviation between the Wigner maps, predicted by bootstrapped models. The Boltzmann distribution inverse temperature β_B is dynamically adjusted during the reconstruction to keep the differential entropy in a range $[-2, 1.5]$. Upper row corresponds to the reconstruction of the initial cat state, middle row - to a displaced cat state, bottom row - to a rotated cat state, see Fig. 5.11 for high-resolution Wigner maps as a reference. The first column shows the first 200 measured points, and the second column shows all 1000 measured points during the reconstruction.

mately equal, leading to $\times 2$ speed-up. However, in this realisation, the choice of the next sampling points is always one step behind, as it relies on measurements acquired in the previous step.

During the first two iterations of the reconstruction, sampling can only be performed from the Gaussians defined in the initial model guess (5.19), since no measurement data are yet available for parameter updates. All subsequent iterations may employ both sampling strategies. However, in the early stages, the predictions of the bootstrapped models can either differ significantly or be nearly identical, depending on the specific bootstrapped datasets and measured values. As a result, the optimal initial temperature for the rejection sampling (see Fig. 5.8) varies considerably between runs. To improve stability, the first 40 sampling points are therefore always drawn from the model distribution 5.19.

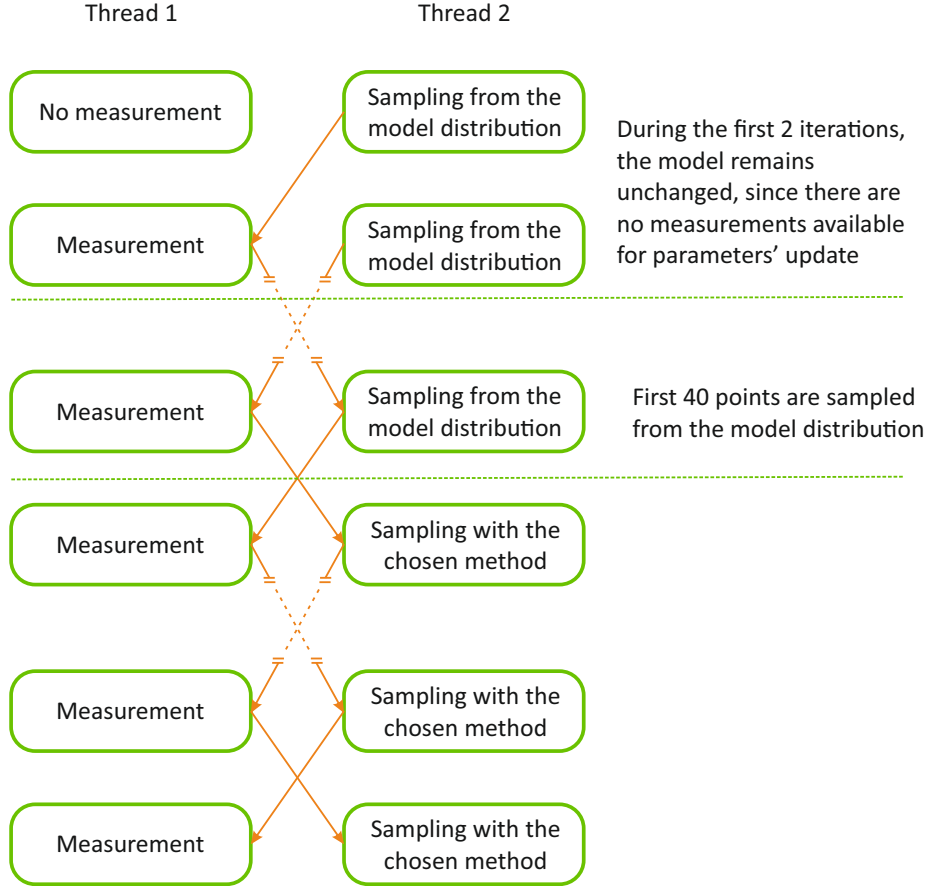


Figure 5.10: The schematics of the data processing pipeline for the reconstruction algorithm.

The reconstruction stops when either the maximum number of iterations is reached or the integrated variance between the different bootstrapped models goes below a specified threshold. The choice of termination reconstruction conditions is parameter- and task-specific and will be discussed in Section 6.1 in more detail.

5.7 Experimental runs

5.7.1 Parameter configuration: priors and bounds

The reconstruction performance is benchmarked using an $\alpha = 3$ cat state. To amplify the Kerr-induced distortions, the disentanglement pulse is deliberately chosen to be significantly longer than required and realised as a 4σ -Gaussian with $\sigma = 40$ ns. Unselective qubit pulses employ a 4σ -Gaussian waveform with $\sigma = 9$ ns, while displacement pulses have a rectangular shape with sine-smoothed edges and take 20 ns in total.

The qcMAP protocol provides a limited flexibility in creating various cat states of the same size. Consequently, displaced and rotated cat states were experimentally realised by

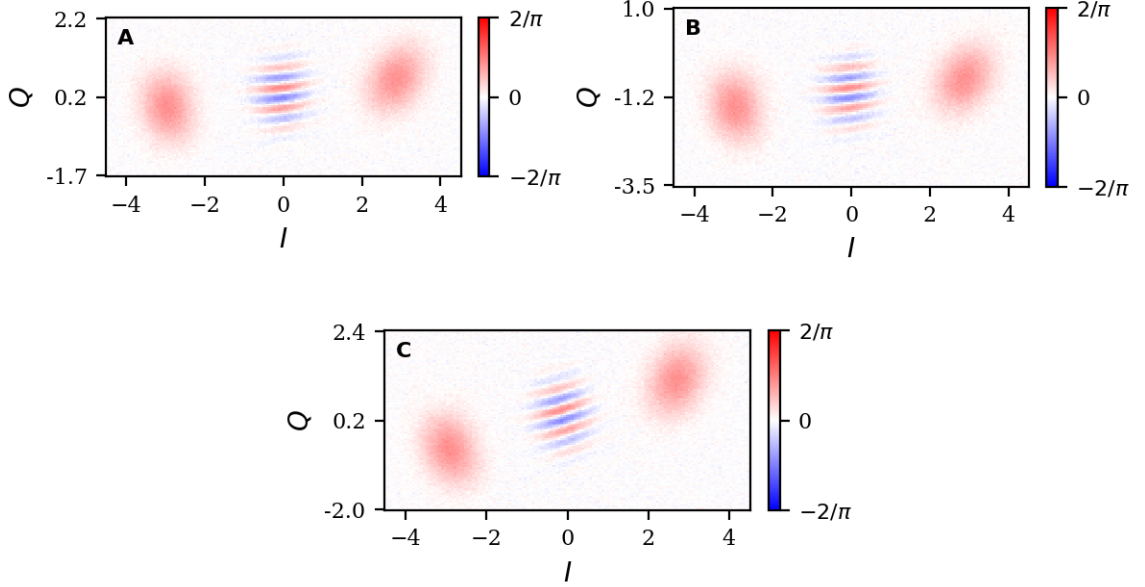


Figure 5.11: Cat states used for the characterisation of the reconstruction method. The Wigner maps are measured with 1000 averages. **A** Cat state created with the qcMAP sequence and $\sigma = 40$ ns disentanglement pulse. **B** Same as **A**, but the phase space axes definition is shifted by $(0, -1)$ on the hardware. **C** Same as **A**, but the phase space axes definition is rotated by 10° on the hardware.

redefining the hardware phase space axes. The corresponding measured cat states used as a reference for the reconstruction are presented in Fig. 5.11.

All the experimental runs share the same prior for the Bayesian model. The centres of the left and right Gaussians are assumed to be located close to their ideal positions at ± 3

$$\begin{aligned}\mu_L &\sim \mathcal{N}(\mu_L^{\text{prior}} = (-3, 0), \Sigma_L^{\text{prior}} = \text{diag}(0.25, 0.25)), \\ \mu_R &\sim \mathcal{N}(\mu_R^{\text{prior}} = (3, 0), \Sigma_R^{\text{prior}} = \text{diag}(0.25, 0.25)).\end{aligned}$$

The covariance priors follow a high-entropy inverse Wishart distribution. It avoids undesirable biasing towards the prior, while keeping the model away from unrealistic predictions on covariance matrices

$$\mathcal{W}_L^{-1} = \mathcal{W}_R^{-1} = \mathcal{W}_C^{-1} = \mathcal{W}^{-1}(\psi = \text{diag}(0.25, 0.25), \nu = 1)$$

To get an intuition on how the scale matrix ψ and degrees of freedom ν change the inverse Wishart distribution, see Fig. B.1 from Appendix B.

In addition, all the runs use the regularisation coefficient⁸ $r = 0.25$ (see Eq. 5.18), and $M = 10$ bootstrapped models for parameter estimation and sampling (unless otherwise specified). During the MAP estimation, the model parameters are forced to be within the following bounds:

$$\begin{aligned}
\mu_{L,R} &\in [-4, 4] \\
\sigma_x^{l,r,c}, \sigma_y^{l,r,c} &\in [0.25, 1] \\
\theta^{l,r,c} &\in [-\frac{\pi}{2} - \epsilon, \frac{\pi}{2} + \epsilon] \\
\varphi &\in [-\pi - \epsilon, \pi + \epsilon] \\
A_L &\in [0.5 - A_{\text{imb}} - \epsilon, 0.5 + A_{\text{imb}} + \epsilon] \\
A_C &\in [0.1, 0.5 + \epsilon]
\end{aligned} \tag{5.25}$$

Here, $A_{\text{imb}} = 0.05$ is an amplitude imbalance between the left and the right Gaussians that can occur, for example, when the disentanglement pulse is not flipping the qubit completely, and $\epsilon = 10^{-3}$ is a small number that allows the minimisation algorithm to reach upper/lower bounds of the intervals.

5.7.2 Results & discussion

The configuration of the reconstruction algorithm is defined by its hyperparameters, which change the performance in terms of accuracy and time consumption. There are three main hyperparameters: number of averages per measurement point, batch size, and number of bootstrapped models. On top of this, there are two available sampling methods and two different ways of initialising bootstrapped models at each iteration.

The benchmarking results for different algorithm configurations are shown in Figs. 5.14 to 5.18. Each configuration is represented by four plots. The first two display the ratios σ_x^l/σ_y^l and σ_x^r/σ_y^r (ellipticities) of the left and right Gaussians, calculated in their principal axis coordinates. The other two plots present the infidelity and integrated residuals between the model fit to the reference Wigner map from Fig. 5.11 and the reconstructed Wigner map. The infidelity utilises the F_2 fidelity definition from the Eq. 3.54. Integrated residuals between the measured Wigner values W_{meas} and the values predicted by the model W_{model} are defined as

$$\begin{aligned}
R &= \iint |W_{\text{meas}}(I, Q) - W_{\text{model}}(I, Q)| dI dQ \\
&\approx \sum_{I, Q} |W_{\text{meas}}(I, Q) - W_{\text{model}}(I, Q)| \delta I \delta Q,
\end{aligned} \tag{5.26}$$

where δI and δQ are map grid steps. An example of the reconstructed Wigner map and residuals between the prediction and observed data is shown in Fig. 5.12.

Each configuration of the reconstruction algorithm in Figs. 5.14 to 5.18 was tested 100 times in the experiment. For a more intuitive representation of the statistics obtained over

⁸This value demonstrated better convergence in the experiment. Alternatively, the convergence behaviour can be adjusted by modifying the parameters of the prior distributions.

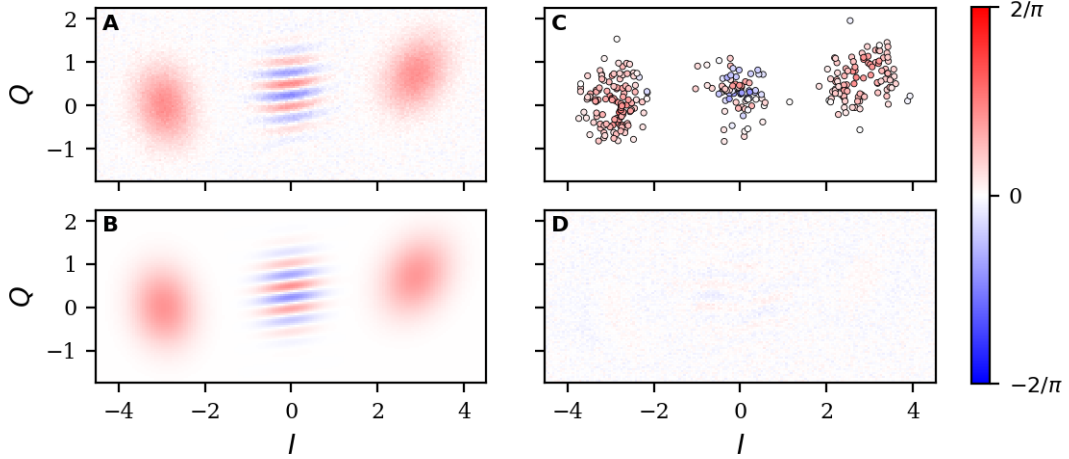


Figure 5.12: Example of the state reconstruction. **A** Reference high-resolution Wigner map. **B** Reconstructed Wigner map. **C** First 300 (out of 1000) measured points during the reconstruction. **D** Residuals between the measured and reconstructed data $R(I, Q) = W_{\text{meas}}(I, Q) - W_{\text{model}}(I, Q)$. Since the model does not account for the bending of the fringes, there is always a small deviation between the model and the observed data in the fringe area.

100 runs for different estimands, plots contain three shades of green. The most intense green colour corresponds to a deviation within σ from the mean, the middle intense green corresponds to a deviation within 2σ , and the most transparent green corresponds to a deviation within 3σ . 3σ -deviation values are particularly important for closed-loop applications: small 3σ -deviation indicates stable state reconstruction that can be reliably integrated into future optimisation runs. The dashed black lines mark the target estimand values for reference.

The number of bootstrapped models is fixed at $M = 10$ due to CPU limitations for all measurements, except those from Fig. 5.17. The list of tested configuration variations includes:

1. Different sampling strategies: sampling from the distribution vs sampling based on the models' deviations, Fig. 5.14

Even though the ratios $\sigma_x^{l,r}/\sigma_y^{l,r}$ converge to similar values, sampling from the model distribution shows a significantly lower $1 - F$ and R values after 1000 measured points. The reason lies in the wrong reconstructed value of the rotation angle for left and right Gaussians θ_L and θ_R . Additionally, the standard deviation for all the estimands is higher when sampling from the model distribution.

The sampling based on the deviation between the bootstrapped models requires the inverse temperature value β_B , which is controlled with an allowed range of the distribution entropy. Fig. 5.14 shows results for two different entropy ranges: $H \in [-2, 1.5]$ and $H \in [-2, 0.1]$. There is no significant difference between the two above-mentioned

regimes. For $H \in [-2, 0.1]$, the infidelity converges to slightly higher values with a lower standard deviation.

2. Different number of averages, Fig. 5.15

The number of averages per measurement point controls the noise of the data and the algorithm's convergence. The configuration employing 100 averages per point demonstrated the most stable and reproducible convergence curves across repeated runs. 10 averages per point always have some unsuccessful runs: 3σ -area is huge for both residuals and infidelity plots. Some runs end with an infidelity close to 1, which means that the reconstruction fails and cannot give reproducible results for such a low number of averages. Note that those outlying runs reveal on 3σ -deviation plots only. Averaging over 1000 points yields the smallest deviation between consecutive reconstruction runs but results in larger deviation in infidelity and residuals. It is important to note that the runs with 1000 averages were terminated after 500 measurement points, whereas all other configurations continued up to 1000 measurements. A typical run with 1000 averages and 500 points takes approximately 684 seconds, which is already about twice as long as the average reconstruction time for other configurations. Running this configuration for 1000 measurement points would require more than a full day of continuous measurements.

At this point, it is important to recall that all the measurements presented in this thesis employed a non-QND high-power readout scheme, which required signal averaging. For a single-shot readout, the convergence behaviour, as well as the reconstruction time, may differ substantially.

3. Different batch size b , Fig. 5.16

The larger the batch size, the faster, but less accurate, the reconstruction process. For 15 measurements per batch, the model appears to be biased after 1000 measurement points for both left and right ratios. In addition, the infidelity remains relatively high. 5 and 10 measurements per batch show compatible results for the estimands $1 - F$ and R . The $b = 5$ configuration shows the least bias for $\sigma_x^{l,r}/\sigma_y^{l,r}$, which comes with a significant increase in the reconstruction time.

The sampling time is approximately 1.24 s and remains independent of the batch size, owing to the implementation of the sampling algorithm. For batch sizes $b \in \{4, 5, 6\}$, the sampling time is comparable to the measurement time (see Fig. 5.13). However, both $b = 5$ and $b = 10$ configurations require a similar total number of measurements to achieve convergence. Consequently, the smaller batch size $b = 5$ necessitates more algorithm iterations to reach the same reconstruction accuracy as $b = 10$, resulting in significantly higher reconstruction time.

4. Different number of bootstrapped models, Fig. 5.17

In the best-case scenario, all bootstrapped models run in parallel. Due to Python GIL, each bootstrapped model should run on a separate CPU core. Therefore, the number of bootstrapped models should not exceed the number of available CPU cores.

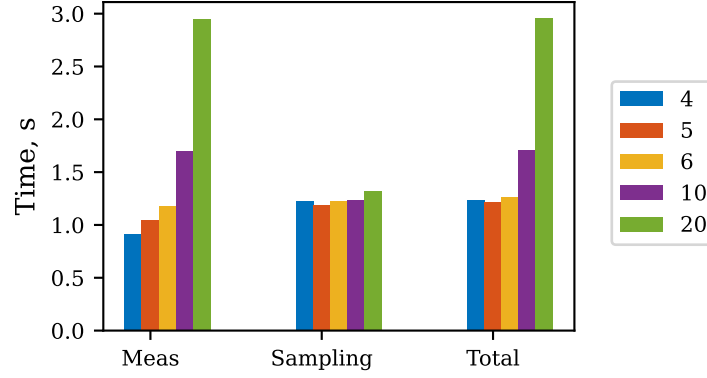


Figure 5.13: Measurement, sampling and total reconstruction time for different batch sizes.

Nevertheless, for completeness, reconstruction convergence for 10 and 20 models is presented in Fig. 5.17. Both configurations converge to compatible values with similar deviation. However, for 10 models, 3σ -curve experiences a jump in infidelity and residuals after approximately 260 measurements. For 20 models, 3σ -area remains restricted.

5. Different strategy for the initialisation of the bootstrapped models, Fig. 5.18

As mentioned in Section 5.5, the models can be initialised with the model parameters from the previous algorithm iteration (in the figure: random initialisation = False), or sampled from the prior distribution (in the figure: random initialisation = True). In both configurations, the mean infidelity goes below 0.1. However, for sampling initial guess from the prior distribution, the deviation for all the estimands is much higher. In addition, such a random initialisation requires much more time for convergence. Note that for this configuration, statistics are plotted for only 32 runs.

Note that for configurations #4 and #5, the measurements were performed in a separate cooldown with a higher base temperature (between 40 and 50 mK), resulting in different convergence behaviour for similar model parameters.

To further explore the convergence stability of the reconstruction algorithm, the reconstruction with a fixed set of hyperparameters was carried out for original, artificially displaced and rotated cat states (see Fig. 5.11 for the Wigner maps of corresponding states). The results are presented in Fig. 5.19. Across all tested states, the reconstruction algorithm demonstrated robust and reproducible convergence, despite a significant mismatch between the prior parameter distributions and initial parameter guesses used for the reconstruction and the target state parameters. This indicates that the reconstruction should remain reliable during closed-loop optimisation, even in scenarios where the optimiser applies large corrections to the cavity and significantly displaces or rotates the cat state in phase space.

For completeness, additional convergence tests were conducted for smaller cat sizes, $\alpha \in [1, 2.5]$, where the algorithm also demonstrated stable, high-fidelity reconstruction. To

maintain a clear narrative flow, the corresponding measurements are presented in the Appendix D.

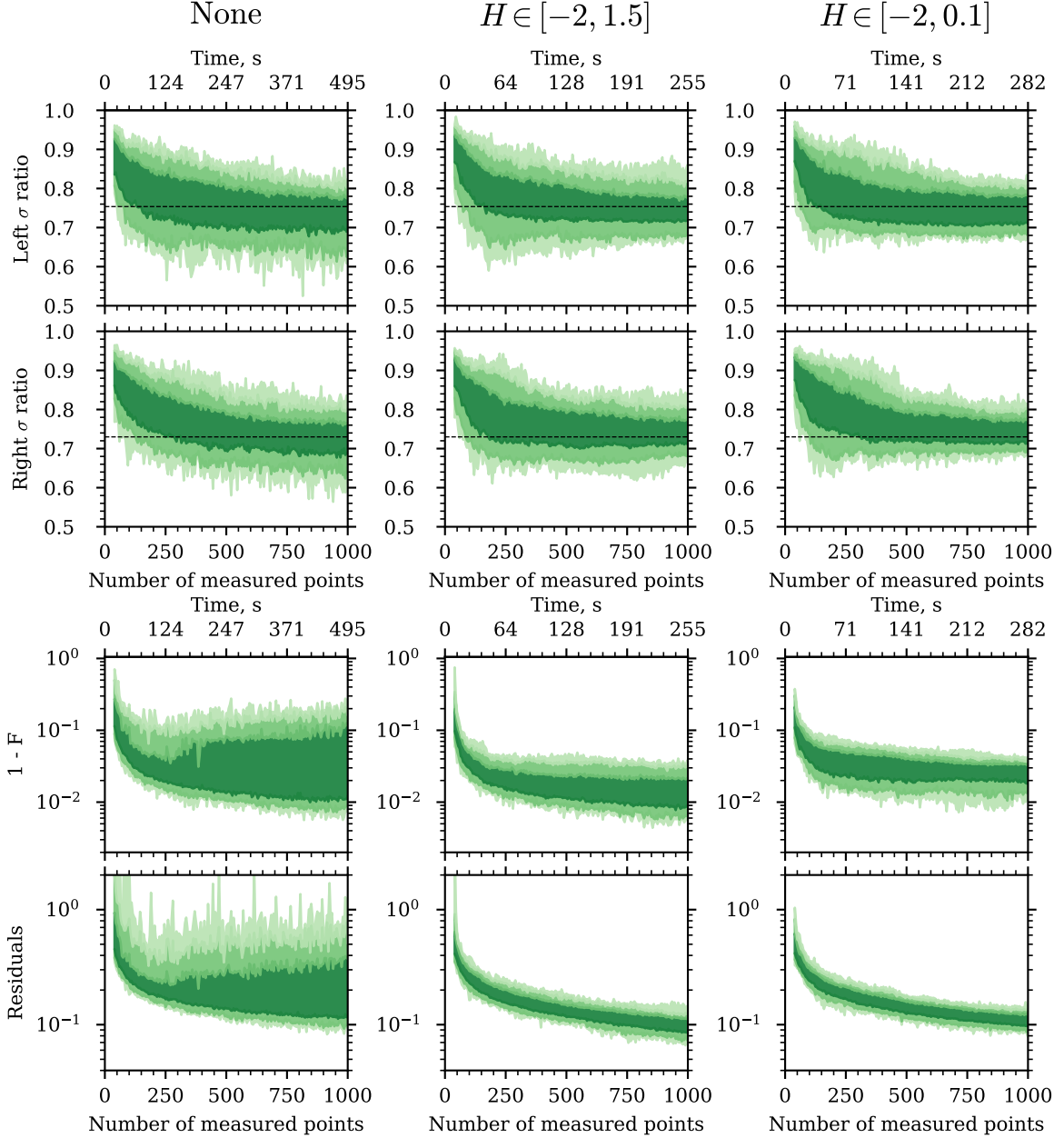


Figure 5.14: Convergence curves for different sampling strategies. The left column corresponds to sampling from the model distribution, the centre and right columns to rejection sampling with different allowed entropy ranges, $H \in [-2, 1.5]$ and $H \in [-2, 0.1]$. Each iteration samples and measures 5 points in phase space with 100 averages. For rejection sampling, initial $\beta = 60$. All the runs reconstruct the uncorrected cat state (see Fig. 5.11 A).

