Machine Learning assisted Superconducting Qubit Readout

by

Benjamin Lienhard

B.S. Elektrotechnik & Informationstechnologie, ETH Zürich (2013)
S.M. Elektrotechnik & Informationstechnologie, ETH Zürich (2015)

Submitted to the
Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2021

© Massachusetts Institute of Technology 2021. All rights reserved.

Author . . . .
Department of Electrical Engineering and Computer Science
August 2, 2021

Certified by . . . .
William D. Oliver
Professor of Electrical Engineering and Computer Science
Professor of Physics
Thesis Supervisor

Accepted by . . . .
Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

Machine Learning assisted Superconducting Qubit Readout

by

Benjamin Lienhard

Submitted to the Department of Electrical Engineering and Computer Science
on August 2, 2021, in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Electrical Engineering and Computer Science

Abstract

Quantum computers hold the promise to solve specific problems significantly faster than classical computers. However, to realize a practical quantum computer, the quantum processor's constituent components, their control, and their readout must be very well-calibrated. Over the last few decades, infrastructure and protocols have been developed to operate small-scale quantum processors efficiently. However, the operation of medium- to large-scale quantum processors presents new engineering challenges. Among those challenges are efficient and high-fidelity multi-qubit control and readout. In particular, qubit-state readout is a significant error source in contemporary superconducting quantum processors. The fidelity of dispersive qubit-state readout depends on the readout pulse shape and frequency as well as the resulting qubit-state discriminator. For a single qubit, fast and high-fidelity readout is achieved with minor changes to the rising and falling edge of a rectangular microwave pulse and a linear matched filter discriminator. However, in resource-efficient, frequency-multiplexed readout of multiple qubits, optimizing the readout pulse shape and discriminator becomes a computationally intensive task.

In this thesis, control and readout hardware and software tools for multiple superconducting qubits are developed. First, I discuss the principles to engineer microwave packages for multiple qubits. I designed and engineered a novel multi-qubit package to enable efficient qubit control and readout and minimize errors due to interactions between the quantum processor and its immediate environment. Second, I demonstrate deep machine learning techniques to improve frequency-multiplexed superconducting qubit readout pulse shapes and discrimination for a five-qubit system. Compared with currently employed readout methods, these novel techniques reduce the required measurement time, the readout resonator reset, and the discrimination error rate by about 20% each. The developed readout techniques are a significant step towards efficient implementations of near-term quantum algorithms based on iterative optimization and quantum error correction protocols necessary for future universal quantum processors.

Thesis Supervisor: William D. Oliver
Title: Professor of Electrical Engineering and Computer Science
Professor of Physics

Acknowledgments

First and foremost, I want to thank my research advisors, William D. Oliver, Terry P. Orlando, and Simon Gustavsson, for their invaluable advice, continuous support, and patience. Will's insightful feedback and attention to detail pushed me to sharpen my thinking and communication skills. Terry's passionate mentorship and guidance profoundly impacted me on a research and personal level. Simon's exceptional experimental wisdom and understanding inspired me to broaden my horizon and planted the aspiration to become familiar with every aspect of the experimental setup.

In addition to my research advisors, I would like to show gratitude to my thesis committee members Paola Cappellaro and Tomaso Poggio, and my academic advisor Karl Berggren. All my advisors' immense knowledge and plentiful experience have inspired and encouraged me in my academic research and daily life.

I wish to express my gratitude to Antti Vepsäläinen, Jochen Braumüller, Luke C. G. Govia, Kevin O'Brien, Cole Hoffer, and Sihao Huang. They were instrumental to the projects underlying this thesis. Moreover, I want to thank my collaborators from the Quantum Engineering and Computing Group, Raytheon BBN Technologies, and in particular, Hari Krovi and Tom Ohki.

I am grateful to all of those I have had the pleasure of working with during my PhD: EQuS postdocs Daniel Campbell, Agustin Di Paolo, Patrick Harrington, Morten Kjaergaard, Philip Krantz, Miuko Tanaka, Joel I. J. Wang, Roni Winik, and Fei Yan; fellow EQuS graduate students Aziza Almanakly, Junyoung An, William Banner, Charlotte Bøttcher, Leon Ding, Qi (Andy) Ding, Ami Greene, Bharath Kannan, Amir Karamlou, Rebecca Qing Li, Chris McNally, Tim Menke, Sarah Muschinske, Yanjie (Jack) Qiu, David Rower, Gabriel O. Samach, and Youngkyu Sung; fellow QPL members Eric Bersin, Darius Bunandar, Edward H. Chen, Kevin Chen, Dirk Englund, Gabriele Grosso, Nicholas C. Harris, Sinan Karaveli, Catherine Lee, Jeff Tsung-Ju Lu, Hyowon Moon, Sara Mouradian, Christopher Panuski, Mihika Prabhu, Cheng Peng, Tim Schröder, Michael Walsh, and Jiabao Zheng. I want to highlight my office mates Jack Qiu and Tim Menke and colleagues Bharath Kannan, Amir Karamlou, Youngkyu Sung, and Francisca Vasconcelos. Moreover, I am grateful for the support of the entire MIT Lincoln Laboratory team, particularly the team responsible for sample fabrication surrounding David Kim and packaging surrounding Wayne Woods and Greg Calusine.

I want to express my appreciation for MIT's staff team, especially Mirabella Pulido and Chihiro Watanabe, for their administrative assistance.

I want to acknowledge the funding agencies and grants that made my research possible: DARPA Polyplexus grant No. HR00112010001; U.S. Army Research Office (ARO) Multidisciplinary University Research Initiative (MURI) W911NF-18-1-0218; Department of Defense via Lincoln Laboratory under Air Force Contract No. FA8721-05-C-0002.

Most importantly, I want to express my deepest appreciation for my family and friends. I wish to thank Darius Bunandar, Matthias Oberli, and Francis Wolenski for their support and advice throughout. I want to express my deepest gratitude to my parents, grandparents, parents-in-law, siblings, and in particular, my wife and son for their continuous support, encouragement, and understanding.

For my Family, Leandro, and Wen

Contents

1 **Introduction** **47**

1.1 Quantum Computing 48

1.2 Superconducting Quantum Computing 49

1.2.1 The Transmon Qubit 49

1.2.2 Superconducting Qubit Readout 51

1.2.3 Measurement Setup and Microwave Control 51

1.3 Package Design for Superconducting Qubits 53

1.3.1 Microwave Package Materials & Geometry 54

1.3.2 Control and Readout Signals 54

1.3.3 Package Modes 55

1.4 Machine Learning 57

1.4.1 Supervised Learning 58

1.4.2 Reinforcement Learning 59

1.5 Superconducting Qubit-State Discrimination using Supervised Learning 59

1.5.1 Superconducting Qubit-State Discriminators 62

1.5.2 Experimental Comparison of Qubit-State Discriminators . . . 63

1.6 Superconducting Qubit-Readout-Pulse Shaping using Deep Reinforcement Learning 64

1.7 Summary & Outline 66

References 69

2 Quantum Computing **77**

2.1 Introduction Into Quantum Computing . . . 78

2.1.1 Qubits . . . 78

2.1.2 Qubit Gates . . . 79

2.1.3 Quantum Parallelism & Interference . . . 82

2.2 Complexity Theory . . . 83

2.2.1 Classical Complexity Theory . . . 83

2.2.2 Quantum Complexity Theory . . . 85

2.3 Computing History . . . 87

2.4 Quantum Computing Challenges . . . 89

2.4.1 Qubit Modalities . . . 90

2.4.2 Qubit Coherence . . . 92

2.4.3 Gate Time and Fidelity . . . 93

2.4.4 Quantum Error Correction . . . 95

2.5 Summary . . . 96

References . . . 98

3 Superconducting Quantum Computing **105**

3.1 Superconducting Qubits . . . 106

3.1.1 From Linear *LC* Circuits to Quantum Harmonic Oscillators . . 106

3.1.2 Josephson Junctions . . . 108

3.1.3 The Transmon Qubit . . . 110

3.1.4 Sample Design and Fabrication Techniques . . . 112

3.1.5 Microwave Regime . . . 113

3.2 Superconducting Qubit Control and Readout . . . 113

3.2.1 Superconducting Qubit Gates . . . 113

3.2.2 Superconducting Qubit Readout . . . 114

3.2.3 Circuit Quantum Electrodynamics . . . 114

3.3 Experimental Infrastructure . . . 116

3.3.1 Microwave Package Engineering . . . 118

3.4 Summary 120
References 121

4 Package Design for Superconducting Qubits 125

4.1 Microwave Packaging 126
4.2 Microwave Package Materials & Geometry 128
4.3 Control and Readout Signals 133
4.3.1 Interposer Design 135
4.3.2 Wirebonding Considerations 137
4.4 Package Modes 142
4.4.1 Box Modes 143
4.4.2 Interposer-Based Modes 144
4.4.3 Chip-Based Modes 145
4.4.4 Mode Characterization 145
4.4.5 Measurement Setup 147
4.4.6 Chip Design 148
4.4.7 Spatial Mode Characterization 150
4.5 Summary & Conclusion 154
References 158

5 Machine Learning 167

5.1 Machine Learning Categories 168
5.1.1 Unsupervised Learning 168
5.1.2 Supervised Learning 168
5.1.3 Reinforcement Learning 169
5.1.4 Supervised Learning for Qubit-State Discrimination 169
5.2 Matched Filter 169
5.3 Support Vector Machine 171
5.3.1 Multi-Class Support Vector Machine 173
5.4 Neural Networks 174
5.4.1 Activation Functions 175

5.4.2 Neural Network Training 176
5.4.3 Neural Network Archetypes 177
5.5 Reinforcement Learning 178
5.5.1 Markov Decision Processes 179
5.5.2 Value Iteration 180
5.5.3 Policy Iteration 181
5.5.4 Exploration and Exploitation 181
5.6 Deep Reinforcement Learning Algorithms 183
5.6.1 AlphaZero 183
5.6.2 Deep-Q Network 184
5.6.3 Deep Deterministic Policy Gradient 184
5.6.4 Proximal Policy Optimization 185
5.7 Summary 186
References 188

6 Superconducting Qubit-State Discrimination using Supervised Learning 191
6.1 Superconducting Qubit-State Discrimination 193
6.2 Superconducting Qubit-State Discriminators 195
6.2.1 Boxcar Filter 196
6.2.2 Matched Filter (MF) Threshold Discriminator 196
6.2.3 Support Vector Machine (SVM) 199
6.2.4 Neural Networks 202
6.2.5 Discriminator Comparison 205
6.3 Experimental Comparison of Qubit-State Discriminators 209
6.3.1 Confusion Matrix 214
6.3.2 Training Effort 218
6.4 Conclusion 219
References 221

7 Superconducting Qubit-Readout-Pulse Shaping using Deep Reinforcement Learning 229

- 7.1 Superconducting Qubit-Readout-Pulse Shaping 230
- 7.2 Superconducting Qubit-Readout-Pulse Shapes 232
  - 7.2.1 Rectangular Qubit-Readout Pulse 232
  - 7.2.2 Cavity Level Excitation and Reset Qubit-Readout Pulse 233
  - 7.2.3 Reinforcement-Learning-Generated Qubit-Readout Pulse . . . 235
- 7.3 Qubit-Readout-Pulse Shaping Simulations 236
  - 7.3.1 Single-Qubit Simulation 239
  - 7.3.2 Multi-Qubit Simulation 241
- 7.4 Conclusion 243
- References 245

8 Summary & Outlook 249

- 8.1 Package Design for Superconducting Qubit 250
- 8.2 Machine Learning assisted Superconducting Qubit Readout 251
  - 8.2.1 Superconducting Qubit-State Discrimination using Supervised Learning 251
  - 8.2.2 Superconducting Qubit-Readout-Pulse Shaping using Deep Reinforcement Learning 252
- 8.3 Global Calibration Module 253
- References 255

A Cavity Bloch Equation Formulation 259

- A.1 Dispersive Jaynes-Cummings Hamiltonian 259
- A.2 Master Equation 260
- A.3 Cavity Bloch Equation 261
- A.4 Frequency-Multiplexed Readout 262
- References 263

B Experimental Infrastructure **265**
B.1 Measurement Setup 265
B.2 Superconducting Chip with Five Transmon Qubits 266
References 268

C Alternative Matched Filter Derivation **269**
C.1 Derivation of Matched Filter 269
C.2 Alternative Derivation of Matched Filter 271
C.3 Matched Filter for Qubit-State Discrimination 272
References 273

List of Figures

1-1 Quantum Harmonic and Anharmonic Oscillator. (a) A simple harmonic resonator formed by a capacitor C and an inductor L is shown in blue. The potential energy is the energy stored in the inductor and assumes a parabolic shape concerning the phase variable ϕ, related to the flux induced by the inductor. The energy levels are equidistant, such that transitions cannot be addressed individually. By replacing the inductor with a Josephson junction, which has a non-linear inductance L_J, the potential becomes anharmonic (shown in red), which isolates two energy levels, $|g\rangle$ and $|e\rangle$, forming the computational basis and hence a superconducting quantum bit. (b) The transmission spectrum of a resonator dispersively coupled to a qubit showing the characteristic qubit-state-dependent dispersive shift of about $1\,\mathrm{MHz}$. 50

1-2 Measurement Setup. (a) Qubit measurement and control setup. Control and readout pulses generated by an arbitrary waveform generator (AWG) and up-converted to gigahertz frequencies using a local oscillator (LO) are sent to the qubit chip through attenuated signal lines. The readout signal upon interaction with the qubit is amplified by a Josephson traveling-wave parametric amplifier (JTWPA), a high-electron-mobility transistor (HEMT) amplifier, and a room-temperature amplifier before being down-converted to megahertz frequencies, in-phase I^{IF} and quadrature Q^{IF} component at intermediate frequencies (IF), and digitized. (b) The chip shown in (a) consists of five superconducting transmon qubits between $4.3\,\mathrm{GHz}$ and $5.2\,\mathrm{GHz}$ frequency-tunable via a global bias. The qubits are capacitively coupled to individual quarter-wave readout resonators at about $7\,\mathrm{GHz}$ that couple inductively to a bandpass (Purcell) filtered feedline line at the top of the chip. 52

1-3 Microwave Packages. (a) The purpose of a microwave package is to shield the enclosed qubit chip from external radiation (purple oscillating arrow) and stray magnetic fields (top; red lines) while providing impedance-matched (left; transmitted green pulse and reflected blue pulse at the input and output), low crosstalk communication channels (center; crosstalk in green at input), and a thermal link to the dilution refrigerator. (b) The presented microwave package consists of a metal enclosing, microwave connectors, an interposer for signal fan out, and a microwave cavity in the center surrounding the quantum chip. 55

1-4 Measurement Data Processing and Multi-Qubit-State Discrimination. (a) Single-qubit matched filter (MF) with subsequent discriminator threshold optimization, (b) Single-qubit linear support vector machine (SQ-LSVM), (c) multi-qubit LSVM (MQ-LSVM), and (d) feedforward neural network (FNN). The MF, as well as the LSVMs assignment fidelity, is maximized if the intermediate frequency (IF) input signal is digitally demodulated down to DC: $I^{\mathrm{IF}} \rightarrow I^{\mathrm{DC}}, Q^{\mathrm{IF}} \rightarrow Q^{\mathrm{DC}}$. During the discriminator training, the single-qubit discriminators (MF, SQ-LSVM) require a relabelling step of the discriminators' training data. The MF tune-up is optimized when each discriminator utilizes training data that contains a single excitation. The MQ-LSVM requires the DC data for each of the five qubits to be stacked and combined into a single data block. The FNN does not require any preprocessing such as digital demodulation, data stacking, label filtering, or data filtering as the complexity of it all is embedded in the neural network itself. 61

1-5 Time-Dependent Qubit-State Assignment Fidelity. Five-qubit geometric mean qubit-state assignment fidelity $\mathcal{F}_{\mathrm{GM}} = (\mathcal{F}_1\mathcal{F}_2\mathcal{F}_3\mathcal{F}_4\mathcal{F}_5)^{1/5}$ versus measurement time for the matched filter (MF), single- and multi-qubit linear SVM (SQ-LSVM and MQ-LSVM), as well as the feedforward neural network (FNN). 63

1-6 Simulation of Readout Pulse Shape versus Measurement Time for a 5-Qubit Chip. (a) Comparison of readout pulse shapes for a single qubit. The typically employed pulse shape, a rectangular pulse shape (gray, dashed), is compared with the cavity level excitation and reset (CLEAR) (in blue) and a pulse shape tune-up using the proximal policy optimization (PPO) algorithm (in red), a deep reinforcement learning algorithm. (b) Comparison of the duration to simultaneously inject photons in 5 different readout resonators, stabilize the photon counts during a waiting period, and return the resonators close to their vacuum states. The waiting period is kept constant for all three methods. In comparison with the rectangular pulse (CLEAR), PPO reduces the injection window by 58% (22%) and reset duration by 46% (17%). 67

1-7 Thesis Overview. Displayed is a microwave package with an enclosed qubit coupled via a resonator to a transmission line. The microwave package defines the immediate qubit environment and serves as a heat sink. The green pulse on the input side illustrates a readout pulse to probe the resonator state. The transmitted signal, colored in orange, is measured and discriminated against. This thesis focus on three aspects towards efficient multi-qubit readout and control: I. The development of microwave packages, II. Machine-learning assisted qubit-state discrimination, and III. Machine-learning assisted qubit-readout-pulse shaping. 68

2-1 Bloch Sphere. The unit sphere depicts the pure states of a qubit. The north pole and south pole along the Z-axis represent the computational basis with $|0\rangle$ and $|1\rangle$. The Bloch vector, shown in red, represents an arbitrary quantum state $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$. α and β are complex probability amplitudes with the constraint $|\alpha|^2 + |\beta|^2 = 1$. Quantum operations can alter the qubit state and thus the Bloch vector. 79

2-2 Computational Complexity Classes. **PSPACE** contains all classes and problems which require a polynomial amount of memory on a conventional computer independent of the number of necessary computational steps to find or verify a solution. **P** describes the class of classically efficient computable problems. Problems part of the complexity class **BQP** are efficiently solvable by a quantum computer. Class **NP** contains problems that are efficiently verifiable. The hardest problems are still efficiently verifiable from complexity class **NP-Complete**. In 2019, Ran Raz and Avishay Tal identified the first problem outside the complexity class **NP** that a quantum computer can solve efficiently. 86

2-3 Physical Realization of Qubits. Shown are some of the most prominent physical quantum mechanical systems used for qubit implementations ranging from photons to electrons to topological states. A photonic qubit can, for example, be encoded as time bins, in the number of photons (#), or using its polarization states. Each quantum mechanical system and its qubit implementation have advantages and disadvantages. For instance, while Majorana qubits promise to be resilient to many noise sources, a physical qubit has not been realized yet. In this thesis, the focus is on superconducting qubits using superconducting loops. Although quantum computing is still in its infancy, the commercialization of quantum information processing hardware and software has grown from a few to many companies over the last decade. Together with trapped ions, superconducting qubits are among the most established and industrially pursued qubit modalities today. 91

3-1 Quantum Harmonic and Anharmonic Oscillator. (a) A simple harmonic resonator formed by a capacitor C and an inductor L is shown in blue. The potential energy is the energy stored in the inductor and assumes a parabolic shape with respect to the phase variable $\hat{\phi}$, related to the flux induced by the inductor. The energy levels are equidistant, such that transitions cannot be addressed individually. By replacing the inductor with a Josephson junction, the potential of the oscillator becomes anharmonic, illustrated in red, which isolates two energy levels to form a computational qubit basis. A Josephson junction is a nonlinear inductor physically realized inserting a about $1\,\mathrm{nm}$-thin insulating barrier (e.g., aluminum oxide AlOx) between two superconductors (e.g., aluminum Al). (b) A well-studied design to enable qubit readout is to off-resonantly, capacitively couple the qubit, an anharmonic oscillator, to a readout resonator, a harmonic oscillator. The readout resonator frequency can be probed to infer the qubit state. (c) and (d) The transmission spectrum of a resonator dispersively coupled to a qubit showing the characteristic qubit-state-dependent dispersive shift of about $1\,\mathrm{MHz}$. The amplitude of the measured signal is shown in (c) and the corresponding signal phase in (d). A qubit-state discriminator tries to distinguish the qubit-states through the evaluation of the readout signal amplitude or phase. . . 109

3-2 Qubit Control and Readout Electronics and Wiring. (a) Depicted is a schematic of a microwave setup, including pulse generation and processing. Microwave pulses are generated by microwave sources and arbitrary waveform generators (AWG). IQ-mixers facilitate phase-sensitive amplitude modulation. The control and readout lines are attenuated and spectrally filtered at various stages. The readout signal is first amplified by a Josephson traveling-wave parametric amplifier (JTWPA) after passing through an isolator. The JTWPA is equipped with a separate microwave source. Before the readout signal is further amplifed by a high-electron-mobility transistor (HEMT) amplifier at the $3\,\mathrm{K}$ stage, the signal passes through additional isolators and is bandpass filtered. At room temperature, the readout signal is amplified once more before it is down-converted, low-pass filtered, and digitized. 117

3-3 Dilution Refrigerator and Microwave Package. (a) Dilution refrigerator with multiple temperature stages holding the qubit chip enclosed in a microwave package. The microwave package interfaced with through microwave lines is mounted on a cold finger in the mixing chamber reaching a base temperature of approximately $10\,\mathrm{mK}$. (b) The next microwave package generation is highlighted with a blue rectangle and discussed in the subsequent Chapter 4. The package consists of a copper casing, a multilayer interposer to perform signal fanout, a shielding cavity in the package center, and 24 microwave connectors. (c) Previous microwave package generation with a gold-plated copper casing and six control lines. 119

4-1 Microwave packages. (a) The purpose of a microwave package is to shield the enclosed qubit chip from external radiation (purple oscillating arrow) and stray magnetic fields (rose lines). Meanwhile, impedance-matched (transmitted green pulse and reflected blue pulse at the input and output), minimal-crosstalk communication channels (crosstalk in green at input), and a thermal link to the dilution refrigerator need to be provided. (b) Small sample package with six control channels. (c) The individual components of the package shown in panel (b). Inserted between the gold-plated copper base and the lid is the main package part. Metallic waveguides for control of the quantum circuit are imprinted on a dielectric circuit board (interposer) and $50\,\Omega$-impedance matched. The interposer routes the coaxial control line signal to the quantum circuit via wirebonds. The connectors to the control elements are coaxial connectors with $50\,\Omega$ impedance enabling the transmission of microwave signals up to about $18\,\mathrm{GHz}$ to $26\,\mathrm{GHz}$. (d) Novel sample package with 24 control lines used to study and establish the reported microwave design principles. 127

4-2 Microwave Simulations of a small Sample Package shown in Figure 4-1(b). (a) Gold-plated copper package (without lid) with wire-bonded superconducting circuit chip. The green arrow indicates the microwave input port. (b) Comparison of the transmission spectra and conductivity loss with a solid pedestal ($\Delta = 0\,\mathrm{mm}$, colored in red), and a drilled out pedestal with four corner posts ($\Delta = 3.8\,\mathrm{mm}$, black). (c) Simulated transmission magnitude spectrum $|S_{21}|$ of the chip holding an interrupted transmission line resonator with a resonance frequency at about $7.7\,\mathrm{GHz}$ mounted inside the package with pedestal (red) and a drilled out pedestal (black). The package with the drilled-out pedestal is free of any undesired package modes in the band of interest. (d) E-field magnitude ($\mathrm{V\,m^{-1}}$) plots in the ZX- and XY-plane at $5.9\,\mathrm{GHz}$ (indicated with an arrow in (c)). The presence of the pedestal enhances the E-field magnitude in the chip and direct vicinity. (e) Simulated conductivity loss $1/Q_{\mathrm{cond}}$ due to the normal conducting ($\sigma = 4.5 \times 10^9\,\mathrm{S\,m^{-1}}$) package, extracted from a transmon qubit (approximated as linear resonator) in the center. The conductivity loss depends on the gap size Δ between the chip and the pedestal (the four corner posts remain in place). 129