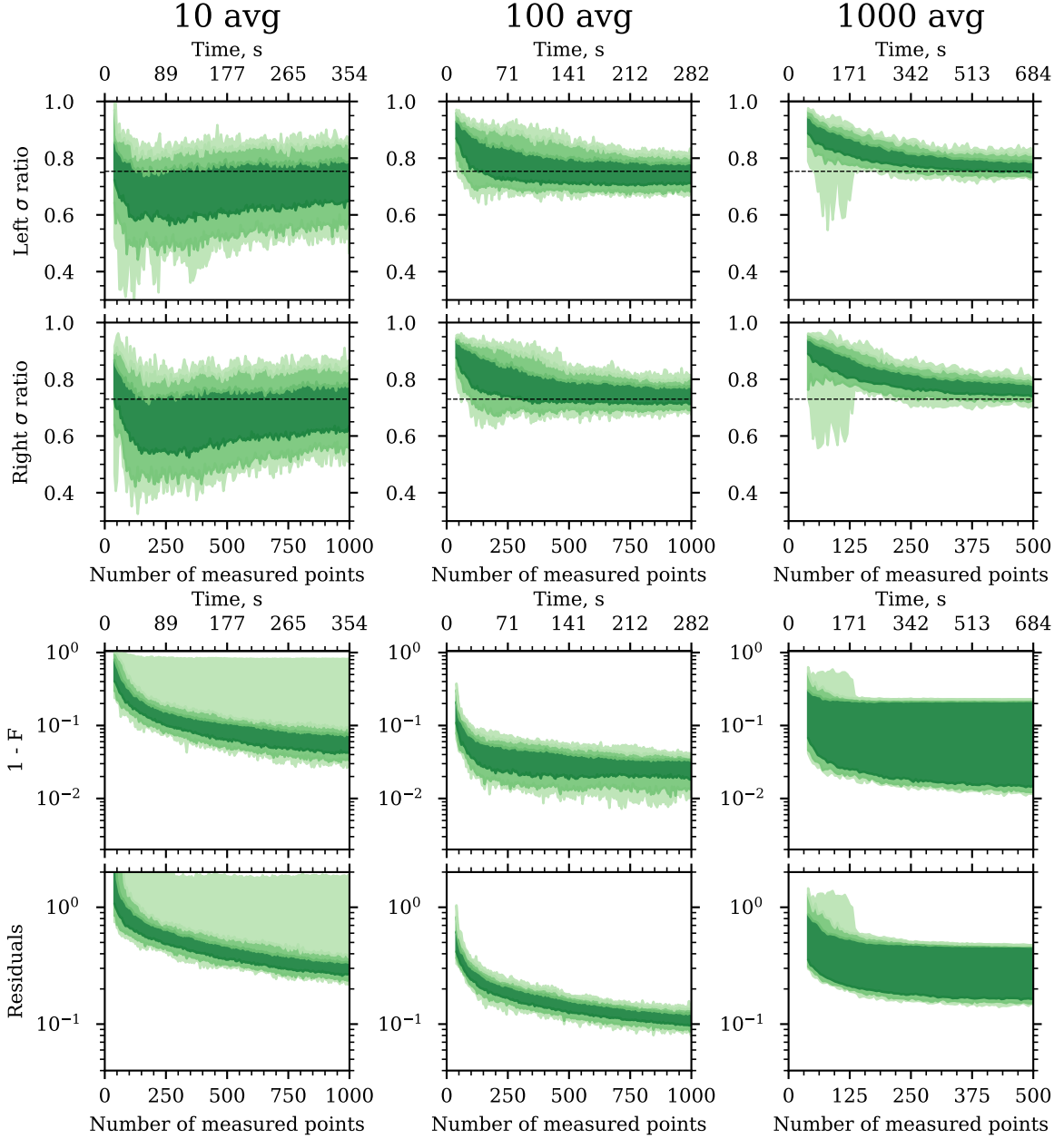


Figure 5.15: Convergence curves for different numbers of averages. The left column corresponds to 10 averages per measurement point, the centre column to 100 averages, and the right column to 1000 averages. All the runs utilise rejection sampling with $H \in [-2, 0.1]$ and initial $\beta = 60$ and reconstruct the uncorrected cat state (see Fig. 5.11 A). The batch size is 5.

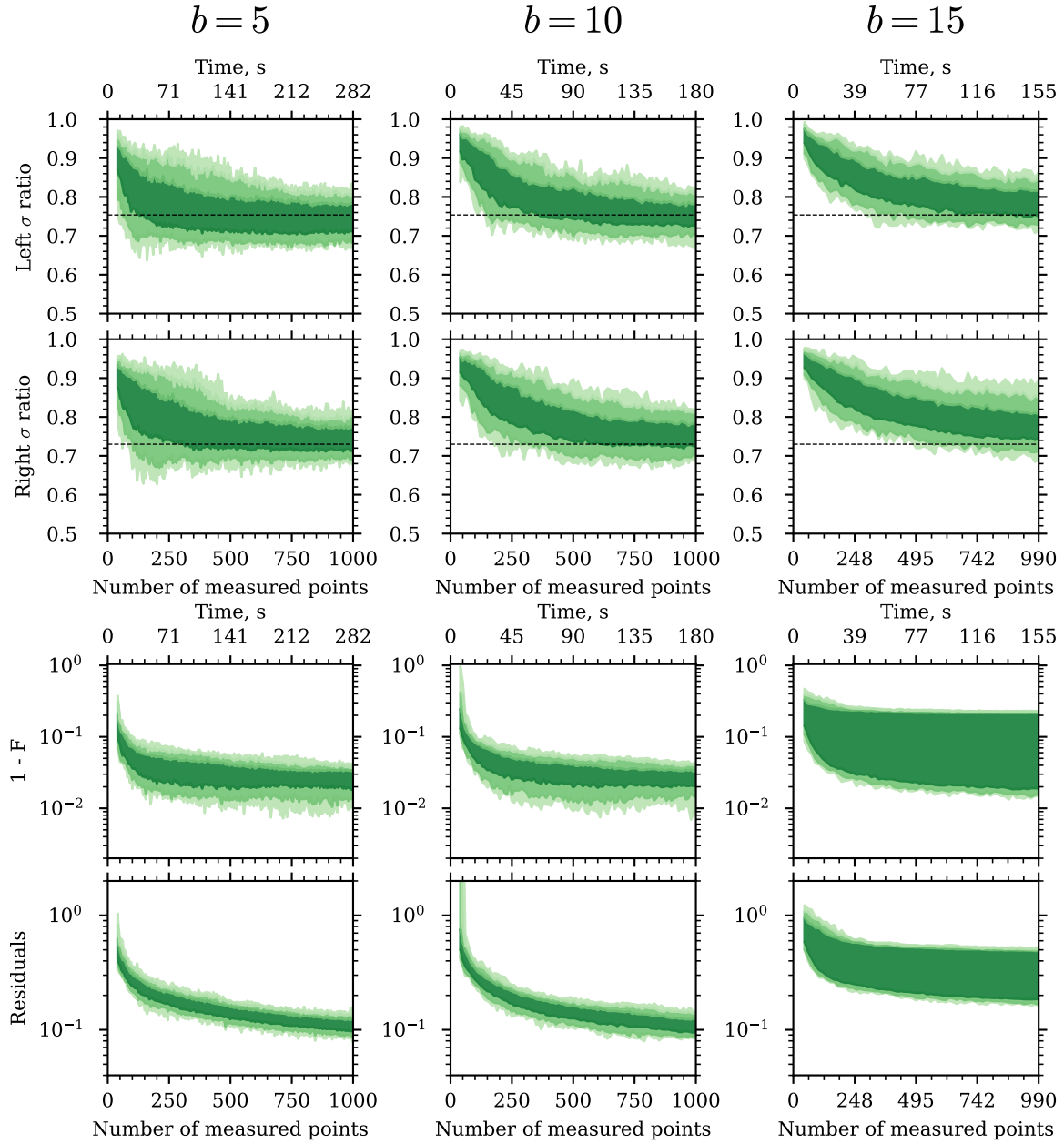


Figure 5.16: Convergence curves for different batch sizes (e.g. number of measured points in one iteration of the reconstruction algorithm). The left column corresponds to 5 samples per batch, the centre column to 10 samples per batch, and the right column to 15 samples per batch. All the runs utilise rejection sampling with $H \in [-2, 0.1]$ and initial $\beta = 60$ and reconstruct the uncorrected cat state (see Fig. 5.11 A). Each measurement point is averaged 100 times.

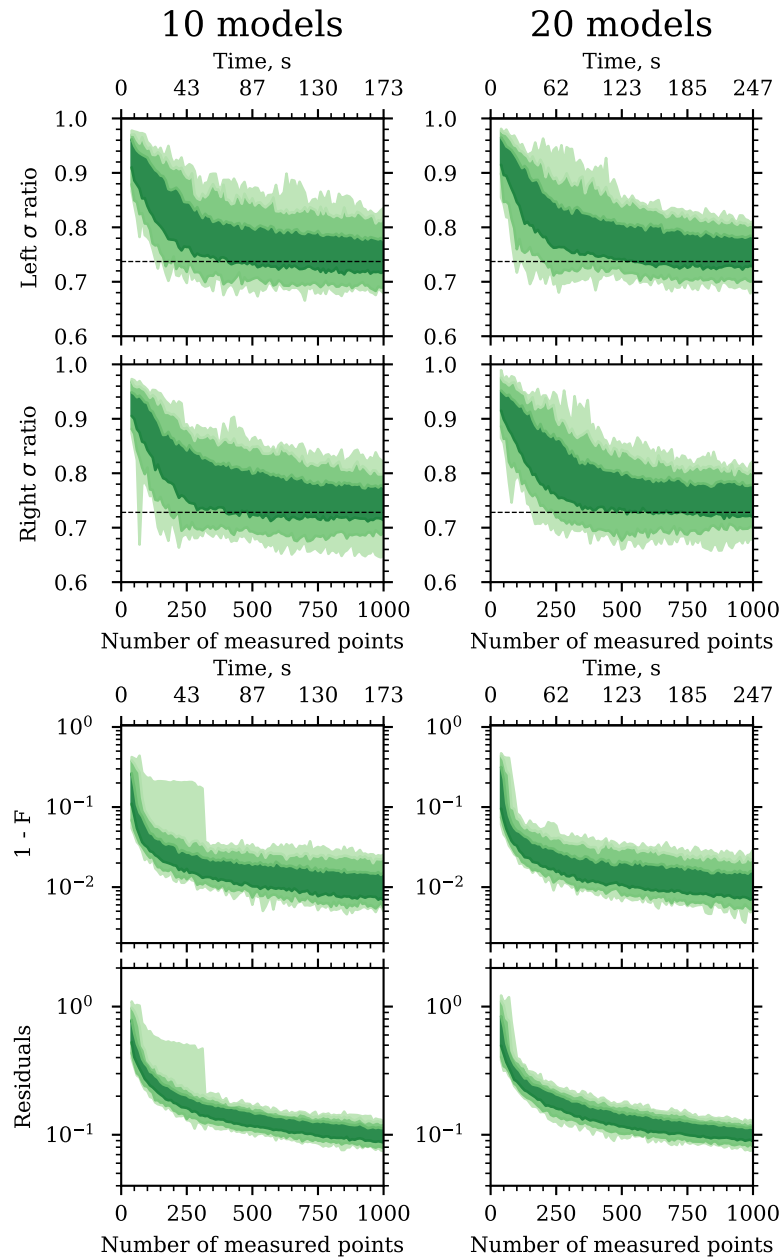


Figure 5.17: Convergence curves for 10 and 20 bootstrapped models. All the runs utilised rejection sampling with $H \in [-2, 1.5]$, initial $\beta = 60$, 100 averages and 10 measurements per batch.

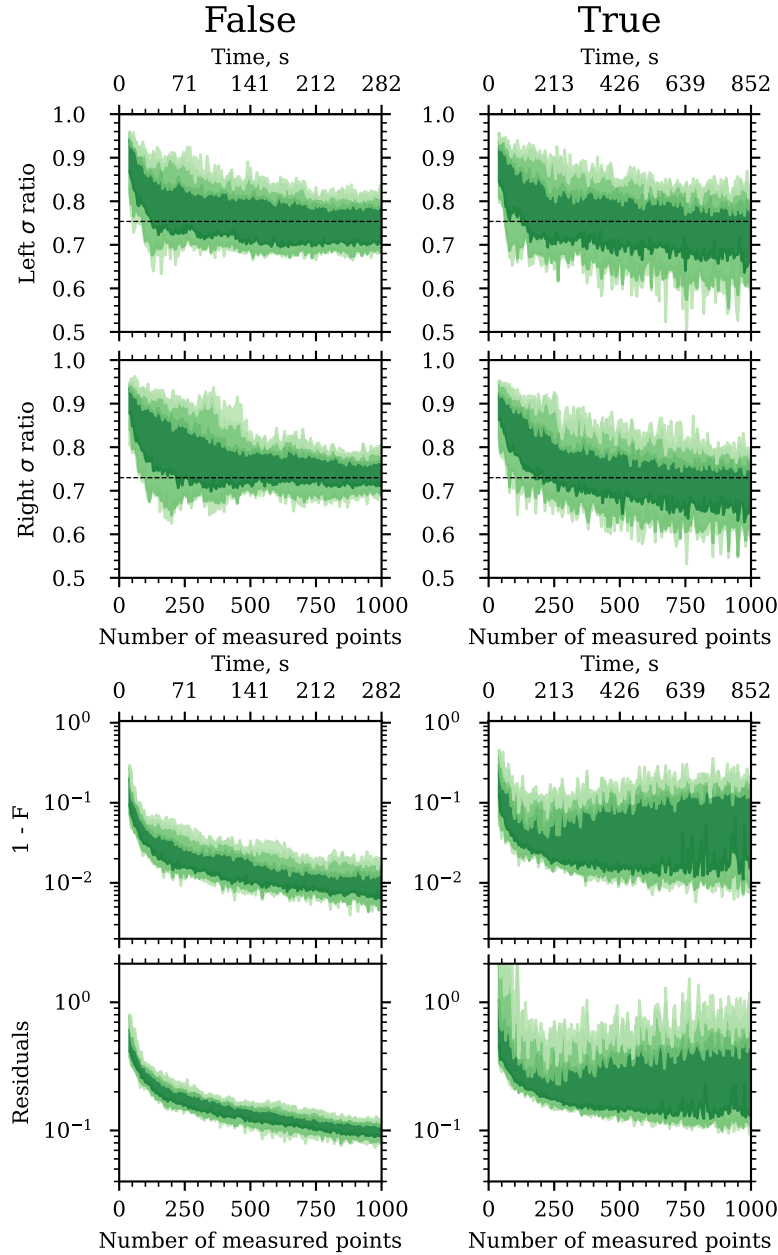


Figure 5.18: Convergence curves for different bootstrapped models' initialisation strategies. The first column corresponds to initialisation with the model parameters from the previous algorithm iteration (random initialisation = False), the second column corresponds to the initialisation with parameters sampled from the prior distribution (random initialisation = True). All the runs utilised rejection sampling with $H \in [-2, 0.1]$, 100 averages and 5 measurements per batch. Note that those plots represent 32 runs for each configuration.

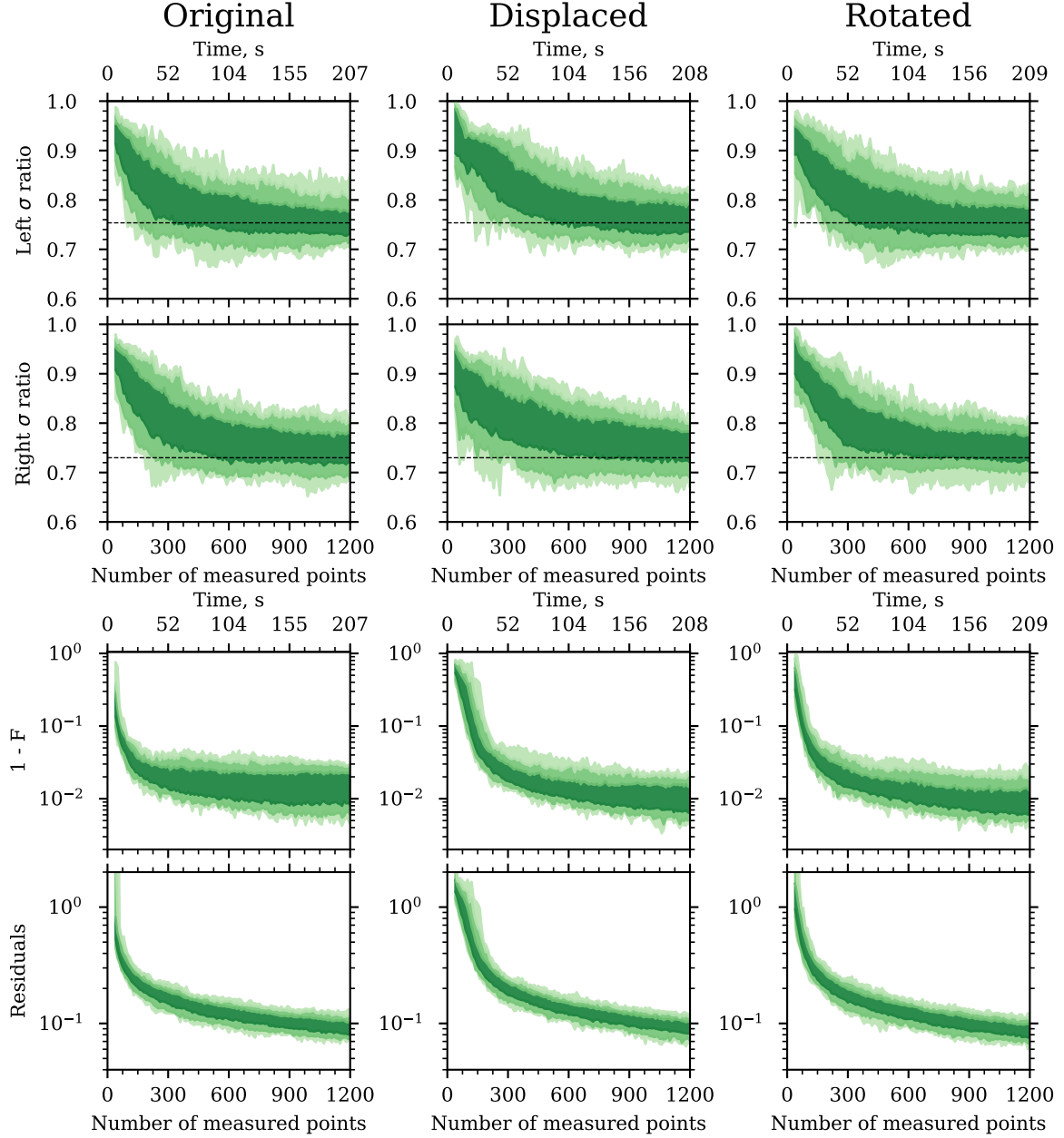


Figure 5.19: Convergence curves for original, displaced and rotated cat states. All the runs utilised rejection sampling with $H \in [-2, 1.5]$, initial $\beta = 60$, 100 averages and 10 measurements per batch.

5.8 Outlook: adjusting algorithm for other types of bosonic states and tomography techniques

Importantly, the reconstruction algorithm is independent of the type of bosonic state. For a new bosonic state, it requires two redefinitions:

1. The Wigner function in the model definition 5.3 should be replaced with the Wigner function of the desired bosonic state. Depending on the state, the number of parameters θ (Eq. 5.8) and their bounds (Eq. 5.25) will vary. More generally, the model need not be represented by the Wigner function. It can be, for example, a characteristic function, which may be useful for number-unresolved systems [65, 79]. Husimi Q may not be optimal for a suggested reconstruction algorithm, since it is not sensitive to quantum features around the phase space centre. Most probably, a stable reconstruction would require a set of M generalised Husimi Q measurements for each point in phase space. In this case, the algorithm needs to perform M MAP estimations, which requires more computational power.
2. The prior in Eq. 5.17 should be adjusted accordingly. A recommended strategy is to bias the model toward an ideal theoretical bosonic state, as implemented in the experiments described in this thesis.

The proposal distribution (Eq. 5.22) should always include the whole phase space area, where the state will most probably be observed. For the $\alpha = 3$ cat state in this thesis, the distribution covers the range $[-4, 4]$ in both I and Q directions. Depending on the bosonic state, the range can be increased or decreased. An excessively high upper bound for the Boltzmann distribution entropy should not lead to a failure of the reconstruction procedure. In the worst-case scenario of excessively high entropy, the Boltzmann distribution approaches a uniform distribution, causing the model to sample more frequently in "blank" regions of phase space, thereby slowing down convergence. Therefore, the reconstruction algorithm can be applied outside of a closed-loop setting without extensive calibration.

From personal experience, cooldowns with varying base plate temperatures (in a 20 to 55 mK range) may require hyperparameter recalibration from cooldown to cooldown. The higher the base plate temperature, the worse the fringe contrast, which affects the number of points sampled from the model distribution at the beginning of the reconstruction process (see first steps on schematics in Fig. 5.10) and the sampling probabilities for the left/right/centre Gaussian in Eq. 5.19. For stable cooldowns with an average base plate temperature of around 20 or 30 mK, no recalibration should be required.

Summary

This Chapter introduces a new reconstruction technique that samples the new measurement points in areas of the highest uncertainty. Even though all the experiments in this Chapter were conducted for the cat states by measuring the Wigner quasi-distribution, the approach can be easily transferred to other types of bosonic states and utilise different kinds of quasi-distributions and characteristic functions.

With a proper configuration of the hyperparameters, the reconstruction algorithm gives a prediction within a few minutes, with the mean infidelity between the prediction and the fit to the full Wigner map below 0.1. The key ingredients of the algorithm include:

- Bayesian inference, that ensures convergence at the early iterations of the algorithm;
- bootstrapping method, that provides an efficient sampling strategy and determines the exit conditions;
- adjustment of the target distribution temperature for rejection sampling, based on the distribution entropy, that keeps the sampling points in the areas of the most uncertainty;
- multithreading and multiprocessing of the code, which leads to a significant speed-up.

Importantly, the reconstruction is not only fast but also stable. 99.72% (3σ deviation) of 100 runs go below 0.03 threshold in infidelity, meaning that even in the case of an unsuccessful reconstruction run, the prediction still has a descent precision.

Even though reconstruction is quite fast, in most cases, it takes more than a minute. At the beginning of the Chapter, the desired upper bound on the reconstruction time was set to 58 s. Increasing this bound by a factor of 2 or 3 potentially leads to day-long closed-loop optimisation runs. It is undesirable, but not fatal, for the first closed-loop applications.

Closed-loop optimisation of cat states

The first attempts to run closed-loop optimisation in Section 4.5 employed the FoM, calculated at fixed points in phase space. The phase space sampling was tailored to a particular state and task. Together with the gradient-free Nelder-Mead algorithm and dCRAB ansatz, this approach demonstrated poor noise resilience and failed to yield significant improvement over the initial cat state.

The closed-loop optimisation improvements could be pursued in two directions: refining the optimisation algorithm or improving the FoM. Any choice of the optimisation algorithm benefits from receiving a fidelity-like, noise-resilient FoM at each iteration. Thus, the first step was to refine the FoM. To this end, the reconstruction algorithm was developed, as described in Chapter 5. The reconstruction algorithm is capable of producing a high-fidelity Wigner map of the cat state within a few minutes. This allows for the calculation of state fidelity and penalty terms based on the final reconstruction parameters.

This Chapter reveals a second attempt to close the loop with gradient-free algorithms, this time based on the reconstructed state fidelity. Such a situation, from an experimental perspective, is the closest to using the full information about the optimised state.

The results presented in this Chapter were carried out in collaboration with Dr Marco Rossignolo¹ and Dr Phila Rembold².

6.1 Reconstruction parameters and stopping criteria

At each closed-loop optimisation step, the optimiser receives an F_2 state fidelity (see Eq. 3.54) from the reconstruction algorithm

$$F_2 = \frac{F_{\text{approx}}^{(r, \text{id})}}{\max \left(F_{\text{approx}}^{(r, r)}, F_{\text{approx}}^{(\text{id}, \text{id})} \right)}, \quad (6.1)$$

¹Qruise GmbH, D-66113 Saarbrücken, Germany

²Atominstitut, TU Wien, Stadionallee 2, 1020 Vienna, Austria

which is calculated based on the fidelity between reconstructed (r) and ideal (id) states, approximated on the phase space grid with steps δI and δQ :

$$F_{\text{approx}}^{(1, 2)} = \pi \sum_i W_i^{(1)} W_i^{(2)} \delta I \delta Q. \quad (6.2)$$

The optimal reconstruction parameters are fixed for all closed-loop optimisation runs and are summarised in Table 6.1. The choice of the optimal parameters is based on the trade-off between reconstruction time and the variance of the reconstructed state fidelity, see Chapter 5 for the detailed convergence analysis.

batch size	10
number of models	10
random initialisation	False
initial β_B	60
R	0.25
angle imbalance	0.2
amplitude imbalance	0.05
H	$[-2, 1.5]$
min # of iterations	25
max # of iterations	100

Table 6.1: Reconstruction parameters used for closed-loop cat state optimisation.

Convergence curves from Chapter 5 show that in the worst-case scenario, 1000 measurement points are enough for a high-quality reconstruction. However, most of the experimental runs can achieve high-quality reconstruction with much fewer measurements. The number of measurement points required to achieve the target reconstruction fidelity varies from run to run. Fixing the number of measurement points to 1000 for all the runs introduces a large time overhead, which significantly increases the total closed-loop optimisation time. Instead, the reconstruction can be stopped earlier by checking an additional metric related to the confidence in the model prediction at the current iteration step. One example of such a metric for the reconstruction algorithm is the integrated variance σ_{int} between M bootstrapped models:

$$\sigma(I, Q) = \sqrt{\sum_{i=1}^M (W_i(I, Q) - \bar{W}(I, Q))^2 / M} \quad (6.3)$$

$$\sigma_{\text{int}} = \int \sigma(I, Q) dI dQ \approx \sum_j \sigma(I_j, Q_j) \delta I \delta Q \quad (6.4)$$

Here, $\bar{W}(I, Q) = \sum_{i=1}^M W_i(I, Q) / M$ is the mean Wigner value of the bootstrapped models' outcomes at phase space coordinate (I, Q) . Once the integrated variance σ_{int} goes below the chosen threshold, the reconstruction process is stopped. The threshold depends on the infidelity and time constraints. For a selected set of reconstruction algorithm hyperparameters, the value $\sigma_{\text{int}}^{\text{th}} = 0.08$ gives a low infidelity in a reasonable amount of time (see Fig. 6.1). The threshold, of course, could be chosen higher, but poor noise resilience of direct search algorithms pushes to play it safe.

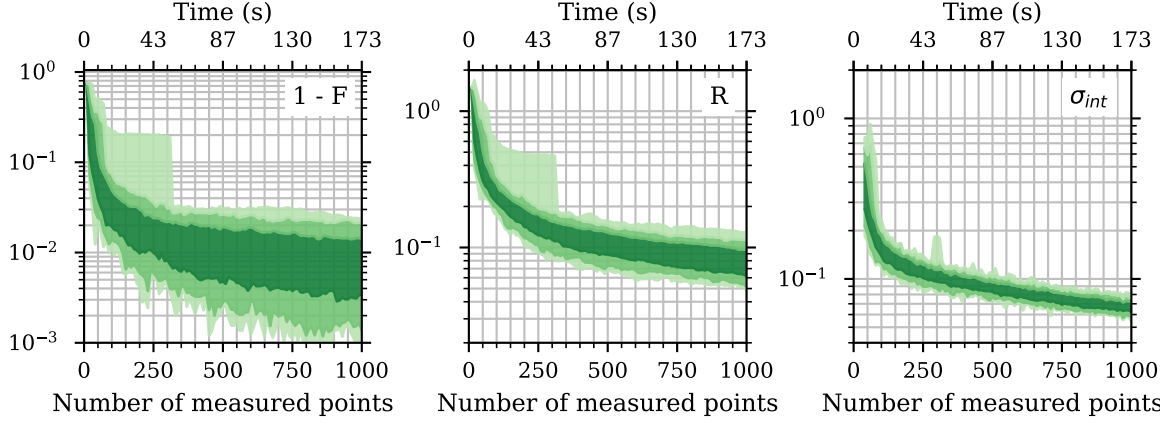


Figure 6.1: Infidelity $1 - F$, residuals R (see Eq. 5.26) and integrated variance σ_{int} (see Eq. 6.4) as a function of the number of sampling points and reconstruction time. Depending on the target infidelity value and time constraints, the integrated variance threshold for the stopping criteria is chosen. Similar to convergence plots from Chapter 5, the plots contain three shades of green. The most intense green colour corresponds to a deviation within σ from the mean, the middle intense green corresponds to a deviation within 2σ , and the most transparent green corresponds to a deviation within 3σ .

By introducing additional checks, the closed-loop optimisation time can be further decreased while improving the overall optimisation stability. During the optimisation, some correction pulses generate states of very low fidelity. In such a case, performing a precise state reconstruction provides little useful information to the optimiser while unnecessarily consuming measurement time. To avoid that, the reconstruction process is terminated if the reconstructed state fidelity after the minimum number of iterations is below the threshold $F_{th} = 0.4$. Another check is related to the minimum number of reconstruction steps. If, at the early reconstruction stages, all bootstrapped models converge to similar parameter values and produce nearly identical Wigner maps with very low integrated variance, the algorithm may prematurely terminate, even though the predicted state remains inaccurate. Such a situation can be avoided by starting to check the exit conditions only after a minimum number of measurement points (iterations), which was fixed to 250 points for the closed-loop runs in this Chapter. This number of measurement points should be enough for either improving the confidence in the current prediction or introducing a larger spread in the models' predictions.

By design, the reconstruction algorithm introduces an additional feature: since the model fits an arbitrary displaced and rotated cat state in phase space, it can account for unwanted displacements and rotations, which generally do not alter the state but affect the fidelity estimation from the Wigner map. The values of parasitic displacements $\Delta I, \Delta Q$ and rotation angle $\Delta\theta$ are straightforward to obtain from the model parameters: $\Delta I = I_C$, $\Delta Q = Q_C$, $\Delta\theta = \theta$. The corrected model parameters

$$\begin{pmatrix} I_{L,R} \\ Q_{L,R} \end{pmatrix} = \begin{pmatrix} \cos(-\Delta\theta) & -\sin(-\Delta\theta) \\ \sin(-\Delta\theta) & \cos(-\Delta\theta) \end{pmatrix} \begin{pmatrix} I_{L,R} - \Delta I \\ Q_{L,R} - \Delta Q \end{pmatrix} \quad (6.5)$$

$$\theta_{L,R,C}^{corr} = \theta_{L,R,C} - \Delta\theta \quad (6.6)$$

are used to estimate a bias-free fidelity. During the optimisation, the reconstruction algorithm is always set to return this corrected state fidelity.

6.2 Closed-loop optimisation configuration

The choice of the optimal optimisation basis is problem-specific and can lead to a potential speedup of up to a factor of 10 [95]. Due to this, an exhaustive search of hyperparameters for various gradient-free algorithms was conducted before running any closed-loop experiments. The cat state optimisation in an open-loop was tested for the (1+1) evolution algorithm [96], CMA-ES [55], and the Nelder-Mead algorithm using the dCRAB formalism. The plots for the Nelder-Mead hyperparameters' search can be found in the Appendix C. The corresponding tests for the other algorithms were conducted by the theory collaborators.

Open-loop hyperparameter search showed that the best state fidelities can be achieved by using the Nelder-Mead algorithm. The optimal configuration of the Nelder-Mead algorithm with dCRAB ansatz is summarised in Table 6.2. The detailed information about the parameters of the optimisation configuration is available at the QuOCS GitHub repository page [97].

basis	sinc with $\omega_{\max} \in [1, 300]$
basis vector number	4
basis max	2
amplitude variation	$0.2 \times \text{qubit_amp}$
superiteration number	2
maximum number of evaluations	1000

Table 6.2: Optimal configuration for closed-loop optimisation with Nelder-Mead algorithm and dCRAB ansatz.

The standard choice for the superiteration number is 3. Due to the limited amount of time available during the cooldown, the number was reduced to 2.

To ensure close reproducibility of the synthesised waveforms by the experimental hardware, the waveforms are scaled by a sigmoid function

$$f_{\text{scaling}}(x) = \frac{1}{1 + be^{-kx}} \quad (6.7)$$

with $k = -4 \log \frac{b^2+3}{2b^2}$ and $b = 10^{-4}$.

Closed-loop optimisations utilise the penalised figure of merit

$$\text{FoM} = F^2 - P = F^2 - \sum_{i \in [l, r]} \left(1 - \left(\frac{\sigma_{\min}^{(i)}}{\sigma_{\max}^{(i)}} \right)^2 \right)^{1/2} \quad (6.8)$$

that corrects for Kerr distortions of the left (l) and right (r) Gaussians. The penalty choice is based on the signal-to-noise ratio (SNR) at the end of the reconstruction process, see Table 6.3 for more details.

	Mean μ	Std σ	SNR
$1 - \frac{\sigma_1}{\sigma_2}$	0.24	0.02	10.60
$\sqrt{1 - \frac{\sigma_1}{\sigma_2}}$	0.49	0.02	20.74
$\sqrt{1 - (\frac{\sigma_1}{\sigma_2})^2}$	0.65	0.03	23.84
$1 - (\frac{\sigma_1}{\sigma_2})^2$	0.43	0.04	12.17
$(1 - (\frac{\sigma_1}{\sigma_2})^2)^2$	0.18	0.03	6.31
$(1 - \frac{\sigma_1}{\sigma_2})^2$	0.06	0.01	5.49

Table 6.3: Mean, standard deviation and signal-to-noise ratio $\text{SNR} = \sigma/\mu$ for different penalty functions after 1000 iterations of reconstruction in configuration from the Table 6.1. The penalty with the largest SNR (3rd row) is the optimal choice for constructing the FoM.

6.3 Closed-loop optimisation runs

Comprehensive analysis of the optimisation parameters in Secs. 6.1 to 6.2 leads to the following optimal configuration for the closed-loop optimisation:

- Optimisation algorithm: Nelder-Mead with dCRAB formalism
- Basis: sinc
- Figure of merit: based on the Wigner map of the reconstructed state, see Eq. 6.8.

The choice of the optimisation parameters was slightly changed compared to the first tests in Section 4.5. Since the reconstruction algorithm accounts for rotation and displacement in phase space, the optimisation of the cavity pulse during the evolution step #3 in the qcMAP protocol is eliminated. The first experimental runs demonstrated that correcting the cat state phase with qubit pulses alone may become complicated for a direct search algorithm, depending on the initial and target states. Therefore, the set of optimisation parameters is as follows:

- Optimised waveforms: I and Q quadratures of qubit pulse during the evolution step #3 of the qcMAP protocol. As an initial guess, I and Q are set to zero, which corresponds to a pure $T_{\pi/\chi}$ operation with the switch closed (see the Appendix A for the experimental setup scheme).
- Optimised pulse parameter: phase φ of the first $\pi/2$ - pulse in the qcMAP sequence.

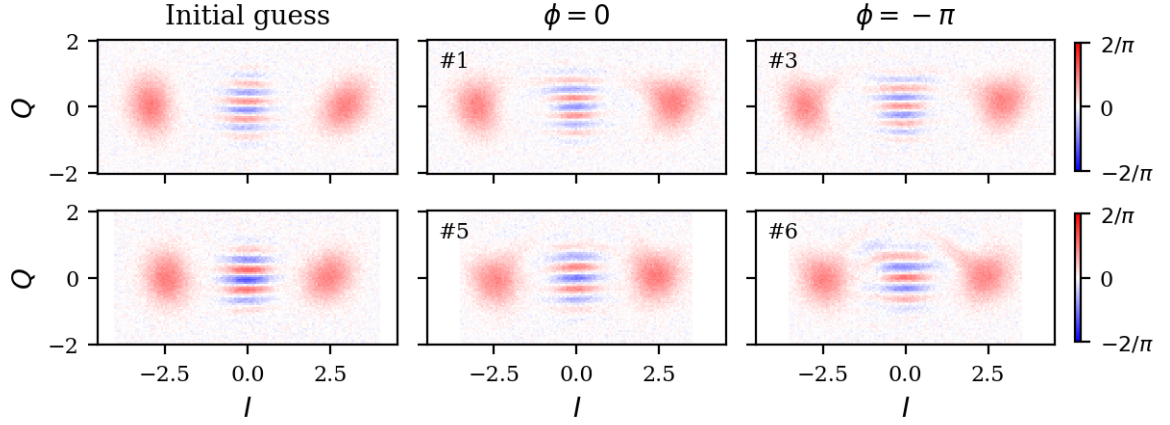


Figure 6.2: Wigner maps for $\alpha = 3$ (upper row) and $\alpha = 2.5$ (bottom row) cat states. Left column shows cat states before the optimisation, centre column - cat states after optimisation for target phase $\varphi = 0$, right column - cat states after optimisation for target phase $\varphi = -\pi$. The number of the optimisation run from Table 6.4 is specified in the corner of the corresponding Wigner map.

Each optimisation run starts from the same corrected cat state (see Eq. 6.6) with the phase of the first $\pi/2$ -pulse $\varphi = 0$ (see Fig. 6.2, first column). Due to Kerr distortions, the phase of the created state is altered, depending on the size of the cat state. Similarly to the first optimisation experiments in Section 4.5, the selectivity of the π -pulse in the qcMAP creation protocol far exceeds the minimum selectivity requirements and enhances the Kerr distortions. The goal is to strengthen the cat imperfections and see whether the optimiser can eliminate them.

Run #	α	φ	FoM	F_2 , init	F_2 , final	(e_l, e_r) , init	(e_l, e_r) , final	Time, h
1	3	0	$F^2 - P$	0.66	0.80	(0.73, 0.73)	(0.79, 0.98)	16.32
2	3	$-\pi/2$	$F^2 - P$	0.31	0.76	(0.73, 0.73)	(0.72, 0.98)	13.03
3	3	$-\pi$	$F^2 - P$	0.43	0.76	(0.73, 0.73)	(0.79, 0.94)	7.64
4	3	$-\pi$	F^2	0.43	0.80	(0.73, 0.73)	(0.73, 0.77)	7.36
5	2.5	0	$F^2 - P$	0.87	0.85	(0.81, 0.82)	(0.96, 0.96)	28.70 ³
6	2.5	$-\pi$	$F^2 - P$	0.22	0.84	(0.81, 0.82)	(0.91, 0.96)	8.35

Table 6.4: Closed-loop optimisation runs for different cat sizes α , target cat phases φ , and FoMs. The last column shows the overall time required to complete the optimisation, measured in hours. Penalty term for the FoM is defined in Eq. 6.8. All the runs for $\alpha = 3$ ($\alpha = 2.5$) start from the same rotated cat, see the first column in Fig. 6.2 for reference. Initial and final states are characterised with F_2 fidelity between the measured and target cat states (6.1). Right and left σ -ratios (ellipticities) are denoted as $e_l = \sigma_{\min}^l / \sigma_{\max}^l$, $e_r = \sigma_{\min}^r / \sigma_{\max}^r$.

The results for different runs are summarised in Table 6.4, examples of cat states before and after the optimisation are presented in Fig. 6.2, and the corresponding convergence curves and waveforms are shown in Fig. 6.3. All the runs exploiting the penalised FoM from

³This is the only run with 3 superiterations, the rest were running for 2 superiterations.

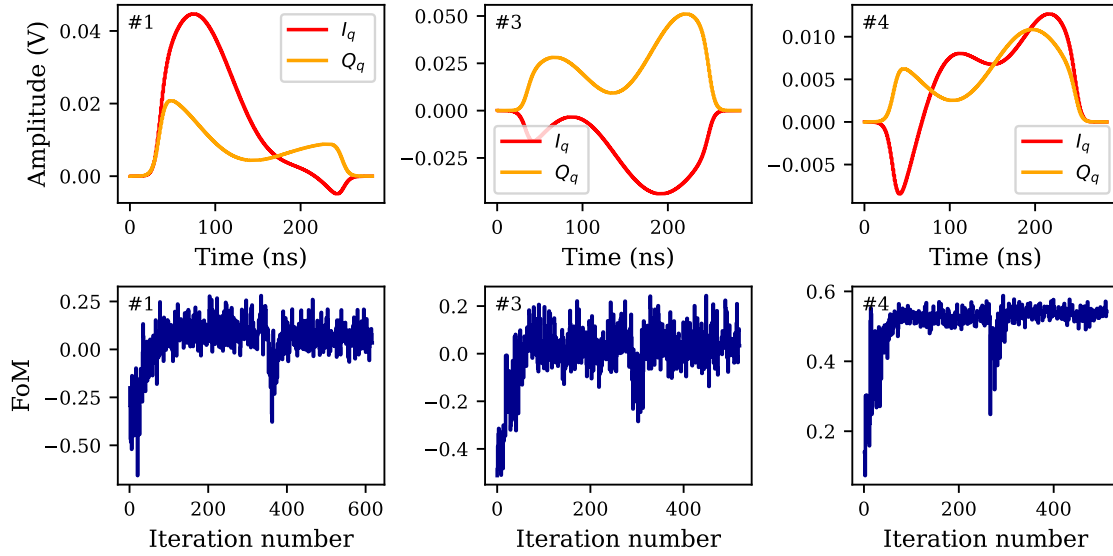


Figure 6.3: Waveforms and FoMs for different optimisation runs. The number of the optimisation run from the Table 6.4 is specified in the corner of each subfigure. Comparison between the run # 3 (FoM = $F^2 - P$) and run # 4 (FoM = F^2) clearly shows that introducing the penalty significantly increases the FoM noise.

Eq. 6.8 systematically decrease σ -ratios for the left and right Gaussians and bring the cat phase φ closer to the target, thereby significantly increasing the state fidelity. However, the correction pulses are partially flipping the qubit population, resulting in a slight decrease in the fringe amplitude. This is the reason for a decrease in fidelity for the run #5, even though σ -ratios got close to the ideal value $e = \sigma_{\min}/\sigma_{\max} = 1$. The comparison between the runs #3 and #4 with the same optimisation and target parameters, but different FoM, shows that introducing the penalty term significantly improves final σ -ratios. Again, the optimiser is not able to find correction pulses that improve the fringes' phase and contrast, while at the same time correcting for Kerr distortions: run #3 with penalty term wins in σ -ratio, but loses in overall fidelity to zero-penalty run #4. Regardless of the FoM and cat state parameters, the optimiser consistently struggles to correct the distortions of the left Gaussian.

For completeness, closed-loop fidelities are compared to the fidelity of the cat state formed with open-loop correction pulses. Open-loop optimisations were performed using the L-BFGS algorithm with GRAPE (see Section 2.1) by the theory collaborators. The measured open-loop $\alpha = 3$ cat state (see Fig. 6.4) has a fidelity of 0.72 and ellipticities $e_l = 0.96$, $e_r = 0.65$. Compared to closed-loop runs, the open-loop waveforms are capable of correcting the Kerr distortions for the left Gaussian. The fidelity is, however, slightly lower than closed-loop results, and is significantly smaller than what simulations predict. The discrepancy may arise from unaccounted higher-order nonlinear terms or hardware limitations.

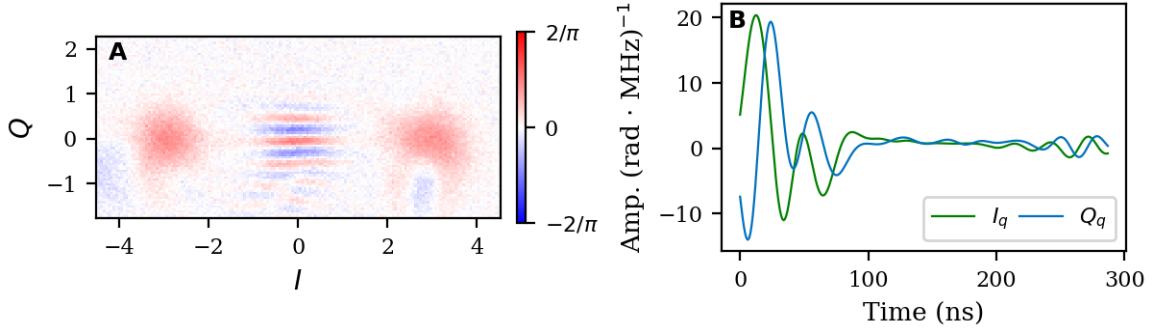


Figure 6.4: Open-loop pulses in the experiment. **A** Wigner map for $\alpha = 3$ cat state with correction pulses from open-loop optimisation. **B** Open-loop qubit correction waveforms.

Summary

A comparison between the results of the first experiments from Section 4.5 and the results from this Chapter clearly shows that access to the state fidelity is necessary for the Nelder-Mead convergence in a closed-loop. By having access to the I and Q quadratures of the qubit pulse during the waiting time in step #3 of the qcMAP protocol and the phase of the first $\pi/2$ -pulse, the optimiser is able to:

- prepare states of different phases (parities);
- significantly improve σ -ratios of the left and right Gaussians, distorted by Kerr.

However, in this setting, the direct search algorithms cannot find complex solutions. As a result, the optimised correction pulses tend to change $|e\rangle$ and $|g\rangle$ state populations, which results in a Kerr correction, but at the same time, an unwanted decrease in fringe amplitude. Even though the results remain suboptimal, they represent the first experimental demonstration of successful closed-loop optimisation of a bosonic state.

In Chapter 4, it was highlighted that there are two possible directions of improvement: tackling the FoM or the optimisation algorithm. The noise-resilient figure of merit, obtained from state reconstruction, significantly improved closed-loop optimisation performance. The next step towards high-fidelity bosonic states is to use a more complex optimisation algorithm. The most promising candidate for that belongs to reinforcement learning.

Conclusion

Quantum optimal and adaptive control has become a vital tool for many research groups pushing forward the field of quantum computing. All recent experiments on quantum error correction use some form of optimal control. For bosonic qubits, any kind of optimisation is challenging due to extensive tomography in continuous-variable phase space. Because of this, experiments on bosonic qubits can use closed-loop only for specific states and figures of merit.

This thesis made a step towards closed-loop optimisation of an arbitrary bosonic state. Although the experiments were conducted for distorted cat states of varying sizes, the results in the thesis are not limited to this family of bosonic states.

As a first step, it was checked whether the full state tomography is necessary for a successful closed-loop cat state optimisation. For discrete variable qubits, there are several examples where the Nelder-Mead algorithm shows significant improvement without access to an actual state or gate fidelity (see Section 2.3). For an arbitrary bosonic state, a fidelity-like figure of merit can be obtained from the measurements at the phase space points chosen in advance. Chapter 4 showed that such a fixed sampling of measurement points together with a direct-search Nelder-Mead algorithm significantly improves the state fidelity in an open-loop, but fails in the closed-loop due to poor noise resilience. Making gradient-free algorithms work in a closed-loop requires a noise-resilient fidelity estimate with high precision.

The tomography technique that is compatible with a closed-loop optimisation needs to satisfy several requirements:

1. the error of the fidelity estimation should be low;
2. the fidelity estimate should be reproducible;
3. the total reconstruction time should not exceed a few minutes.

Current bosonic state tomography techniques do not satisfy all the criteria. Chapter 5 introduced an efficient reconstruction technique that fulfils all the requirements listed above. Importantly, the reconstruction approach is independent of the type of bosonic state and the type of quasi-distribution or characteristic function that describes it.

The reconstruction technique proved itself effective both inside and outside the closed-loop. In this thesis, most of the closed-loop runs from the Chapter 6 did not exceed the targeted overnight time constraint. For future applications that require even faster latency times, the algorithm can be potentially sped up even more. Most of the reconstruction time on the CPU is spent on rejection sampling. As mentioned in Chapter 5, this is not the most effective way of drawing samples from a complex probability distribution. One possible solution is to replace rejection sampling with Markov Chain Monte Carlo [43]. Speeding up the maximum a posteriori estimation part would require extensive work in rewriting the Python code in C++. Another possible direction for improvement is related to a hardware configuration update in a closed-loop. Instead of uploading the whole new waveform at each iteration, one can upload the coefficients for basis wavefunctions that were loaded at the beginning of the experiment, as was done in [3].

In the Chapter 6, closed-loop optimisation of an arbitrary bosonic state by example of $\alpha = 3$ and $\alpha = 2.5$ cat states was demonstrated. The Nelder-Mead algorithm, used in a closed-loop, showed an improvement in the target state phase and Kerr distortion correction, sometimes at the cost of reduced fidelity. This demonstrates that gradient-free algorithms struggle to find complex solutions for open systems in experiments, although they are still able to achieve fidelity improvements. The first closed-loop results look promising and encourage future experiments based on more advanced optimisation algorithms with access to the fidelity of the reconstructed states.

To enhance the performance of future closed-loop optimisations, it is highly beneficial to apply slight modifications to the quantum system design and experimental setup. Reducing the qubit-cavity dispersive shift would suppress phase space-dependent shifts in the measured positive and negative parity signal levels, which arise from too selective $\pi/2$ pulses in the Wigner measurement sequence. Together with a quantum amplifier, it enables single-shot parity or Wigner measurements, thereby improving noise resilience and drastically reducing measurement time. These improvements would substantially accelerate the reconstruction process while enhancing its convergence stability.

Either way, the most promising approach for future closed-loop optimisation is to employ reinforcement learning (RL) algorithms. As discussed in Section 2.3, RL has already been successfully applied in a closed-loop in the break-even experiment [11], and has demonstrated stable convergence in noise-free open-loop scenarios for bosonic state generation tasks when the state fidelity is accessible [61]. This motivates the RL application to correct bosonic state imperfections, as explored in this thesis. A potential implementation would consist of two stages:

1. training the RL agent in simulation, using experimentally characterised Hamiltonian parameters;
2. fine-tuning the RL agent directly in the experiment.