4-3 Qubit Full-Wave EM Simulations. (a) Lithographically defined superconducting qubit using aluminum and aluminum oxide on a pristine silicon chip. Blue inset shows the superconducting qubit with the Josephson junction in red, aluminum in turquoise, and silicon in grey. (b) Illustrated are the qubit's electric (E-field, left) and magnetic (H-field, right) fields. The electric field is more prominent than the magnetic field 131

4-4 Interposer Design. (a) Interposer layout for a 24-line package fabricated out of a three-layer Rogers 4350™controlled impedance glass-reinforced ceramic laminate. Stripline-based waveguides with dense via shielding are utilized to reduce signal crosstalk. (b) Schematic diagram of the interposer stack configuration (not to scale). The board consists of two Rogers 4350™cores bonded by a layer of FEP film, indicated in purple. Through vias are used for grounding and shielding, whereas blind vias are utilized for signal routing to minimize parasitic resonances. (c) Picture of a $5\,\mathrm{mm}$ by $5\,\mathrm{mm}$ qubit chip mount with signal launches in the periphery. Wirebonds are used to provide signal connections and ground the device. (d) Single-ended TDR measurement of the SMP-to-interposer connector transition. The reference plane of the connector is located at $0\,\mathrm{ns}$. The wirebond launch is left open (no chip connected), resulting in a steep increase in impedance. (e) and (f) Measured nearest neighbor and second nearest signal crosstalk. A high-port count network analyzer using Keysight M9374A PXIe modules was used to obtain the full scattering matrix. All transmission parameters corresponding to signal crosstalk information are overlaid on the plots. The purple background fading out with increasing frequency indicates the decreasing relevance of modes as their frequency separation to the qubit transition frequency increases. Note that the separate group of traces with increased isolation, as highlighted in (e) and (f), correspond to crosstalk that is reduced at corners of the package, as indicated by the pairs of blue lines in (a). 138

4-5 Wirebond Design. (a) Diagram of a chip (gray) with four transmission lines wirebonded to an interposer (beige) surrounding it. The instigating line, indicated in purple, can lead to near-end (NEXT) and far-end (FEXT) crosstalk in the disturbed line, shaded red and blue. (b) Schematic representation of a wirebond interconnect and its lumped element model. The flares located on the ends of the interposer and chip's transmission lines, as highlighted by the dashed box in (a), correspond to the two tuning capacitors to the ground, while the wirebond forms the series inductance. Note that the wirebonds used to connect the signal lines are spread out in a V-shape to minimize mutual inductance. (c) Plot of maximum inductance compensation versus cutoff frequency for Butterworth and Chebyshev filter designs, with $1\,\mathrm{nH}$ roughly corresponding to $1\,\mathrm{mm}$ in wirebond length. (d) and (e) The effect of wirebond configuration on signal crosstalk. A simplified model with the exact spatial dimensions as the 24-pin design is employed to simulate the relative reduction of signal wirebond crosstalk with more grounding wirebonds. As the number of grounding wirebonds increases, the electric field strength between the chip ground and the package cavity rapidly decreases, resulting in the suppressed coupling between adjacent signal wirebonds. . . . 141

4-6 Package Mode Measurements and Simulation. (a) Package modes are probed at liquid nitrogen temperature using a multi-port vector network analyzer. The scattering matrix elements corresponding to transmission across the package cavity are measured via the ports marked green in (b). The results are overlaid (gray) and averaged (purple), with the relevant frequency range for the qubit indicated by the fading purple background shading. (b) EM simulations reveal eigenmodes at $11.1\,\mathrm{GHz}$ (orange) and $18.1\,\mathrm{GHz}$ (red) respectively, which correspond well with the measured peaks as marked by the correspondingly colored vertical lines in (a). 142

4-7 Setup to Identify Package Modes. (a) Measurement setup used to obtain the package mode profile. The control, readout, and probe signals are combined and sent down the dilution refrigerator. The readout signal is amplified by a Josephson traveling-wave parametric amplifier (JTWPA), a high-electron-mobility transistor (HEMT) amplifier, and an amplifier at room temperature before being down-converted and subsequently digitized. Note, the JTWPA can be bypassed if necessary. (b) Superconducting qubit chip with six fixed-frequency transmon qubits with individual readout resonators coupling to a common transmission line. In the following panels, the results of the qubit and resonator indicated with a yellow dashed rectangle are explicated. (c) The qubit has an average qubit lifetime of $T_1 \approx 121.4\,\mu\mathrm{s}$ and coherence time of $T_2^* \approx 53.2\,\mu\mathrm{s}$ measured in intervals across a period of 12 hours. 146

4-8 Experimental Package Mode Evaluation using four Qubits. The Ramsey spectra are taken simultaneously from four different qubits while a probe tone is injected through the central transmission line. Note the similarity between the spurious modes (indicated by yellow arrows) measured by qubits 1 and 3, as well as qubits 2 and 4. The orange lines indicate qubit-related features, which shift depending on the sensor qubit used. 149

4-9 Experimental Package Mode Characterization. (a) and (b) The microwave environment between $2\,\mathrm{GHz}$ and $20\,\mathrm{GHz}$ registered by the qubit is mapped out using the qubit itself as a sensor. A continuous-wave probe tone—added to the qubit readout and control signal—excites the qubit microwave environment. Probe tone-dependent indirect and direct effects on the qubit are recorded using Ramsey spectroscopy. In the time-domain (a) and frequency-domain (b) panels, the qubits, readout resonators, and the ground to third excited state transition are identifiable as shifts in the Ramsey frequency. Furthermore, four features between $11\,\mathrm{GHz}$ and $18\,\mathrm{GHz}$ can be identified. The following mode characterization procedure is exemplified on the fourth mode, the expected package cavity mode shown augmented again in (c) at $17.18\,\mathrm{GHz}$. (d) The linewidth of a mode $\kappa/(2\pi)$ is the full width at half maximum (FWHM) of the Fourier transformed frequency-dependent Ramsey scan, here $20\,\mathrm{MHz}$. (e) The Ramsey frequency change as the power of the probe tone—parked at the mode's resonance frequency of $17.189\,\mathrm{GHz}$—is varied. The power is measured at the signal generator. (f) The qubit dephasing resulting from the varying probe tone power is extrapolated by performing fits to T_2^* experiments. The presented analysis yields a coupling rate of $g/2\pi \approx 17.73\,\mathrm{MHz}$ for Mode IV. Note, the Josephson traveling-wave parametric amplifier (JTWPA) was bypassed for this experiment to prevent interference with the probe tone. 152

4-10 Next-Generation Packaging. (a) Blueprint for scaling up the package design to support NISQ devices. By dividing the cavity into subcells corresponding to $5\,\mathrm{mm}{\times}5\,\mathrm{mm}$ chip areas, each with the same number of signal launches, the crosstalk, and mode performance can be preserved while increasing the number of pins. Qubit control and readout are conducted via multi-coax Ardent connectors with eight channels each. (b) Fins or spring-loaded pins above the qubit device suppress spurious modes. 155

5-1 Matched Filter Threshold Discriminator. (a) A matched filter projects the input data to a single dimension. The projection is optimized in terms of signal-to-noise ratio. (b) The matched filter trace for a qubit measurement. (b) Windowed matched filter (MF) kernel using a boxcar filter (BC) to reduce the impact of qubit-state transitions. (c) Histogram of matched-filtered qubit-state measurements. The discrimination boundary is situated at the origin for a two-class system complying with the assumed noise model for the described linear matched filter. 171

5-2 Support Vector Machines (SVM). (a) The linear decision boundary can be mathematically described as $wx - b = 0$. The margin width—the distance between the two dashed lines is equal to $2/||w||$. Depending on which side of the discrimination boundary, the label is either 1 or -1. The Linear SVM. (b) Linear SVM and (c) Nonlinear SVM are used to discriminate the states of a single qubit. 172

5-3 Multi-Class Support Vector Machines (SVM). (a) A linear and (b) nonlinear SVM are used to discriminate the states of three qubits (eight states). (c) A SVM tasked to discriminate more than two classes can show a diminished discrimination accuracy due to ambiguous areas. For three classes, a plane should show three distinct areas. However, dividing a plane with two lines generally results in four areas, meaning one area is not assigned to a single class. 173

5-4 Neural Network Architecture. (a) Basic feedforward neural network architecture with an input layer, two hidden layers, and an output layer. (b) Neural network node values, apart from input layer nodes, are determined by calculating the product between the previous layer's node values x^{l-1} and their associated weights w plus a bias value b. An activation function further processes the resulting value. 174

5-5 Activation Functions. Activation functions $f(x)$ are shown in blue, and their respective derivatives $f'(x)$ are shown in orange. (a) Linear, (b) Rectified Linear Unit (ReLU), (c) Scaled Exponential Linear Unit (SELU), and (d) Sigmoid are common activation functions used in machine learning. 175

5-6 Markov Decision Process. Given a current state s_t and reward r_t at time t, an agent decided on a next action a_t. Based upon the action the state changes to s_{t+1} and a reward r_{t+1} can be associated with the new state. The cycle continues until a specific number of cycles or a particular reward is reached. To minimize computational effort, the cycle can also be stopped once the agent's progress has plateaued. 179

5-7 Proximal Policy Optimization Architecture. The algorithm is composed of two deep feedforward neural networks. One neural network (the policy network) generates a new action, whereas the second neural network (value network) tries to associate an expected reward to the proposed new action. 186

6-1 Measurement Setup and Chip. (a) Schematic of superconducting qubit control and readout. The control and readout pulses, generated by an arbitrary waveform generator (AWG) and up-converted to GHz frequencies using a local oscillator (LO), are sent through attenuated signal lines to the readout resonator on the five-qubit chip. The transmitted readout signal is amplified by a Josephson traveling-wave parametric amplifier (JTWPA), a high-electron-mobility transistor (HEMT), and a room-temperature amplifier. Subsequently, the signal is down-converted to MHz frequencies and digitized—in-phase $I^{\mathrm{IF}}[n]$ and quadrature $Q^{\mathrm{IF}}[n]$ sequences at intermediate frequencies (IF). Colored optical micrograph (b) and associated circuit schematic (c) comprising five superconducting transmon qubits. The qubit transition frequencies are tuned via a global flux bias. Each qubit is capacitively coupled to a quarter-wave readout resonator that couples inductively to a bandpass (Purcell) filtered feedline. (d) The resonator frequencies $\omega_{\mathrm{Res}}/2\pi$ are near $7\,\mathrm{GHz}$ with $\chi/\kappa_{\mathrm{eff}}$ ratios ranging from 0.12 to 0.19, where χ and κ_{eff} are respectively the dispersive shift and the effective resonator decay rate through the feedline. Table of the qubit lifetimes (T_1) and operating frequencies ($\omega_{\mathrm{Qubit}}/2\pi$). Qubit color indicate the qubit operating frequency: red (purple) $\rightarrow$ lowest (highest) operating frequency. 194

6-2 Readout Data Statistics. (a) Magnitude of the time-bin weights of the qubit-specific matched filter shapes derived using prepared ground and excited states. A rectangular window (RW) is applied to each matched filter kernel to reduce the impact of qubit-energy decays and maximize qubit-state-assignment fidelities. The resulting matched filter windows are shaded in gray. (b) Shown are the histograms of the qubit-state-readout single-shot traces after applying the optimized $1\,\mu\mathrm{s}$-long matched filter. The dashed lines represent the optimized thresholds with the states to the right attributed to the ground state and left to the excited state. Using bimodal Gaussian fit functions for the ground state (green) and trimodal Gaussian fit functions for the excited state (blue) provides insight into the underlying dynamics such as thermal excitation or qubit-energy decays (see Table 6.1). 197

6-3 Time-dependent Qubit-State-Assignment Fidelity. Matched filter discriminator for each qubit versus measurement time. The maximum assignment fidelity $\mathcal{F}_i(t_i)$ for each qubit i is reached after $t_1 = 1\,\mu\mathrm{s}$, $t_2 = 2\,\mu\mathrm{s}$, $t_3 = 0.5\,\mu\mathrm{s}$, $t_4 = 0.8\,\mu\mathrm{s}$, and $t_5 = 0.5\,\mu\mathrm{s}$. 198

6-4 Support Vector Machine. Plotted are boxcar filtered single-shot traces of ground (black) and excited states (gray) in the $\mathcal{IQ}$-plane. A linear support vector machine trained on the two-dimensional data generates the qubit-specific colored discrimination boundary. 201

6-5 Architecture and Training of Fully-Connected Feedforward Neural Network (FNN). (a) The FNN architecture used here comprises an input layer, three hidden layers, and an output layer. For a $1\,\mu$s-long measurement time, the input layer consists of 1,000 nodes. 1,000, 500, and 250 nodes form the first, second, and third hidden layer (the shown number of nodes per layer does not reflect the reported implementation). The output layer scales as 2^N (N denotes the number of qubits). For five qubits, the output layer encompasses 32 nodes. The inset at the bottom shows a node as a part of the hidden layer l. The node's function depends on the following parameter inputs: the output values x_n^{l-1} of the prior layer $l-1$ and a node-specific bias b. The output value x_m^l of node m corresponds to the weighted (weights w_n) sum of the inputs x_n^{l-1} and the bias b after passing through an activation function, here a scaled exponential linear unit (SELU), shown in blue. (b) Shown is the training performance for an FNN tasked to discriminate N qubits with $N = 1, 2, \ldots, 5$. The generalization—the ratio of the geometric mean test $\mathcal{F}_{\mathrm{GM}}^{\mathrm{test}}$ and training qubit-state-assignment fidelity $\mathcal{F}_{\mathrm{GM}}^{\mathrm{train}}$—as the number of epochs increases is shown in black using the left y-axis. The associated standard deviation of the generalization is indicated in gray. The number of epochs to achieve the maximum qubit-state-assignment fidelity is indicated with a red vertical bar. The learning rate η, shown in blue and using the right y-axis, is gradually reduced as the number of epochs increases. 204

6-6 Measurement Data Processing and Discrimination. (a) M-dimensional data (z^{IF}[n]) processing for single-qubit (SQ) and multi-qubit (MQ) discrimination. For single-qubit discrimination, z^{IF}[n] is digitally demodulated at the intermediate frequency of a resonator i. Thereafter, the signal z_i[n] can either be simplified with a boxcar filter (BF) [$\frac{1}{M}\sum_n z_i$[n] $= \bar{\mathcal{I}}_i + j\bar{\mathcal{Q}}$] or kept as sequences $\mathcal{I}_i$[n] and $\mathcal{Q}_i$[n]. The discriminators can either be trained with the spectator qubits exclusively in their ground state (denoted by $\emptyset$) or in either their ground or excited state (denoted by $*$). For multi-qubit discriminators, the digitally demodulated signals z_i[n] at all resonator frequencies i are stacked up. The resulting data block is subsequently used for the discriminator training. Alternatively, the discriminator can be tasked to discriminate z^{IF}[n] directly without any digital preprocessing. (b) Comparison of the geometric mean qubit-state-assignment fidelity for five qubits after a $1\,\mu\mathrm{s}$-long measurement and 10,000 training instances per qubit-state configuration. All single-qubit discriminators are evaluated using training data with the spectator qubits in the ground and all ground and excited state combinations. The matched filter (MF) threshold discriminator [the matched filter is part of the discriminator and thus not shown in (a)] is shown in two configurations; the threshold set to 0 and the threshold optimized. The linear support vector machine (SVM) is applied to boxcar-filtered (BF) and time-trace data of $\mathcal{I}_i$[n] and $\mathcal{Q}_i$[n]. The multi-qubit discriminators are evaluated utilizing digitally demodulated and unprocessed data. Shown are a multi-qubit linear SVM, a recurrent neural network (NN), a convolutional NN, and feedforward NN. 206

6-7 Single-Qubit and Three-Qubit Discrimination. (a) Comparison of the qubit-state-assignment fidelity for qubit 1. (b) Comparison of the geometric mean qubit-state-assignment fidelity for qubit 1, 3, and 5. The analysis in (a) and (b) is conducted for a $1\,\mu\mathrm{s}$-long measurement and 10,000 training instances per qubit-state configuration. All single-qubit discriminators are evaluated using training data with the spectator qubits in the ground and all ground and excited state combinations. The matched filter (MF) threshold discriminator is evaluated in two configurations; the threshold set to 0 and the threshold optimized. The linear support vector machine (SVM) is applied to boxcar-filtered (BF) and time-trace data of $\mathcal{I}_i[\mathrm{n}]$ and $\mathcal{Q}_i[\mathrm{n}]$. The multi-qubit discriminators are evaluated utilizing digitally demodulated and unprocessed data. Shown are a multi-qubit linear SVM, a recurrent neural network (NN), a convolutional NN, and feedforward NN. 207

6-8 Focused Measurement Data Processing and Discrimination. (a) Superconducting qubit-state discrimination can be accomplished using a single-qubit matched filter (MF) with kernel $k_i[\mathrm{n}]$ which serves as a windowing function that projects the readout signals to a single axis and subsequent discriminator threshold optimization (no pulse applied, denoted by $\emptyset$, qubit initialized in the ground state: $\emptyset \rightarrow |0\rangle$ and labeled as 0; π-pulse applied, denoted by π, qubit initialized in the excited state: $\pi \rightarrow |1\rangle$ and labeled as 1). (b) Single-qubit linear support vector machines (SQ-LSVM), (c) multi-qubit LSVMs (MQ-LSVM), and (d) fully-connected feedforward neural networks (NN) are analyzed as alternatives to MFs. The qubit-state-assignment fidelity of the MF and LSVM is maximized if the intermediate frequency signal ($z^{\mathrm{IF}}[\mathrm{n}] = I^{\mathrm{IF}}[\mathrm{n}] + \mathrm{j}Q^{\mathrm{IF}}[\mathrm{n}]$) is digitally demodulated (e.g., for resonator 1: $z^{\mathrm{IF}}[\mathrm{n}].*^{-\mathrm{j}\omega_1^{\mathrm{IF}}\mathrm{n}} = \mathcal{I}_1[\mathrm{n}] + \mathrm{j}\mathcal{Q}_1[\mathrm{n}]$ with $.*$ indicating an element-wise multiplication). The training data is relabelled to train five parallel single-qubit discriminators (MF, SQ-LSVM). The training data can either be limited to measurements during which spectator qubits are kept in their ground state (denoted by $\emptyset$) or in all combinations of the ground and excited state (symbolized by $*$. The MQ-LSVM as a single multi-qubit discriminator requires the digitally demodulated data stacked up to form a single data block. The feedforward neural networks does not require any digital demodulation or preprocessing. 208

6-9 Qubit-State-Assignment Fidelity. (a) Geometric mean qubit-state-assignment fidelity $\mathcal{F}_{\mathrm{GM}}$ (Equation 6.4) for five qubits versus measurement time for the matched filter (MF), single-qubit linear support vector machine (SQ-LSVM), multi-qubit linear SVM (MQ-LSVM), and the fully-connected feedforward neural network (FNN). (b) $\mathcal{F}_{\mathrm{GM}}$ versus the number of training instances for each of the 32 qubit-state configurations evaluated after a measurement time of $1\,\mu\mathrm{s}$ [vertical dashed-dotted line in (a)]. (c) Achievable assignment fidelity $\mathcal{F}_{\mathrm{assignment}}$ per qubit when $N = \{1, 2, \ldots, 5\}$ qubits are simultaneously discriminated after a $1\,\mu\mathrm{s}$-measurement time. For each N-qubit discrimination task, the spectator qubits are initialized in their ground state. Single-qubit discrimination ($N = 1$): the first data point of each of the five panels represents the single-qubit $\mathcal{F}_{\mathrm{assignment}}$ defined by Equation 6.3. At the same time, the states of the four spectator qubits are initialized in their ground state and not discriminated. When employed as single-qubit discriminators, all methods perform similarly. Two-qubit discrimination ($N = 2$): The following four data points show $\mathcal{F}_{\mathrm{assignment}}$ when the state of each panel's qubit is simultaneously discriminated with the state of one other qubit. N-qubit discrimination ($N > 2$): the state of each panel's qubit is simultaneously discriminated with the states of $N - 1$ other qubits. For each N-qubit discrimination task, the non-spectator qubits are indicated with a colored square at the graph bottom. 210

6-10 Confusion Matrix. Confusion (assignment probability) matrix of the feedforward neural network (FNN) (a) and matched filter (MF) (b). The rows of the confusion matrix encompass the probability distribution of the discriminator to assign each of the 32 qubit-state configurations to the row's prepared qubit-state configuration (no pulse applied, qubit initialized in the ground state: $\emptyset \rightarrow 0$; π-pulse applied, qubit initialized in the excited state: $\pi \rightarrow 1$). The probabilities of correctly classified states–on the diagonal–are shown in blue, whereas the misclassification probabilities–the off-diagonals–are shown in red. $\mathcal{F}_{\mathrm{N}}$, introduced in Equation 6.5, represents a metric to indicate the overlap between the confusion matrix and an identity matrix (the ideal confusion matrix). $\mathcal{F}_{\mathrm{N}} = 1$ if the confusion matrix is an identity matrix. 215

6-11 Qubit-State-Assignment Fidelity Analysis. (a) Difference between the confusion (assignment probability) matrix of the feedforward neural network (FNN) $\mathcal{P}_{\mathrm{assign}}^{\mathrm{FNN}}$ and of the matched filter (MF) $\mathcal{P}_{\mathrm{assign}}^{\mathrm{MF}}$ [shown in Figure 6-10]. The rows of the confusion matrix encompass the discriminator's probability distribution to assign each of the 32 qubit-state configurations to the row's prepared qubit-state configuration (no pulse applied, qubit initialized in the ground state: $\emptyset \rightarrow 0$; π-pulse applied, qubit initialized in the excited state: $\pi \rightarrow 1$). An increase (decrease) in the relative state-assignment probability is marked in red (blue). Red diagonal and blue off-diagonal elements indicate an improvement of the FNN over MF discrimination performance. (b) The cutouts [bold frame in the lower right corner of (a)] of the FNN, MF, and resulting relative confusion matrix display the most prominent pattern that arises due to the discrimination of qubit 2. (c) The FNN and MF cross-fidelity matrices, as defined in Equation 6.7, indicate the discrimination correlation. The off-diagonals are ideally 0. A positive (negative) matrix off-diagonal entry indicates qubit-state assignment to be correlated (anti-correlated). 216

6-12 Training Wall-Clock Time. The data set acquisition for the matched filter (MF) scales linearly and exponentially for the single-qubit LSVM (SQ-LSVM), multi-qubit LSVM (MQ-LSVM), and feedforward neural network (FNN). The discriminator training remains linear for the MF while being exponential for the other three. 218

7-1 Readout Pulse Shaping Introduction. (a) A basic rectangular readout pulse (black dashed) with a normalized drive power p_{norm} contrasted by the cavity level excitation and reset (CLEAR) pulse (blue). The CLEAR pulse distinguishes itself from a rectangular pulse by the rising and falling edges. A two-amplitude segment replaces each edge to inject and deplete photons more rapidly, facilitating a shorter measurement time. (b) Shown are the resulting photon counts for the resonators in a five-qubit system for the rectangular pulse, shown in (a), frequency-multiplexed probing the quantum system. The three qubit-readout-pulse-shape sections (injection, stabilization, and reset) are indicated. The injection and reset sections are subject to an optimization routine to minimize their duration. The length of the middle part, the stabilization part, with a constant amplitude of $\sqrt{p}$ depends on how much of the signal needs to be acquired to discriminate the qubit states. As shown in (b), the photon number from a rectangular pulse oscillates at first. In (c) and (d) are two possible pathways presented to optimize readout pulse shapes. The readout pulse shape can either be optimized by enforcing to reach a specific photon number or delivering the maximally possible discrimination fidelity as quickly as possible. In (c), the readout pulse shape is optimized by targeting a specific photon number inferred by evaluating the qubit's dephasing rate with a Ramsey measurement. (d) Alternatively, the injection and stabilization part can be optimized by targeting the maximally possible discrimination rate. To optimize the reset, however, the photon number needs to be monitored. 234