Once trained, the RL agent should be able to compensate for system drifts significantly faster than during the initial training phase, offering a major advantage for rapid recalibration in future quantum computing and quantum simulation experiments.

As experiments become increasingly application-oriented, there is a growing need for more precise quantum optimal control methods. These new control tools are being developed alongside advances in machine learning, deep learning, and reinforcement learning, opening a broad spectrum of possibilities and enabling the automation and simplification of calibration procedures. It is an exciting time for both physicists and computer scientists; interdisciplinary collaborations have the potential to push the boundaries of knowledge in both fields, offer insights into longstanding challenges, and give rise to even more complex ones.

Bibliography

- [1] B. Vlastakis, G. Kirchmair, Z. Leghtas, S. E. Nigg, L. Frunzio, S. M. Girvin, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, “Deterministically encoding quantum information using 100-photon schrödinger cat states”, *Science* **342**, 607 (2013).
- [2] G. Kirchmair, B. Vlastakis, Z. Leghtas, S. E. Nigg, H. Paik, E. Ginossar, M. Mirrahimi, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, “Observation of quantum state collapse and revival due to the single-photon kerr effect”, *Nature* **495**, 205–209 (2013).
- [3] H. Hutin, P. Bilous, C. Ye, S. Abdollahi, L. Cros, T. Dvir, T. Shah, Y. Cohen, A. Binenfain, F. Marquardt, and B. Huard, “Preparing schrödinger cat states in a microwave cavity using a neural network”, *PRX Quantum* **6**, 10.1103/prxquantum.6.010321 (2025).
- [4] T. Krisnanda, C. Y. Fontaine, A. Copetudo, P. Song, K. X. Lee, N.-N. Huang, F. Valadares, T. C. Liew, and Y. Y. Gao, “Demonstrating efficient and robust bosonic state reconstruction via optimized excitation counting”, *PRX Quantum* **6**, 10.1103/prxquantum.6.010303 (2025).
- [5] V. G. Matsos, C. H. Valahu, T. Navickas, A. D. Rao, M. J. Millican, X. C. Kolesnikow, M. J. Biercuk, and T. R. Tan, “Robust and deterministic preparation of bosonic logical states in a trapped ion”, *Phys. Rev. Lett.* **133**, 050602 (2024).
- [6] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, “Quantum metrology with nonclassical states of atomic ensembles”, *Rev. Mod. Phys.* **90**, 035005 (2018).
- [7] M. Chen, N. C. Menicucci, and O. Pfister, “Experimental realization of multipartite entanglement of 60 modes of a quantum optical frequency comb”, *Phys. Rev. Lett.* **112**, 120505 (2014).
- [8] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, “Circuit quantum electrodynamics”, *Rev. Mod. Phys.* **93**, 025005 (2021).
- [9] Google Quantum AI team, “Quantum error correction below the surface code threshold”, *Nature* **638**, 920–926 (2024).
- [10] Z. Ni, S. Li, X. Deng, Y. Cai, L. Zhang, W. Wang, Z.-B. Yang, H. Yu, F. Yan, S. Liu, C.-L. Zou, L. Sun, S.-B. Zheng, Y. Xu, and D. Yu, “Beating the break-even point with a discrete-variable-encoded logical qubit”, *Nature* **616**, 56–60 (2023).

- [11] V. V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. L. Brock, A. Z. Ding, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, “Real-time quantum error correction beyond break-even.”, *Nature* **616**, 50 (2023).
- [12] B. L. Brock, S. Singh, A. Eickbusch, V. V. Sivak, A. Z. Ding, L. Frunzio, S. M. Girvin, and M. H. Devoret, “Quantum error correction of qudits beyond break-even”, *Nature* **641**, 612–618 (2025).
- [13] Google Quantum AI team, “Suppressing quantum errors by scaling a surface code logical qubit”, *Nature* **614**, 676–681 (2023).
- [14] K. Chou, “Teleported operations between logical qubits in circuit quantum electrodynamics”, Ph.D. thesis, Yale University (2018).
- [15] B. M. Vlastakis, “Controlling coherent state superpositions with superconducting circuits”, Ph.D. thesis, Yale University, Connecticut (2015).
- [16] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, “A quantum engineer’s guide to superconducting qubits”, *Applied Physics Reviews* **6**, 021318 (2019).
- [17] P. Rembold, “Quantum optimal control of spin systems and trapped atoms”, Ph.D. thesis, Universität zu Köln (2022).
- [18] P. Reinhold, “Controlling error-correctable bosonic qubits”, Ph.D. thesis, Yale University (2019).
- [19] Y. Nakamura, Y. A. Pashkin, and J. S. Tsai, “Coherent control of macroscopic quantum states in a single-cooper-pair box”, *Nature* **398**, 10.1038/19718 (1999).
- [20] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, “Superconducting qubits: Current state of play”, *Annual Review of Condensed Matter Physics* **11**, 369 (2020).
- [21] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, “Surface codes: Towards practical large-scale quantum computation”, *Phys. Rev. A* **86**, 032324 (2012).
- [22] A. Joshi, K. Noh, and Y. Y. Gao, “Quantum information processing with bosonic qubits in circuit qed”, *Quantum Science and Technology* **6**, 033001 (2021).
- [23] C. P. Koch, U. Boscain, T. Calarco, G. Dirr, S. Filipp, S. J. Glaser, R. Kosloff, S. Montangero, T. Schulte-Herbrüggen, D. Sugny, and F. K. Wilhelm, “Quantum optimal control in quantum technologies. strategic report on current status, visions and goals for research in europe”, *PJ Quantum Technol.* **9**, 10.1140/epjqt/s40507-022-00138-x (2022).

- [24] T. Mahesh, P. Batra, and M. Ram, “Quantum optimal control: Practical aspects and diverse methods”, *J Indian Inst Sci* **103**, [10.1007/s41745-022-00311-2](#) (2023).
- [25] Z. Zong, Z. Sun, Z. Dong, C. Run, L. Xiang, Z. Zhan, Q. Wang, Y. Fei, Y. Wu, W. Jin, C. Xiao, Z. Jia, P. Duan, J. Wu, Y. Yin, and G. Guo, “Optimization of a controlled- z gate with data-driven gradient-ascent pulse engineering in a superconducting-qubit system”, *Phys. Rev. Appl.* **15**, 064005 (2021).
- [26] J. Johansson, P. Nation, and F. Nori, “Qutip: An open-source python framework for the dynamics of open quantum systems”, *Computer Physics Communications* **183**, 1760 (2012).
- [27] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, “Simple pulses for elimination of leakage in weakly nonlinear qubits”, *Phys. Rev. Lett.* **103**, 110501 (2009).
- [28] J. M. Chow, L. DiCarlo, J. M. Gambetta, F. Motzoi, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, “Optimized driving of superconducting artificial atoms for improved single-qubit gates”, *Phys. Rev. A* **82**, 040305 (2010).
- [29] F. Setiawan, P. Groszkowski, H. Ribeiro, and A. A. Clerk, “Analytic design of accelerated adiabatic gates in realistic qubits: General theory and applications to superconducting circuits”, *PRX Quantum* **2**, 030306 (2021).
- [30] J. García-Ripoll, A. Ruiz-Chamorro, and E. Torrontegui, “Quantum control of frequency-tunable transmon superconducting qubits”, *Phys. Rev. Appl.* **14**, 044035 (2020).
- [31] R. W. Heeres, B. Vlastakis, E. Holland, S. Krastanov, V. V. Albert, L. Frunzio, L. Jiang, and R. J. Schoelkopf, “Cavity state manipulation using photon-number selective phase gates”, *Phys. Rev. Lett.* **115**, 137002 (2015).
- [32] E. Hyyppä, A. Vepsäläinen, M. Papič, C. F. Chan, S. Inel, A. Landra, W. Liu, J. Luus, F. Marxer, C. Ockeloen-Korppi, S. Orbell, B. Tarasinski, and J. Heinsoo, “Reducing leakage of single-qubit gates for superconducting quantum processors using analytical control pulse envelopes” (2024), [arXiv:2402.17757 \[quant-ph\]](#) .
- [33] B. Chiaro and Y. Zhang, “Active leakage cancellation in single qubit gates” (2025), [arXiv:2503.14731 \[quant-ph\]](#) .
- [34] R. W. Heeres, P. Reinhold, N. Ofek, L. Frunzio, L. Jiang, M. H. Devoret, and R. J. Schoelkopf, “Implementing a universal gate set on a logical qubit encoded in an oscillator”, *Nat Commun* **8**, [10.1038/s41467-017-00045-1](#) (2017).
- [35] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, “Optimal control of coupled spin dynamics: design of nmr pulse sequences by gradient ascent algorithms”, *Journal of Magnetic Resonance* **172**, 296 (2005).

- [36] P. Rembold, N. Oshnik, M. M. Müller, S. Montangero, T. Calarco, and E. Neu, “Introduction to quantum optimal control for quantum sensing with nitrogen-vacancy centers in diamond”, *AVS Quantum Sci* **2**, doi.org/10.1116/5.0006785 (2020).
- [37] A. I. Konnov and V. F. Krotov, “On global methods for the successive improvement of control processes”, *Avtomat. i Telemekh.* **60**, 77 (1999).
- [38] V. F. Krotov, “Global methods in optimal control theory”, in *Advances in Non-linear Dynamics and Control: A Report from Russia*, edited by A. B. Kurzhanski (Birkhäuser Boston, Boston, MA, 1993) pp. 74–121.
- [39] N. Leung, M. Abdelhafez, J. Koch, and D. Schuster, “Speedup for quantum optimal control from automatic differentiation based on graphics processing units”, *Physical Review A* **95**, [10.1103/physreva.95.042318](https://doi.org/10.1103/physreva.95.042318) (2017).
- [40] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, “Charge-insensitive qubit design derived from the cooper pair box”, *Physical Review A* **76**, [10.1103/physreva.76.042319](https://doi.org/10.1103/physreva.76.042319) (2007).
- [41] V. E. Manucharyan, J. Koch, L. I. Glazman, and M. H. Devoret, “Fluxonium: Single cooper-pair circuit free of charge offsets”, *Science* **326**, 113 (2009), <https://www.science.org/doi/pdf/10.1126/science.1175552> .
- [42] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, “Automatic differentiation in machine learning: a survey” (2018), [arXiv:1502.05767](https://arxiv.org/abs/1502.05767) [cs.SC] .
- [43] I. Goodfellow, Y. Bengio, and A. Courville, “Deep learning” (MIT Press, 2016) <http://www.deeplearningbook.org>.
- [44] J. Nocedal and S. Wright, “Numerical optimization”, 2nd ed., Springer series in operations research and financial engineering (Springer, New York, NY, 2006).
- [45] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization” (2017), [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) [cs.LG] .
- [46] J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, I.-C. Hoi, E. Jeffrey, A. Megrant, J. Mutus, C. Neill, P. J. J. O’Malley, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Cleland, and J. M. Martinis, “Optimal quantum control using randomized benchmarking”, *Phys. Rev. Lett.* **112**, 240504 (2014).
- [47] M. A. Rol, C. C. Bultink, T. E. O’Brien, S. R. de Jong, L. S. Theis, X. Fu, F. Luthi, R. F. L. Vermeulen, J. C. de Sterke, A. Bruno, D. Deurloo, R. N. Schouten, F. K. Wilhelm, and L. DiCarlo, “Restless tuneup of high-fidelity qubit gates”, *Phys. Rev. Appl.* **7**, 041001 (2017).

- [48] M. Werninghaus, D. J. Egger, F. Roy, S. Machnes, F. K. Wilhelm, and S. Filipp, “Leakage reduction in fast superconducting qubit gates via optimal control”, npj Quantum Information **7**, [10.1038/s41534-020-00346-2](https://doi.org/10.1038/s41534-020-00346-2) (2021).
- [49] F. Frank, T. Unden, J. Zoller, R. S. Said, T. Calarco, S. Montangero, B. Naydenov, and F. Jelezko, “Autonomous calibration of single spin qubit operations”, npj Quantum Information **3**, [10.1038/s41534-017-0049-8](https://doi.org/10.1038/s41534-017-0049-8) (2017).
- [50] J. A. Nelder and R. Mead, “A Simplex Method for Function Minimization”, *The Computer Journal* **7**, 308 (1965), <https://academic.oup.com/comjnl/article-pdf/7/4/308/1013182/7-4-308.pdf> .
- [51] A. R. Conn, K. Scheinberg, and L. N. Vicente, “*Introduction to derivative-free optimization*” (Society for Industrial and Applied Mathematics, 2009) <https://epubs.siam.org/doi/pdf/10.1137/1.9780898718768> .
- [52] M. Emmerich, O. M. Shir, and H. Wang, “Evolution strategies”, in *Handbook of Heuristics*, edited by R. Martí, P. Panos, and M. G. C. Resende (Springer International Publishing, Cham, 2018) pp. 1–31.
- [53] R. Strock, “Covariance matrix adaptation evolution strategy (cma-es) implemented with tensorflow”, <https://pypi.org/project/cma-es/1.5.0/> (2021).
- [54] N. Hansen and A. Ostermeier, “Completely derandomized self-adaptation in evolution strategies”, *Evolutionary Computation* **9**, 159 (2001).
- [55] N. Hansen, “The cma evolution strategy: a tutorial” (2023), [arXiv:1604.00772](https://arxiv.org/abs/1604.00772) .
- [56] T. Caneva, T. Calarco, and S. Montangero, “Chopped random-basis quantum optimization”, *Phys. Rev. A* **84**, 022326 (2011).
- [57] N. Rach, M. M. Müller, T. Calarco, and S. Montangero, “Dressing the chopped-random-basis optimization: A bandwidth-limited access to the trap-free landscape”, *Phys. Rev. A* **92**, 062343 (2015).
- [58] N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. M. Girvin, L. Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, “Extending the lifetime of a quantum bit with error correction in superconducting circuits”, *Nature* **536**, [10.1038/nature18949](https://doi.org/10.1038/nature18949) (2016).
- [59] L. Ella, L. Leandro, O. Wertheim, Y. Romach, L. Schlipf, R. Szmuk, Y. Knol, N. Ofek, I. Sivan, and Y. Cohen, “Quantum-classical processing and benchmarking at the pulse-level” (2023), [arXiv:2303.03816 \[quant-ph\]](https://arxiv.org/abs/2303.03816) .
- [60] Y. Liu, “Quantum feedback control of multiple superconducting qubits”, Ph.D. thesis, Yale University (2016).

- [61] V. V. Sivak, A. Eickbusch, H. Liu, B. Royer, I. Tsioutsios, and M. H. Devoret, “Model-free quantum control with reinforcement learning”, *Phys. Rev. X* **12**, 011059 (2022).
- [62] K. Reuer, J. Landgraf, T. Fösel, J. O’Sullivan, L. Beltrán, A. Akin, G. J. Norris, A. Remm, M. Kerschbaum, J.-C. Besse, F. Marquardt, A. Wallraff, and C. Eichler, “Realizing a deep reinforcement learning agent for real-time quantum feedback.”, *Nat Commun* **14**, 10.1038/s41467-023-42901-3 (2023).
- [63] Y. Baum, M. Amico, S. Howell, M. Hush, M. Liuzzi, P. Mundada, T. Merkh, A. R. Carvalho, and M. J. Biercuk, “Experimental deep reinforcement learning for error-robust gate-set design on a superconducting quantum computer”, *PRX Quantum* **2**, 040324 (2021).
- [64] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf, L. Frunzio, M. Mirrahimi, and M. H. Devoret, “Quantum error correction of a qubit encoded in grid states of an oscillator.”, *Nature* **584**, 368 (2020).
- [65] A. Eickbusch, V. Sivak, A. Z. Ding, S. S. Elder, S. R. Jha, J. Venkatraman, B. Royer, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, “Fast universal control of an oscillator with weak dispersive coupling to a qubit.”, *Nat. Phys.* **18**, 1464 (2022).
- [66] S. Rosenblum, Y. Y. Gao, P. Reinhold, C. Wang, C. J. Axline, L. Frunzio, S. M. Girvin, L. Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, “A cnot gate between multiphoton qubits encoded in two cavities.”, *Nat Commun* **6**, 10.1038/s41467-018-03059-5 (2018).
- [67] K. S. Chou, J. Z. Blumoff, C. S. Wang, P. C. Reinhold, C. J. Axline, Y. Y. Gao, L. Frunzio, M. H. Devoret, L. Jiang, and R. J. Schoelkopf, “Deterministic teleportation of a quantum gate between two logical qubits”, *Nature* **561**, 368–373 (2018).
- [68] S. Touzard, A. Grimm, Z. Leghtas, S. O. Mundhada, P. Reinhold, C. Axline, M. Reagor, K. Chou, J. Blumoff, K. M. Sliwa, S. Shankar, L. Frunzio, R. J. Schoelkopf, M. Mirrahimi, and M. H. Devoret, “Coherent oscillations inside a quantum manifold stabilized by dissipation”, *Physical Review X* **8**, 10.1103/physrevx.8.021005 (2018).
- [69] L. Hu, Y. Ma, W. Cai, X. Mu, Y. Xu, W. Wang, Y. Wu, H. Wang, Y. P. Song, C.-L. Zou, S. M. Girvin, L.-M. Duan, and L. Sun, “Quantum error correction and universal gate set operation on a binomial bosonic logical qubit.”, *Nat. Phys.* **15**, 503 (2019).
- [70] M. Kudra, M. Kervinen, I. Strandberg, S. Ahmed, M. Scigliuzzo, A. Osman, D. P. Lozano, M. O. Tholén, R. Borgani, D. B. Haviland, G. Ferrini, J. Bylander, A. F. Kockum, F. Quijandría, P. Delsing, and S. Gasparinetti, “Robust preparation of wigner-negative states with optimized snap-displacement sequences”, *PRX Quantum* **3**, 030301 (2022).

- [71] Y. Ma, Y. Xu, X. Mu, W. Cai, L. Hu, W. Wang, X. Pan, H. Wang, Y. P. Song, C.-L. Zou, and L. Sun, “Error-transparent operations on a logical qubit protected by quantum error correction”, *Nature Physics* **16**, 827–831 (2020).
- [72] M. Aspelmeyer, T. J. Kippenberg, and F. Marquardt, “Cavity optomechanics”, *Rev. Mod. Phys.* **86**, 1391 (2014).
- [73] S. Lloyd and S. L. Braunstein, “Quantum computation over continuous variables”, *Phys. Rev. Lett.* **82**, 1784 (1999).
- [74] A. Ciani, D. P. DiVincenzo, and B. M. Terhal, “Lecture notes on quantum electrical circuits” (TU Delft OPEN Publishing, 2024).
- [75] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, “Circuit quantum electrodynamics”, *Reviews of Modern Physics* **93**, 10.1103/revmodphys.93.025005 (2021).
- [76] V. Ambegaokar and A. Baratoff, “Tunneling between superconductors”, *Phys. Rev. Lett.* **10**, 486 (1963).
- [77] S. P. Chitta, T. Zhao, Z. Huang, I. Mondragon-Shem, and J. Koch, “Computer-aided quantization and numerical analysis of superconducting circuits”, *New Journal of Physics* **24**, 103020 (2022).
- [78] R. M. Foster, “A reactance theorem”, *The Bell System Technical Journal* **3**, 259 (1924).
- [79] D. Atanasova, I. Yang, T. Hönigl-Decrinis, D. Gusenkova, I. Pop, and G. Kirchmair, “In situ tunable interaction with an invertible sign between a fluxonium and a post cavity”, *PRX Quantum* **6**, 10.1103/prxquantum.6.020318 (2025).
- [80] I. W. Yang, “Quantum information processing with superconducting circuits and high q cavities”, Ph.D. thesis, University of Innsbruck (2024).
- [81] M. D. Reed, L. DiCarlo, B. R. Johnson, L. Sun, D. I. Schuster, L. Frunzio, and R. J. Schoelkopf, “High-fidelity readout in circuit quantum electrodynamics using the jaynes-cummings nonlinearity”, *Physical Review Letters* **105**, 10.1103/physrevlett.105.173601 (2010).
- [82] K. S. Chou, “Teleported operations between logical qubits in circuit quantum electrodynamics”, *Ph.D. thesis*, Yale University (2018).
- [83] C. Eichler, “Experimental characterization of quantum microwave radiation and its entanglement with a superconducting qubit”, *Ph.D. thesis*, ETH (2013).
- [84] K. E. Cahill and R. J. Glauber, “Density operators and quasiprobability distributions”, *Phys. Rev.* **177**, 1882 (1969).

- [85] S. Haroche and J.-M. Raimond, “[Exploring the Quantum](#)” (Oxford University Press, 2006).
- [86] Y.-C. Liang, Y.-H. Yeh, P. E. M. F. Mendonça, R. Y. Teh, M. D. Reid, and P. D. Drummond, “Quantum fidelity measures for mixed states”, [Reports on Progress in Physics](#) **82**, 076001 (2019).
- [87] J. Zhang, Y.-x. Liu, R.-B. Wu, K. Jacobs, and F. Nori, “Quantum feedback: Theory, experiments, and applications”, [Physics Reports](#) **679**, 1–60 (2017).
- [88] I. Yang, T. Agrenius, V. Usova, O. Romero-Isart, and G. Kirchmair, “Hot schrödinger cat states”, [Science Advances](#) **11**, 10.1126/sciadv.adr4492 (2025).
- [89] D. F. Walls and G. J. Milburn, “[Quantum optics](#)”, 3rd ed. (Springer Cham, 2025).
- [90] S. Krastanov, V. V. Albert, C. Shen, C.-L. Zou, R. W. Heeres, B. Vlastakis, R. J. Schoelkopf, and L. Jiang, “Universal control of an oscillator with dispersive coupling to a qubit”, [Phys. Rev. A](#) **92**, 040303 (2015).
- [91] Y. Zhang, J. C. Curtis, C. S. Wang, R. J. Schoelkopf, and S. M. Girvin, “Drive-induced nonlinearities of cavity modes coupled to a transmon ancilla”, [Physical Review A](#) **105**, 10.1103/physreva.105.022423 (2022).
- [92] M. Rossignolo, T. Reisser, A. Marshall, P. Rembold, A. Pagano, P. J. Vetter, R. S. Said, M. M. Müller, F. Motzoi, T. Calarco, F. Jelezko, and S. Montangero, “Quocs: The quantum optimal control suite”, [Computer Physics Communications](#) **291**, 108782 (2023).
- [93] C. M. Bishop and H. Bishop, “Deep learning: Foundations and concepts” (Springer Cham, 2023).
- [94] B. Efron, “Bootstrap Methods: Another Look at the Jackknife”, [The Annals of Statistics](#) **7**, 1 (1979).
- [95] A. Pagano, M. M. Müller, T. Calarco, S. Montangero, and P. Rembold, “Role of bases in quantum optimal control”, [Phys. Rev. A](#) **110**, 062608 (2024).
- [96] S. Droste, T. Jansen, and I. Wegener, “On the analysis of the (1+1) evolutionary algorithm”, [Theoretical Computer Science](#) **276**, 51 (2002).
- [97] “[Settings in the optimization dictionary](#)”.
- [98] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, “Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation”, [Phys. Rev. A](#) **69**, 062320 (2004).

- [99] H. Pichler, S. Choi, P. Zoller, and M. D. Lukin, “Universal photonic quantum computation via time-delayed feedback”, *Proceedings of the National Academy of Sciences* **114**, 11362 (2017), <https://www.pnas.org/doi/pdf/10.1073/pnas.1711003114> .
- [100] A. Ask and G. Johansson, “Non-Markovian steady states of a driven two-level system”, *Phys. Rev. Lett.* **128**, 10.1103/PhysRevLett.128.083603 (2022), [arXiv:2102.11140](https://arxiv.org/abs/2102.11140) [quant-ph] .
- [101] R. Raussendorf, D. E. Browne, and H. J. Briegel, “Measurement-based quantum computation on cluster states”, *Phys. Rev. A* **68**, 022312 (2003).
- [102] H. T. Su, Y. Wang, F. Huang, and M. J. Lancaster, “Superconducting delay lines”, *Journal of Superconductivity and Novel Magnetism* **21**, 10.1007/s10948-007-0239-2 (2008).
- [103] N. Maleeva, L. Grünhaupt, T. Klein, F. Levy-Bertrand, O. Dupre, M. Calvo, F. Valenti, P. Winkel, F. Friedrich, W. Wernsdorfer, A. V. Ustinov, H. Rotzinger, A. Monfardini, M. V. Fistul, and I. M. Pop, “Circuit quantum electrodynamics of granular aluminum resonators”, *Nature Communications* **9**, 10.1038/s41467-018-06386-9 (2018).
- [104] Y. Krupko, V. D. Nguyen, T. Weiß, Dumur, J. Puertas, R. Dassonneville, C. Naud, F. W. J. Hekking, D. M. Basko, O. Buisson, N. Roch, and W. Hasch-Guichard, “Kerr nonlinearity in a superconducting josephson metamaterial”, *Physical Review B* **98**, 10.1103/physrevb.98.094516 (2018).
- [105] A. Palacios-Laloy, “Superconducting qubit in a resonator: test of the leggett-garg inequality and single-shot readout”, Ph.D. thesis, SPEWC-CEA Saclay (2010).
- [106] D. G. Atanasova, *Fabrication and characterization of a Dimer Josephson Junction Array Amplifier*, Master’s thesis, University of Innsbruck (2020).
- [107] C. X. Yu, S. Zihlmann, G. Troncoso Fernández-Bada, J.-L. Thomassin, F. Gustavo, Dumur, and R. Maurand, “Magnetic field resilient high kinetic inductance superconducting niobium nitride coplanar waveguide resonators”, *Applied Physics Letters* **118**, 10.1063/5.0039945 (2021).