7-2 Single-Qubit Readout Pulse Shapes. (a), (b), and (c) show the pulse shapes and photon number for the photon injection part. (d), (e), and (f) present the pulse shapes and photon numbers for the photon depletion section. (a) and (d) show the CLEAR pulse. (b) and (e) the PPO pulse. (c) and (f) contains the expected photon numbers of the rectangular, CLEAR, and PPO readout-resonator pulse. For a single-qubit readout, both CLEAR and PPO outperform a rectangular readout pulse. The pulse generated by CLEAR and PPO is virtually indistinguishable. 240

7-3 Multi-Qubit Readout Pulse Shapes. (a) Injection segment of resonator-readout-pulse shapes and (b) Reset segment of resonator-readout-pulse shapes for five qubits. (c) Photon numbers versus pulse length for a rectangular, CLEAR, and PPO-generate readout pulse. 241

7-4 IQ-plane resonator trajectories for each resonator and each pulse shape: rectangular, CLEAR, PPO-generated. Ground-state trajectories are shown in purple and orange for the excited state. The dashed grey circle marks the target photon number of four photons. The inner dashed circle indicates 0.1 photons. The 'X'-markers are spaced $20\,\mathrm{ns}$ apart. Each trajectory starts and returns to its origin. Depending on the photon population dynamics, the trajectory changes its shape. 242

7-5 Simulation of Readout Pulse Shape versus Measurement Time for a Five-Qubit Chip. (a) Comparison of readout pulse shapes for a single qubit. The typically employed pulse shape, a rectangular pulse shape (gray, dashed), is compared with the cavity level excitation and reset (CLEAR) (in blue) and a pulse shape tune-up using the proximal policy optimization (PPO) algorithm (in orange), a deep reinforcement learning algorithm. (b) Comparison of the duration to simultaneously inject photons in 5 different readout resonators, stabilize the photon counts during a waiting period, and return the resonators close to their vacuum states. The waiting period is kept constant for all three methods. In comparison with the rectangular pulse (CLEAR), PPO reduces the injection window by 58% (22%) and reset duration by 46% (17%). 244

List of Tables

2.1 Quantum Gates. A universal quantum gate set can be formed of single- and two-qubit gates. Common single-qubit gates are the three Pauli-gates, the Hadamard gate, and the phase shifting gates S and T. A frequently used two-qubit gate is the Controlled-NOT (CNOT) gate. The computational basis states are expressed as $|0\rangle = \begin{bmatrix}1 & 0\end{bmatrix}^T$ and $|1\rangle = \begin{bmatrix}0 & 1\end{bmatrix}^T$. 81

2.2 Computing Milestones Per Decade. Key decade milestones in the chronology of classical and quantum computing development between 1900 and 2020. 88

4.1 Material Induced Losses. Comparison of qubit losses induced by different casing materials for the presented package geometry. The values are obtained with EM simulations of a $\nu = 5\,\mathrm{GHz}$ transmon qubit. The qubit is located in the corner of a $5\,\mathrm{mm} \times 5\,\mathrm{mm}$ chip to estimate the maximum impact of package materials-induced losses. The T_1-limit is estimated as $1/(1/Q_c + 1/Q_d) \cdot 1/(2\pi\nu)$. The following material-dependent characteristics are assumed: Copper (Cu) has a conductivity of $5 \times 10^{10}\,\mathrm{S\,m}$ and an oxide thickness of $10\,\mathrm{nm}$; Aluminum (Al) has a conductivity of infinity and an oxide thickness of $2\,\mathrm{nm}$; Gold (Au) has a conductivity of $5 \times 10^{9}\,\mathrm{S\,m}$ and no oxide layer. 132

4.2 Identified Package Modes. The four identified package modes between $2\,\mathrm{GHz}$ and $20\,\mathrm{GHz}$. The table contains the mode's resonance frequency $\omega_m/(2\pi)$, the mode linewidth $\kappa/(2\pi)$, coupling strength to the qubit $g/(2\pi)$ as extrapolated using Equation 4.5, and the resulting qubit energy relaxation time $T_{\mathrm{pkg\ mode}}^{\mathrm{Purcell}} = 1/\gamma_{\mathrm{pkg\ mode}}^{\mathrm{Purcell}}$ (Equation 4.2). 151

4.3 Package Design Considerations. Summary of package design considerations with a focus on qubit-energy-loss channels, impedance matching, signal crosstalk, and package-mode profile. 157

6.1 Matched Filter Statistics. Numerical values extracted from Gaussian fits to readout data distribution after a $1\,\mu\mathrm{s}$-measurement time using a matched filter, as illustrated in Figure 6-2(a,b). The peak ratio of bimodal Gaussian fits (with equal variance) to the readout-traces histograms of qubits initialized in the ground state (no pulse applied: $\emptyset$) provide insight in the thermal excitation probability $\mathcal{P}(1|\emptyset)$. Comparing the peak ratios for trimodal Gaussian fits to the readout-traces histograms of qubits initialized in the excited state (π-pulse applied: π) indicate the conditional probability for qubit-energy decays $\mathcal{P}(0|\pi)$ and second-excited state population $\mathcal{P}(2|\pi)$. $\mathcal{F}_{\mathrm{label}} = 1-(\mathcal{P}(1|\emptyset)+\mathcal{P}(0|\pi))/2$ denotes a lower boundary for the initialization fidelity and thus the label accuracy using the conditional state transition rates. $\mathcal{F}_{\pi}$ represents the fitted π-pulse fidelities resulting in the preparation fidelities $\mathcal{F}_{\mathrm{prep}} = (1+[1-2\mathcal{P}(1|\emptyset)]\mathcal{F}_{\pi})/2$. $\langle S_0\rangle$, $\langle S_1\rangle$, and var(S) are the mean ground state, mean excited state, and variance of both states used to derive the Fisher criterion R and achievable assignment fidelity $\mathcal{F}_{\mathrm{ach}}$ (see Equation 6.1, 6.2). $\mathcal{F}^{\overline{\mathrm{MF}}}$, the product of $\mathcal{F}_{\mathrm{label}}$ and $\mathcal{F}_{\mathrm{ach}}$, is an estimate for an upper qubit-state-assignment fidelity bound for a classifier composed of a matched filter and the subsequent optimized threshold, here referred to as MF. 200

6.2 Qubit-Assignment Fidelity Statistics. Qubit-assignment fidelity if discriminated individually, $\mathcal{F}_i^{1\mathrm{Q}}$, and in parallel with all other qubits, $\mathcal{F}_i^{5\mathrm{Q}}$. The last five columns present the assignment fidelity for an N-qubit discrimination process with $N = \{1, 2, \ldots, 5\}$. $\langle \mathcal{F}^{N\mathrm{Q}} \rangle$ represents the mean assignment fidelity of all qubit permutations. The single-qubit assignment fidelity is similar for all discriminator approaches. For a two-qubit discrimination task, the SQ-LSVM and FNN outperform the MF and MQ-LSVM. For N-discrimination tasks with $N > 2$, the FNN outperforms all other methods. 213

6.3 Readout Crosstalk. Mean absolute value, $\langle | \cdot | \rangle$, of the qubit-state-assignment correlations between readout resonators i and j $(i \neq j)$ extracted from the cross-fidelity matrix $\mathcal{F}^{\mathrm{CF}}$ when using a MF or FNN discriminator. 217

7.1 Summary of Simulation Results. Listed is the injection and reset pulse length for the case of single-qubit and five-qubit readout. The investigated readout pulse shapes are rectangular, CLEAR, and PPO-generated. 243

B.1 Superconducting Chip—Qubit Details. Chip comprising five superconducting frequency-tunable transmon qubits with alternating transition frequencies. A normalized magnetic flux bias Φ/Φ_0 (magnetic flux quantum Φ_0) detunes the qubits from their idling to their operating frequency. The qubit anharmonicities α are in the moderate transmon regime. The qubit lifetimes T_1, Ramsey coherence times $T_{2\mathrm{R}}$, and spin-echo relaxation times $T_{2\mathrm{E}}$ are measured at the qubit operating frequency. 267

B.2 Superconducting Chip—Resonator Details. Chip comprising five superconducting readout resonators at bare resonance frequencies of about $7\,\mathrm{GHz}$. Signals are up-converted from MHz intermediate frequencies (IF) using a common local oscillator at $\omega_{\mathrm{LO}}/2\pi = 7.127\,\mathrm{GHz}$. Each resonator couples to a designated qubit with strength g, leading to a dispersive shift χ. The effective resonator decay rate through the Purcell filter is κ_{eff}. The qubit-resonator interaction remains in the dispersive regime for readout resonator photon populations below the critical photon number n_{crit}. 267

C.1 Definition of used Terminology. Derivation of a matched filter to discriminate qubit states. The noise source affecting the measurement is assumed to be additive and stochastic. 271

Chapter 1

Introduction

Quantum computers hold the promise to perform specific computational problems more efficiently than classical computers. The challenge of developing quantum computer hardware is to balance the system's isolation while maintaining efficient control and readout channels. Towards the realization of quantum computers, many milestones have already been achieved. However, many engineering challenges remain, and realizing a functional large-scale quantum computer is likely years away. The focus of this thesis is on developing superconducting quantum processor control hardware and software techniques using microwave design principles and machine learning to maintain single-qubit operation performance for multi-qubit systems. The thesis is divided into two parts:

I. The concepts and infrastructure of superconducting quantum computing

II. Machine learning tools and their applicability to superconducting qubit readout.

This introduction is an overview of the main concepts and results for these two parts of the thesis. The chapter begins with a brief introduction to superconducting quantum computing, then gives the main results of the design and characterization of a superconducting microwave cavity, and concludes with an overview of machine learning as applied to qubit readout and control. These topics are elaborated on in the subsequent chapters of the thesis.

1.1 Quantum Computing

Quantum computers are anticipated to solve particular computational tasks substantially faster than conventional computers [1, 2]. Depending on the task, such quantum computers need to be composed of hundreds to millions of quantum bits (qubits), the principal building blocks of a quantum processor. Today's quantum processors comprise tens of qubits and are not yet able to execute any meaningful computations. However, within the past two years, a quantum processor with 53 working superconducting qubits [3] and a device with 76 photons [4] achieved a major quantum computing milestone, termed quantum supremacy. Quantum supremacy describes completing a computational task on a quantum processor that cannot be performed on existing classical computers in a reasonable amount of time [5]. Despite the importance of this milestone, the required technology to build a universal and fault-tolerant quantum processor is likely at least a decade away. Meanwhile, today's noisy intermediate-scale quantum devices [5] offer proof of concept demonstrations such as the simulation of molecular electronic configurations [6]. An increase in the number of qubits has generally been accompanied by an increase in errors that inhibits achieving a quantum advantage [5]. While remarkable progress has been made in reducing qubit-error rates, closed-loop error correction protocols are likely necessary to implement fault-tolerant quantum computation in large qubit networks [7]. Typically, these quantum error correction protocols use syndrome measurements to detect if and where an error occurred. Detected errors can then be corrected using a 'feed-forward' approach, in which pulses are applied to the errant qubits to correct them. Among the most imminent challenges towards demonstrating such error-corrected quantum processors are accurate and resource-efficient qubit-state readout to identify and correct errant qubits.

1.2 Superconducting Quantum Computing

Various qubit modalities such as superconducting qubits [8], trapped ions [9], trapped neutral atoms [10], spin qubits [11], or solid-state quantum emitters [12] are actively pursued by academic and industrial teams to build quantum devices. Each of these modalities has different applications with individual strengths and drawbacks. This thesis focuses on superconducting qubits, particularly their microwave environment, and how to efficiently measure them using machine learning as the number of qubits is increased.

Over the past two decades, superconducting qubits have emerged as one of the leading candidates for building larger-scale quantum systems [8, 13]. Today, superconducting qubits routinely reach coherence times in the range of $100\,\mu\mathrm{s}$ [14], achieve gate times of a few tens of nanoseconds [3], and have exhibited single- and two-qubit gates with fidelities exceeding the threshold for the most lenient quantum error correction codes [15]. Moreover, the characteristics of qubits and their surrounding circuit elements for control and readout can be engineered with high precision using well-established fabrication techniques [16] leveraging standard tools developed for the semiconductor industry.

1.2.1 The Transmon Qubit

The transmon qubit [17] has emerged as one of the most popular qubit designs due to its robust fabrication process, demonstrated operation and readout, and reproducible coherence times around $100\,\mu\mathrm{s}$ [14]. It is closely related to a quantum harmonic LC-oscillator, which features equidistant energy levels. Coherent control requires an isolated pair of energy levels that form a computational qubit basis [16], and this motivates the need for anharmonic oscillators. The needed anharmonicity is provided by a Josephson junction, a lithographically defined tunnel barrier between two superconducting electrodes, which behaves as a non-linear inductor without any significant dissipation [18]. The schematic transmon circuit is depicted in Figure 1-1(a).

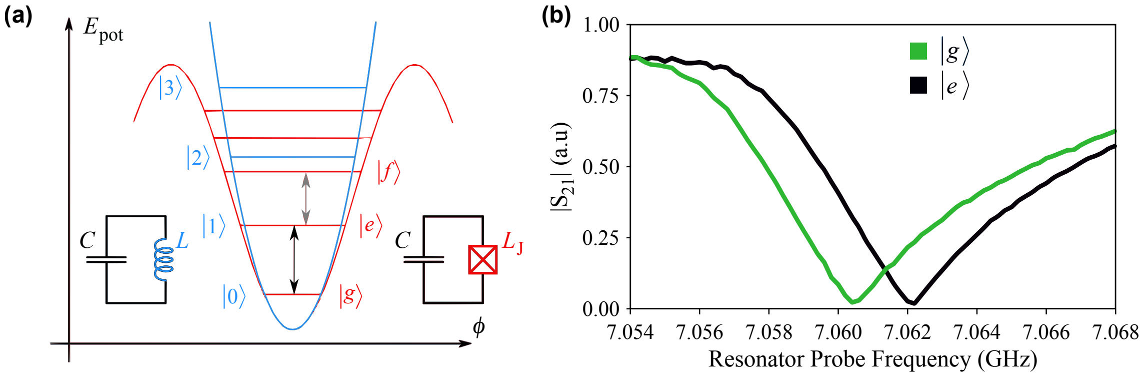


Figure 1-1: Quantum Harmonic and Anharmonic Oscillator. (a) A simple harmonic resonator formed by a capacitor C and an inductor L is shown in blue. The potential energy is the energy stored in the inductor and assumes a parabolic shape concerning the phase variable ϕ, related to the flux induced by the inductor. The energy levels are equidistant, such that transitions cannot be addressed individually. By replacing the inductor with a Josephson junction, which has a non-linear inductance L_{J}, the potential becomes anharmonic (shown in red), which isolates two energy levels, $|g\rangle$ and $|e\rangle$, forming the computational basis and hence a superconducting quantum bit. (b) The transmission spectrum of a resonator dispersively coupled to a qubit showing the characteristic qubit-state-dependent dispersive shift of about $1\,\mathrm{MHz}$.

1.2.2 Superconducting Qubit Readout

Superconducting qubit-state readout is most commonly performed today using a dispersive readout scheme [19, 20]. The readout device is a resonator weakly coupled to the qubit at a detuned frequency, typically a few gigahertz. Due to a qubit-state-dependent 'dressing' of the readout resonator, the qubit state can be inferred by spectroscopically probing a dispersive shift. For example, in Figure 1-1(b), the dispersive shift is about $1\,\mathrm{MHz}$ for a qubit-resonator detuning of about $2\,\mathrm{GHz}$. After a sequence of quantum gate operations or free evolution, a measurement process is initiated by driving the readout resonator with a microwave pulse. After transmission or reflection off the resonator, the signal acquires a qubit-state-dependent phase shift or change in amplitude that can later be associated with a specific qubit state. For a single superconducting qubit, this type of qubit-state discrimination has been demonstrated to reach a classification accuracy of $99\,\%$ [21, 22]. However, resource-efficient multi-qubit readout has not yet been demonstrated to reach such a high accuracy due to an increase in nonidealities such as readout crosstalk.

1.2.3 Measurement Setup and Microwave Control

A typical microwave control and measurement setup is schematically depicted and described in Figure 1-2(a). Room-temperature electronics such as arbitrary waveform generators and local oscillators feed microwave pulses to the cryostat through coaxial cables. The measurement pulses that are transmitted through the qubit chip are passed through a chain of amplifiers at different temperature stages: a quantum-limited traveling-wave parametric amplifier [23] at about $20\,\mathrm{mK}$, a low-noise high electron mobility amplifier at about $3\,\mathrm{K}$, and room-temperature amplifiers. Subsequently, the signal is down-converted to intermediate frequencies in the megahertz range and digitized.

Quantum gates are performed by applying microwave drive pulses at or close to the qubit transition frequency. As a result, the qubit undergoes coherent oscillations between its two fundamental basis states, referred to as Rabi oscillations [8]. This

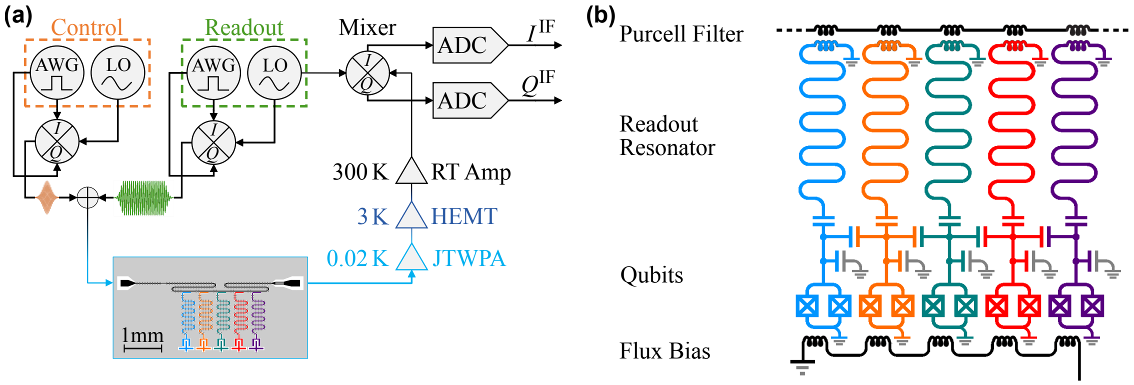

(a)
Control
Readout
AWG
LO
AWG
LO
I
Q
I
Q
Mixer
I
Q
ADC
ADC
I^{IF}
Q^{IF}
300 K
RT Amp
3 K
HEMT
0.02 K
JTWPA
1mm
(b)
Purcell Filter
Readout Resonator
Qubits
Flux Bias

Figure 1-2: Measurement Setup. (a) Qubit measurement and control setup. Control and readout pulses generated by an arbitrary waveform generator (AWG) and up-converted to gigahertz frequencies using a local oscillator (LO) are sent to the qubit chip through attenuated signal lines. The readout signal upon interaction with the qubit is amplified by a Josephson traveling-wave parametric amplifier (JTWPA), a high-electron-mobility transistor (HEMT) amplifier, and a room-temperature amplifier before being down-converted to megahertz frequencies, in-phase I^{IF} and quadrature Q^{IF} component at intermediate frequencies (IF), and digitized. (b) The chip shown in (a) consists of five superconducting transmon qubits between $4.3\,\mathrm{GHz}$ and $5.2\,\mathrm{GHz}$ frequency-tunable via a global bias. The qubits are capacitively coupled to individual quarter-wave readout resonators at about $7\,\mathrm{GHz}$ that couple inductively to a bandpass (Purcell) filtered feedline line at the top of the chip.

oscillation can be stopped at any time to prepare the desired superposition state. Microwave pulses inducing single-qubit rotations have a typical duration of about $20\,\mathrm{ns}$.

1.3 Package Design for Superconducting Qubits

A microwave package defines the immediate qubit chip environment and facilitates qubit control and readout channels. An approach to multi-port packaging and relevant microwave design principles are discussed using a novel 24-port package with enclosed transmon qubits. After a characterization of the package materials, signal transmission, and mode profile, the package's mode profile is experimentally inspected through a 'hidden-mode' survey [24] using four transmon qubits with lifetimes reaching $120\,\mu\mathrm{s}$. The considered loss channels enable the novel package to support transmon qubits with lifetimes above $350\,\mu\mathrm{s}$. While the qubits employed in the following characterization are not lifetime-limited by the package, the limit is within the same order of magnitude as the lifetime of state-of-the-art transmons. This underscores the importance of further improvements in microwave engineering to minimize qubit energy loss channels. The subsequently discussed engineering principles provide tools for the development of improved packaging for near-term quantum processors [5].

The purpose of a microwave package is threefold: first, the package casing suppresses the coupling of the qubits to decoherence channels external to the package, such as environmental electromagnetic noise; second, the package accommodates qubit control channels to and from the enclosed quantum processor; and finally, the package sinks excess thermal energy due to qubit control and readout operations. However, a microwave package—the immediate qubit environment—can induce losses of its own if not carefully engineered. Thus, designing a microwave package—from material choices to a signal interface—requires considering both the suppression of external and package-induced loss channels.

1.3.1 Microwave Package Materials & Geometry

Material-dependent losses can be of magnetic, conductive, or dielectric origin [25]. Energy loss channels couple to the qubit through its electric or magnetic dipole moment. For transmon qubits, the electric dipole moment presently dominates the loss [17]. Qubits are fabricated using high-Q materials and substrates to reduce loss. In addition, the device geometry is designed to reduce the electric field density in lossy regions, such as surfaces and interfaces [26].

Commonly employed package materials include superconducting aluminum, copper, and gold-plated copper. Superconducting aluminum forms a thin oxide layer of approximately $2\,\mathrm{nm}$ [27], inducing some dielectric losses while keeping the conductivity losses at a minimum. Like aluminum, copper forms an oxide layer [28] leading to dielectric as well as conductivity loss due to its non-zero resistance [29]. Gold-plating limits the oxide formation at the cost of an increase in conductivity losses by up to one order of magnitude [30].

The device package shown in Figure 1-3(b) is composed of a base and lid, both milled from oxygen-free high-conductivity (OFHC) copper. To increase the package fundamental mode frequency and suppress material-induced losses, the qubit chip is suspended by at least $3\,\mathrm{mm}$ to form a cavity above and below it [26]. A layer of aluminum with a target thickness of $500\,\mathrm{nm}$ is evaporated on the lid center cavity surface to reduce conductivity losses and to help shield magnetic fields. Full-wave EM simulations (COMSOL Multiphysics®) indicate the layer of aluminum on the center cavity surface to reduce the material-induced loss channels by three orders of magnitude. As such, the material-induced losses of the presented package are negligible, enabling it to support qubits with lifetimes up to seconds.

1.3.2 Control and Readout Signals

The signal paths comprised in a package introduce various nonidealities, including a distorted step response, insertion loss, and crosstalk, which can have a bearing on qubit control. Good impedance matching leads to lower insertion losses and

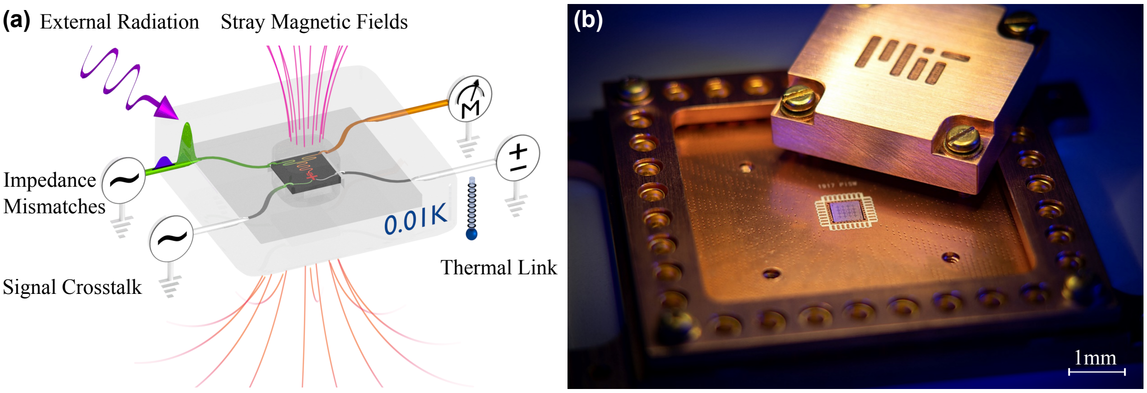


Figure 1-3: Microwave Packages. (a) The purpose of a microwave package is to shield the enclosed qubit chip from external radiation (purple oscillating arrow) and stray magnetic fields (top; red lines) while providing impedance-matched (left; transmitted green pulse and reflected blue pulse at the input and output), low crosstalk communication channels (center; crosstalk in green at input), and a thermal link to the dilution refrigerator. (b) The presented microwave package consists of a metal enclosing, microwave connectors, an interposer for signal fan out, and a microwave cavity in the center surrounding the quantum chip.

improved signal integrity, critical for high-fidelity control and readout.

The developed package's interposer, the communication interface between the ports mounted on the package enclosing and the qubit chip, is constructed using a three-layer, low-loss Rogers 4350™laminate composed of glass-woven hydrocarbon and ceramics. The interposer uses symmetric striplines embedded in the printed circuit board dielectric protected by via fences. EM simulations and time-domain modeling are used to minimize impedance mismatches. The simulations ensure that microwave connector transitions, composed of a grounding cage and a signal via, as well as the wirebond launches, are properly impedance matched [31–33]. In the designed package, the crosstalk between directly neighboring control and readout lines is suppressed to below $-40\,\mathrm{dB}$ and next-nearest neighbors to $-60\,\mathrm{dB}$ up to $10\,\mathrm{GHz}$.

1.3.3 Package Modes

In addition to crosstalk, suppressing package modes is key to a successful microwave package design. Package modes can arise due to the enclosing metal cavity, slotlines, two metallic planes separated by a dielectric gap, or the chip substrate itself. These resonance modes can reduce the qubit lifetime and induce decoherence. A two-level quantum system can model the coupling interaction between the qubit and an EM cavity (the resonant mode) with a rate g. For a small detuning $\Delta = |\omega_\mathrm{q} - \omega_\mathrm{m}| \ll g$ between the qubit angular frequency ω_q and the package mode angular frequency ω_m, their energy levels hybridize, and excitations are coherently swapped between the qubit and the mode. However, since the package modes are often lossy and have a low-quality factor Q_m, they lead to a reduction in the qubit lifetime.

The dispersive Jaynes-Cummings model describes the coupling between a far-detuned mode and a qubit in the dispersive approximation [8]. A mode coupling to a qubit ac-Stark shifts the qubit transition frequency by an amount proportional to the average number of photons present in the mode. Photon-number fluctuations of a mode within a few GHz of the qubit transition frequency lead to pure dephasing. Modes on the order of MHz detuning within the linewidth of the mode and qubit can result in qubit energy decay due to the Purcell effect, thus reducing the qubit lifetime.

To ensure high-fidelity qubit coherence, either the coupling of package modes to the qubit needs to be suppressed, or their resonance frequency must be far detuned from the qubit operational frequency spectrum.

A qubit-based technique to probe package modes is the hidden-mode experiment [24], where a fixed-frequency qubit is used as a mode sensor. It is preferable to use qubits with long coherence times and stable baseline Ramsey oscillations to resolve subtler mode structures. A continuous-wave probe tone is injected into the package, either through the readout line or a dedicated port, and swept through the frequency range of interest. For each probe frequency, a T_2 measurement is performed using Ramsey interferometry on a fixed-frequency qubit. As the probe frequency sweeps in resonance with a package mode, the package mode will be populated by coupling the transmission line and the mode itself. Depending on the mode photon number fluctuations, the coupling to the qubit, and its detuning, the qubit will dephase. The degree of induced dephasing can be inferred with a T_2 measurement. Due to the wide-band nature of the mode-induced qubit energy level shift, the ac-Stark effect, this technique provides the advantage of facilitating mode measurements across a broad frequency range, typically several tens of GHz.

In the package shown in Figure 1-3(b), four package modes are identified between $2\,\mathrm{GHz}$ and $20\,\mathrm{GHz}$. The designed package with these modes supports lifetimes for transmon qubits in excess of $350\,\mu\mathrm{s}$. While this is sufficient for current devices with coherence times in the range of $100\,\mu\mathrm{s}$, this result is problematic because, despite the absence of strong spurious modes up to $11\,\mathrm{GHz}$, a comprehensive survey reveals that the higher frequency modes can still have a significant limiting effect on qubit lifetime [34]. Furthermore, these limits will likely be saturated soon as qubit lifetimes increase, underscoring the need for further package design improvements.

Looking forward, package design will become increasingly critical for larger quantum devices due to their increased complexity. As the number of qubits in today's noisy intermediate-scale quantum devices [5] increases, the precise characterization and suppression of electromagnetic modes and signal crosstalk become even more relevant. These established principles for superconducting qubit packages are similarly pertinent for future work as packaging techniques are being advanced for systems in the range of 100 to 1000 qubits.

1.4 Machine Learning

The study of computational algorithms that can improve through experience is typically referred to as machine learning (ML) [35]. Without explicit instructions, these algorithms strive to identify patterns in sample data and create an approximate model of an underlying decision process. While many machine learning ideas are decades old, they only recently became widely applicable due to the development of sufficient classical computational resources. Machine learning is applied today in applications such as image processing [36] or advanced games such as chess [37].

Machine learning can be broadly divided into three categories: unsupervised, supervised, and reinforcement learning. Unsupervised learning describes methods tasked to extract an often defined number of patterns in a data set. Methods requiring training data composed of input-label pairs (e.g., class or number as labels) are coined supervised learning methods. Reinforcement learning describes a class of tasks where an agent optimizes a task in an environment through interaction. Therefore, the environment's role is more of an 'oracle' than that of a 'supervisor.'

This thesis focuses on supervised learning methods to improve superconducting qubit state discrimination and use reinforcement learning to generate system-dependent readout pulse shapes.

1.4.1 Supervised Learning

Methods that learn an input-output mapping function using a trustworthy set of input-output pairs (training set) are generally called supervised learning methods. Typically, a set of input-output pairs for training is acquired by the 'supervisor,' hence the terminology. The quality of the learned mapping function can be probed utilizing a second set of trusted input-output pairs (test set). Comparing the method's performance on the training versus the test set is termed 'generalization.'

Support vector machines (SVM) [38] or artificial neural networks (ANN) [39] are typically trained in a supervised fashion. SVMs are quadratic programs to maximize the distance between each data point and the decision boundaries separating distinct classes. ANNs are mapping functions composed of connected units, neurons, similar to biological neural networks. Customizing the neural network architecture, the arrangement of neurons, and the connections between them, enables tailoring for specific tasks ranging from image processing [36] to natural language processing [40].

1.4.2 Reinforcement Learning

Reinforcement learning is a subset of machine learning in which an agent aims to maximize a reward by choosing the appropriate action at a given state. Rather than comparing predictions with labeled data sets like in supervised learning applications (e.g., image classification), reinforcement learning algorithms ‘learn’ through the process of their agents directly interacting with the environment. Additionally, reinforcement algorithms do not need to know how the underlying system functions and can thus be model-free. Instead, reinforcement algorithms compute the reward that a specific state-action pair will produce. The growth of reinforcement learning research and applications has been shown through success in complex games such as Chess and Go [37] and robot automation and planning [41].

This thesis demonstrates that a proximal policy optimization (PPO) algorithm [42] proves advantageous over algorithms such as Deep-Q Network (DQN) algorithms [43], Deep Deterministic Policy Gradient (DDPG) algorithms [44], or AlphaZero [37]. PPO algorithms identify the maximum possible improvement without unnecessary risks by evaluating the acquired data. In contrast to a similar type of algorithm, trust region policy optimization [45], PPO algorithms computationally simplified the optimization task to first-order methods such as gradient descent. Two neural networks underlay this algorithm: one serves as a policy and one as a reward estimator, which creates a new action, while the second estimates the associated reward. These processes enable computationally efficient exploration of the surrounding area to determine the next action-state pair.

1.5 Superconducting Qubit-State Discrimination using Supervised Learning

As systems grow, simultaneous multi-qubit readout becomes a necessity to limit system resource overhead. Many state-of-the-art qubit-state discriminators assume ideal operating conditions or require considerable computational effort, limiting their applicability for high-fidelity multi-qubit readout. A novel approach to multi-qubit readout using neural networks as multi-qubit state discriminators is presented in this thesis. A comparison is made between the approach with contemporary state-of-the-art methods employed on a 5-qubit superconducting qubit chip with frequency-multiplexed readout. A fully connected feedforward neural network (FNN) is found to significantly reduce the readout assignment error rate for the utilized multi-qubit system by up to 20%. This work demonstrates an advantageous building block to scaling up quantum processors while maintaining high-fidelity operations and will impact the design of future fault-tolerant systems and near-term devices of moderate size.

Superconducting qubit readout is generally performed under the paradigm of circuit quantum electrodynamics (cQED) in the dispersive regime [19]. Here, the qubit is coupled to a far-detuned cavity, such that the leading-order effect on the cavity is the qubit-state-dependent frequency shift. As a result, a coherent microwave signal incident on the cavity will acquire a qubit-state-dependent phase shift or change in amplitude upon transmission or reflection. In the dispersive regime, the signal-to-noise ratio (SNR) of the phase-shifted signal is maximized at a few tens of photons populating the readout cavity. It is not feasible for room-temperature equipment to measure such weak signals without significant averaging, which makes the need for amplification a necessity. The amplifier chain significantly increases the signal-to-noise ratio. Subsequent heterodyne detection and digitization of the amplified signal projects the information of the qubit state on the in-phase and quadrature components of the output signal.

For a qubit with static coupling to its readout resonator, spontaneous decay and excitation during the readout itself are typically the primary sources of qubit measurement errors. Moving to multi-qubit systems, it becomes impractical for each qubit to have its own dedicated set of readout electronics. Instead, qubit readout must be multiplexed, such that the readout signal contains state information on multiple qubits, often spread in frequency. This can introduce errors where the state of other qubits can alter the measurement signal from one qubit through crosstalk.

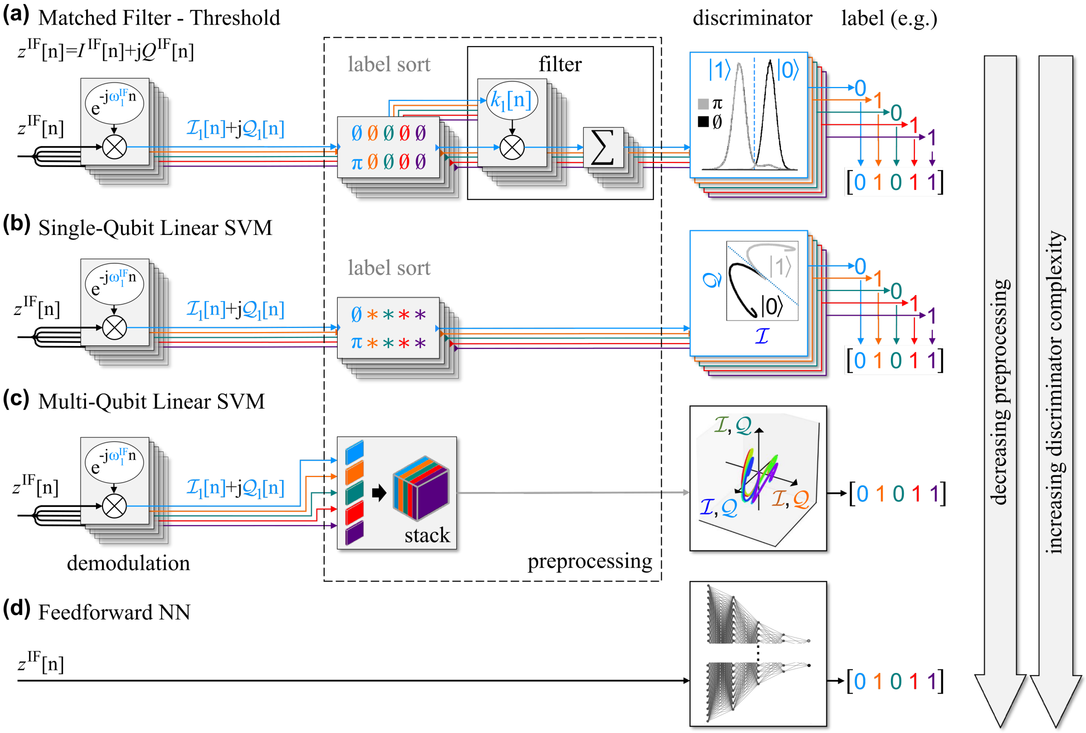

(a) Matched Filter - Threshold
$z^{IF}[n]=I^{IF}[n]+jQ^{IF}[n]$
discriminator
label (e.g.)
label sort
filter
$k_1[n]$
$\mathcal{I}_1[n]+j\mathcal{Q}_1[n]$
|1⟩
|0⟩
π
∅
[0 1 0 1 1]
(b) Single-Qubit Linear SVM
label sort
(c) Multi-Qubit Linear SVM
stack
demodulation
preprocessing
(d) Feedforward NN
$z^{IF}[n]$
decreasing preprocessing
increasing discriminator complexity

Figure 1-4: Measurement Data Processing and Multi-Qubit-State Discrimination. (a) Single-qubit matched filter (MF) with subsequent discriminator threshold optimization, (b) Single-qubit linear support vector machine (SQ-LSVM), (c) multi-qubit LSVM (MQ-LSVM), and (d) feedforward neural network (FNN). The MF, as well as the LSVMs assignment fidelity, is maximized if the intermediate frequency (IF) input signal is digitally demodulated down to DC: $I^{\text{IF}} \rightarrow I^{\text{DC}}, Q^{\text{IF}} \rightarrow Q^{\text{DC}}$. During the discriminator training, the single-qubit discriminators (MF, SQ-LSVM) require a relabelling step of the discriminators' training data. The MF tune-up is optimized when each discriminator utilizes training data that contains a single excitation. The MQ-LSVM requires the DC data for each of the five qubits to be stacked and combined into a single data block. The FNN does not require any preprocessing such as digital demodulation, data stacking, label filtering, or data filtering as the complexity of it all is embedded in the neural network itself.

Crosstalk errors can also occur in multi-qubit systems due to intrinsic interactions between the qubits themselves or between the qubits and nearby readout resonators belonging to other qubits.

As a result of crosstalk, spontaneous decay/excitation, and other nonidealities, multi-qubit heterodyne signals are difficult to distinguish. There has been significant progress in reducing error rates and measurement times [22, 46]; however, managing, classifying, and extracting useful information from the heterodyne signal in a scaleable manner remains an essential challenge towards improving readout as systems grow in size.

1.5.1 Superconducting Qubit-State Discriminators

This thesis focuses on state discrimination for multiple qubits connected to the same microwave transmission line. The labels are 5-bit strings, with each bit representing the prepared state of a single qubit. For multi-qubit readout, the individual readout tones are superposed using the same local oscillator in conjunction with a unique intermediate frequency (IF). The transmitted or reflected multi-qubit signal is then down-converted back to intermediate frequencies with $I-Q$ components ($I^{\mathrm{IF}}, Q^{\mathrm{IF}}$), optionally preprocessed, and classified as described in Figure 1-4.

Considerable signal processing goes into determining the qubit-state-dependent phase because it is the main avenue to improved qubit readout fidelity. In the standard approach, an integration kernel (or filter) is applied. The integrated signal has a value dependent on the signal phase, and hence, qubit states can be discriminated by the integrated signal. For stationary noise, the optimal filter in terms of SNR is a generalized matched filter [47], which creates a linear partition of the integrated signal that distinguishes the qubit states. This is the approach used in contemporary, state-of-the-art single-qubit readout schemes [3, 46–48]. However, realistic noise sources often do not satisfy the conditions required for optimal matched filtering. Consequently, recent approaches have explored using support vector machines (SVM) to learn the optimal, non-linear filter [21].

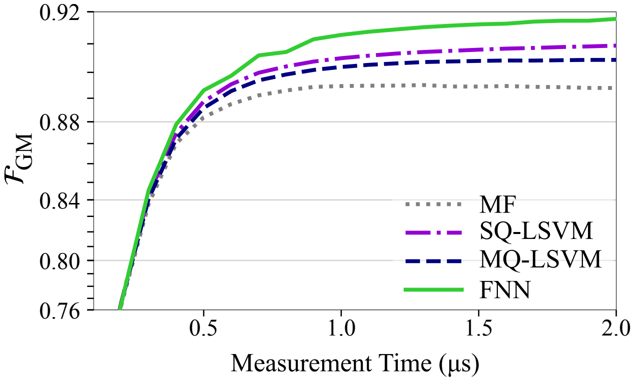


Figure 1-5: Time-Dependent Qubit-State Assignment Fidelity. Five-qubit geometric mean qubit-state assignment fidelity $\mathcal{F}_{\mathrm{GM}} = (\mathcal{F}_1\mathcal{F}_2\mathcal{F}_3\mathcal{F}_4\mathcal{F}_5)^{1/5}$ versus measurement time for the matched filter (MF), single- and multi-qubit linear SVM (SQ-LSVM and MQ-LSVM), as well as the feedforward neural network (FNN).

In contrast to the matched filter and SVMs, a neural network can discriminate the frequency-multiplexed signal directly. Training the network on the frequency-multiplexed multi-qubit signal bypasses the need for digital demodulation and other data processing stages, making more efficient use of the measurement output as illustrated in Figure 1-4. Neural net classifiers have not yet been applied to superconducting qubit readout, although there exist demonstrations [49, 50] for trapped-ion qubits, where an improvement was shown. However, trapped-ion qubits do not share the same readout line as is typically the case for superconducting qubits.

1.5.2 Experimental Comparison of Qubit-State Discriminators

The focus is on classifiers used in a supervised learning scheme, trained to determine the state of all qubits from the output signal data. The data to train and evaluate the discriminator performance was acquired using a 5-qubit chip shown in Figure 1-2(b). For five qubits, all 32 possible qubit configurations (2^{N}, where N is the number of qubits) are sequentially initialized and recorded. The generated data set contains 50,000 single-shot time-traces recorded over $2\,\mu\mathrm{s}$ for each qubit-state configuration. The recorded data set is subsequently divided into a training and test set (15,000 traces per qubit-state configuration for training and 35,000 for testing). The training set serves to train the discriminators, and the test data set indicates how well the trained discriminator generalizes to unseen data.

Figure 1-5 presents the performance results of four discriminators; two single-qubit discriminators—matched filter and single-qubit LSVM—as well as two multi-qubit discriminators—multi-qubit LSVM and FNN. The geometric mean qubit-state assignment fidelity is defined as $\mathcal{F}_{\mathrm{GM}} = (\mathcal{F}_1\mathcal{F}_2\mathcal{F}_3\mathcal{F}_4\mathcal{F}_5)^{1/5}$ where $\mathcal{F}_i = 1-[P(g_i|e_i)+P(e_i|g_i)]/2$ is the qubit-state assignment fidelity. $P(g_i|e_i)$ denotes the conditional probability that qubit i is assigned to the ground state g_i when initialized in the excited state e_i. Similarly, $P(e_i|g_i)$ represents the conditional probability that qubit i is assigned to the excited state e_i when initialized in the ground state g_i. The FNN excels in assignment fidelity independent of the measurement duration. The FNN reduces the qubit-state assignment error rate by up to 20% per qubit relative to the matched filter, the most commonly employed qubit-state discrimination method.

The assignment fidelity is evaluated for different numbers of training samples per qubit configuration. While single-qubit discriminators reach their optimal performance at a few hundred training samples per qubit-state configuration, multi-qubit discriminators are more demanding. The employed FNN starts saturating at around 5,000 training samples per state configuration.

As the number of qubits per discriminator increases, the assignment fidelity does not degrade as quickly for the FNN relative to the other methods. The neural network appears to compensate for crosstalk and state transitions more effectively, explaining the observed assignment fidelity improvement of neural network discriminators. These results are published in Reference [51] and further described in chapter 6.

1.6 Superconducting Qubit-Readout-Pulse Shaping using Deep Reinforcement Learning

For a qubit dispersively coupled to a resonator, quick resonator ring-up and ring-down ensure fast readout and reduces the residual photon number in the resonator, reducing qubit dephasing in future operations. In an efficient, frequency-multiplexed readout of multiple qubits, effects such as drive crosstalk increase the complexity of optimal readout pulse shapes, requiring computationally intensive methods to discover high-fidelity readout pulse shapes. In this thesis, a pulse shaping optimization module using deep reinforcement learning (DRL) is investigated. Using a multi-qubit simulator, the DRL pulse optimization module reduces the readout resonator population and reset time by about 20% relative to current methods. Finally, the DRL method is experimentally compared with conventional readout pulse shaping techniques applied to multi-qubit devices.

A superconducting qubit measurement can be divided into three components: photon injection, waiting period, and resonator reset. The injection phase describes the process of populating a readout cavity initially in the vacuum state with a target photon number. The waiting period of a measurement depends on the strength of the measurement (the qubit-resonator coupling), noise present in the system, and the qubit-state discriminator. Finally, the resonator reset phase encompasses the process to return the populated readout cavity to a photon number close to zero. Quick readout resonator population and reset are significant in feedback applications such as for quantum error correction protocols and many near-term quantum algorithms [52, 53]. Generally, readout resonator reset is realized by waiting for several resonator decay time constants, which is a slow and inexact process.

Progress has been made by engineering readout pulses able to speed up the process to both populate the resonator and evacuate the photons post measurement. The cavity level excitation and reset (CLEAR) pulse can depopulate the resonator two resonator decay time constants faster than the conventional approach [54]. Further improvements can be achieved using gradient ascent pulse engineering (GRAPE) to optimize the readout pulse [55]. However, these methods mainly rely on a simple rectangular pulse with modified flanks or a gradient-based optimization approach that relies on an accurate description of the system and a robust initial guess that is often difficult to identify [55].

Here, a readout pulse shaping module using deep reinforcement learning (DRL) is being developed. Among the considered DRL algorithms are DQN, DDPG, PPO, and AlphaZero. DQN algorithms require a discrete action space as the network produces a predicted Q-value for each possible action when provided a state. The action space can be directly mapped to the discrete output space of the AWG. However, the vast action set size also makes the neural network approximation more difficult and time-consuming to train due to the high number of output nodes.

In contrast, the DDPG algorithm, a model-free algorithm with a continuous action space, can learn optimal policies in high-dimensional spaces more efficiently. DDPG, however, can suffer from slow optimization due to vanishing or exploding gradients. PPO is a more robust alternative to DDPG and can benefit from inexpensive sampling costs of the underlying system. AlphaZero was recently used for the optimization of two-qubit gates [56]. However, the power of this algorithm is accompanied by a significant computational effort. In a simulation, the PPO algorithm outperformed the alternatively investigated approaches in simplicity and robustness.

The readout measurement duration of contemporary and the PPO pulse shape optimization module are compared on a simulated 5-qubit chip. The resulting pulse shapes and measurement duration are presented in Figure 1-6. The PPO pulse optimization module yielded a reduction of the injection time by 22% and a 17% shorter reset time compared to the currently most advanced readout pulse shapes, the cavity level excitation, and reset pulse.