Experimental setup and readout

A.1 Setup

The setup is similar to the one in [80] and shown in Fig. A.1. All the details on the theory of up- and downconversion can be found in [80].

A.2 High-power readout

At low drive powers, the resonator frequency experiences a qubit state-dependent dispersive shift [98]. This observation formed the basis for the first idea of the qubit readout. By applying a low-power drive to the readout resonator at the frequency corresponding to the qubit in the ground state, it is possible to distinguish the $|g\rangle$ and $|e\rangle$ qubit states from the resonator response in transmission (or reflection). The lower-power readout is QND, but typically has a low SNR and preferably requires an additional quantum-limited amplifier.

Increasing the drive power leads to a more complex response of the system [81]: the system enters a bistable region, where it alternates between "dim" and "bright" semiclassical solutions. In the transmission measurements, this bistable region looks like a nonlinear dip that shifts closer to the bare resonator frequency with increasing power. Above a certain power threshold, the system exits the bistable region and collapses to a "bright" solution, where the largest transmission dip appears at the bare resonator frequency. The minimum drive power for the excitation of the "bright" state depends on the qubit state (see Fig. A.2 A, B), which allows for the qubit readout. The best readout contrast can be found by sweeping the amplitude of the readout signal applied at a bare resonator frequency, as shown in Fig. A.2 C. The advantage of this scheme is a significant improvement in SNR compared to a standard low-power readout scheme. However, even though it can be used as a single-shot (see Fig. A.2 D), it is not QND.

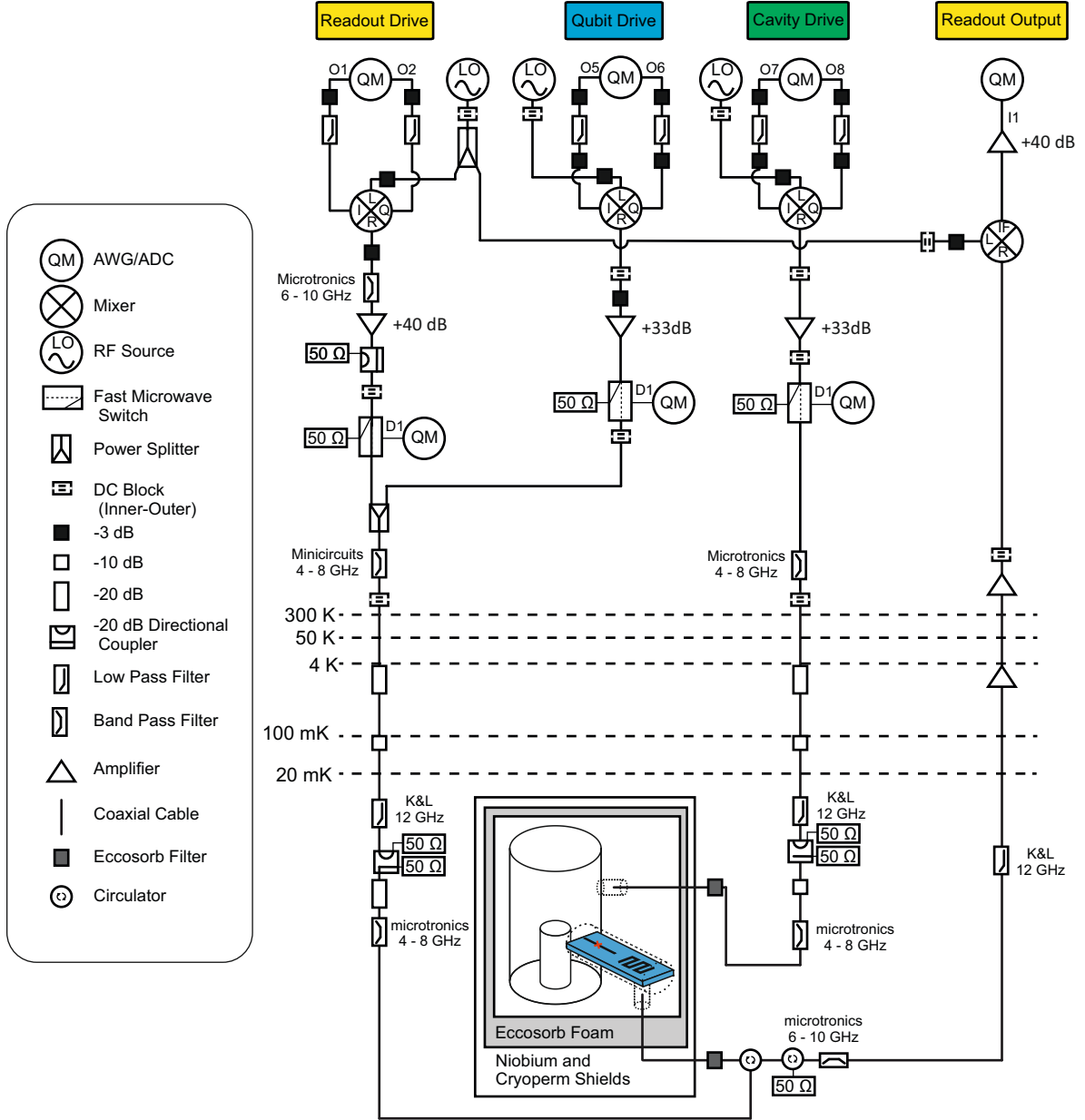


Figure A.1: Experimental microwave setup. The drive and readout pulses were created with IQ mixing up- and downconversion. Quantum Machines (QM) analog outputs have additional low-pass filters with 105 MHz cut-off for readout and 225 MHz cut-off for qubit and cavity (since the cut-off is very smooth, all the used intermediate frequencies pass through without significant attenuation). 3 dB attenuators are additionally introduced in the following cases: 1. before IQ mixers' inputs, to suppress the excitation of standing waves in short cables; 2. before reflective lowpass and bandpass filters. Since the QM OPX+ has filters inside, it is crucial to put an attenuator between analog output of the QM OPX+ and the subsequent lowpass filter. Each drive channel has its own amplifier to obtain enough power for short excitation pulses and high-power readout. To decouple the sample from amplified noise in the absence of the excitation drive, each drive channel has a homemade microwave switch. The switch introduces distortions to the pulses, which are subsequently suppressed by bandpass filters further along the microwave chain. Inner-outer DC blocks are introduced in the setup to avoid ground loops, which affect the remaining flux-tunable experiments in the cryostat.

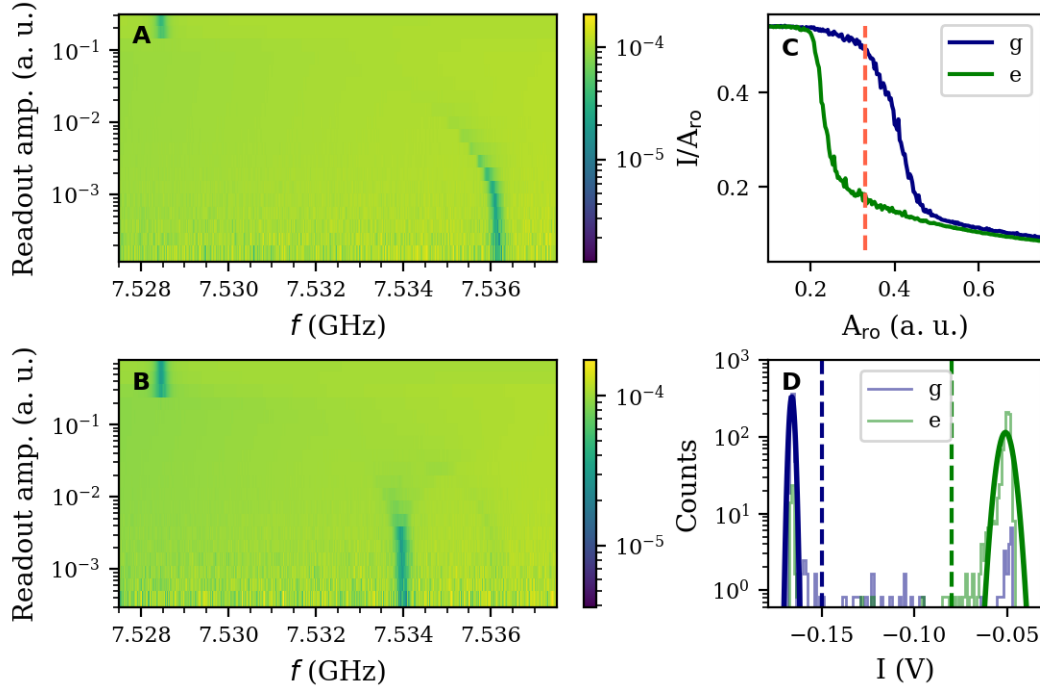


Figure A.2: High power readout. The readout amplitude is reported in units relative to the full-scale value of 0.4 V. For plots **C** and **D**, the readout signal is rotated so that all the information is contained in the I quadrature. **A**, **B** Resonator spectroscopy as a function of the readout pulse amplitude A_{ro} for qubit in $|g\rangle$ and $|e\rangle$. The linecut for each readout amplitude is normalised. **C** Response at bare resonator frequency (7.5284898(4) GHz) as a function of readout pulse amplitude for the qubit prepared in $|g\rangle$ and $|e\rangle$. The optimal readout amplitude corresponds to the largest readout signal difference, marked with a dashed line. This readout configuration was used for all the measurements in the thesis. Typically, 100 averages yield a good readout contrast. **D** Single-shot readout histogram for 1000 measurements. Dashed blue and green lines correspond to discrimination thresholds. Solid lines correspond to the fit of $|g\rangle$ data below the $|g\rangle$ threshold and $|e\rangle$ data above the $|e\rangle$ threshold with a normal distribution. For the chosen thresholds, readout single-shot fidelity is about 0.93. For single-shot Wigner measurements, the readout signal experienced significant bias at various points in phase space due to inefficient unselectivity of $\pi/2$ -pulses, which is why conventional averaging of the readout signal was used.

Notes on priors and inverse Wishart

As a reminder, the objective of Bayesian inference is to obtain the posterior distribution $p(\theta|X)$, which is updated after observing the data X , based on the prior distribution (prior knowledge or belief) $p(\theta)$:

$$p(\theta|X) = \frac{p(X|\theta)p(\theta)}{p(X)} \propto p(X|\theta)p(\theta) \quad (\text{B.1})$$

posterior $\propto$ likelihood $\times$ prior

If the prior distribution $p(\theta)$ belongs to some parametric family $\mathcal{P}$, the posterior distribution $p(\cdot|X)$ does not necessarily belong to the same parametric family. When posterior and prior belong to different parametric distribution families, the closed-form (algebraic) expression of the posterior may not exist at all. Then the posterior requires numerical computation. It becomes particularly problematic for dynamical or feedback problems, where the algorithm receives new data at each iteration, calculates the posterior and uses it as a prior at the next iteration. This approach is called sequential Bayesian update, or recursive Bayesian estimation.

Either way, it is beneficial to have a posterior in a closed-form expression, which can be derived if the prior distribution is conjugate. The prior family $\mathcal{P}$ is conjugate to the likelihood family $\mathcal{F}$ if

$$\forall \theta : p(\theta) \in \mathcal{P}, p(\cdot|\theta) \in \mathcal{F} \rightarrow p(\cdot|X) \in \mathcal{P} \quad (\text{B.2})$$

There is a list of conjugate prior distributions that correspond to likelihood distributions. In most cases, likelihood and conjugate prior families do not overlap. The conjugate prior for the multivariate normal distribution is an inverse Wishart distribution. The inverse Wishart distribution is defined on positive definite $p \times p$ matrices Σ and has a probability density

$$\mathcal{W}^{-1}(\Sigma; \psi, \nu) = \frac{|\psi|^{\nu/2}}{2^{\nu p/2} \Gamma_p(\frac{\nu}{2})} |\Sigma|^{-(\nu+p+1)/2} \exp\left(-\frac{1}{2} \text{tr}(\psi \Sigma^{-1})\right), \quad (\text{B.3})$$

which is parametrised with a $p \times p$ scale positive definite matrix ψ (scale matrix) and real number $\nu > p - 1$ (degrees of freedom). Here, Γ_p is a multivariate Gamma function, and

$|\cdot|$ is the determinant. ν and ψ control the mode (e.g. the value that appears the most) and the variance of the distribution

$$\Sigma_{\text{mode}} = \frac{\psi}{\nu + p + 1} \quad (\text{B.4})$$

The probability distributions for different values of ν and ψ are shown in Fig. B.1.

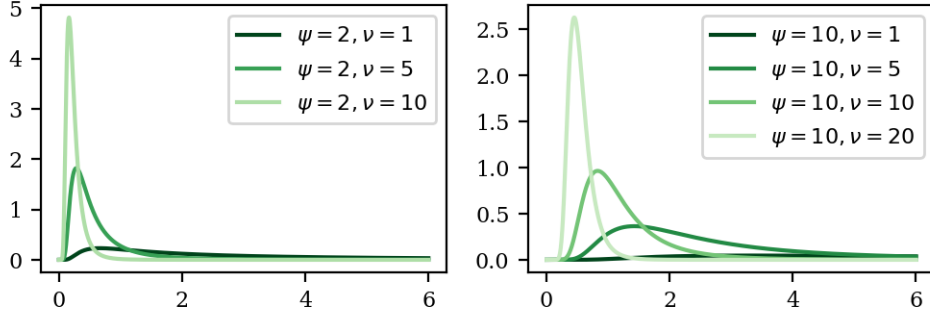


Figure B.1: Inverse Wishart distribution for different scale and degrees of freedom values.

However, if the prior distribution is an inverse Wishart, the posterior of the Bayesian model 5.3 does not have a closed-form expression. This may seem a bit counterintuitive, since the model consists of a sum of three two-dimensional Gaussians. The answer lies in the problem statement (see Fig. B.2). In the density estimation problem, the objective is to find the model that can generate the observed data by sampling from it,

$$\beta \sim \mathcal{N}(\mu, \Sigma). \quad (\text{B.5})$$

Here, β is a sample in a two-dimensional phase space, for consistency with the main text. In this setting, all the reasoning from above holds. For the reconstruction problem, however, the problem statement is very different. The repetitive measurement at some point β in two-dimensional phase space gives

$$W \propto \mathcal{N}(\beta|\mu, \Sigma), \quad (\text{B.6})$$

which is then fitted with the model 5.3 (here, for simplicity, $A_C = A_R = 0$, which corresponds to coherent/vacuum state). The measurement gives a deterministic (in a statistical sense) answer for a point in phase space, while the density estimation samples a point from the distribution.

Since there is no closed-form solution, MAP estimation is calculated numerically with `scipy.minimize`, see 5.18 for the loss function.

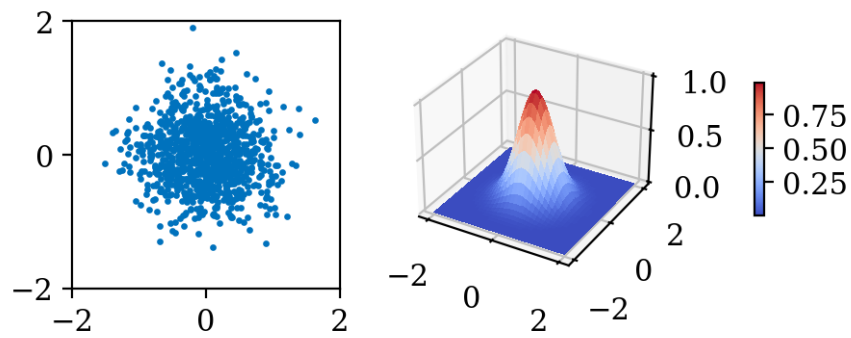


Figure B.2: Density estimation problem (left) fits the observed data distribution with the most suitable model distribution. Wigner measurement (right) returns a quasiprobability at phase space points, which is passed to the model. Both problems are defined on a two-dimensional phase space.

Nelder-Mead hyperparameter search

Hyperparameter search is done by sweeping one parameter in the optimisation dictionary with respect to a "base" configuration, in which

- basis vector number: 5
- basis max: 10
- amplitude variation: $0.2 \times \text{qubit_amp}$
- basis: Sinc with $\omega_{\max} \in [1, 300]$
- superiteration number: 2
- maximum number of evaluations: 1000

Each configuration is tested on 5 runs, each run is plotted with a different green line. Blue lines for each run represent a maximum running FoM value. Red line is an average maximum running FoM over 5 runs.

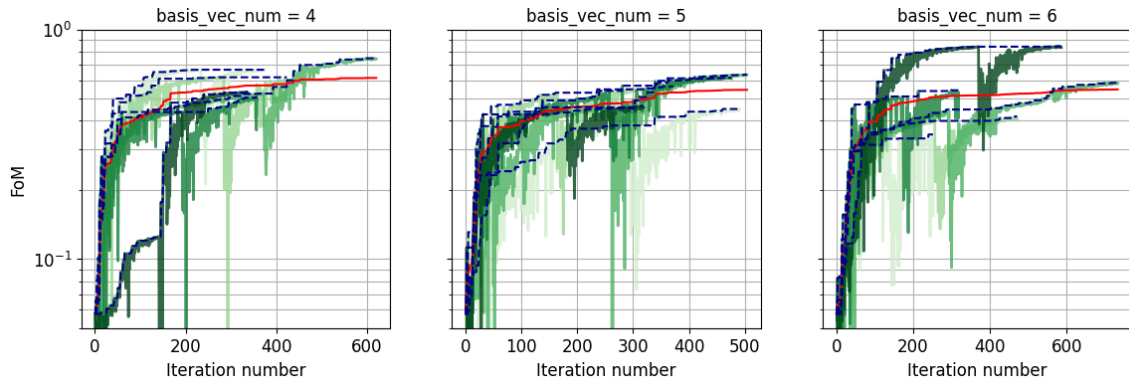


Figure C.1: Convergence curves for different numbers of basis vectors.

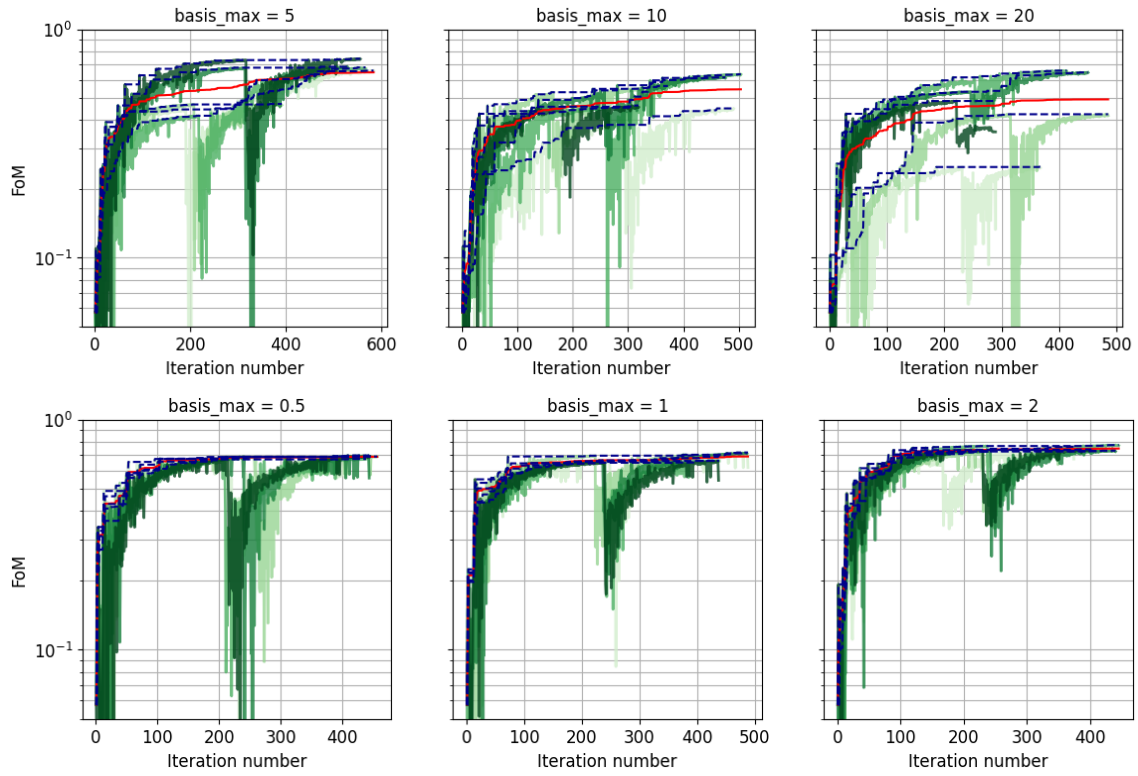


Figure C.2: Convergence curves for different `basis_max` parameter, that determines the width of the sigmoid.

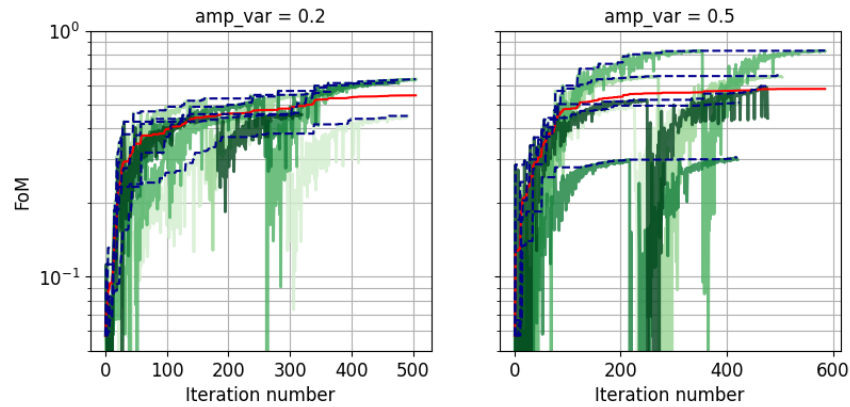


Figure C.3: Convergence curves for different amplitude variation values.

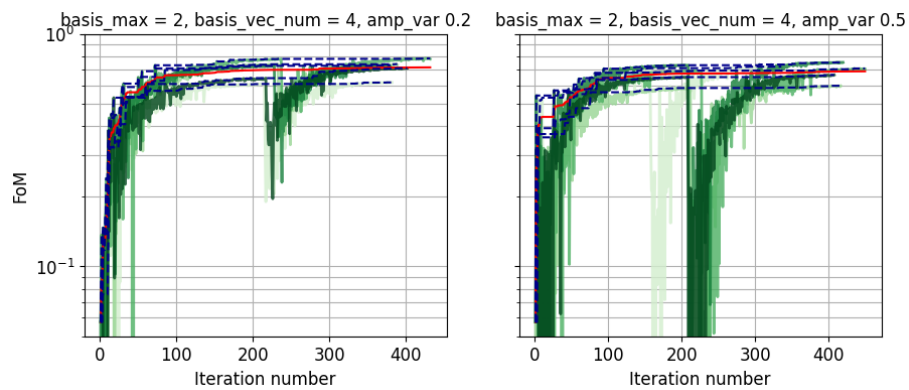


Figure C.4: Convergence curves for the two most prominent parameter sets.

Reconstruction convergence for smaller cats

This Appendix contains additional reconstruction data (Figs. D.2 to D.4 and D.7) for smaller cats shown in Fig. D.1. Interestingly, for $\alpha = 1.5$ cat, sampling with tight entropy ranges $[-2, 0.1]$ and $[-2, 0.5]$ results in 0.2–0.3 reconstruction infidelity: the model predicts a wrong cat rotation angle (see Fig. D.5). It seems that smaller states require more sampling randomness to achieve low infidelity (see Fig. D.6).

All the hyperparameters not specified under the plots are the same as for the $\alpha = 3$ cat state in the main text.