1.7 Summary & Outline

Solid-state qubits with transition frequencies in the microwave regime, such as superconducting qubits, are at the forefront of quantum computing. However, high-fidelity, simultaneous control of superconducting qubits at even a moderate scale remains a challenge. Among the significant error sources of contemporary superconducting quantum processors is qubit-state readout. For a single qubit, fast and high-fidelity readout has been demonstrated. However, in resource-efficient, frequency-multiplexed readout of multiple qubits, optimizing the readout pulse shape and discriminator becomes a computationally intensive task.

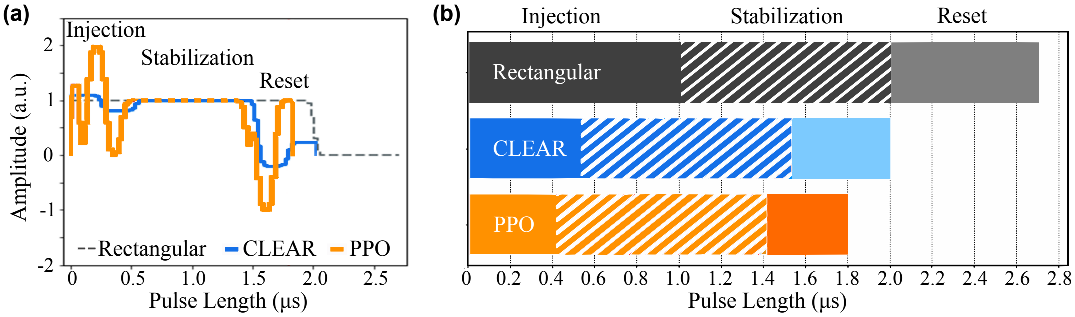


Figure 1-6: Simulation of Readout Pulse Shape versus Measurement Time for a 5-Qubit Chip. (a) Comparison of readout pulse shapes for a single qubit. The typically employed pulse shape, a rectangular pulse shape (gray, dashed), is compared with the cavity level excitation and reset (CLEAR) (in blue) and a pulse shape tune-up using the proximal policy optimization (PPO) algorithm (in red), a deep reinforcement learning algorithm. (b) Comparison of the duration to simultaneously inject photons in 5 different readout resonators, stabilize the photon counts during a waiting period, and return the resonators close to their vacuum states. The waiting period is kept constant for all three methods. In comparison with the rectangular pulse (CLEAR), PPO reduces the injection window by 58% (22%) and reset duration by 46% (17%).

This thesis focuses on hardware design and software techniques using microwave design principles and machine learning to achieve multi-qubit readout performing on the level of single-qubit operations, as illustrated in Figure 1-7. Principles to engineer microwave packages for multiple qubits to enable efficient control and readout while minimizing environment-induced errors are demonstrated using a novel multi-qubit package. Deep machine learning techniques to improve superconducting qubit readout pulse shapes and discrimination are experimentally demonstrated and compared to conventional methods. The presented techniques are a potential building block step towards implementing quantum error correction protocols and practical quantum computers.

After the concepts of quantum computing, Chapter 2, and in particular superconducting quantum computing, Chapter 3, have been introduced, microwave design principles to engineer packages with efficient qubit control and readout are discussed in Chapter 4. The focus shifts in Chapter 5 to the basic methods and ideas of machine learning before they are applied to readout discrimination in Chapter 6 and readout pulse shaping in Chapter 7. The thesis is concluded in Chapter 8.

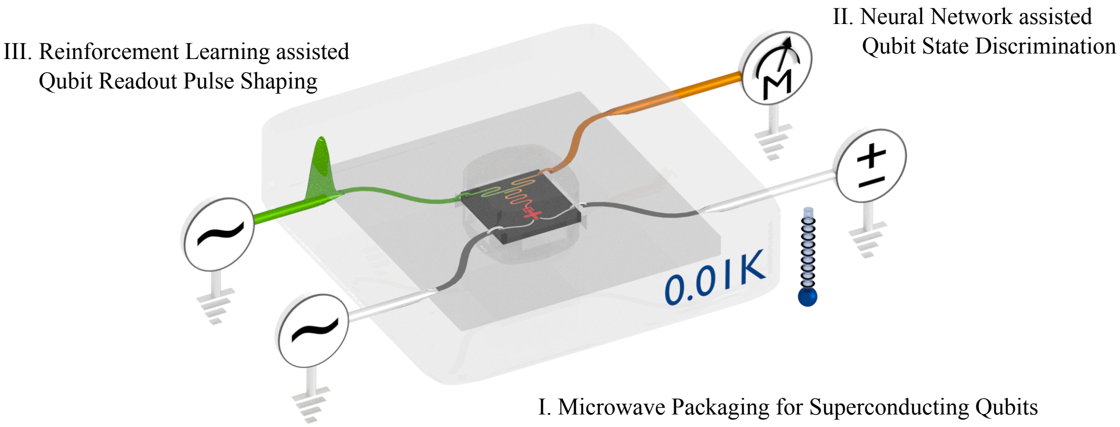

III. Reinforcement Learning assisted
Qubit Readout Pulse Shaping
II. Neural Network assisted
Qubit State Discrimination
0.01K
I. Microwave Packaging for Superconducting Qubits

Figure 1-7: Thesis Overview. Displayed is a microwave package with an enclosed qubit coupled via a resonator to a transmission line. The microwave package defines the immediate qubit environment and serves as a heat sink. The green pulse on the input side illustrates a readout pulse to probe the resonator state. The transmitted signal, colored in orange, is measured and discriminated against. This thesis focus on three aspects towards efficient multi-qubit readout and control: I. The development of microwave packages, II. Machine-learning assisted qubit-state discrimination, and III. Machine-learning assisted qubit-readout-pulse shaping.

References

[1] L. K. Grover, "A fast quantum mechanical algorithm for database search," Proceedings, 28th Annual ACM Symposium on the Theory of Computing, 212 (1996).

[2] P. W. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," Proceedings 35th annual symposium on foundations of computer science (1994), 124–134.

[3] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, "Quantum supremacy using a programmable superconducting processor," Nature **574**, 505–510 (2019).

[4] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, "Quantum computational advantage using photons," Science **370**, 1460–1463 (2020).

[5] J. Preskill, "Quantum Computing in the NISQ era and beyond," Quantum **2**, (2018).

[6] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, D. A. Buell, B. Burkett, N. Bushnell, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, S. Demura, A. Dunsworth, E. Farhi, A. Fowler, B. Foxen, C. Gidney, M. Giustina, R. Graff, S. Habegger, M. P. Harrigan, A. Ho, S. Hong, T. Huang, W. J. Huggins, L. Ioffe, S. V. Isakov, E. Jeffrey, Z. Jiang, C. Jones, D. Kafri, K. Kechedzhi, J. Kelly, S. Kim, P. V. Klimov, A. Korotkov, F. Kostritsa, D. Landhuis, P. Laptev, M. Lindmark, E. Lucero, O. Martin, J. M. Martinis, J. R. McClean, M. McEwen, A. Megrant, X. Mi, M. Mohseni, W. Mruczkiewicz, J. Mutus, O. Naaman, M. Neeley, C. Neill, H. Neven, M. Y. Niu, T. E. O'Brien, E. Ostby, A. Petukhov, H. Putterman, C. Quintana, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, D. Strain, K. J. Sung, M. Szalay, T. Y. Takeshita, A. Vainsencher, T. White, N. Wiebe, Z. J. Yao, P. Yeh, and A. Zalcman, "Hartree-fock on a superconducting qubit quantum computer," Science **369**, 1084–1089 (2020).

[7] J. M. Gambetta, J. M. Chow, and M. Steffen, "Building logical qubits in a superconducting quantum computing system," npj Quantum Information **3**, (2017).

[8] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "A quantum engineer's guide to superconducting qubits," Applied Physical Review **6**, (2019).

[9] C. D. Bruzewicz, J. Chiaverini, R. McConnell, and J. M. Sage, "Trapped-ion quantum computing: progress and challenges," Applied Physics Reviews **6**, 021314 (2019).

[10] M. Saffman, "Quantum computing with neutral atoms," National Science Review **6**, 24–25 (2018).

[11] D. M. Zajac, T. M. Hazard, X. Mi, E. Nielsen, and J. R. Petta, "Scalable gate architecture for a one-dimensional array of semiconductor spin qubits," Physical Review Applied **6**, 054013 (2016).

[12] N. H. Wan, T.-J. Lu, K. C. Chen, M. P. Walsh, M. E. Trusheim, L. De Santis, E. A. Bersin, I. B. Harris, S. L. Mouradian, I. R. Christen, E. S. Bielejec, and D. Englund, "Large-scale integration of artificial atoms in hybrid photonic circuits," Nature **583**, (2020).

[13] X. Gu, A. F. Kockum, A. Miranowicz, Y.-X. Liu, and F. Nori, "Microwave photonics with superconducting quantum circuits," Physics Reports **718-719**, 1–102 (2017).

[14] X. Y. Jin, A. Kamal, A. P. Sears, T. Gudmundsen, D. Hover, J. Miloshi, R. Slattery, F. Yan, J. Yoder, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "Thermal and residual excited-state population in a 3d transmon qubit," Physical Review Letters **114**, 240501 (2015).

[15] 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**, (2019).

[16] R. J. Schoelkopf and S. M. Girvin, "Wiring up quantum systems," Nature **451**, 664–669 (2008).

[17] 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**, 042319 (2007).

[18] W. D. Oliver and P. B. Welander, "Materials in Superconducting Qubits," MRS Bulletin **38**, 816–825 (2013).

[19] 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," Physical Review A **69**, 1–14 (2004).

[20] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, R. S. Huang, J. Majer, S. Kumar, S. M. Girvin, and R. J. Schoelkopf, "Strong coupling of a single photon

to a superconducting qubit using circuit quantum electrodynamics," Nature **431**, 162–167 (2004).

[21] E. Magesan, J. M. Gambetta, A. Còrcoles, and J. M. Chow, "Machine Learning for Discriminating Quantum Measurement Trajectories and Improving Readout," Physical Review Letters **114**, 200501 (2015).

[22] T. Walter, P. Kurpiers, S. Gasparinetti, P. Magnard, A. Potočnik, Y. Salathé, M. Pechal, M. Mondal, M. Oppliger, C. Eichler, and A. Wallraff, "Rapid High-Fidelity Single-Shot Dispersive Readout of Superconducting Qubits," Physical Review Applied **7**, 1–11 (2017).

[23] C. Macklin, K. O'Brien, D. Hover, M. E. Schwartz, V. Bolkhovsky, X. Zhang, W. D. Oliver, and I. Siddiqi, "A near–quantum-limited josephson traveling-wave parametric amplifier," Science **350**, 307–310 (2015).

[24] S. Sheldon, M. Sandberg, H. Paik, B. Abdo, J. M. Chow, M. Steffen, and J. M. Gambetta, "Characterization of hidden modes in networks of superconducting qubits," Applied Physics Letters **111**, 222601 (2017).

[25] D. M. Pozar, *Microwave engineering, 4th edition* (Wiley, 2012).

[26] B. Lienhard, J. Braumüller, W. Woods, D. Rosenberg, G. Calusine, S. Weber, A. Vepsäläinen, K. O'Brien, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "Microwave packaging for superconducting qubits," 2019 IEEE MTT-S International Microwave Symposium (2019).

[27] C. M. Quintana, A. Megrant, Z. Chen, A. Dunsworth, B. Chiaro, R. Barends, B. Campbell, Y. Chen, I.-C. Hoi, E. Jeffrey, J. Kelly, J. Y. Mutus, P. J. J. O'Malley, C. Neill, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Cleland, and J. M. Martinisa, "Characterization and reduction of microfabrication-induced decoherence in superconducting quantum circuits," Applied Physics Letters **105**, 062601 (2014).

[28] N. Fredj and T. D. Burleigh, “Transpassive Dissolution of Copper and Rapid Formation of Brilliant Colored Copper Oxide Films,” Journal of the Electrochemical Society **158**, C104–C110 (2011).

[29] J. E. Campbell, E. A. Eldridge, and J. K. Thompson, *Handbook on Materials for Superconducting Machinery* (Metals and Ceramics Information Center, 1974).

[30] D. R. Lide, *Handbook of Chemistry and Physics* (CRC Press, 1996), 11–41.

[31] A. Isidoro-Munoz, R. Torres-Torres, M. A. Tlaxcalteco-Matus, and G. Hernandez-Sosa, “Scalable models to represent the via-pad capacitance and via-traces inductance in multilayer pcb high-speed interconnects,” 2017 International Caribbean Conference on Devices, Circuits and Systems (ICCDCS), (2017).

[32] J.-E. Roh, J.-W. Li, B.-C. Ahn, C.-S. Park, and E.-J. Cha, “Novel approach to optimizing a broadband right-angle coaxial-to-microstrip transition,” Microwave and Optical Technology Letters **49**, 451–456 (2006).

[33] *Demystifying Vias in High-Speed PCB Design*, Keysight HSD Seminar, Mastering SI & PI Design, 2017.

[34] S. Huang, B. Lienhard, G. Calusine, A. Vepsäläinen, J. Braumüller, D. K. Kim, A. J. Melville, B. M. Niedzielski, J. L. Yoder, B. Kannan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, “Microwave package design for superconducting quantum processors,” (2021).

[35] C. M. Bishop, *Pattern recognition and machine learning (information science and statistics)* (Springer-Verlag, Berlin, Heidelberg, 2006).

[36] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” Proceedings of the IEEE **86**, 2278–2324 (1998).

[37] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, and D. Hassabis, “A general reinforcement learning algorithm that masters chess, shogi, and go through self-play,” Science **362**, 1140–1144 (2018).

[38] B. E. Boser, I. M. Guyon, and V. N. Vapnik, “A training algorithm for optimal margin classifiers,” Proceedings of the fifth annual workshop on computational learning theory, COLT ’92 (1992), 144–152.

[39] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning* (The MIT Press, 2016).

[40] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “Bert: pre-training of deep bidirectional transformers for language understanding,” (2019).

[41] S. Gu, E. Holly, T. Lillicrap, and S. Levine, “Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates,” 2017 ieee international conference on robotics and automation (icra) (2017), 3389–3396.

[42] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” (2017).

[43] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, “Human-level control through deep reinforcement learning,” Nature **518**, 529–533 (2015).

[44] T. Lillicrap, J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, “Continuous control with deep reinforcement learning,” CoRR **abs/1509.02971**, (2016).

[45] J. Schulman, S. Levine, P. Moritz, M. I. Jordan, and P. Abbeel, “Trust region policy optimization,” (2017).

[46] J. Heinsoo, C. K. Andersen, A. Remm, S. Krinner, T. Walter, Y. Salathé, S. Gasparinetti, J. C. Besse, A. Potočnik, A. Wallraff, and C. Eichler, “Rapid High-fidelity Multiplexed Readout of Superconducting Qubits,” Physical Review Applied **10**, 1–14 (2018).

[47] C. A. Ryan, B. R. Johnson, J. M. Gambetta, J. M. Chow, M. P. Da Silva, O. E. Dial, and T. A. Ohki, “Tomography via correlation of noisy measurement records,” Physical Review A **91**, 1–7 (2015).

[48] N. T. Bronn, B. Abdo, K. Inoue, S. Lekuch, A. D. Còrcoles, J. B. Hertzberg, M. Takita, L. S. Bishop, J. M. Gambetta, and J. M. Chow, “Fast, high-fidelity readout of multiple qubitss,” Journal of Physics: Conference Series **834**, 012003 (2017).

[49] A. Seif, K. A. Landsman, N. M. Linke, C. Figgatt, C. Monroe, and M. Hafezi, “Machine learning assisted readout of trapped-ion qubits,” Journal of Physics B: Atomic, Molecular and Optical Physics **51**, 174006 (2018).

[50] Z.-H. Ding, J.-M. Cui, Y.-F. Huang, C.-F. Li, T. Tu, and G.-C. Guo, “Fast High-Fidelity Readout of a Single Trapped-Ion Qubit via Machine-Learning Methods,” Physical Review Applied **12**, 014038 (2019).

[51] B. Lienhard, A. Vepsäläinen, L. C. G. Govia, C. R. Hoffer, J. Y. Qiu, D. Ristè, M. Ware, D. Kim, R. Winik, A. Melville, B. Niedzielski, J. Yoder, G. J. Ribeill, T. A. Ohki, H. K. Krovi, T. P. Orlando, S. Gustavsson, and W. D. Oliver, “Deep neural network discrimination of multiplexed superconducting qubit states,” (2021).

[52] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, and A. A.-G. amd Jeremy L. O’Brien, “A variational eigenvalue solver on a photonic quantum processors,” Nature Communications **4213**, (2014).

[53] E. Farhi, J. Goldstone, and S. Gutmann, “A quantum approximate optimization algorithms,” (2014).

[54] D. T. McClure, H. Paik, L. S. Bishop, M. Steffen, J. M. Chow, and J. M. Gambetta, “Rapid Driven Reset of a Qubit Readout Resonator,” Physical Review Applied **5**, 1–5 (2016).

[55] S. Boutin, C. K. Andersen, J. Venkatraman, A. J. Ferris, and A. Blais, "Resonator reset in circuit QED by optimal control for large open quantum systems," Physical Review A **96**, 1–11 (2017).

[56] M. Dalgaard, F. Motzoi, J. J. Sørensen, and J. Sherson, "Global optimization of quantum dynamics with alphazero deep exploration," npj Quantum Information **6**, 6 (2020).

Chapter 2

Quantum Computing

The realization of practical large-scale quantum computers promises a new era of computational capabilities. Quantum computers process information in a fundamentally different way than classical computers. The promise of quantum computation lies in using unique features of quantum mechanics to provide enhanced computational performance over classical computers for specific tasks, such as simulations of small-scale physical, chemical, or biological systems, financial services and technology, or machine learning and optimization problems.

This chapter provides a brief introduction to quantum computing in Section 2.1. After that, the computational power of quantum computers is put in context in Section 2.2 before its development is compared with classical computers in Section 2.3. Finally, a brief introduction to some of the most prominent qubit modalities and imminent challenges follows in Section 2.4.

2.1 Introduction Into Quantum Computing

Quantum mechanics is a framework able to describe specific phenomena, such as quantum entanglement. Classical physics cannot describe that. The idea to harness the peculiar effects of quantum physics to process information originated in the early 1980s [1, 2]. Intensive research since revealed that such quantum computing devices promise to solve tasks such as factoring large numbers [3], combinatorial searching [4], or simulations of quantum systems [5–7] significantly more efficient than present-day Boolean logic-based computers (henceforward referred to as classical computers).

2.1.1 Qubits

The basic information unit of classical computing is a **b**inary dig**it**, or simply a bit. A bit is an entity with two discrete states: 0 or a 1. A common physical implementation to store a bit is a transistor [8]. State 0 can mean the transistor is off, and no current is present. Consequently, state 1 means a current is flowing. These discrete states are robust and can be measured.

Quantum bits (qubits) are the classical bit's quantum analog. However, unlike bits, qubits can be in a superposition of 0 and 1. The quantum state of a single qubit can be written as a wave function

$$|\Psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle\,, \tag{2.1}$$

with complex probability amplitudes α and β constrained by $|\alpha|^2 + |\beta|^2 = 1$. The computational states $|0\rangle$ and $|1\rangle$ in bracket notation can be expressed as vectors $[10]^T$ and $[01]^T$. Graphically, such a quantum state can be represented by a Bloch vector on the Bloch sphere [9] as illustrated in Figure 2-1.

For a quantum system with N qubits, a complex superposition of all 2^N permu-

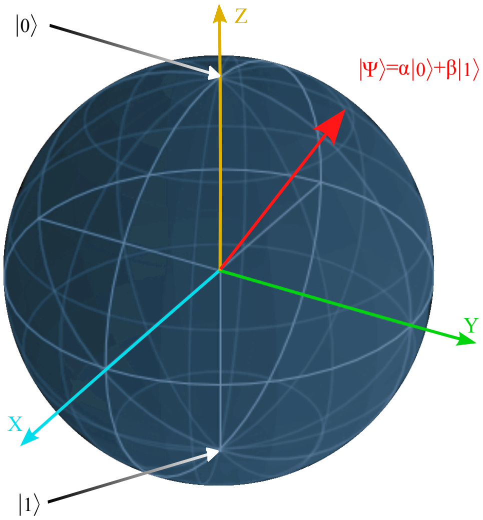


Figure 2-1: Bloch Sphere. The unit sphere depicts the pure states of a qubit. The north pole and south pole along the Z-axis represent the computational basis with $|0\rangle$ and $|1\rangle$. The Bloch vector, shown in red, represents an arbitrary quantum state $|\Psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$. α and β are complex probability amplitudes with the constraint $|\alpha|^2 + |\beta|^2 = 1$. Quantum operations can alter the qubit state and thus the Bloch vector.

tations, results in a wave function of the system's entire state

$$|\Psi\rangle = c_1\,|0\ldots00\rangle + c_2\,|0\ldots01\rangle + \ldots. + c_{2^N}\,|1\ldots11\rangle \tag{2.2}$$

and $\sum_i |c_i|^2 = 1$. $|0\ldots00\rangle$ denotes the tensor product of N qubits $|0\rangle \otimes \ldots |0\rangle \otimes |0\rangle$.

Measuring a qubit converts the quantum information into classical information. The result of a measurement is a probabilistic projection of the qubit onto its computational states 0 or 1, typically along the Z-axis. For an equal superposition, the probability to measure 0 or 1 is $50\,\%$ for an equal superposition. Unlike in classical computing, the quantum information stored in the qubit cannot generally be copied [10]. Hence, to infer the qubit's probabilistic distribution and thus the degree of superposition, multiple identically prepared states have to be measured. Consequently, quantum computers rely on encoding information in fundamentally different ways than classical computers.

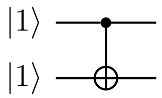
2.1.2 Qubit Gates

To encode arbitrary quantum algorithms, a set of quantum operations are necessary. A universal gate set represents a finite set of operations with which any possible computational operation can be constructed. For example, classical computers can perform arbitrary Boolean logic [11] with merely a few types of logic gates. A universal gate set can be formed with the NOT gate, a single-bit gate inverting a bit, and the AND gate, a two-bit gate outputting 1 if both inputs are 1, and 0 otherwise.

Universal quantum computing requires such a universal gate set as well. However, while some quantum gates have a direct classical analog, most gates differ from classical gates. A key difference is that all quantum gates are reversible. They are so-called unitary gates. The quantum analog of the classical NOT gate is an X-gate, swapping the states $|0\rangle$ to $|1\rangle$. On the Bloch sphere, this operation can be visualized with a 180-degree rotation around the X-axis and is thus often referred to as a π-pulse. Consequently, 180-degree rotations around the Y-axis and Z-axis are called Y-gate and Z-gate. The X-, Y-, and Z-gate are commonly referred to as Pauli gates [12].

In addition to the Pauli gates, Hadamard, S-, and T-gates are common quantum algorithms as single-qubit gates [12]. Hadamard gates put a qubit initialized in a computational basis state into an equal superposition. Similar to the Z-gate, S- and T-gates are phase shift gates. Phase shift gates add a phase of $\exp\{i\phi\}$ to $|1\rangle$ while $|0\rangle$ remains unaffected. ϕ is equal to π for the Z-gate, $\pi/2$ for the S-gate, and $\pi/4$ for the T-gate.

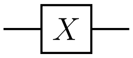
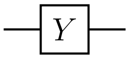
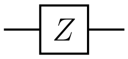
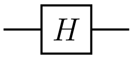
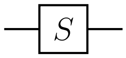
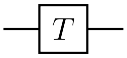
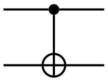
To have a universal gate set, single-bit or single-qubit gates are not sufficient. The quantum analog to the classical exclusive OR, abbreviated XOR-gate, is the Controlled-NOT gate, or CNOT-gate, composed of a controlling qubit and a target qubit [12]. If the controlling qubit is in state $|0\rangle$, then the target qubit is not affected. On the contrary, if the controlling qubit is in state $|1\rangle$, then the target qubit experiences an X-gate.



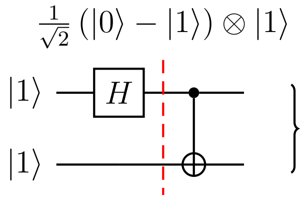
$$\Big\}\ |10\rangle = |1\rangle \otimes |0\rangle \quad \Rightarrow \quad \text{product state} \tag{2.3}$$

Note that the control and target qubit is entangled after a CNOT-gate is applied for a control qubit in a superposition state. Entanglement is a unique feature of

Table 2.1: Quantum Gates [13]. A universal quantum gate set can be formed of single- and two-qubit gates. Common single-qubit gates are the three Pauli-gates, the Hadamard gate, and the phase shifting gates S and T. A frequently used two-qubit gate is the Controlled-NOT (CNOT) gate. The computational basis states are expressed as $|0\rangle = \begin{bmatrix}1 & 0\end{bmatrix}^T$ and $|1\rangle = \begin{bmatrix}0 & 1\end{bmatrix}^T$.

	Gate Name	Gate Symbol	Matrix
Single Qubit Gates	Pauli-X (X)		$\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$
	Pauli-Y (Y)		$\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}$
	Pauli-Z (Z)		$\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$
	Hadamard (H)		$\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{bmatrix}$
	Phase (S)		$\begin{bmatrix}1 & 0\\ 0 & i\end{bmatrix}$
	$\pi/8$ (T)		$\begin{bmatrix}1 & 0\\ 0 & e^{i\pi/4}\end{bmatrix}$
Two-qubit Gate	Controlled-NOT (CNOT)		$\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{bmatrix}$

quantum physics. Thus, the previously separable quantum state can no longer be factorized into two separate qubit components, as seen by comparing Equation 2.3 and Equation 2.4.


$\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \otimes |1\rangle$
$|1\rangle$
H
$|1\rangle$

$$\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \quad \Rightarrow \quad \text{entangled state} \tag{2.4}$$

Universal quantum computation can be built from a small subset of these types of single and two-qubit gates. Table 2.1 lists a set of quantum gates to create a universal gate set. A commonly used universal gate set comprises single-qubit rotations around X, Y, and Z, a single-qubit phase shift gate, and a CNOT-gate as the two-qubit gate or CNOT-, H-, S- and T-gates [14].

2.1.3 Quantum Parallelism & Interference

Two subsequent fundamental principles substantiate the power of quantum computation and distinguish itself from classical computers: Quantum parallelism and interference.

Quantum Parallelism

A classical N-bit state is uniquely defined. Processing two distinct N-bit input states can either be done subsequently or in parallel. Therefore, the computational effort either doubles in time or hardware resources. The qubits in a quantum processor, on the other hand, can be set into a single superposition state that simultaneously carries aspects of all these 2^N components. Consequently, no additional resources are needed. This effect is referred to as quantum parallelism.

Quantum Interference

Quantum interference describes the effect of constructive and destructive interference between quantum states and alters the probability amplitudes. Therefore, specific quantum states become more likely to be measured than others. A quantum algorithm intends to enhance the probability to measure the desired state—the state that represents the best answer to a computational task—and thus biases the measurement using the effect of quantum interference.

2.2 Complexity Theory

Computational problems are classified based on the best algorithms known today. The different complexity classes arise due to the different scaling laws as a function of problem size for the physical resources and the time requirements for these best-known algorithms. For instance, computing devices comprise a fixed number of physical memory units and a processor that can perform a certain number of computational steps per time unit. The physical memory of a computing device places an upper bound on the maximum, manageable problem complexity. The number of elementary computational steps required to solve a task can be used as a quantitative measure of time. In general, computational tasks can be classified by their memory consumption, their required temporal effort to evaluate a solution, and the time needed to verify the proposed solution.

2.2.1 Classical Complexity Theory

Usually, one desires algorithms that scale polynomial in time and memory as a particular computational task's complexity increases. The complexity class describing problems with polynomial memory requirements that are independent of the respective time complexity is referred to as **PSPACE** [15]. Computational problems with algorithms able to compute solutions in polynomial time constitute the computational complexity class **P** [15]. Problems in class **P** can be efficiently computed on a classical computer.

A computational task without a known algorithm that scales polynomial in time is the traveling salesman problem. The traveling salesman describes the challenge of identifying the shortest route between multiple cities to visit each city only once. This task's time consumption scales exponentially in the number of cities.

A classical computer is typically unable to solve problems with exponential time requirements efficiently. However, once a solution is presented, it may only take polynomial computational time to confirm the solution. For instance, factoring an integer into two prime numbers is an exponentially hard problem, in general. However, verifying the proposed solution involves a single multiplication that scales polynomially in time. Problems with solutions that can be verified in polynomial computational time form the **NP** complexity class (NP stands for **N**on-deterministic **P**olynomial time) [15]. Among the problems that can be efficiently evaluated are the problems in class **P**. If an efficient algorithm proposes a solution to a problem, an efficient algorithm exists to verify the solution.

The most complex computational problems in **NP** are those for which the computation is exponential in time. Therefore, the task is not in **P** and instead assigned to the **NP-Complete** complexity class. For example, the traveling salesman is an **NP-Complete** problem [15]. It is exponentially hard to find a solution, but it only takes polynomial time to evaluate the proposed solution. A proposed solution can be easily verified by depicting the proposed solution on a map and verifying that the conditions are not violated.

It is an open question if algorithms exist that can solve **NP-Complete** problems in polynomial time. If such algorithms exist, then **NP** would consequently collapse into **P**, meaning they are the same complexity classes, **N=NP** [16]. The search for such algorithms has continued for more than half a century. The inability to find such an algorithm leads most computer scientists to believe that indeed **N≠NP** [17]. Figure 2-2 summarizes the hierarchy of the four introduced complexity classes. Classical computers can efficiently solve computational tasks part of **P**.

2.2.2 Quantum Complexity Theory

Classical computers can efficiently solve problems in the computational complexity class **P**. A computational device able to solve exponentially challenging problems in **NP-Complete** and even more complex problems could be tremendously beneficial for scientific discovery and, in particular to optimization tasks. In the following, the class of problems a quantum computer can solve efficiently is embedded in the framework of the complexity classes **P**, **NP**, **NP-Complete**, and **PSPACE**.

Quantum computers outperform classical computers for specific problems, such as factoring large integer numbers [3], a problem that is believed to be part of **NP** [18]. A quantum computer's advantage arises from exploiting the problem's underlying mathematical structure differently from classical computers. In addition, the quantum computer can determine the prime factors of a large integer number in polynomial time, which is exponentially faster than it takes classical computers to perform the same task.

For example, Shor's factoring algorithm shows that quantum computers can outperform classical computers on problems that feature a mathematical structure that the working principles of quantum computers can exploit [19]. **NP-Complete** problems do not seem to exhibit these sort of favorable mathematical structures. There is no evidence to date that suggests quantum algorithms exist that could outperform classical algorithms for **NP-Complete** problems.

There exist other quantum algorithms able to generate minor speedups over the best known classical algorithm for problems in **N** and **NP** [4]. Nevertheless, it remains an open question whether quantum computers can efficiently solve **NP-Complete** problems. There are proposals for quantum algorithms that may be able to verify a proposed solution in polynomial time for problems outside of **NP** [20]. However, there is no evidence that quantum computers can outperform classical computers on problems outside of **PSPACE** [21]. The problems quantum computers can solve efficiently form the computational complexity class **BQP**, with **BQP** being an abbreviation for ***B**ounded-error **Q**uantum **P**olynomial time* [12].

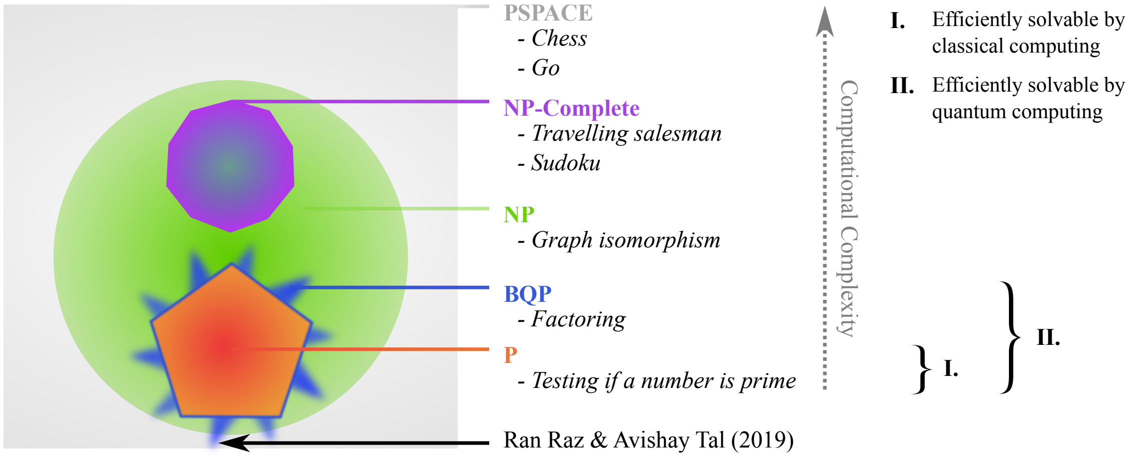

PSPACE
- Chess
- Go
NP-Complete
- Travelling salesman
- Sudoku
NP
- Graph isomorphism
BQP
- Factoring
P
- Testing if a number is prime
Ran Raz & Avishay Tal (2019)
Computational Complexity
I. Efficiently solvable by classical computing
II. Efficiently solvable by quantum computing
I.
II.

Figure 2-2: Computational Complexity Classes. **PSPACE** contains all classes and problems which require a polynomial amount of memory on a conventional computer independent of the number of necessary computational steps to find or verify a solution. **P** describes the class of classically efficient computable problems. Problems part of the complexity class **BQP** are efficiently solvable by a quantum computer. Class **NP** contains problems that are efficiently verifiable. The hardest problems are still efficiently verifiable from complexity class **NP-Complete**. In 2019, Ran Raz and Avishay Tal identified the first problem outside the complexity class **NP** that a quantum computer can solve efficiently.

Quantum computers are more powerful than classical computers for specific problems. However, there is no formal proof that the hierarchy shown in Figure 2-2 is manifestly true. Meaning, there does not exist proof yet describing the limits of classical algorithms. Thus, a believed quantum advantage today may not be a quantum advantage tomorrow.

2.3 Computing History

A closer look at the history of classical computing can help gauge the future development of quantum computers. An overview listing some of the most relevant milestones per decade can be found in Table 2.2. The development of modern electronic computers started with the invention of the vacuum tube [22] at the beginning of the twentieth century that eventually led to the construction of a large-scale computational system, the Electronic Numerical Integrator and Computer (ENIAC) [23], about 40 years later. The first large-scale transistor-based computer, the Transistorized Experimental Computer Zero (TX-0) [24], was developed only ten years after the conception of the transistor in 1947 [8]. In 1959, the first silicon-based integrated circuit was demonstrated that led to the commercialization of the first monolithic processor with 2300 transistors [25]; however, only two years passed until the next generation doubled the number of transistors. This doubling of the number of transistors approximately every two years continued and became known as Moore's Law [26]. As a result, today's computer chips are composed of more than a billion transistors. Although performance increases had previously followed from this Moore's law type scaling, physical limitations have slowed down these improvements over the last decade.

In contrast, quantum computing is a much more recent technology. In the early 1980s, Richard Feynman suggested that quantum systems are fundamentally better suited to simulate quantum systems than classical computers [2]. Researchers have explored different algorithmic avenues that could provide such a quantum advan-

Table 2.2: Computing Milestones Per Decade. Key decade milestones in the chronology of classical and quantum computing development between 1900 and 2020.

Classical Computing

1900	Invention of vacuum tube [22]
1910	First flip-flop circuit design [27]
1920	Demonstration of AND logic gate [28]
1930	First mechanical binary programmable computer (Z1) [29]
1940	First electronic programmable computer (ENIAC) [23] & invention of transistor
1950	First transistor based computer (TX-0) [24]
1960	Transistor-transistor logic developed [30]
1970	2025 Transistors (Intel 4004) [25]
1980	10^6 Transistors (Intel 80486) [31]
1990	10^7 Transistors (Hitachi SH-4) [31]
2000	10^8 Transistor, dual-core (Intel Core 2 Duo) [31]
2010	10^9 Transistor, quad-core (Intel Core i7) [31]
2020	10^{10} Transistors, octa-core (Apple M1) [31]

Quantum Computing

1980	Proposal of a quantum computer [1, 2]
1990	Factoring algorithm [3] & quantum error correction [32]
2000	12-qubit quantum information processor [33]
2010	Quantum supremacy with 53 superconducting qubits [34]
2020	Quantum supremacy with up to 76 photons [35]

tage and the role of quantum states' inherent fragility. David Deutsch proposed the first simple quantum algorithm in 1985 [36]. A fundamental breakthrough was Peter Shor's proposed algorithm to factorize large numbers [3, 37], a challenging problem for classical computers. Shortly after, again, Peter Shor and Robert Calderbank, and Andrew Steane developed the first protocols to detect and correct qubit errors [32, 38, 39]. For the last five years, about a dozen industrial and academic institutions have demonstrated quantum devices with more than ten qubits [34, 35, 40]. In 2019, a quantum advantage was demonstrated for the first time by a team at Google [34]. A quantum advantage is achieved if a problem can be solved by a quantum computer significantly faster than a classical computer [34].

Current quantum processors comprise tens of qubits that are not yet of the quality to execute valuable computations. However, in 2019 one of the largest programmable quantum processors with 53 working superconducting qubits achieved a significant quantum computing milestone, quantum supremacy [34]. Towards the end of 2020, a photonic quantum processor platform using up to 76 photons reached the same milestone [35]. Quantum supremacy describes completing a computational task on a quantum computer that cannot be performed on existing classical computers in a reasonable amount of time [41]. Despite the importance of this milestone, the required technology to build a universal and fault-tolerant quantum processor is likely at least a decade away. In the meantime, today's noisy intermediate-scale quantum devices, often referred to as NISQ devices [41], offer proof of concept demonstrations such as the simulation of molecular electronic configurations [42].

2.4 Quantum Computing Challenges

A viable physical qubit consists of a pair of addressable quantum levels, offers channels to initialize and prepare specific quantum states, techniques to induce a unitary evolution, and methods to measure the final qubit state. Furthermore, to build a quantum processor, multiple qubits have to be coupled. Finally, control and readout techniques need to be sufficiently sophisticated to maintain quantum coherence, such that the quantum information is not degraded and lost.

As illustrated in Figure 2-3, nature provides numerous examples of quantum mechanical two-level systems that potentially could serve as qubits. Examples of quantum mechanical two-level systems or artificial atoms include the electronic states of an ion, photon polarization states, or nuclear spins [12]. David DiVincenzo articulated five fundamental requirements for any qubit technology to be a suitable physical implementation for large-scale quantum computation [43]:

1. The physical system provides well-defined and characterized qubits.
2. The system allows the initialization of individual qubits.
3. There exists a mechanism to read out qubits.
4. A universal set of gate operations is available and can be performed.
5. The qubits can robustly present and store quantum information.

These criteria, known as the DiVincenzo Criteria, are in many ways adapted from the conditions for operational classical computers and summarize the fundamental requirements qubit technologies need to fulfill to guarantee successful quantum computation.

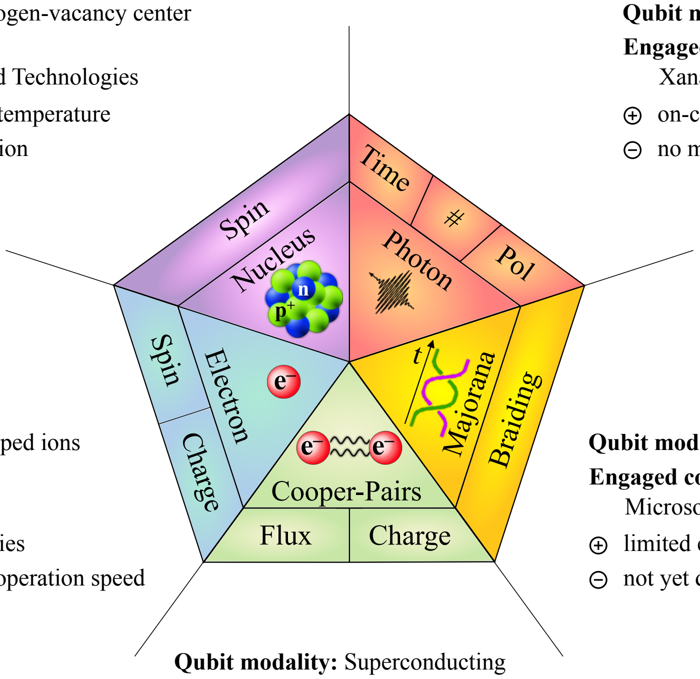
2.4.1 Qubit Modalities

Various qubit modalities comply with DiVinceno's Criteria, such as superconducting qubits [44], trapped ions [45], trapped neutral atoms [46], spin qubits [47], or solid-state quantum emitters [48]. Figure 2-3 shows some of the most prominent qubit modalities pursued by research and industry. Each of these modalities has its strength and drawbacks and hence applications.

For instance, atoms and atom-like quantum emitters play central roles in many areas of photonics-based quantum information processing. In recent years, there has been tremendous progress in developing quantum emitter systems based on

Qubit modality: Nitrogen-vacancy center
Engaged companies:
Quantum Diamond Technologies
⊕ operation at room temperature
⊖ deterministic creation

Qubit modality: Photonic
Engaged companies:
Xanadu, PsiQuantum
⊕ on-chip integration
⊖ no memory



Qubit modality: Trapped ions
Engaged companies:
IonQ, Honeywell
⊕ highest gate fidelities
⊖ microsecond gate operation speed

Qubit modality: Topological
Engaged companies:
Microsoft
⊕ limited error susceptibility
⊖ not yet demonstrated

Qubit modality: Superconducting
Engaged companies:
Google, IBM, Rigetti, ...
⊕ nanosecond gate operation speed
⊖ noise susceptibility

Figure 2-3: Physical Realization of Qubits. Shown are some of the most prominent physical quantum mechanical systems used for qubit implementations ranging from photons to electrons to topological states. A photonic qubit can, for example, be encoded as time bins, in the number of photons (#), or using its polarization states. Each quantum mechanical system and its qubit implementation have advantages and disadvantages. For instance, while Majorana qubits promise to be resilient to many noise sources, a physical qubit has not been realized yet. In this thesis, the focus is on superconducting qubits using superconducting loops. Although quantum computing is still in its infancy, the commercialization of quantum information processing hardware and software has grown from a few to many companies over the last decade. Together with trapped ions, superconducting qubits are among the most established and industrially pursued qubit modalities today.

crystallographic defects in wide-bandgap semiconductors ranging from diamond [49, 50] to silicon carbide [51] to III-nitrides [52, 53]. Today, solid-state quantum emitters embedded in photonic structures [54, 55] or circuits [48] are among the most promising candidates for quantum repeaters, nanoscale sensors, and single-photon sources.

Superconducting qubits are at the forefront of large-scale quantum computers [34]. Over the past two decades, superconducting qubits have emerged as a leading quantum computing platform [44, 56]. Today, individual qubits with coherence times exceeding $100\,\mu\mathrm{s}$ [57], gate times of a few tens of nanoseconds [34], and individual single- and two-qubit gate operation fidelities above the most lenient thresholds for quantum error correction have been demonstrated for devices with up to 50 qubits [34, 58]. However, considerable work is still needed to retain and even further improve these fidelities as systems increase in size and complexity [59]. This thesis focuses on superconducting qubits, particularly their microwave environment, and how to efficiently interact with them and interpret their response using machine learning as the number of qubits is scaled up.

2.4.2 Qubit Coherence

The DiVincenzo Criteria articulated the requirements for a qubit technology to qualify as a viable candidate for the physical implementation of a quantum computer. To satisfy DiVincenco's Criterion 2-4, a communication channel between the qubit and a control instance is needed to execute qubit operations and measurements. The requirement of this channel limits the degree to which unintentional interactions between the qubit and its environment can be suppressed. Therefore, a measurable qubit always couples to the environment to some extent. The coupling causes the qubit at some point in time to lose its initial well-defined quantum information to the environment, thus resulting in an unrecognizable qubit state. Consequently, an initialized qubit is unable to store quantum information for an indefinite amount of time. DiVincenzo's fifth criterion describes the minimum required time window a qubit can be read out without losing the quantum information completely.

There are two fundamental ways in which a qubit loses quantum information [12]. The first is energy relaxation. A qubit in an excited state relaxes eventually back to the ground state and loses its energy to the environment. The relaxation process can be graphically illustrated on the Bloch sphere as a flip of the Bloch vector to the ground state $|0\rangle$, the north pole. This time is referred to as T_1 or longitudinal relaxation rate $\Gamma_1 \triangleq T_1^{-1}$. Secondly, a qubit can lose quantum information is through a loss of phase coherence, also known as decoherence. Decoherence, the transverse relaxation rate Γ_2, is a combination of pure dephasing Γ_ϕ and energy relaxation—a phase breaking process. Γ_ϕ is the average time after which environmental noise moved the qubit state around the equator on the Bloch sphere in a random non-repetitive way to an unrecognizable extent. For two-level systems weakly coupled to noise sources, the dynamics can be described by $\Gamma_2 \triangleq 1/T_2^* = \Gamma_1/2 + \Gamma_\phi$ following the Bloch-Redfield formalism [60, 61]. In conclusion, a qubit can lose quantum information via two mechanisms, energy relaxation and loss of phase coherence after time T_1 and T_2^*, respectively. The fifth DiVincenzo Criterion requires these two mechanisms to be sufficiently suppressed to guarantee the manipulation of the quantum information via gates.

2.4.3 Gate Time and Fidelity

Just as with classical computers, quantum computers require a sufficiently high clock speed, the time required to perform a quantum operation, the gate time with which we can operate the quantum computer. A key figure of merit is the number of gates one can perform within the qubit's lifetime. The more gates one can implement before an error occurs, the larger an algorithm one can run. A qubit with a long lifetime and little interaction with the environment is not as easy to control and causes longer gate times. In contrast, qubits that are strongly interacting with the controlling fields tend to interact more strongly with the environment and are short-lived. The number of gates one can perform, on average, before an error occurs may not differ much between these two cases. However, one of these qubit modalities may have a much faster clock speed than the other one, which is, in general, the preferred configuration.

Decoherence is not the only limitation qubits can experience but also errors due to control errors and imperfections in the pulses used to drive a gate operation. The concept of gate fidelity is a rigorous, more general characterization of how well a gate operation on a qubit works. The gate fidelity is a comparison of the expected ideal with the resulting output state [12]. Gate errors may be fundamentally different, depending on the initialized qubit's starting point on the Bloch sphere. Therefore, the gate performance needs to be measured concerning the initial qubit state set that spans the entire qubit state space, in the case of a single qubit gate, the Bloch sphere. Any location on the Bloch sphere, as a two-dimensional surface, can represent two distinct vector components. Therefore, it is sufficient to map out the single-qubit gate fidelity with two orthogonal qubit state inputs [12].

Generally, N qubits acting in 2^{2N} dimensions require 2^{2N} orthogonal qubit states. To evaluate the gate fidelity of an N-qubit gate, it is sufficient to apply the N-qubit gate on 2^{2N} orthogonal qubit states, measure them out and compare the resulting the expected ideal output state. The gate fidelity is then the average performance over the entire qubit state space. It is referred to as process tomography [12], representing a complete description of errors during a gate operation. Although single and coupled qubit gates are the only gates of importance, reality requires that process tomography needs to consider all present qubits due to leakage between them. Considering that the implementation of process tomography of an N-qubit gate requires 2^{2N} input states and hence the same amount of measurements, it becomes evident that due to the poor scaling with the number of qubits, process tomography can be impractical. Furthermore, process tomography is sensitive to all errors and does not solely determine the gate operation's quality.