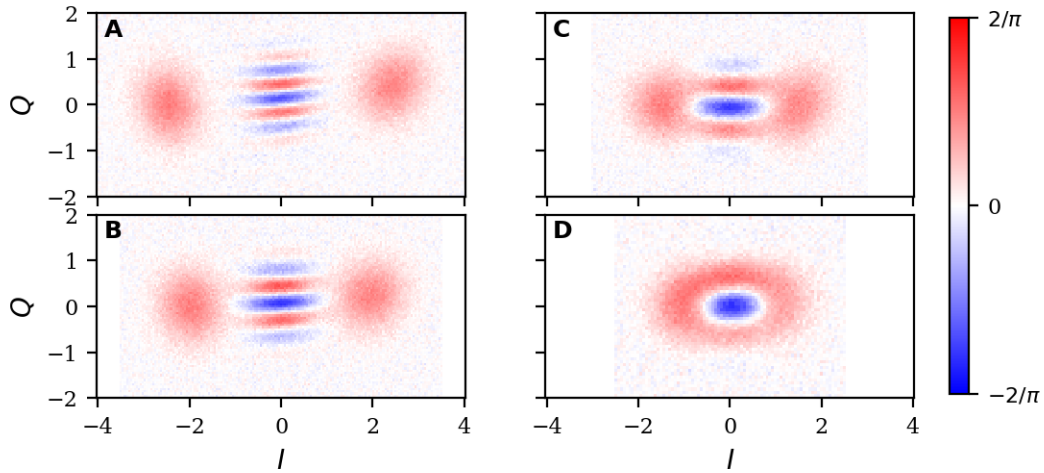


Figure D.1: Wigner maps for cat states of different sizes: **A** $\alpha = 2.5$, **B** $\alpha = 2$, **C** $\alpha = 1.5$, **D** $\alpha = 1$.

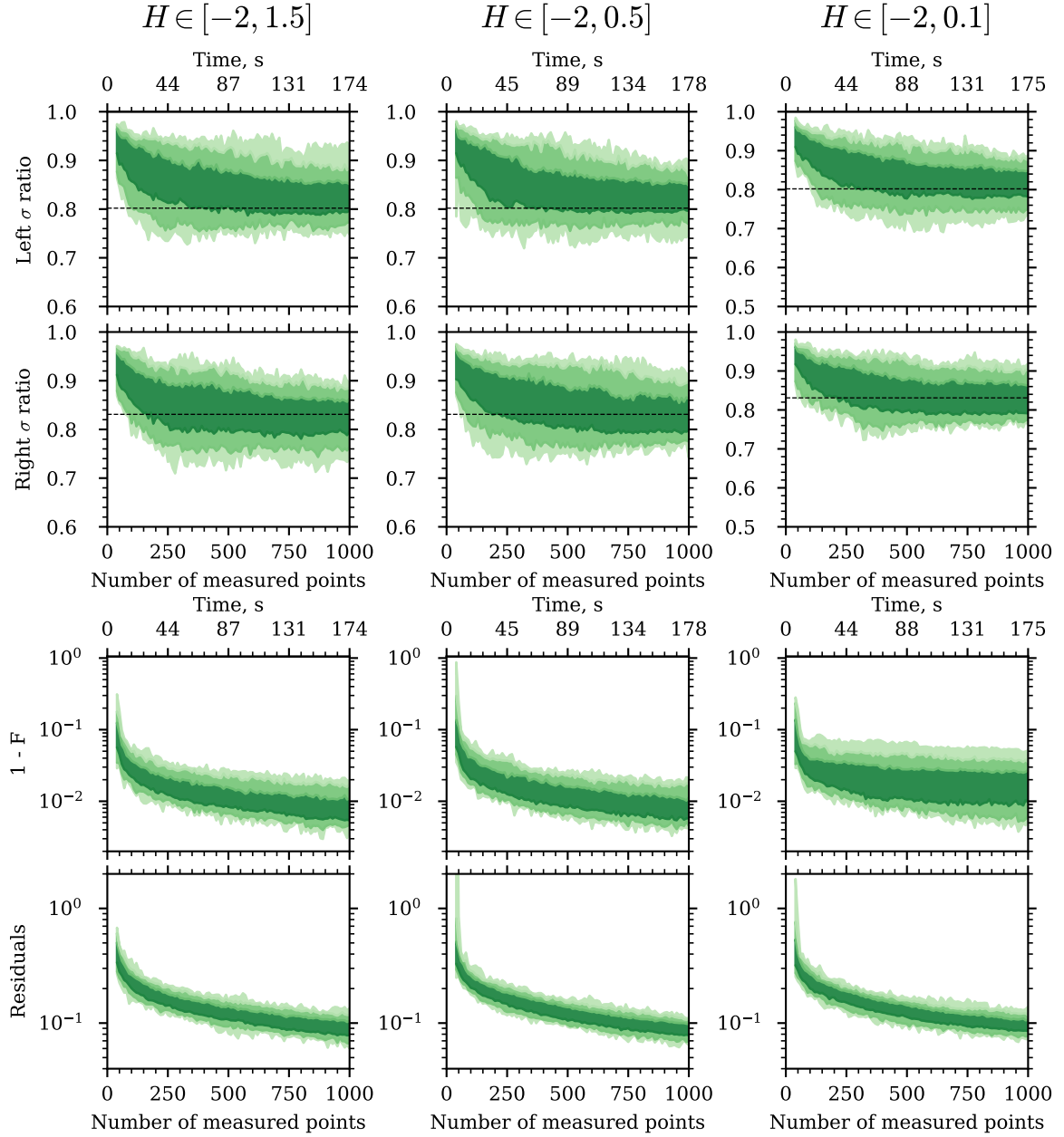


Figure D.2: Convergence curves for different entropy ranges for rejection sampling and $\alpha = 2.5$ cat state (see Fig. D.1 A). The left column corresponds to $H \in [-2, 1.5]$, the centre and right columns to rejection sampling with different entropy ranges, $H \in [-2, 0.5]$ and $H \in [-2, 0.1]$. All the runs utilised an initial $\beta_B = 60$, 100 averages and 10 measurements per batch.

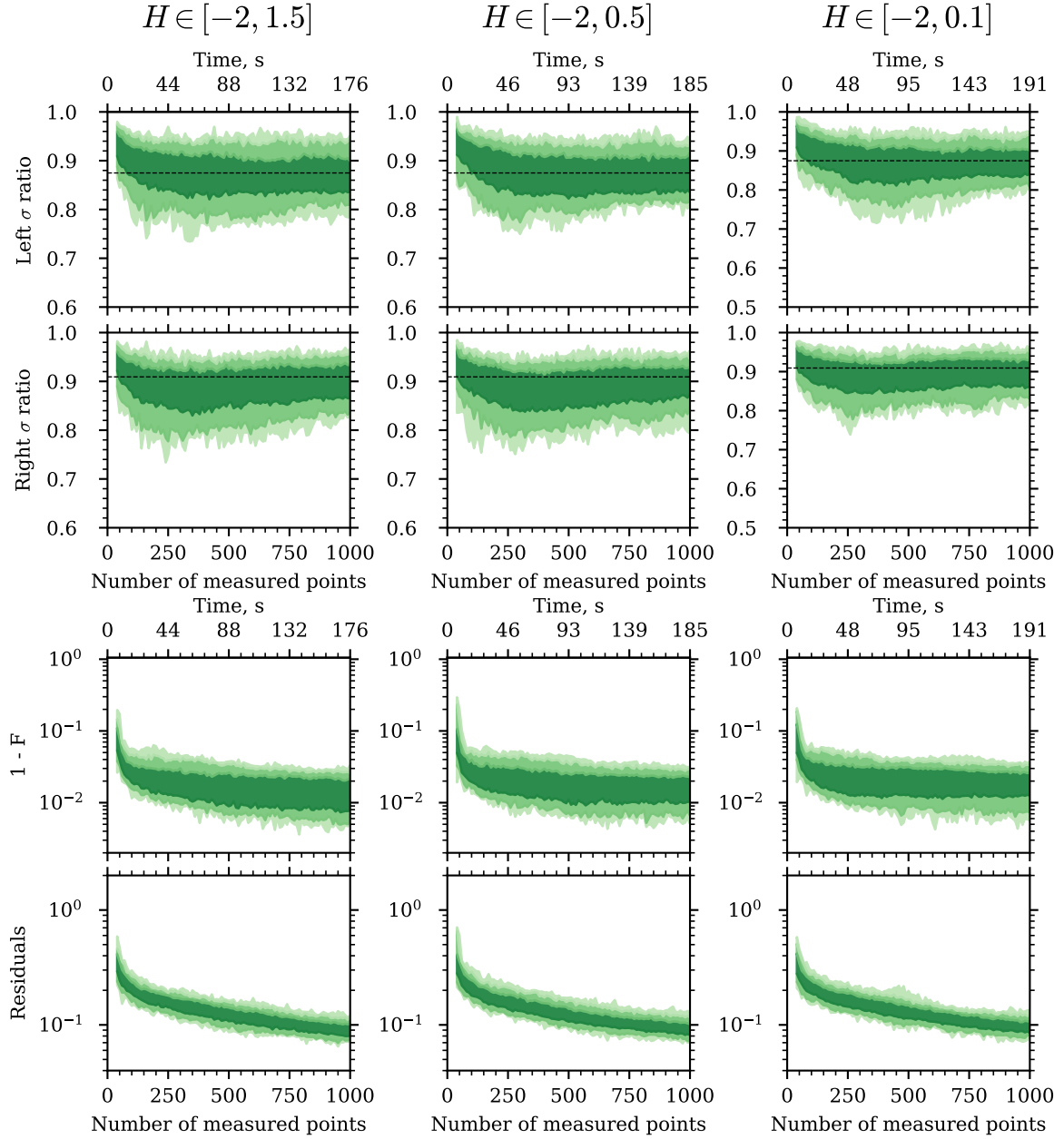


Figure D.3: Convergence curves for different entropy ranges for rejection sampling and $\alpha = 2$ cat state (see Fig. D.1 B). The left column corresponds to $H \in [-2, 1.5]$, the centre and right columns to rejection sampling with different entropy ranges, $H \in [-2, 0.5]$ and $H \in [-2, 0.1]$. All the runs utilised an initial $\beta_B = 60$, 100 averages and 10 measurements per batch.

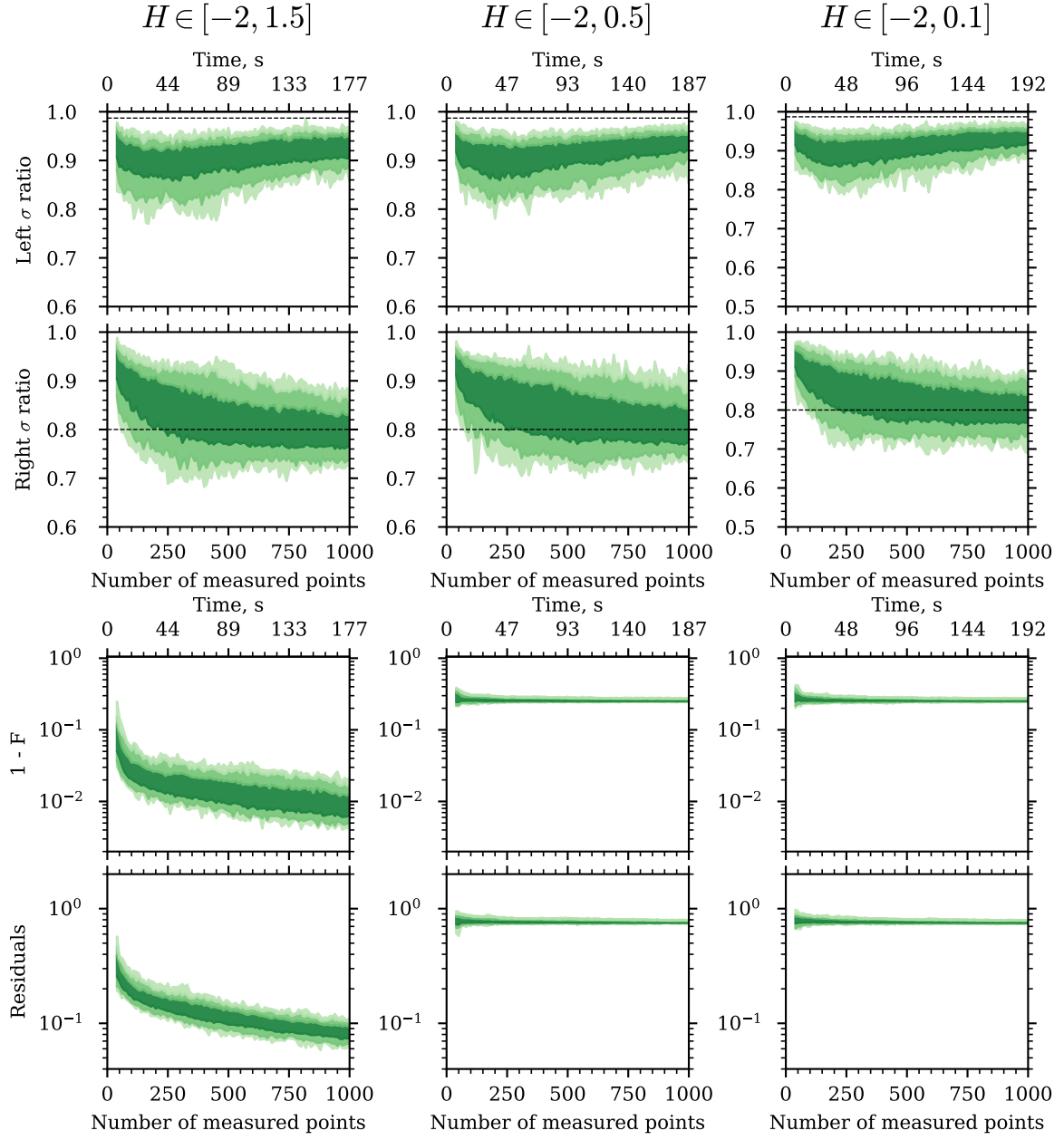


Figure D.4: Convergence curves for different entropy ranges for rejection sampling and $\alpha = 1.5$ cat state (see Fig. D.1 C). The left column corresponds to $H \in [-2, 1.5]$, the centre and right columns to rejection sampling with different entropy ranges, $H \in [-2, 0.5]$ and $H \in [-2, 0.1]$. All the runs utilised an initial $\beta_B = 60$, 100 averages and 10 measurements per batch.

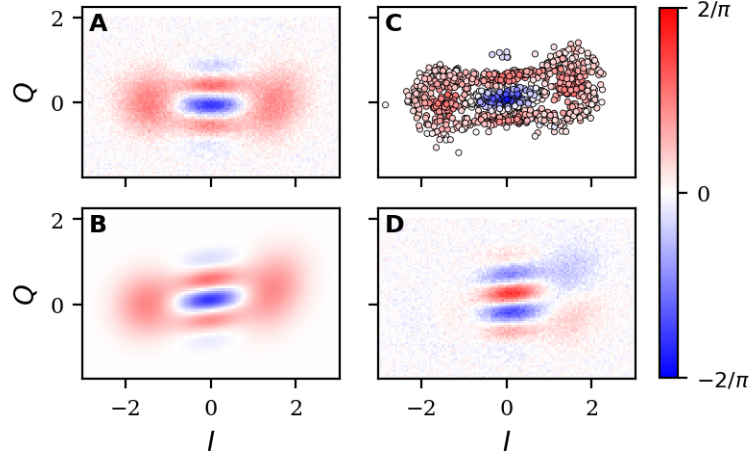


Figure D.5: Example of an unsuccessful $\alpha = 1.5$ cat reconstruction for $H \in [-2, 0.5]$. The model is biased towards the wrong cat rotation angle. **A** Reference high-resolution Wigner map. **B** Reconstructed Wigner map. **C** 1000 measured points during the reconstruction. **D** Residuals between the measured and reconstructed data $R(I, Q) = W_{\text{meas}}(I, Q) - W_{\text{model}}(I, Q)$.

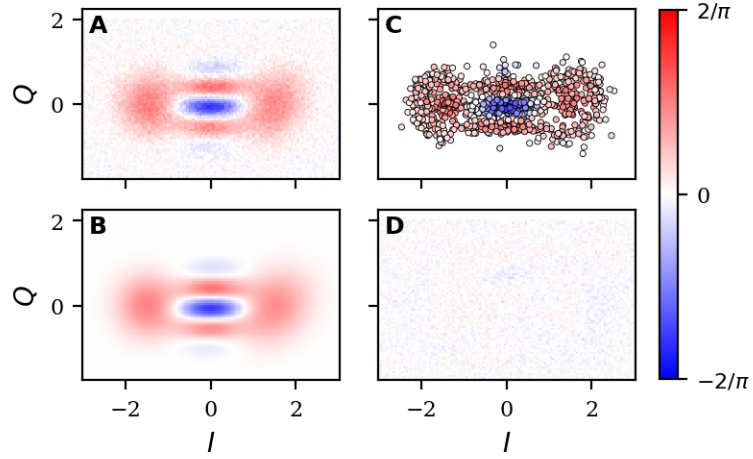


Figure D.6: Example of a successful $\alpha = 1.5$ cat reconstruction for $H \in [-2, 1.5]$. **A** Reference high-resolution Wigner map. **B** Reconstructed Wigner map. **C** 1000 measured points during the reconstruction. **D** Residuals between the measured and reconstructed data $R(I, Q) = W_{\text{meas}}(I, Q) - W_{\text{model}}(I, Q)$.

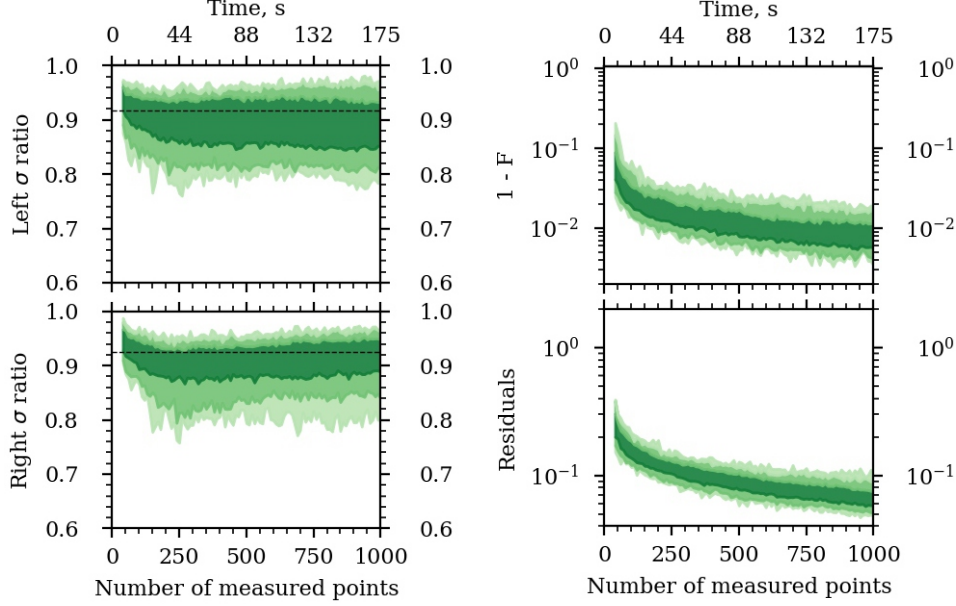


Figure D.7: Convergence curves for $\alpha = 1$ cat state (see Fig. D.1 D) reconstruction with rejection sampling, $H \in [-2, 1.5]$. For the $\alpha = 1$ cat state, only one set of hyperparameters was characterised.

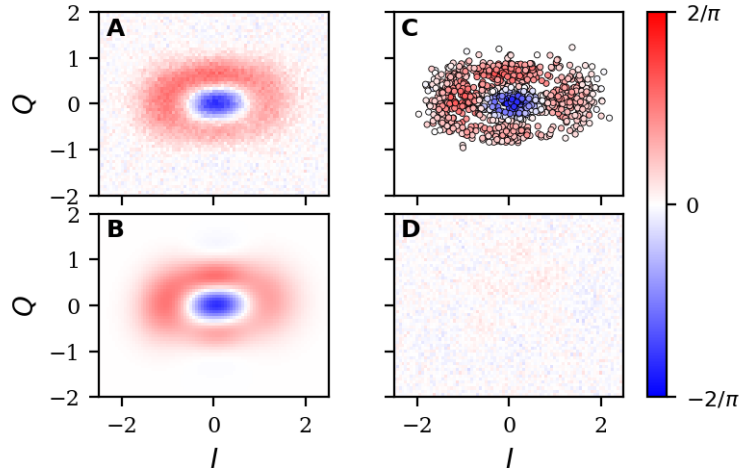


Figure D.8: Example of a successful $\alpha = 1.5$ cat reconstruction for $H \in [-2, 1.5]$. **A** Reference high-resolution Wigner map. **B** Reconstructed Wigner map. **C** 1000 measured points during the reconstruction. **D** Residuals between the measured and reconstructed data $R(I, Q) = W_{\text{meas}}(I, Q) - W_{\text{model}}(I, Q)$.

GrAl delay line

E.1 Motivation

The initial goal of the PhD was to create a system consisting of a quantum emitter interacting with the photons via a time-delayed feedback [99]. Such a system would serve as a good platform for observing non-Markovian effects [100] and creating 2D cluster states, which form the basis for measurement-based quantum computation [101]. In the generation of $N \times N$ cluster states with time-delayed feedback (see Fig. E.1 and reference [99] for the protocol details), the quantum emitter interacts with the emitted photon one more time after $N - 1$ more photons were emitted. Those $N - 1$ photons are meanwhile stored in the delay line, which determines the minimum required delay for the protocol. One cycle of the protocol was estimated to take at least 250 ns (depending on the quantum emitter coupling and photon shaping). Thus a 3×3 cluster state requires at least 750 ns delay. One way to achieve such a long delay on a chip is to use a high kinetic inductance material, such as granular Aluminium (grAl), for the delay line. Even for grAl with high resistivity $\rho = 1170 \mu\Omega \cdot \text{cm}$, the required delay line length is on the order of meters.

Given the extensive four-year fabrication effort, this Appendix presents only the most significant steps that were crucial in establishing the final fabrication recipe.