An alternative approach is randomized benchmarking [62]. Randomized benchmarking characterizes the performance of the gate operation by including a random but known assortment of other gates in the evaluation as a type of normalization.

The system is characterized by applying a random but known gate sequence and comparing the expected ideal with the resulting output state. After that, adding the gate of interest to the random gate sequence and the information of the performance of the entirely random gate sequence allows deducting the error rate contribution due to the gate operation of interest. Repeating this approach with an increasing number of gates yields an improved estimate of the average error gate rate and the requested gate fidelity.

Randomized benchmarking is, in comparison with process tomography, more efficient and insensitive to initialization and measurement errors but, in return, provides only a net error rate without any information on specific error channels. A perfect match of the actual operation and the flawless operation results in a gate fidelity of 100%, such as an exact π-rotation around the x-axis of the Bloch sphere in case of an X-gate. It is not possible to achieve perfect gate fidelities in general. Nevertheless, it is required to improve the gate performance to surpass a minimum value, called the threshold, to enable efficient quantum error correction and limit the required resource overhead.

2.4.4 Quantum Error Correction

Error-free transmission of information is physically impossible. Therefore, to minimize the influence of errors, signal loss and noise are required to be minimized. In classical communication, loss of information is a minor issue since the number of information carriers can be arbitrarily increased. Furthermore, the influence of noise is exceedingly negligible in digital communication because signals are constantly corrected to the closest discrete bit value (0 or 1).

Quantum information processing is particularly prone to errors due to the high sensitivity to the environment. The detrimental interaction between the qubit and the environment can be mitigated with error correction protocols. For a practical quantum processor, the error probability per gate and qubit must be kept below a certain threshold for quantum error correction protocols to work. This threshold is reached when on average less than 10,000 operations are necessary to encounter the first error [12].

Motivated by classical error correction, the redundancy of information repetition codes is the most fundamental error correction code for quantum information processing. The no-cloning theorem [10] forbids simple copying of information; hence a serial, repetitive transmission of the same quantum information and a final statistical analysis yields with increased probability the correct result, so-called majority voting [32]. In addition to such active error correction, passive error correction can reduce the susceptibility of qubit errors, particularly systematic and coherent stochastic errors [12].

Typically, active quantum error correction protocols use syndrome measurements to detect if and where an error occurred. Detected errors can then be corrected using a feed-forward approach, in which pulses are applied to the errant qubits to correct the error [12]. Therefore, accurate and resource-efficient qubit-state readout is fundamental to identifying errant qubits and thus quantum error correction. Moreover, qubit readout needs to be fast relative to the qubit lifetime such that several rounds of error correction can be conducted. Most qubit modalities do not yet offer methods to determine a qubit state with sufficiently high accuracy and speed.

2.5 Summary

Universal quantum computation aims to perform specific computational problems such as integer factorization [3] or database search [4] in a significantly more efficient way than classical computers. The challenge of developing quantum computer hardware is to balance the system's isolation while maintaining efficient control and readout channels. Towards the realization of quantum computers, many milestones have already been achieved. However, many engineering challenges remain, and realizing a functional large-scale quantum computer is likely years away. This thesis focuses on scaling up quantum processors' control and readout using microwave design principles and machine learning. The developed readout techniques are a significant step towards efficient implementations of near-term quantum algorithms based on iterative optimization and quantum error correction protocols necessary for future practical universal quantum processors.

References

[1] Y. I. Manin, *Vychislimoe i nevychislimoe [Computable and Noncomputable] (in Russian). Sov.Radio.* 1980.

[2] R. P. Feynman, “Simulating physics with computers,” International Journal of Theoretical Physics **21**, 467–488 (1982).

[3] P. W. Shor, “Algorithms for quantum computation: discrete logarithms and factoring,” Proceedings 35th annual symposium on foundations of computer science (1994), 124–134.

[4] L. K. Grover, “A fast quantum mechanical algorithm for database search,” Proceedings, 28th Annual ACM Symposium on the Theory of Computing, 212 (1996).

[5] S. Lloyd, “Universal quantum simulators,” Science **273**, 1073–1078 (1996).

[6] I. M. Georgescu, S. Ashhab, and F. Nori, “Quantum simulation,” Reviews of Modern Physics **86**, 153–185 (2014).

[7] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, and M. Troyer, “Elucidating reaction mechanisms on quantum computers,” Proceedings of the National Academy of Sciences **114**, 7555–7560 (2017).

[8] W. D. Hackmann, *Transistor* (Oxford University Press, 2003).

[9] F. Bloch, “Nuclear induction,” Physical Review **70**, 460–474 (1946).

[10] W. K. Wootters and W. H. Zurek, “A single quantum cannot be cloned,” Nature **299**, 1476–4687 (1982).

[11] G. Boole, *The mathematical analysis of logic: being an essay towards a calculus of deductive reasoning*, Cambridge Library Collection - Mathematics (Cambridge University Press, 2009).

[12] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary edition* (Cambridge University Press, USA, 2011).

[13] A. Kay, “Tutorial on the quantikz package,” (2018).

[14] C. P. Williams, *Quantum gates* (Springer London, London, 2011), 51–122.

[15] J. Watrous, *Quantum computational complexity*, edited by R. A. Meyers (Springer New York, New York, NY, 2009), 7174–7201.

[16] S. A. Cook, "The complexity of theorem-proving procedures," Proceedings of the third annual acm symposium on theory of computing, STOC '71 (1971), 151–158.

[17] S. Aaronson, *P=np*, https://www.scottaaronson.com/papers/pnp.pdf, `https://www.scottaaronson.com/papers/pnp.pdf`.

[18] T. Gowers, J. Barrow-Green, and I. Leader, *The princeton companion to mathematics*, 1st ed. (Princeton University Press, Princeton, NJ, 2008, 2008).

[19] S. Aaronson, "Quantum computing and hidden variables," Phys. Rev. A **71**, 032325 (2005).

[20] R. Raz and A. Tal, "Oracle separation of bqp and ph," Proceedings of the 51st annual acm sigact symposium on theory of computing, STOC 2019 (2019), 13–23.

[21] E. Bernstein and U. Vazirani, "Quantum complexity theory," In proc. 25th annual acm symposium on theory of computing, acm (1993), 11–20.

[22] R. von Lieben, *Patent nr. 179807 - kathodenstrahlenrelais*, Kaiserliches Patentamt. November 19 1906.

[23] J. J. P. Eckert and J. W. Mauchly, *Patent nr. us3120606a - electronic numerical integrator and computer*, US Patent Office, June 26 1947.

[24] J. L. Mitchell and K. H. Olsen, "Tx-0, a transistor computer with a 256 by 256 memory," Papers and discussions presented at the december 10-12, 1956, eastern joint computer conference: new developments in computers, AIEE-IRE '56 (Eastern) (1956), 93–101.

[25] Intel, *Https://www.intel.co.uk/content/www/uk/en/history/museum-story-of-intel-4004.html*, 2021.

[26] G. E. Moore, “Cramming more components onto integrated circuits, reprinted from electronics,” IEEE Solid-State Circuits Society Newsletter **11**, 33–35 (2006).

[27] F. W. Jordan and W. H. Eccles, *Patent nr. gb148582a - improvements in ionic relays*, GB Patent Office, June 21 1918.

[28] B. Rossi, “Method of registering multiple simultaneous impulses of several geiger’s counters,” Nature **125**, 636–636 (1930).

[29] F. L. Bauer, *Origins and foundations of computing: in cooperation with heinz nixdorf museumsforum*, 2010th edition (Springer, Berlin Heidelberg, 2010, 2009).

[30] B. J. L., *Patent nr. us3283170a - coupling transistor logic and other circuits*, US Patent Office, September 8 1961.

[31] Wikipedia, *Transistor count*, 2021, `https://en.wikipedia.org/wiki/Transistor_count`.

[32] P. W. Shor, “Scheme for reducing decoherence in quantum computer memory,” Phys. Rev. A **52**, R2493–R2496 (1995).

[33] C. Negrevergne, T. S. Mahesh, C. A. Ryan, M. Ditty, F. Cyr-Racine, W. Power, N. Boulant, T. Havel, D. G. Cory, and R. Laflamme, “Benchmarking quantum control methods on a 12-qubit system,” Physical Review Letters **96**, 170501 (2006).

[34] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus,

O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, "Quantum supremacy using a programmable superconducting processor," Nature **574**, 505–510 (2019).

[35] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, "Quantum computational advantage using photons," Science **370**, 1460–1463 (2020).

[36] D. Deutsch, "Quantum theory as a universal physical theory," International Journal of Theoretical Physics **24**, 1572–9575 (1985).

[37] P. Shor, "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer," Proceedings of the 37th Annual Symposium on Foundations of Computer Science (FOCS) (IEEE Press, Burlington, VT, (1996).

[38] A. M. Steane, "Simple quantum error-correcting codes," Physical Review A **54**, 4741–4751 (1996).

[39] A. R. Calderbank and P. W. Shor, "Good quantum error-correcting codes exist," Physical Review A **54**, 1098–1105 (1996).

[40] J. S. Otterbach, R. Manenti, N. Alidoust, A. Bestwick, M. Block, B. Bloom, S. Caldwell, N. Didier, E. S. Fried, S. Hong, P. Karalekas, C. B. Osborn, A. Papageorge, E. C. Peterson, G. Prawiroatmodjo, N. Rubin, C. A. Ryan, D. Scarabelli, M. Scheer, E. A. Sete, P. Sivarajah, R. S. Smith, A. Staley, N. Tezak, W. J. Zeng, A. Hudson, B. R. Johnson, M. Reagor, M. P. da Silva, and C. Rigetti, "Unsupervised machine learning on a hybrid quantum computer," (2017).

[41] J. Preskill, “Quantum Computing in the NISQ era and beyond,” Quantum **2**, (2018).

[42] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, D. A. Buell, B. Burkett, N. Bushnell, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, S. Demura, A. Dunsworth, E. Farhi, A. Fowler, B. Foxen, C. Gidney, M. Giustina, R. Graff, S. Habegger, M. P. Harrigan, A. Ho, S. Hong, T. Huang, W. J. Huggins, L. Ioffe, S. V. Isakov, E. Jeffrey, Z. Jiang, C. Jones, D. Kafri, K. Kechedzhi, J. Kelly, S. Kim, P. V. Klimov, A. Korotkov, F. Kostritsa, D. Landhuis, P. Laptev, M. Lindmark, E. Lucero, O. Martin, J. M. Martinis, J. R. McClean, M. McEwen, A. Megrant, X. Mi, M. Mohseni, W. Mruczkiewicz, J. Mutus, O. Naaman, M. Neeley, C. Neill, H. Neven, M. Y. Niu, T. E. O’Brien, E. Ostby, A. Petukhov, H. Putterman, C. Quintana, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, D. Strain, K. J. Sung, M. Szalay, T. Y. Takeshita, A. Vainsencher, T. White, N. Wiebe, Z. J. Yao, P. Yeh, and A. Zalcman, “Hartree-fock on a superconducting qubit quantum computer,” Science **369**, 1084–1089 (2020).

[43] D. P. DiVincenzo, “The physical implementation of quantum computation,” (2000).

[44] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, “A quantum engineer’s guide to superconducting qubits,” Applied Physical Review **6**, (2019).

[45] C. D. Bruzewicz, J. Chiaverini, R. McConnell, and J. M. Sage, “Trapped-ion quantum computing: progress and challenges,” Applied Physics Reviews **6**, 021314 (2019).

[46] M. Saffman, “Quantum computing with neutral atoms,” National Science Review **6**, 24–25 (2018).

[47] D. M. Zajac, T. M. Hazard, X. Mi, E. Nielsen, and J. R. Petta, “Scalable gate architecture for a one-dimensional array of semiconductor spin qubits,” Physical Review Applied **6**, 054013 (2016).

[48] N. H. Wan, T.-J. Lu, K. C. Chen, M. P. Walsh, M. E. Trusheim, L. De Santis, E. A. Bersin, I. B. Harris, S. L. Mouradian, I. R. Christen, E. S. Bielejec, and D. Englund, "Large-scale integration of artificial atoms in hybrid photonic circuits," Nature **583**, (2020).

[49] M. E. Trusheim, N. H. Wan, K. C. Chen, C. J. Ciccarino, J. Flick, R. Sundararaman, G. Malladi, E. Bersin, M. Walsh, B. Lienhard, H. Bakhru, P. Narang, and D. Englund, "Lead-related quantum emitters in diamond," Physical Review B **99**, 075430 (2019).

[50] M. E. Trusheim, B. Pingault, N. H. Wan, M. Gündoğan, L. De Santis, R. Debroux, D. Gangloff, C. Purser, K. C. Chen, M. Walsh, J. J. Rose, Becker, B. Lienhard, E. Bersin, I. Paradeisanos, G. Wang, D. Lyzwa, A. R. Montblanch, G. Malladi, H. Bakhru, A. C. Ferrari, I. A. Walmsley, M. Atatüre, and D. Englund, "Transform-limited photons from a coherent tin-vacancy spin in diamond," Physical Review Letters **124**, 023602 (2020).

[51] B. Lienhard, T. Schröder, S. Mouradian, F. Dolde, T. T. Tran, I. Aharonovich, and D. Englund, "Bright and photostable single-photon emitter in silicon carbide," Optica **3**, 768–774 (2016).

[52] G. Grosso, H. Moon, B. Lienhard, S. Ali, D. K. Efetov, M. M. Furchi, P. Jarillo-Herrero, M. J. Ford, I. Aharonovich, and D. Englund, "Tunable and high-purity room temperature single-photon emission from atomic defects in hexagonal boron nitride," Nature Communications **8**, 1–8 (2017).

[53] B. Lienhard, T.-J. Lu, K.-Y. Jeong, H. Moon, A. Iranmanesh, G. Grosso, and D. Englund, "Bright high-purity quantum emitters in aluminium nitride integrated photonics," ACS Photonics, (2020).

[54] N. H. Wan, B. J. Shields, D. Kim, S. Mouradian, B. Lienhard, M. Walsh, H. Bakhru, T. Schröder, and D. Englund, "Efficient extraction of light from a nitrogen-vacancy center in a diamond parabolic reflector," Nano Letters **18**, 2787–2793 (2018).

[55] J. Zheng, B. Lienhard, G. Doerk, M. Cotlet, E. Bersin, H. S. Kim, Y.-C. Byun, C.-Y. Nam, J. Kim, C. T. Black, and D. Englund, “Top-down fabrication of high-uniformity nanodiamonds by self-assembled block copolymer masks,” Scientific Reports **9**, 1–6 (2019).

[56] X. Gu, A. F. Kockum, A. Miranowicz, Y.-X. Liu, and F. Nori, “Microwave photonics with superconducting quantum circuits,” Physics Reports **718-719**, 1–102 (2017).

[57] X. Y. Jin, A. Kamal, A. P. Sears, T. Gudmundsen, D. Hover, J. Miloshi, R. Slattery, F. Yan, J. Yoder, T. P. Orlando, S. Gustavsson, and W. D. Oliver, “Thermal and residual excited-state population in a 3d transmon qubit,” Physical Review Letters **114**, 240501 (2015).

[58] 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**, (2019).

[59] J. M. Gambetta, J. M. Chow, and M. Steffen, “Building logical qubits in a superconducting quantum computing system,” npj Quantum Information **3**, (2017).

[60] F. Bloch, “Generalized theory of relaxation,” Physical Review **105**, 1206–1222 (1957).

[61] A. G. Redfield, “On the theory of relaxation processes,” IBM Journal of Research and Development **1**, 19–31 (1957).

[62] J. Emerson, R. Alicki, and K. Życzkowski, “Scalable noise estimation with random unitary operators,” Journal of Optics B: Quantum and Semiclassical Optics **7**, S347–S352 (2005).

Chapter 3

Superconducting Quantum Computing

The superconducting qubit modality is a leading candidate towards the realization of quantum information processors. Over the last two decades, the fabrication, design, and control of superconducting qubits have considerably improved, resulting in exponential improvements of coherence properties [1]. The increase in qubit performance has enabled the demonstration of several significant milestones in the pursuit of scalable quantum computation. Among others, multi-qubit control and entanglement techniques [2, 3], improved quantum gate fidelities [4], and better readout schemes [5, 6] have enabled the demonstration of a quantum advantage using a quantum processor with 53 operational qubits [7].

Here, we introduce the fundamentals of superconducting circuits and discuss commonly employed experimental techniques and operation principles in the microwave regime. First, the basics of superconducting qubits are discussed in Section 3.1 before their readout and control are covered in Section 3.2. Finally, the experimental infrastructure to operate superconducting quantum circuits is presented in Section 3.3.

3.1 Superconducting Qubits

The building blocks of superconducting quantum computing hardware are superconducting qubits, solid-state artificial atoms with level transitions in the microwave regime [8]. The transmon qubit [9] has emerged as one of the most popular qubit designs due to its robust fabrication process, demonstrated operation and readout, and reproducible lifetimes and coherence times in the order of several tens of microsenconds [10]. It is closely related to a harmonic LC-oscillator, which features equidistant energy levels, illustrated in Figure 3-1(a). Coherent control requires an isolated pair of energy levels that form a computational qubit basis [11], and this motivates the need for anharmonic oscillators. The necessary anharmonicity is provided by the Josephson junction—a lithographically defined tunnel barrier between two superconducting electrodes—that behaves as a nonlinear inductor without any significant dissipation [12].

3.1.1 From Linear *LC* Circuits to Quantum Harmonic Oscillators

A classical linear LC resonant circuit is a simple harmonic oscillator. The system energy oscillates between electrical energy stored in the capacitor C and magnetic energy in the inductor L. The energy terms of each element depend on the voltage $V(t)$ and current $I(t)$ of the capacitor and inductor and are calculated as

$$E(t) = \int_{-\infty}^{t} V(t')I(t')dt'. \tag{3.1}$$

The energy terms for the capacitor and inductor expressed in terms of branch fluxes $\Phi(t) = \int_{-\infty}^{t} V(t')dt'$ or the charge on the capacitor $Q(t) = C\dot{\Phi}(t)$ are

$$\mathcal{T}_C = \frac{Q^2}{2C}, \tag{3.2}$$

and

$$\mathcal{U}_L = \frac{\Phi^2}{2L}. \tag{3.3}$$

The Hamiltonian describes the total energy of a system. Consequently, the Hamiltonian for a classical LC circuit is the sum of the electrical and magnetic energy and defined as

$$H = \frac{Q^2}{2C} + \frac{\Phi^2}{2L}. \tag{3.4}$$

A similar mathematical construct describes a quantum LC oscillator but with the charge Q and flux Φ promoted to quantum operators $\hat{Q}$ and $\hat{\Phi}$.

$$H = \frac{\hat{Q}^2}{2C} + \frac{\hat{\Phi}^2}{2L}. \tag{3.5}$$

Note, quantum operators are indicated by a 'hat.' $\hat{Q}$ and $\hat{\Phi}$ do not commute $[\hat{\Phi}, \hat{Q}] = \hat{\Phi}\hat{Q} - \hat{Q}\hat{\Phi} = i\hbar$. Therefore, it is not possible to simultaneously measure the flux and charge of a quantum circuit with absolute precision.

Superconducting Quantum Harmonic Oscillator

A superconductor is a material that has no DC (direct current) resistance. Metals such as aluminum with a nonzero resistance at room temperature turn superconducting below its critical temperature T_{c}, which is about $1.2\,\mathrm{K}$ for aluminum. Metals in a non-superconducting state are called 'normal.' The carriers of such lossless currents are Cooper-pairs comprising two electrons each [13]. An LC circuit is, by definition, lossless. The Hamiltonian in Equation 3.5 can be expressed in terms of its physical implementation using superconductors. Defining a reduced flux operator $\hat{\phi} = 2\pi\hat{\Phi}/\Phi_0$ with the superconducting magnetic flux quantum $\Phi_0 = \frac{h}{2e}$ and the reduced charge $\hat{n} = \hat{Q}/2e$ with the elementary charge e, the quantum-mechanical Hamiltonian follows as

$$H = 4E_{\mathrm{C}}\hat{n}^2 + \frac{1}{2}E_{\mathrm{L}}\hat{\phi}^2. \tag{3.6}$$

with the inductive $E_{\mathrm{L}} = (\Phi_0/2\pi)^2/L$ and charging energy $E_{\mathrm{C}} = e^2/(2C)$. The charging energy represents the energy necessary to add a single electron to a superconducting island.

The Hamiltonian describing a quantum harmonic oscillator (QHO), with a single particle (e.g., a photon) in a one-dimensional quadratic potential, is equivalent to Equation 3.6. $\hat{n}$ and $\hat{\phi}$ can be transformed into a creation $\hat{a}^\dagger$ and annihilation operator $\hat{a}$. Adding a particle to a system is described by the creation operator. Consequently, the creation operator's complex conjugate, the annihilation operator, captures the system's particle number reduction. The quantum harmonic oscillator for a resonance frequency $\omega_\mathrm{r} = 1/\sqrt{LC}$ is defined as

$$H = \hbar\omega_r \left(a^\dagger a + \frac{1}{2} \right). \tag{3.7}$$

with the Planck constant $h = 2\pi\hbar$. Each particle added to the system requires an energy $\hbar\omega_\mathrm{r}$.

Two energy levels need to be uniquely addressable to employ a quantum system as a qubit: meaning, the energy levels of a multi-level quantum system need to differ such that $\hbar\omega_{0\to1} \neq \hbar\omega_{1\to2}$. As depicted in Figure 3-1, Quantum harmonic oscillators have equidistant energy levels with an energy spacing of $\hbar\omega_\mathrm{r}$, and thus, are impractical qubits.