E.2 Optical lithography

The first set of resonators was fabricated on a prime silicon wafer using optical lithography with a laser writer from Microtech. After a standard wafer cleaning in an ultrasonic bath with acetone and isopropanol (IPA), a 20 nm grAl layer was evaporated in a Plassys system. The evaporation recipe included a sample cleaning (descum) step, Ti gettering for improving the vacuum in the evaporation chamber, and grAl target cleaning. Even though grAl evaporation parameters were recalibrated each time right before the fabrication of the new sample, the grAl resistivity remained unstable. Since it was difficult to obtain reproducible grAl resistivity in the Plassys, grAl evaporation was later replaced with sputtering in the AJA system in O_2/Ar atmosphere (see recipes in Section E.3). To improve optical resist adhesion, the adhesion promoter AR 300-80 was spun onto a wafer, which was preliminarily cleaned with O_2 plasma in Sentech, before the optical resist AZ 1505 was applied. The resist was exposed in a Contour In mode of the laser writer and

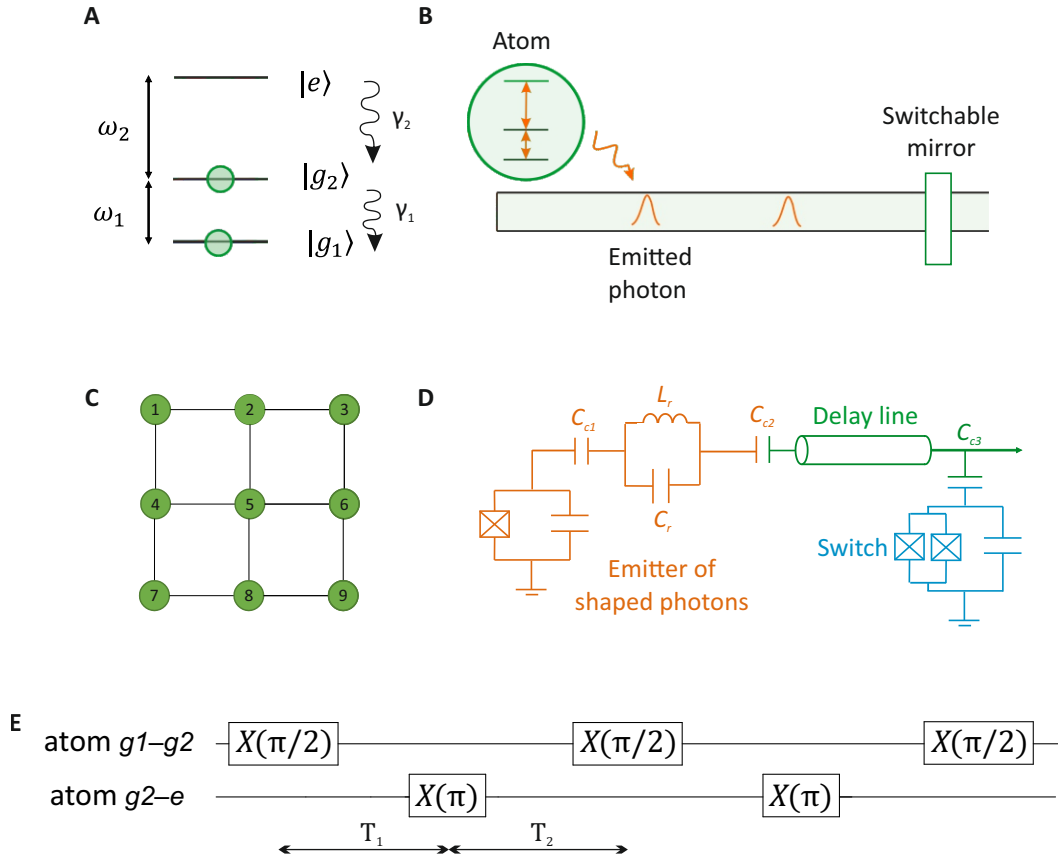


Figure E.1: 2D cluster state generation with time-delayed feedback. The qubits in a cluster state are encoded via the presence or absence of a photon in a pulse. **A** The quantum emitter (atom) has two metastable states $|g_1\rangle$ and $|g_2\rangle$, and an excitation state $|e\rangle$. The protocol starts with the initialisation of the atom in the superposition of metastable states, $(|g_1\rangle + |g_2\rangle)/\sqrt{2}$. **B** The atom emits photons into the waveguide. Depending on the state of the switchable mirror, the photons either leave the waveguide or reflect and interact with an atom for a second time. As a result of this interaction, the atom picks a π phase. **C** Example of a 3×3 cluster state. The photons #1 – 3 are emitted by an atom one after another and reflected from the switch back to the atom. If the photonic mode #1 is not empty, it interacts with the atom before the emission of the photon #4, and the atom picks up a π -phase. **D** Possible experimental realisation for a superconducting circuits platform. **E** Pulse sequence for a 2D cluster state generation. After the initialisation of an atom, the protocol consists of repeating the following steps until the desired state dimension is reached:

- Wait for a time $T_1 = 1/\gamma_1$, so that the photon that was generated several steps ago and returned to the atom interacts with it (CPASE gate).
- Apply π -pulse at transition $|g_2\rangle - |e\rangle$.
- Wait for a time $T_2 = 1/\gamma_2$, such that the state $|e\rangle$ completely decays into the waveguide.
- Apply $\pi/2$ -pulse at transition $|g_1\rangle - |g_2\rangle$.

grAl evaporation, Plassys	
wafer cleaning	acetone (5 min) + IPA (3 min) in US bath (35 kHz, 100%)
descum + rot	O ₂ /Ar plasma (5/10 sccm) V = 200 V, I = 10 mA, V _{acc} = 50 V
Ti gettering	0.2 nm/s, 2 min
grAl evaporation ¹	1 nm/s, O ₂ flow depends on the calibration (3 - 5 sccm)
grAl target cleaning	
Optical lithography	
plasma cleaning, Sentech	O ₂ , 10 min 5 Pa, 30 sccm, RF 100 W, ICP 0 W
adhesion promoter spinning	AR 300-80 new, 4000 rpm, 100 s
adhesion promoter baking	105°C, 2 min
resist spinning	AZ 1505, 4000 rpm, 100 s
resist baking	100°C, 1 min
resist exposure	Contour In, ~ 50 mJ/cm ²
development	AZ developer:H ₂ O, 1:1 for 1 minute
plasma etching, Sentech	BCl ₃ /Ar (30/3 sccm, RF 40W, ICP 200W), 35 s
resist removal	acetone (10 min) + IPA in US bath at 35 kHz, 100 %

Table E.1: Fabrication recipe for long CPW grAl resonators with optical lithography.

developed in AZ developer. In the end, the samples were etched with BCl₃/Ar plasma in the Sentech reactive ion etching system. Finally, the resist was removed with acetone in a heated ultrasonic (US) bath at 35 kHz for 10 min. The fabrication parameters are presented in Table E.1.

To reduce the excitation of slotline modes, ground planes between the turns were connected with wire bonds. However, the slotline modes remained too strong. Since both the central conductor and the ground plane were made of grAl, the resonator and slotline modes exhibited similar power-dependent behaviour, making them difficult to distinguish. Changing the resonator geometry to achieve partial cancellation of the slotline modes did not yield any significant improvement. Sonnet simulations showed that the only way to eliminate strongly coupled slotline modes was to change the material of the ground plane to Aluminium, as shown in Fig. E.2. Combining two different materials on the same chip would have required at least one additional fabrication step involving precise alignment across the wafer. However, the available laser writer lacked an interferometric stage, which already made the patterning of long resonators particularly challenging. To ensure the reproducibility of the CPW resonator's central conductor under these constraints, the resonator layout was divided into smaller sections, each containing two continuous lines (CPW gaps). The lines within each section were then patterned sequentially to minimise stitching errors and ensure geometric consistency across the wafer. This approach helped reduce the effect of motor backlash, but the alignment between the two fabrication layers would have made the situation even trickier. Due to this reason, the fabrication process was shifted to electron beam lithography (e-beam).

¹All the grAl evaporation recipes include the target cleaning step before the evaporation

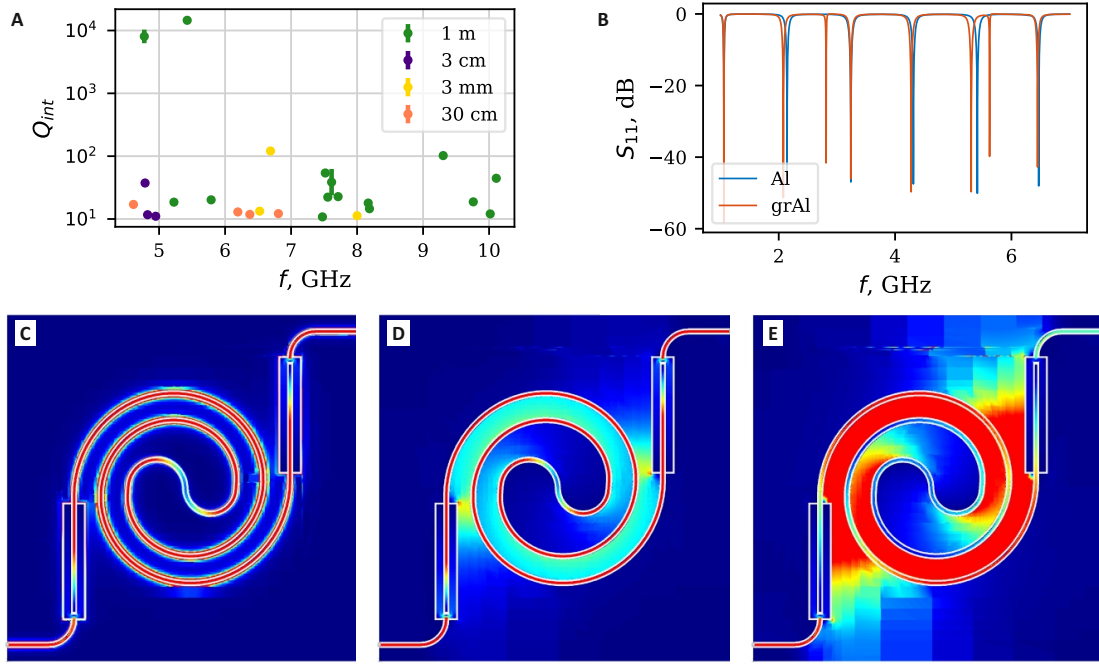


Figure E.2: Tackling the slotline problem for high kinetic inductance resonators. **A** Measured internal quality factors for the resonators with grAl central conductor and ground plane. Due to strong hybridisation between the slotline and resonator modes, it was not possible to distinguish them from transmission/reflection measurements. **B** - **E** The simulations for a $2237 \mu\text{m}$ long grAl CPW resonator performed in Sonnet electromagnetic analysis software. The spiral design is intended to help reduce the coupling between adjacent lines [102]. **B** Simulated reflection coefficient S_{11} from port 1 for the spiral resonator with Al (blue) and grAl (orange) ground planes. Al ground plane effectively eliminates slotline modes. **C** Current magnitude distribution for a mode at ≈ 2.159 GHz for a resonator with Al ground plane. The current magnitude varies from 0 (blue) to 87.1 (red) Amps/meter. **D** Current magnitude distribution for a mode at ≈ 2.074 GHz for a resonator with grAl ground plane. The simulation shows a partial excitation of slotline modes in the ground plane. The current magnitude varies from 0 (blue) to 83.2 (red) Amps/meter. **E** Current magnitude distribution for a slotline mode at ≈ 2.8255 GHz for a resonator with grAl ground plane. The current magnitude varies from 0 (blue) to 131 (red) Amps/meter.

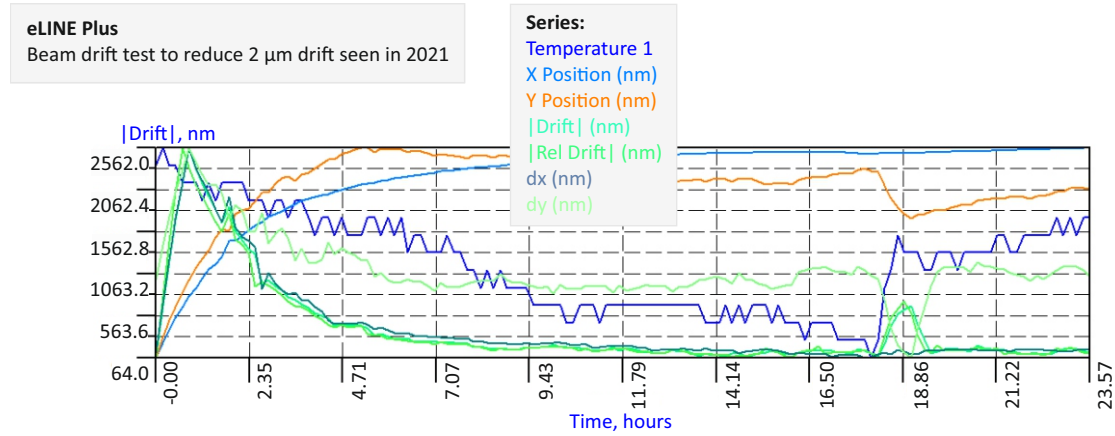


Figure E.3: Beam drift on Raith e-Line after loading the sample holder (credits to: Markus Weiss). To minimise drifts due to thermalisation issues, the sample holder was permanently kept inside the loading module when the e-Line was not used for patterning. To eliminate the drifts, one needs either to keep the sample inside the e-Line chamber for a few hours before the patterning or to realign during the exposure.

E.3 Electron beam lithography

To eliminate possible stitching problems in beam deflection (BD) mode and decrease the total patterning time, the resonators were patterned in FBMS mode, where the lithographer moves the stage while keeping an electron beam on. The e-beam lithographer accumulates 2 – 2.5 μm drifts in about 5 hours (see Fig. E.3), which is excessive for precise submicrone patterning over the whole wafer. Therefore, each resonator was divided into subsections, similar to the design for the laser writer (see Fig. E.4). Each subsection had its own alignment mark. Initially, a linescan alignment was used. However, it appeared to be unstable for accurate adjustments: 10 subsequent alignments on the same mark in a linescan mode led to complete misalignment and even loss of focus. The linescan alignment was later replaced with a GDSII-based mark scan, which compares the scanned image of the mark with the designed one (see Fig. E.4 C).

Over the years, the fabrication recipe has undergone substantial modifications. The final version is summarised in Table E.2 and schematically illustrated in Fig. E.6. To suppress slotline modes, all resonators were manually wirebonded across the signal lines. As a subsequent development, a design incorporating airbridges was proposed for implementation via greyscale lithography. However, this project was not completed due to unforeseen issues encountered during the fabrication of long resonators. Many wafers exhibited particle contamination that disrupted the resonator continuity. Notably, long test grAl lines fabricated on high-resistivity silicon did not exhibit this issue. The underlying cause remained unclear until the end of the project. A notable observation made afterwards suggested that particle contamination might have originated from the oil pump used in the gold evaporation machine, which could have emitted particles during operation. Although this explanation could account for the inconsistent contamination levels, it was never experimentally verified.

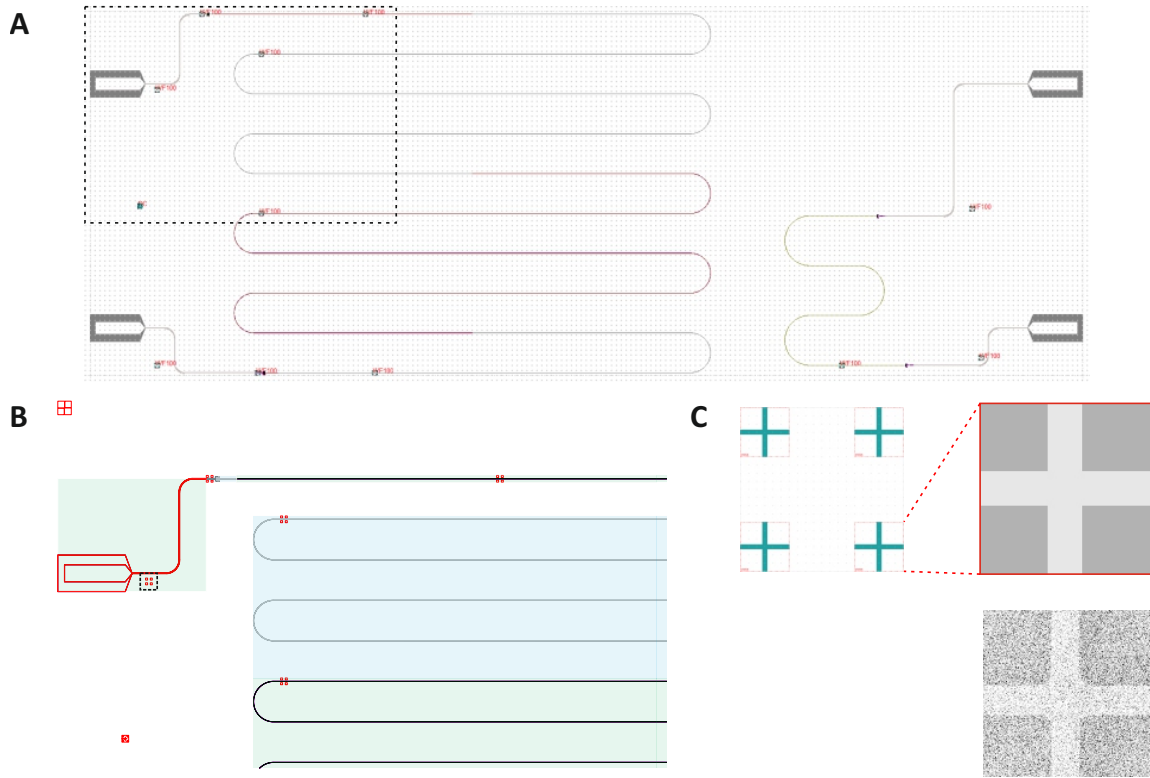


Figure E.4: The long resonator design. The resonators could be measured either in transmission or in reflection from each side. The coupling on both sides was designed to be equal. **A** One of the samples on the wafer. The 2-inch wafer contained several samples with the resonators of varying lengths, ranging from 1 mm to 30 cm. **B** Zoom in the dashed rectangle area from Fig. **A**. The long resonator design is subdivided into subsections, coloured in blue and green for guidance. Each subsection has its own working area and alignment mark, which is scanned with an image alignment mode to compensate for thermal drifts. The subsections have a slight overlap to guarantee continuity of the resonator's central conductor. The design and e-Line position list are generated using Python code. Realignment at each part of the resonator ensures reproducibility of CPW resonator dimensions over the wafer. **C** Zoom in on the dashed rectangle area from Fig. **B**, showing the alignment mark. For realignment, the e-Line software compares an image of each corner of the alignment mark to its design.

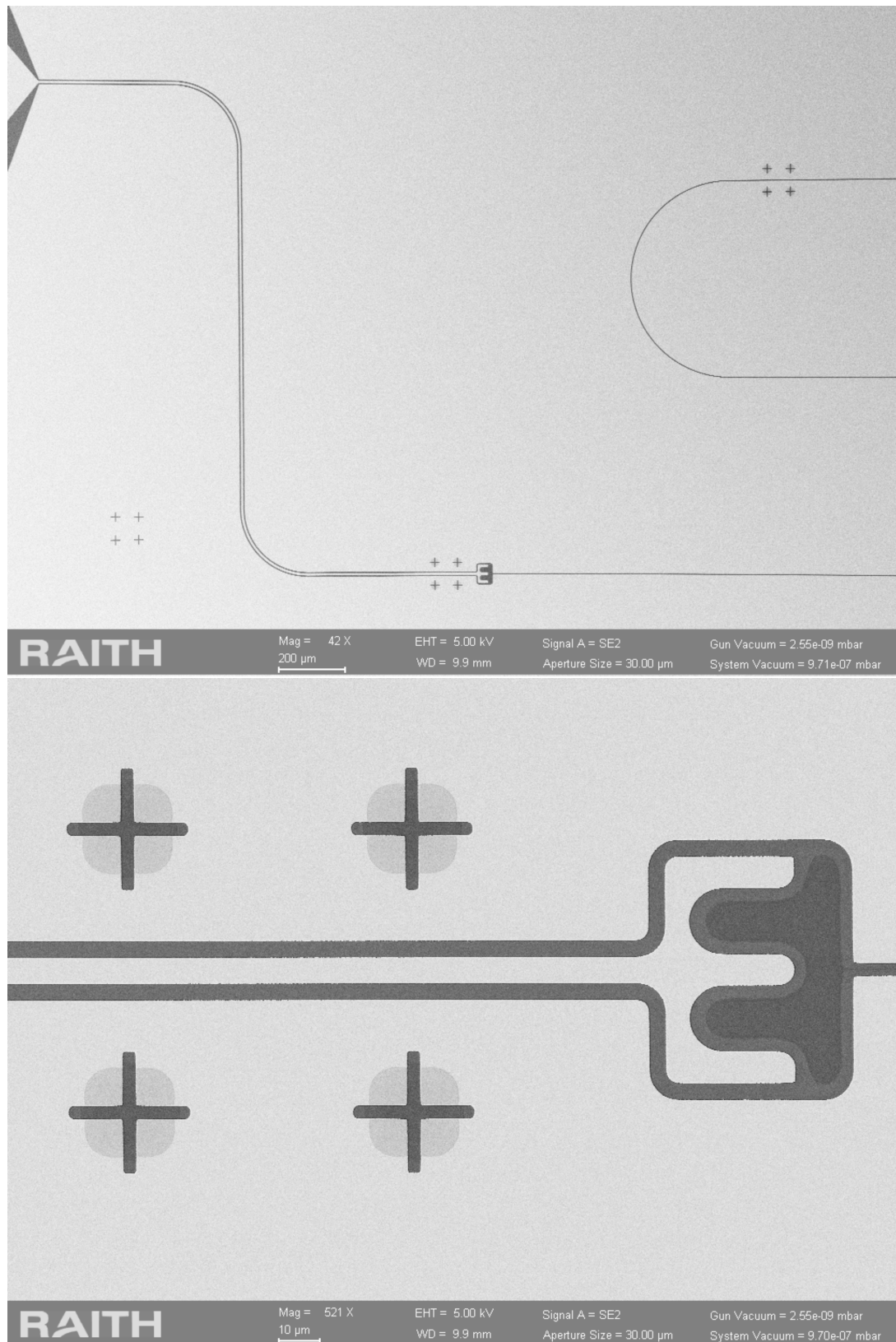


Figure E.5: SEM images of one of the fabricated long resonators. The capacitance claw and central conductor of the resonator are fabricated out of grAl and have a dark grey colour. The transmission line and the ground plane are made of Al, which looks light grey on SEM images. The designed resonator central conductor width is $0.6 \mu\text{m}$.

The oil pump may not have been the only contributor to the contamination. While tackling the particle contamination problem, it was found that cleaning wafers with various chemicals in the yellow room could itself introduce particle contamination of the wafer. The test results were inconclusive, as it was impossible to identify the exact source of contamination — whether from the beakers, storage chemical bottles, or the delivered chemicals. At that stage, the least contaminated wafers used for grAl sputtering were consistently those that skipped the wet cleaning step.

The final recipe appeared promising in terms of avoiding potential resonator interruptions during the fabrication. Examining the sample created with the final recipe from Table E.2 did not reveal severe particle problems; however, all the central conductors appeared to be heavily overetched, resulting in approximately 100 nm wide lines for a designed width of 600 nm. Most probably, overetching led to interruptions in the long central conductors. Despite the optimistic insight into the final recipe, fabrication was stopped after the first set of resonators was fabricated, as the project was taking way too long.

²negative resist

Alignment marks layer	
resist spinning	EL13 at 1500 rpm
resist baking	5 min on hotplate at 200°C
resist spinning	A4 at 2000 rpm
resist baking	5 min on hotplate at 200°C
e-Line, 30 kV, 30 μm aperture	BD dose: 300 $\mu\text{C}/\text{cm}^2$, dwell time: 0.002836
development	IPA:H ₂ O 3:1 at 6°C for 2 min 15 s
Nb evaporation, Plassys	descum: Ar/O ₂ 10/5, 200 V, 10 mA, 2 min
	Ti gettering 0.2 nm/s, 2 min
	Nb 70 nm
liftoff	NMP (5 min) + acetone (5 min) + IPA
	40°C, US 130 kHz 50%
grAl sputtering, AJA	
chamber preparation	O ₂ , 20 sccm, 50 mTorr, 10 min
target cleaning	5 min, 160 V
grAl sputtering, 20 nm	O ₂ /Ar 10 sccm, Ar 20 sccm, 6 min, 205 V
Nb ground plane	
resist spinning	MaN ² 2401 at 3000 rpm for 45 s
baking	1 min on hotplate at 90°
	dose: 180 $\mu\text{C}/\text{cm}^2$
e-Line, 30 kV, 20 μm aperture	FBMS step 0.05 for 0.6 μm , defl cycle 238 μs
(A)	speed: 0.156879 mm/s
	BD dose: 140 $\mu\text{C}/\text{cm}^2$, dwell time: 0.0033
development + etching (B)	MaD 525 (TMAH) for 1 min, transfer to water $\times 2$
	descum: Ar/O ₂ 10/10 sccm, 30 mTorr, 240 s
Nb evaporation, AJA (C)	Nb sputtering, 70 nm: 7 min, 372 V
Transmission line central conductor and Al capping of the ground plane	
resist spinning	MaN 2403 at 3000 rpm for 45 s
baking	1 min on hotplate at 90°
	dose: 220 $\mu\text{C}/\text{cm}^2$
e-Line, 30 kV, 60 μm aperture	FBMS step 0.3 for 3 μm , defl cycle 174 μs
(E)	speed 0.215772 mm/s
	BD dose: 240 $\mu\text{C}/\text{cm}^2$, dwell time 0.000674
development (F)	MaD 525 (TMAH) for 1 min 30 s, transfer to water $\times 2$
etching oxide layer, Plassys	Ar 5 sccm, 400 V, 20 mA, 2 min
Al evaporation, Plassys (G)	20 nm
	acetone, 5 min US 130 kHz 20%
liftoff (H)	NMP, 60°C, 5 min US 130 kHz 20%
chamber preparation, Sentech	O ₂ 50 sccm, 400/5W ICP/RF, 10 min
etching, Sentech (K)	CF ₄ /O ₂ , 20/2 sccm, ICP/RF 50/50, 1 min 56 s

Table E.2: The last version of the fabrication recipe for long CPW grAl resonators. The corresponding fabrication steps in Fig. E.6 are marked with circled letters.

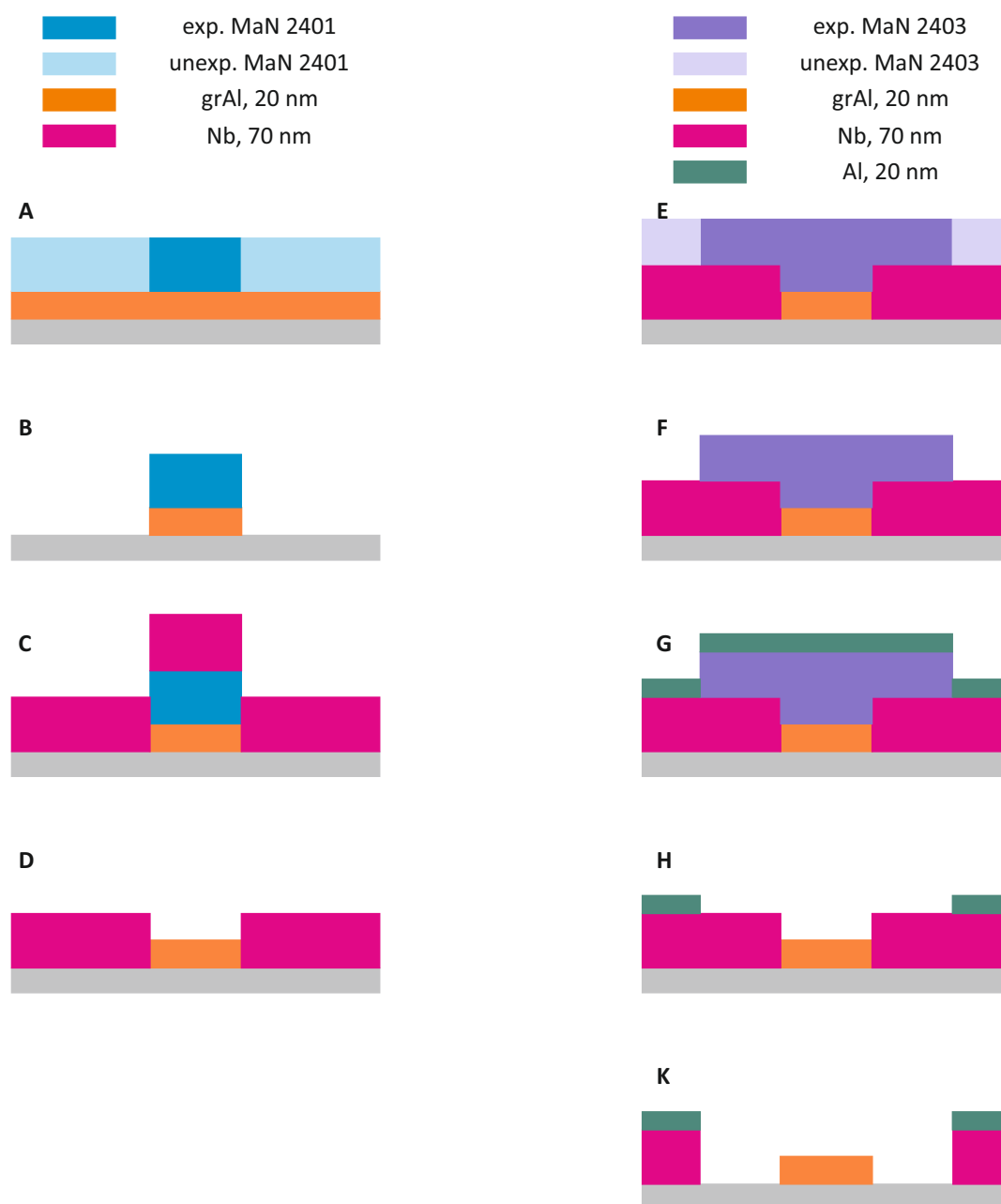


Figure E.6: Schematic representation of the fabrication process. Dimensions are not to scale.

E.4 Transmission and reflection measurements

Out of all the fabrication runs, only one had a few uninterrupted resonators. Fig. E.7 shows transmission and reflection measurements for 1 cm and 1 mm long resonators (longer resonators were interrupted). The recipe for this fabrication run differs from the final recipe (Table E.2) and is presented in Table E.4. To be more specific, it used liftoff for the grAl central conductor before the particle contamination problem was properly investigated.

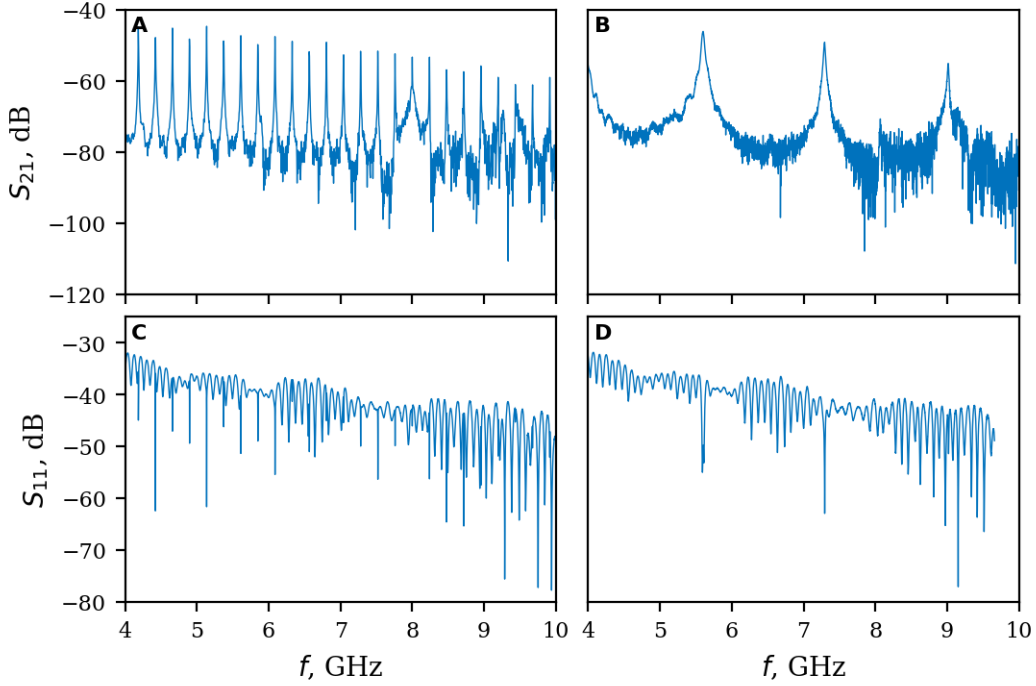


Figure E.7: Transmission measurements for 1 cm (A) and 1 mm (B) resonators at -45 dBm input power. Reflection measurements for 1 cm (C) and 1 mm (D) resonators at -35 dBm input power.

Due to a too strong coupling, it was not possible to extract the internal quality factor of the resonators. Loaded and coupling quality factors of ports 1 and 2 are presented in Fig. E.8.

Large ripples in reflection (arising from a degraded Eccosorb filter) forbade the circle fit procedure from converging for a 1 mm long resonator S_{11} data. From here on, all the calculations and plots are presented only for a 1 cm long resonator.

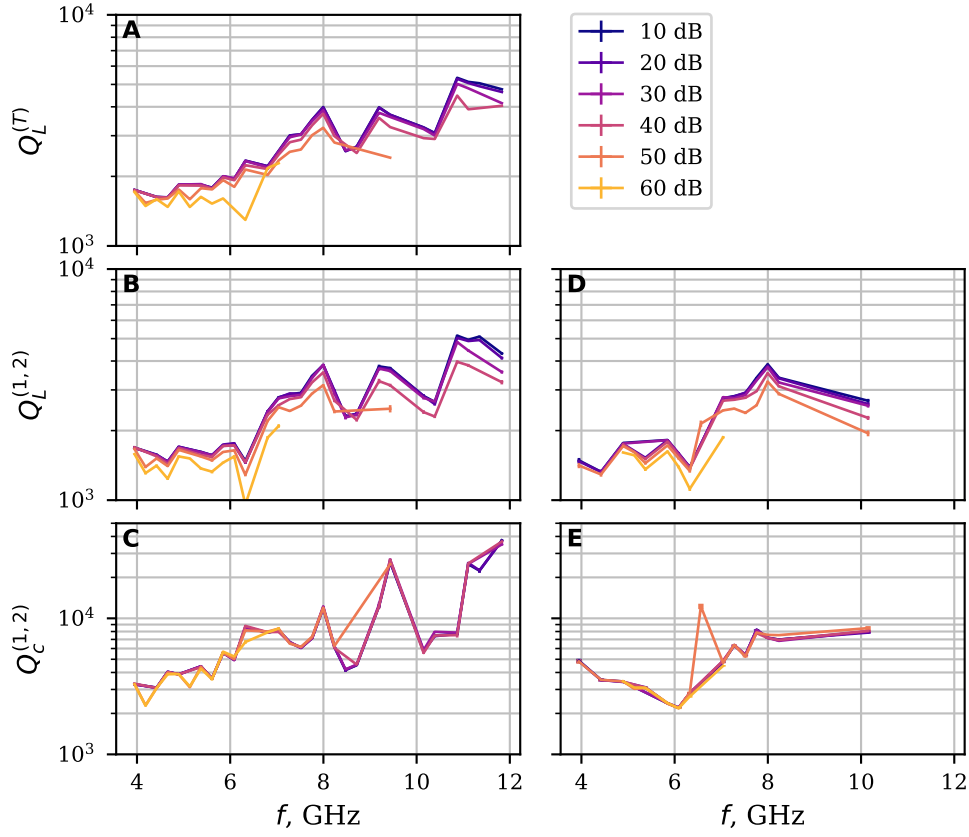


Figure E.8: Quality factors, extracted from a circle fit, for a 1 cm long resonator for different attenuation. The plots include only successful fit results. The VNA power was set to -5 dBm. **A** Loaded quality factors, obtained from transmission measurements. **B** and **D**: loaded quality factors, obtained from reflection measurements from ports 1 and 2. **C** and **E**: coupling quality factors, obtained from reflection measurements from ports 1 and 2.

E.5 Kerr measurements

Granular Aluminium consists of superconducting Aluminium islands, separated by oxide barriers, and, therefore, can be interpreted as a JJ chain [103]. According to [104], the k -th linear mode frequency ω_k of the JJ chain with N junctions undergoes a shift

$$\omega'_k = \omega_k - \sum_{k'} K_{kk'}/2 \quad (\text{E.1})$$

$$K_{kk'} = \left(\frac{1}{2} + \frac{\delta_{kk'}}{8} \right) \frac{\hbar^2 \omega_k \omega_{k'}}{2NE_J} \quad (\text{E.2})$$

where $K_{kk'}$ is a cross-Kerr between the modes k and k' with frequencies ω_k and $\omega_{k'}$, and E_J is the island Josephson junction energy. Therefore, tracking the mode k frequency shift by applying a tone on resonance with the mode k (k') allows for measuring self-Kerr (cross-Kerr) of the mode.

At low power, the frequency shift scales linearly with the number of photons, allowing for the determination of self-Kerr coefficients from linear fits of the low-power data. At high power, the behaviour becomes nonlinear, and the weak nonlinearity theory is no longer applicable. The resonances become bistable, which results in a non-Lorentzian shape, as shown in Fig. E.9. The mean photon number for the mode k is calculated via the relation [105]

$$n_k = \frac{4Q_L^2}{\hbar\omega_k Q_C} P_{\text{in}}, \quad (\text{E.3})$$

where ω_k is the mode frequency, P_{in} is an input power, Q_L and Q_C are the mode loaded and coupling quality factors, calculated in Section E.4. The fridge attenuation is about -83 dB ([106, p. 49]).

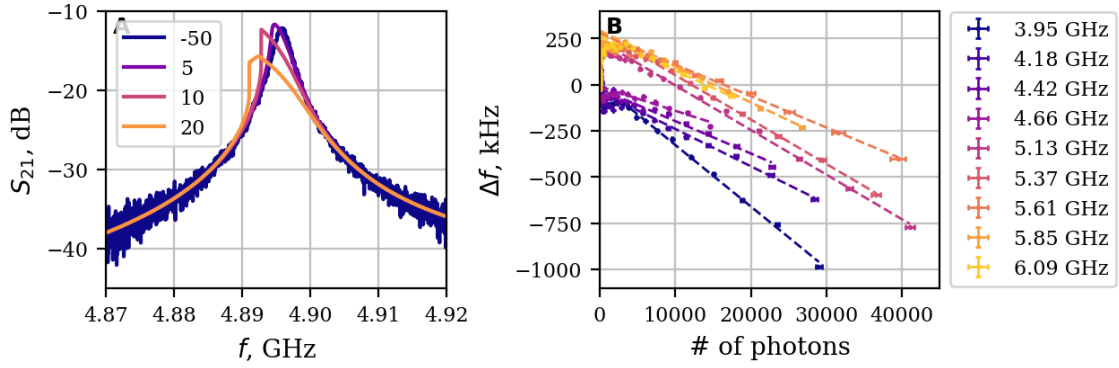


Figure E.9: **A** Resonance behaviour of the mode #21 at different applied power (in dBm). Due to self-Kerr, the mode undergoes a frequency shift with applied power. This mode is a "probe" mode for cross-Kerr and two-tone measurements in the following experiments. **B** Frequency shift of the mode # 21 due to cross-Kerr. Cross-Kerr coefficients are obtained from a linear fit (dashed lines) of the data (dots).

Cross-Kerr measurements between the probe and target modes are based on a modified two-tone spectroscopy [104], which consists of the following steps:

- find frequencies for the probe mode k (mode #21) and target mode k' as a function of applied power to find self-Kerr coefficients $K_{k,k}$ and $K_{k',k'}$
- pump the resonator at target and probe frequencies at the same time and make a sweep over the power of the target signal. The probe tone frequency and power should correspond to the bare, unshifted probe resonance found at the previous step. Track the frequency shift of the probe mode with power. The target signal frequency for each power should be determined from the measurements taken in the previous step to eliminate the self-Kerr contribution to the frequency shift.

Cross-Kerr coefficients are derived from linear fits of low-power data (see the data for the mode #21 in Fig. E.9 B) and summarised in Table E.3.

#	f , GHz	cross-Kerr $K_{i,21}$, Hz/ph	self-Kerr $K_{i,i}$, Hz/ph
17	3.95	6.67 ± 0.12	15.05 ± 0.45
18	4.18	4.05 ± 0.10	9.71 ± 0.31
19	4.42	3.51 ± 0.17	8.78 ± 0.36
20	4.66	2.63 ± 0.25	5.27 ± 0.35
21	4.89		9.44 ± 0.37
22	5.13	4.8 ± 0.10	15.72 ± 0.33
23	5.37	4.81 ± 0.07	15.46 ± 0.38
24	5.61	3.41 ± 0.07	10.66 ± 0.32
25	5.85	3.89 ± 0.11	10.64 ± 0.33
26	6.09	3.76 ± 0.24	11.34 ± 0.58

Table E.3: Self-Kerr for modes #17 - 26 and cross-Kerr between modes #17 - 26 and mode #21.

E.6 Dispersion relation and kinetic inductance

One of the easiest and most robust ways to extract kinetic inductance is to extract it from two-tone spectroscopy (see Fig. E.10), which relies on the cross-Kerr shift of the resonance frequency. The first tone is fixed at the bare probe mode (# 21) frequency, and the second tone frequency is swept within the desired range. Once the second tone frequency coincides with the resonator mode, the frequency of the probe tone shifts due to cross-Kerr effect, resulting in reduced transmission.

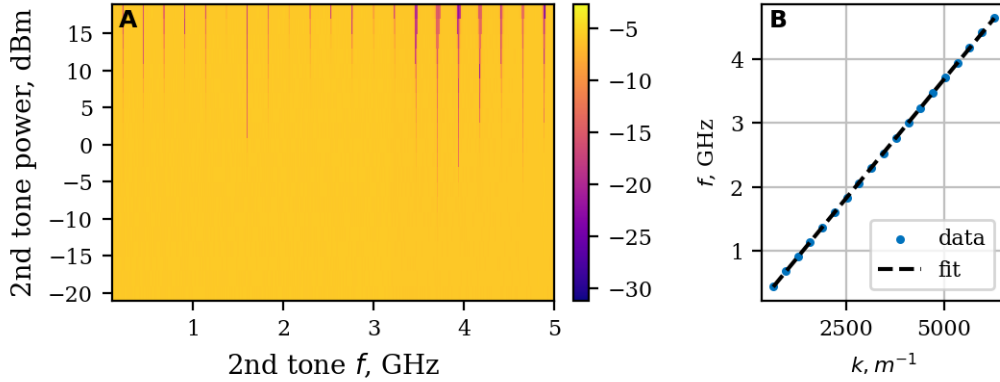


Figure E.10: **A** Two-tone spectroscopy of the modes below the cut-off of the measurement setup. **B** Dispersion relation. Bare mode frequencies are extracted from the two-tone spectroscopy **A**.

Each resonance dip, induced by cross-Kerr, was measured with higher resolution in a low-power regime to determine the bare frequency. The dispersion relation was extracted from the linear fit of bare mode frequencies [107]. This gives the speed of light and the kinetic inductance:

$$v_{\text{ph}} = \frac{\omega_n}{k_n} = \frac{1}{\sqrt{C_l(L_l^m + L_l^{\text{kin}})}}$$

$$\frac{v_{\text{ph}}}{c} = (155.73 \pm 0.22)10^{-4}$$

$$L_{\text{kin}} = 238.9 \pm 0.7 \text{ pH/sq}$$

Therefore, this set of fabrication parameters yields a delay of approximately 2 ns/cm and an impedance of $Z = 1773 \Omega$.

³positive resist

Al ground plane and Nb alignment marks (Plassys)	
descum + rot	descum: Ar/O ₂ 10/5, 200 V, 10 mA, 2 min
Ti gettering	0.2 nm/s, 2 min
Al evaporation, Plassys	20 nm
resist spinning	EL13 at 1500 rpm
resist baking	5 min on hotplate at 200°C
resist spinning	A4 at 2000 rpm
resist baking	5 min on hotplate at 200°C
e-Line, 30 kV, 30 μm aperture	BD dose: 300 $\mu C/cm^2$
development	IPA:H ₂ O 3:1 at 6°C for 2 min 15 s
Nb evaporation, Plassys	descum: Ar/O ₂ 10/5, 200 V, 10 mA, 2 min Ti gettering 0.2 nm/s, 2 min Nb 100 nm
liftoff	acetone + IPA, 40°C, US 130 kHz 50%
CPW gaps	
resist spinning	CSAR ³ at 2000 rpm
resist baking	1 min on hotplate at 150°C dose: 120 $\mu C/cm^2$
e-Line, 30 kV, 60 μm aperture	FBMS step 0.1504 for 2 μm speed 0.512585 mm/s
development	BD dose: 120 $\mu C/cm^2$, dwell time 0.000249 AR-P 600-546 for 1 min
resist baking	H ₂ O US 35% 5 min, IPA US 30% 10 min for 1 min 130°C
plasma etching, Sentech	BCl ₃ /Ar 30/3 sccm, RF 40W, ICP 200W, 8 min 30 s vacuum grease for thermalisation to avoid resist burning
resist and vacuum grease removal	NMP (5 min)×2, acetone (5 min)×2 at 60°C, transfer to IPA
resist removal, Sentech	US bath at 35 kHz, 100 % O ₂ 60 sccm, 10 Pa, RF 50 W, ICP 50 W, 5 min
grAl central conductor	
resist spinning	EL13 at 1500 rpm
resist baking	5 min on hotplate at 200°C
resist spinning	A4 at 2000 rpm
resist baking	5 min on hotplate at 200°C dose: 1000 $\mu C/cm^2$
e-Line, 30 kV, 30 μm aperture	FBMS step 0.0496 for 0.2 μm speed: 0.1769 mm/s
development	BD dose: 300 $\mu C/cm^2$, dwell time: 0.00217
descum + rot	IPA:H ₂ O 3:1, 1.5 min O ₂ /Ar plasma (5/10 sccm) V = 200 V, I = 10 mA, V _{acc} = 50 V
Ti gettering	0.2 nm/s, 2 min
grAl evaporation, Plassys	0.5 nm/s 2.2 sccm
grAl target cleaning	
liftoff	NMP 80°C, 10 min

Table E.4: The version of the fabrication recipe used for CPW grAl resonators measured in sections appendices E.4 to E.6.

Acknowledgements

Every research project is built upon the results of previous generations, but it is also shaped by the environment in which we work: the lab atmosphere, on- and off-topic discussions, colleagues' support, and the tools and techniques available. Although the lab rules and atmosphere are influenced by the people working at a given time, they are built and continuously developed by the group leader. Building a community of open-minded, curious people is a difficult task, but in my opinion, it was brilliantly solved by my supervisor Prof. **Gerhard Kirchmair**. Your enthusiasm, passion for science, and willingness to dive deeply into the unknown motivate everyone in the group to push the boundaries of knowledge further and further. You are always ready to help with any experimental or theoretical challenge we encounter, and I deeply appreciate that. I am grateful for your positive attitude towards any wild ideas that come to my mind, for the countless scientific discussions, and for the encouraging atmosphere you create.

Towards the end of my PhD, I was fortunate to meet Prof. **Samuele Tosatto**, with whom we worked on applying ML techniques to the state reconstruction. Without you, the reconstruction would have never come close to the benchmarks we were aiming for, if it had been completed at all. I appreciate your availability to discuss and help with any ML-related issues I encountered, even though it was a side project on top of your already busy schedule. Through our discussions, I learned a lot beyond the usual physicist's toolbox, and I truly enjoyed working with you.

While working on the cat optimisation project, I spent most of my time with the **high Q cavity team**. A special gratitude goes to **Ian Yang**, who introduced me to the world of bosonic qubit experiments, kindly handed over the experimental setup from the "hot cats" project and introduced me to one of his countless side projects. Thanks to you, I was able to transition into a completely different research topic smoothly and finally wrap up the project in a reasonable time. **Desislava Atanasova**, I really enjoyed our scientific and non-scientific discussions together. Even though our projects barely overlapped, you were always ready to help and, which is no less important, exchange complaints about the hardware. For me, the friendship we built over the years of the PhD is the most important outcome of it, and I hope we will keep the connection in the future, no matter how far we are from each other. **Lucien Québaud**, you have a heavy burden on your shoulders: you are the one and only who can graduate on time from our team. All the team members believe in you! You've already made a good progress, and I am excited to see your future results.

Romain Albert, you were the one who extended a helping hand when the fabrication of grAl delay lines reached a dead end. I am deeply grateful for your support, guidance, and the countless hours you spent with me debugging and improving the process in the cleanroom. Even though you were not officially assigned to my project, you continued to help and supervise me all the way until the end of your postdoc in Innsbruck. You broadened my horizons not only in fabrication, but also in many areas of physics and

programming. Your breadth and depth of knowledge across so many fields have always impressed me.

Of course, without the support of the rest of the Kirchmair team, it would have been a completely different experience. I want to thank **Christian Schlager** for having exciting times together in the cleanroom and drastic support outside of it, **Teresa Hönigl-Decrinis** for helping with my transition to e-beam lithography and e-beam recipes, and **Julian Daser** for assisting with etching and for ice-cream sessions. The members of the Mechanics team, **Lukas Deeg**, **Christian Dejaco**, **Bhargava Thyagarajan**, **Raamamurthy Sathyanarayanan**, our newcomer in the TruePA team **Tanmoy Bera**, and our already former team member **Philipp Straub**: even though we had quite diverse research directions, it was always a lot of fun to chat with you about physics and life, play board games and even win a second place in a Schrödinger cup. **Christian Dejaco**, I am particularly grateful for the Vorarlberg cultural experience! Of course, my deep gratitude goes to the former group members who built the lab, established fabrication processes from scratch, and wrote thousands of lines of code: **Christian Schneider**, **Maximillian Zanner**, **David Zöpfel**, **Oscar Gargiulo**, **Mathieu Juan**, **Aleksei Sharafiev**, **Phani Muppalla** and **Stefan Oleschko**.

The closed-loop optimisation project grew out of a collaboration that extended far beyond IQOQI. **Phila Rembold** and **Marco Rossignolo**, thanks a lot for introducing me to the world of optimal control, for patiently waiting for experimental results, and for running open-loop optimisations. **Phila**, thank you so much for always keeping track of all my unpredictable changes in the reconstruction and closed-loop experiments, and for actively and consistently engaging in discussions throughout the project.

Smooth research progress is possible only with reliable administrative support. I am grateful to the **administrative teams** of IQOQI and the Institute of Experimental Physics at the University of Innsbruck for their openness and readiness to help with any request or issue. My special thanks go to **Elisabeth Huck**, **Klaus Falschlunger**, **Claudia Mevenkamp**, and **Elke Wölflmaier**, with whom I interacted most frequently. And, naturally, many thanks to the IQOQI **mechanical workshop team**, who machined any parts we needed, from sample holders and mounts to microwave cavities.

My beta-readers **Desislava Atanasova**, **Phila Rembold**, **Ian Yang**, **Samuele Tosatto**, and **Christian Schlager**: thanks a lot for your constructive feedback!

Без поддержки моей семьи мой путь в аспирантуре сложился бы совершенно иначе. Мои мама, папа, сестренка, бабушки и дедушка: огромное спасибо за поддержку во всех моих начинаниях и увлечениях, понимание и безоговорочную веру в меня. Что бы ни случилось, я всегда и во всём могу на вас положиться, и, благодаря этому, я могу уверенно смотреть в будущее.