3.1.2 Josephson Junctions

Introducing a nonlinearity in an LC circuit changes the description from a harmonic to an anharmonic oscillator. The quantized energy levels of an anharmonic oscillator are no longer equidistant. A nonlinearity can be introduced using Josephson junctions [14]. A Josephson junction is a dissipation-free circuit element that is physically formed between two superconducting materials separated by a thin insulating barrier of about $1\,\mathrm{nm}$, as illustrated in Figure 3-1. The thickness enables Cooper pairs to tunnel between the superconductors coherently. The Josephson equations for $I(t)$, the current through the junction, and $V(t)$, the voltage across the junction, are

$$I(t) = I_\mathrm{c} \sin(\phi(t)), \tag{3.8}$$

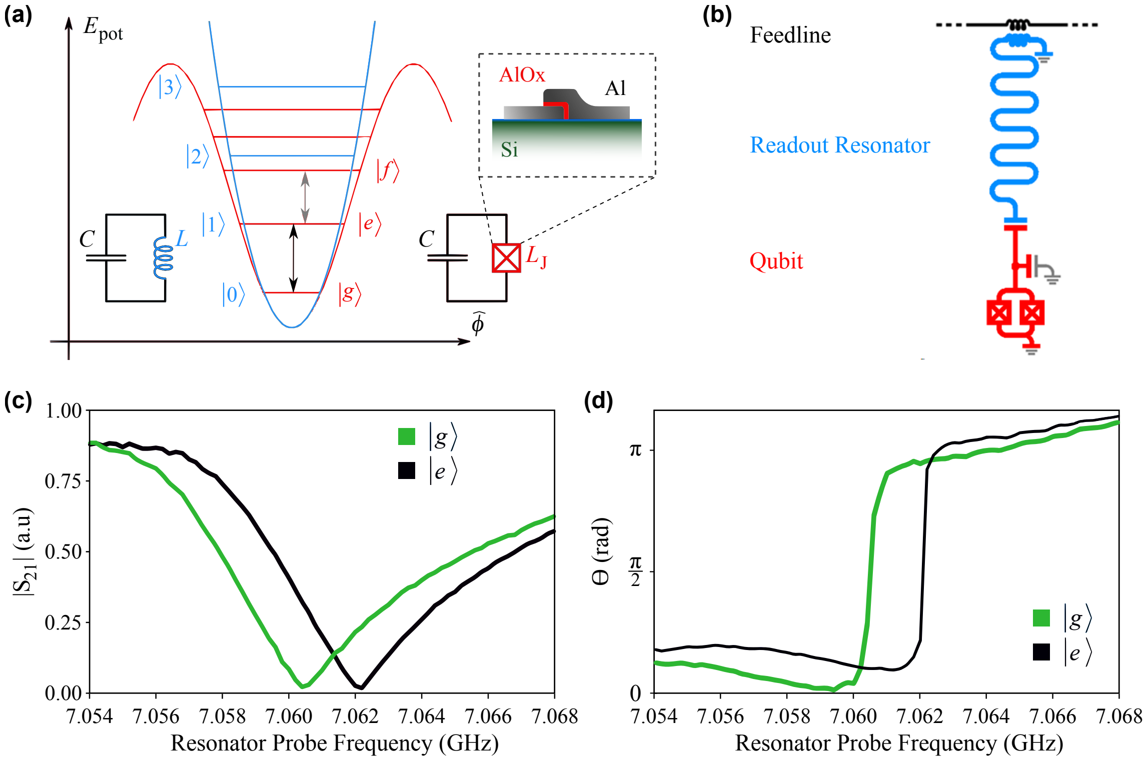


Figure 3-1: Quantum Harmonic and Anharmonic Oscillator. (a) A simple harmonic resonator formed by a capacitor C and an inductor L is shown in blue. The potential energy is the energy stored in the inductor and assumes a parabolic shape with respect to the phase variable $\hat{\phi}$, related to the flux induced by the inductor. The energy levels are equidistant, such that transitions cannot be addressed individually. By replacing the inductor with a Josephson junction, the potential of the oscillator becomes anharmonic, illustrated in red, which isolates two energy levels to form a computational qubit basis. A Josephson junction is a nonlinear inductor physically realized inserting a about $1\,\mathrm{nm}$-thin insulating barrier (e.g., aluminum oxide AlOx) between two superconductors (e.g., aluminum Al). (b) A well-studied design to enable qubit readout is to off-resonantly, capacitively couple the qubit, an anharmonic oscillator, to a readout resonator, a harmonic oscillator. The readout resonator frequency can be probed to infer the qubit state. (c) and (d) The transmission spectrum of a resonator dispersively coupled to a qubit showing the characteristic qubit-state-dependent dispersive shift of about $1\,\mathrm{MHz}$. The amplitude of the measured signal is shown in (c) and the corresponding signal phase in (d). A qubit-state discriminator tries to distinguish the qubit-states through the evaluation of the readout signal amplitude or phase.

and

$$V(t) = \frac{\hbar}{2e}\frac{\partial \phi}{\partial t}. \tag{3.9}$$

where $I_\mathrm{c} = 2eE_\mathrm{J}/\hbar$ is defined as the critical current of the junction and $\phi(t)$ is the phase difference across the junction. The zero-voltage current is bounded by I_c. The Josephson energy E_J depends on the superconductor material, temperature, and applied magnetic field. Even without a voltage difference across the junction, a current flows for a non-zero phase difference. This effect is known as the DC Josephson effect. Applying a DC voltage results in an oscillatory current, referred to as the AC Josephson effect.

The inductance due to the Josephson effect can be derived using the relation $V = L\left(\partial I/\partial t\right)$ resulting in

$$L_\mathrm{J} = \frac{\Phi_0}{2\pi I_\mathrm{c}\cos(\phi)} = \frac{L_\mathrm{J0}}{\cos(\phi)} \tag{3.10}$$

with the Josephson inductance $L_\mathrm{J0} = \Phi_0/2\pi I_\mathrm{c}$. Using Equation 3.1, the energy of a Josephson junction yields an energy of $E(\phi) = -E_\mathrm{J}\cos(\phi)$.

3.1.3 The Transmon Qubit

The Cooper pair box is a relatively simple superconducting qubit modality. A Cooper pair box consists of a superconducting island connect to a reservoir via a Josephson junction [15]. The states of a Cooper pair box are associated with the number of excess Cooper pairs on the island. The Hamiltonian is defined as

$$H = 4E_\mathrm{C}(\hat{n} - n_g)^2 - E_\mathrm{J}\cos(\phi), \tag{3.11}$$

with the Cooper pair number operator $\hat{n}$ accounting for the number of excess Cooper pairs on one of the islands, the effective offset charge n_g, and the reduced flux ϕ. The capacitive energy $E_\mathrm{C} = e^2/2C_\Sigma$ where $C_\Sigma = C_\mathrm{S} + C_\mathrm{J}$ represents the total capacitance which is the sum of the cpacitance formed due to the Josephson junction C_J and the shunt capacitance. In the charge basis, the Hamiltonian can be written as

$$H = 4E_{\mathrm{C}}(\hat{n} - n_g)^2 - \frac{E_{\mathrm{J}}}{2}\sum_n (|n\rangle\langle n+1| + |n+1\rangle\langle n|). \tag{3.12}$$

The operation regime for a Cooper pair box is $E_{\mathrm{J}} \ll E_{\mathrm{C}}$. Cooper pair boxes are sensitive to charge noise [16]. Operating the qubit at so-called 'sweet spots' $n_g = \pm 0.5$ where the charge dispersion slope is zero can suppress the effects of flux noise [17]. While the coherence times are significantly improved if the Cooper pair box is operated at the sweet spot, the necessary flux bias can vary due to system fluctuations. Adding a large shunt capacitor reduces the sensitivity to charge noise [9]. A Cooper pair box with a large shunt capacitance is called a transmission line shunted plasma oscillation (transmon) qubit. The Hamiltonian for the transmon qubit is equivalent to the one for the Cooper pair box. The large shunt capacitance alters the qubit's capacitive energy and thus the operation regime. The operation regime $E_{\mathrm{J}}/E_{\mathrm{C}}$ is significantly increased. The sensitivity to charge noise decreases exponentially as the ratio $E_{\mathrm{J}}/E_{\mathrm{C}}$ increases. However, as the ratio is being increased the anharmonicity $\alpha = \omega_{1\to2} - \omega_{0\to1}$ shrinks linearly.

The transition frequency of a transmon is $\omega_{\mathrm{q}} = 1/\sqrt{L_{\mathrm{J}}C_{\Sigma}} \approx \sqrt{8E_{\mathrm{J}}E_{\mathrm{C}}}$. The Josephson junction can be replaced with a pair of parallel Josephson junctions to alter the transition frequency. Such a layout is known as a DC superconducting quantum interference device (SQUID). The effective critical current and consequently the Josephson energy can be changed by an applied magnetic flux Φ_{ext} through the SQUID. The junctions can be the same or intentionally different. A tunable transmon with differing Josephson junctions is an asymmetric transmon and is less sensitive to flux noise relative to its symmetric counterpart [18]. The degree of asymmetry is captured by $d = (E_{\mathrm{J2}}/E_{\mathrm{J1}} - 1)/(E_{\mathrm{J2}}/E_{\mathrm{J1}} + 1)$. The Josephson energy for an asymmetric transmon is

$$E_{\mathrm{J}}' = (E_{\mathrm{J1}} + E_{\mathrm{J2}})\sqrt{\cos^2\left(\frac{\pi\Phi_{\text{ext}}}{\Phi_0}\right) + d^2\sin^2\left(\frac{\pi\Phi_{\text{ext}}}{\Phi_0}\right)}. \tag{3.13}$$

leading to the following Hamiltonian

$$H = 4E_{\mathrm{C}}(\hat{n} - n_g)^2 - E'_{\mathrm{J}} \cos\phi \tag{3.14}$$

describing a frequency tunable asymmetric transmon qubit. The external flux Φ_{ext} alters the transmon transition frequency as follows

$$\omega_{\mathrm{q}}(\Phi_{\text{ext}}) = \sqrt{8E'_{\mathrm{J}}E_{\mathrm{C}}}. \tag{3.15}$$

The insensitivity to noise makes the transmon one of the most frequently-used superconducting qubit modalities. However, the relatively low anharmonicity of $100\,\mathrm{MHz}$ to $300\,\mathrm{MHz}$ can result in undesired excitations of the second excited state.

3.1.4 Sample Design and Fabrication Techniques

By varying the relative strengths of the energies associated with the inductance, capacitance, and tunnel elements in the circuit, various architectures of superconducting qubits can be realized [19], each featuring their unique noise susceptibility and operation regime [8].

Superconducting quantum circuits are generally fabricated on hundreds of micrometer thick, commercially available, high-purity, and low-loss silicon or sapphire substrates [12]. The chip metallization (superconducting materials, such as aluminum, niobium, or titanium nitride) is deposited by thermal evaporation or magnetron sputtering and structured by optical and electron-beam lithography. The AlOx dielectric for Josephson junctions is formed by controlled in-situ oxidization of an aluminum film. The properties and mutual couplings of circuit elements can be individually tailored and fabricated in a reproducible manner due to their macroscopic physical sizes (millimeter-scale) [11].

3.1.5 Microwave Regime

Superconducting circuits behave quantum coherently when cooled to milli-Kelvin temperatures. This is mainly due to the absence of conductivity losses in the superconductor at these very low temperatures. Superconducting circuits are fabricated with elementary superconductors such as aluminum, niobium, or related compounds such as NbN with critical temperatures between $1\,\mathrm{K}$ and $16\,\mathrm{K}$. The circuit operation temperature is small compared to the superconducting gap ($\geq 50\,\mathrm{GHz}$), which further suppresses the losses induced by residual unpaired electrons (quasiparticles). The sample operation temperature is $T \sim 10\,\mathrm{mK}$, achieved by ^{4}He/^{3}He-dilution refrigerators, further discussed in Section 3.3. It corresponds to a frequency of $f = k_\mathrm{B}T/h \sim 0.2\,\mathrm{GHz}$ (where k_B and h are the Boltzmann and Planck constant, respectively), such that frequency transitions in the $5\,\mathrm{GHz}$ regime are only weakly thermally populated. The circuit can approximately be considered to remain in its ground-state in the absence of any controls.

3.2 Superconducting Qubit Control and Readout

Superconducting qubits are typically controlled through a resonant interaction with an electromagnetic pulse. However, in the dispersive regime, qubit readout is conducted off-resonantly relative to the qubit but on resonance with the associated readout resonator. The basis of superconducting qubit control and readout are discussed in the following subsections.

3.2.1 Superconducting Qubit Gates

Qubit excitation and quantum gates are performed by applying microwave drive pulses at or close to the qubit transition frequency. The qubit undergoes coherent oscillations between its two fundamental basis states, referred to as Rabi oscillations, which can be stopped at any time to prepare the desired superposition state. By tuning the phase of the applied microwave pulses, rotations around any axis in the equatorial plane of the Bloch sphere can be initiated. Access to rotations around the vertical Z-axis of the Bloch sphere can be performed by tuning the qubit frequency or simply waiting and making use of the Larmor precession, which is a rotation of the Bloch sphere to the intrinsic time evolution of the qubit state [20]. Microwave pulses inducing single-qubit rotations have a typical duration of about $20\,\mathrm{ns}$. They are frequently amplitude modulated by a Gaussian envelope to achieve a localized pulse in Fourier space.

3.2.2 Superconducting Qubit Readout

Today, superconducting qubit state measurements are most commonly performed using a dispersive readout scheme [5, 6]. The readout device is a resonator that is weakly coupled to the qubit at a detuned frequency. Due to a qubit state-dependent 'dressing' of the readout resonator, the qubit state can be inferred by spectroscopically probing a dispersive shift of about $1\,\mathrm{MHz}$ in its resonance frequency [6]. This scheme enables a quantum non-demolition measurement, where the qubit is mapped onto one of its basis states that correspond to the measurement outcome [21].

Readout pulses with a carrier frequency at or close to the resonance frequency are typically about $0.1\,\mu\mathrm{s}$ to $2\,\mu\mathrm{s}$ long and lead to a population of the resonator with up to several tens of photons. The readout signal acquires a qubit-state-dependent dispersive shift upon interaction with the readout resonator. The dispersively-shifted signal carries the information of the qubit-state and can be classically processed to identify the qubit state [see Chapter 6 for more details].

3.2.3 Circuit Quantum Electrodynamics

The Jaynes-Cummings Hamiltonian can describe the dynamics of a two-level system located in a cavity by [5]

$$H_{JC} = \omega_r \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) + \frac{\omega_a}{2} \hat{\sigma}_z + g(\hat{a}\hat{\sigma}_+ + \hat{a}^\dagger \hat{\sigma}_-), \tag{3.16}$$

with the two-level transition frequency ω_a, the cavity frequency ω_r, and the Pauli-Z operator $\hat{\sigma}_z$. The coherent interaction between the two-level system and cavity is described by the last term $g(\hat{a}\hat{\sigma}_+ + \hat{a}^\dagger\hat{\sigma}_-)$. The coupling strength between the system and cavity is expressed by g. The oscillation of energy between the two-level system and the cavity is $\hat{\sigma}_+ = |0\rangle\langle 1|$ and $\hat{\sigma}_- = |1\rangle\langle 0|$. Leakage from the cavity is expressed by the cavity leakage rate κ. Similarly, the rate at which energy is emitted from the two-level system and not absorbed by the cavity is referred to as γ.

For a situation in which the coupling between the cavity and two-level system exceeds the leakage rates of the cavity and two-level system, the coupled system is said to be in the strong coupling regime [$g \gg \kappa$ and $g \gg \gamma$]. In the strong coupling regime, energy oscillates between the two-level system and cavity at a frequency of $g/2\pi$.

If the cavity and the transition frequency of the two-level system are sufficiently detuned $\Delta = |\omega_a - \omega_r| \gg g$, their coupling is in the dispersive limit. Expanding in g to the second-order, the Jaynes-Cummings Hamiltonian in the dispersive regime can be approximated as

$$H_{\text{JC, disp}} \approx \omega_r \left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) + \left[\omega_a + 2\frac{g^2}{\Delta}\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)\right]\frac{1}{2}\hat{\sigma}_z. \tag{3.17}$$

The cavity experiences a frequency shift by $\chi = g^2/\Delta$ depending on the state of the two-level system. For an excited two-level system, the shift is positive. In contrast, for the two-level system in the ground state, a negative shift follows. Additional details on the dispersive Jaynes-Cummings Hamiltonian and mathematical concepts to describe the system dynamics are in Appendix A.

The Hamiltonian for a general cavity and two-level system can be directly applied to the situation of a qubit coupling to a readout resonator inducing a dispersive shift $\chi = g^2/\Delta$. Finally, to remain in the dispersive regime, the number of photons populating the readout resonator should not exceed a so-called critical photon number. The critical photon number is defined as $n_\text{crit} = \frac{\Delta^2}{4g^2}$ and depends on the frequency separation between the resonator and qubit $\Delta = |\omega_r - \omega_q|$ and their coupling strength g.

3.3 Experimental Infrastructure

The material composition and thermal environment of a qubit define its feasible operational frequency range. Superconducting qubits are generally designed to have transition frequencies between $2\,\mathrm{GHz}$ and $10\,\mathrm{GHz}$ [12]. The qubits are shielded and cooled in a ^{3}He-^{4}He dilution refrigerator to minimize thermal excitations. Today's commercially available dilution refrigerators reach a base temperature around $10\,\mathrm{mK}$, well below the temperature corresponding to the transition frequency of the qubit (which is around $240\,\mathrm{mK}$ for a $5\,\mathrm{GHz}$ qubit) and the critical temperature of superconductors used for qubit design (e.g., $1.2\,\mathrm{K}$ for aluminum). The device package is mounted on the mixing chamber plate, the coldest stage in the refrigerator.

In thermal equilibrium, the excitation probability of a qubit can be expressed with the Boltzmann factor $\exp\{(-h\nu/k_\mathrm{B}T)\}$ (ν: qubit transition frequency, T: qubit temperature, h: Planck constant, k_B: Boltzmann constant). However, thermal equilibrium with $10\,\mathrm{mK}$ is generally not reached due to the influx of thermal photons from higher temperature stages of the refrigerator via the signal lines. With state-of-the-art attenuation and filtering [22], superconducting qubits have achieved effective temperatures of $35\,\mathrm{mK}$, corresponding to an excited state thermal population of 0.1% at $5\,\mathrm{GHz}$ [10].

A typical microwave measurement setup is schematically depicted in Figure 3-2. Room temperature electronics feed microwave pulses to the cryostat through coaxial cables with a characteristic impedance matched to $50\,\Omega$. The signals pass a series of attenuators that are thermally anchored to the different temperature stages, typically at $3\,\mathrm{K}$, $800\,\mathrm{mK}$, and at the base temperature of $\sim 10\,\mathrm{mK}$ respectively, to sequentially reduce the room-temperature Johnson-Nyquist noise, which forms a decoherence channel for qubits.

The control and readout signal can either be generated with advanced arbitrary waveform generators (AWG) at the target frequency directly or generated at an

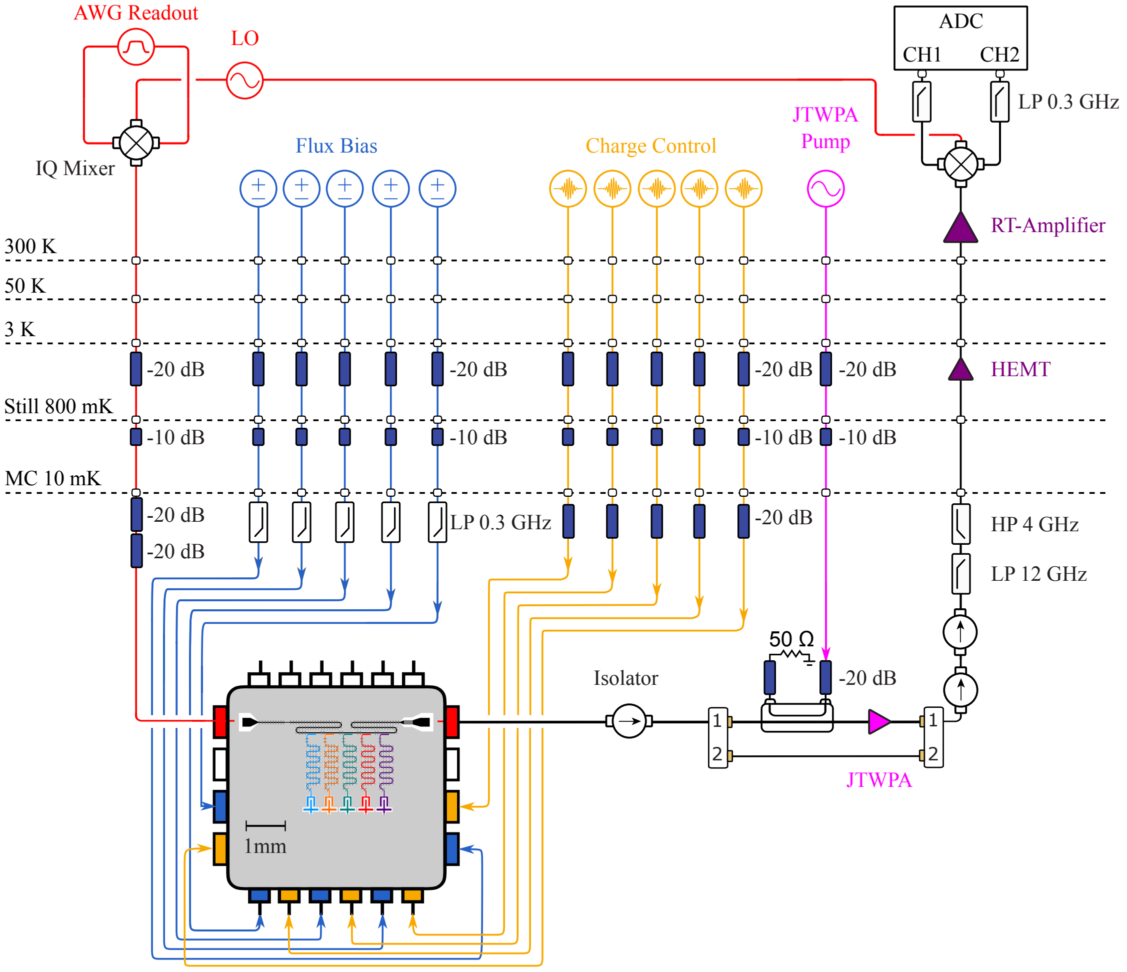

AWG Readout
LO
ADC
CH1
CH2
LP 0.3 GHz
JTWPA
Pump
Flux Bias
Charge Control
IQ Mixer
RT-Amplifier
300 K
50 K
3 K
-20 dB
-20 dB
-20 dB
-20 dB
HEMT
Still 800 mK
-10 dB
-10 dB
-10 dB
-10 dB
MC 10 mK
-20 dB
-20 dB
LP 0.3 GHz
-20 dB
HP 4 GHz
LP 12 GHz
50 Ω
-20 dB
Isolator
1
2
1
2
JTWPA
1mm

Figure 3-2: Qubit Control and Readout Electronics and Wiring. (a) Depicted is a schematic of a microwave setup, including pulse generation and processing. Microwave pulses are generated by microwave sources and arbitrary waveform generators (AWG). IQ-mixers facilitate phase-sensitive amplitude modulation. The control and readout lines are attenuated and spectrally filtered at various stages. The readout signal is first amplified by a Josephson traveling-wave parametric amplifier (JTWPA) after passing through an isolator. The JTWPA is equipped with a separate microwave source. Before the readout signal is further amplifed by a high-electron-mobility transistor (HEMT) amplifier at the $3\,\mathrm{K}$ stage, the signal passes through additional isolators and is bandpass filtered. At room temperature, the readout signal is amplified once more before it is down-converted, low-pass filtered, and digitized.

intermediate frequency and upconverted via single-sideband modulation. In Figure 3-2, the qubit control signals are directly generated, whereas the readout signal is upconverted using a local oscillator (LO). For a single pulse, separate real and imaginary parts of the complex readout signal, $R_I(t)$ and $R_Q(t)$, are fed into an IQ-mixer along with a local oscillator frequency ω_{LO}. The output of the mixer and readout signal is

$$R_{\text{in}} = R_I(t)\cos(\omega_{\mathrm{LO}}t) - R_Q(t)\sin(\omega_{\mathrm{LO}}t) = \mathrm{Re}\big\{(R_I(t) + iR_Q(t))e^{i\omega_{\mathrm{LO}}t}\big\}. \tag{3.18}$$

The readout probe tone signals transmitted through the sample are passed through microwave isolators before reaching a quantum-limited amplifier such as a traveling wave parametric amplifier (JTWPA) [23]. The JTWPA amplifies microwave signals of individual photons by about $20\,\mathrm{dB}$ in broadband of $\sim 2\,\mathrm{GHz}$. It works close to a regime where only the minimum amount of noise dictated by quantum mechanics is added to the amplified signal, known as quantum-limited amplification. Such sensitive amplification enables a single-shot readout of a set of qubits and facilitates the implementation of real-time quantum feedback [24]. The isolators prevent leakage of the pump tone required to operate the JTWPA back to the sample. The JTWPA is operated through a separate microwave source. After passing a low-noise high electron-mobility amplifier (HEMT) at the outputs, thermalized to $3\,\mathrm{K}$, the signal is further amplified at room temperature before being processed in room temperature microwave electronics. The transmitted readout signal is downconverted to intermediate frequencies using the same LO that was used for upconversion. The readout signal at intermediate frequencies is subsequently low-pass filtered and digitized. Finally, the digitized readout signal is classically discriminated to assign a qubit state label.

3.3.1 Microwave Package Engineering

Superconducting quantum chips are mounted into a microwave package thermally attached to the cold stage of a dilution refrigerator. It defines the immediate elec-

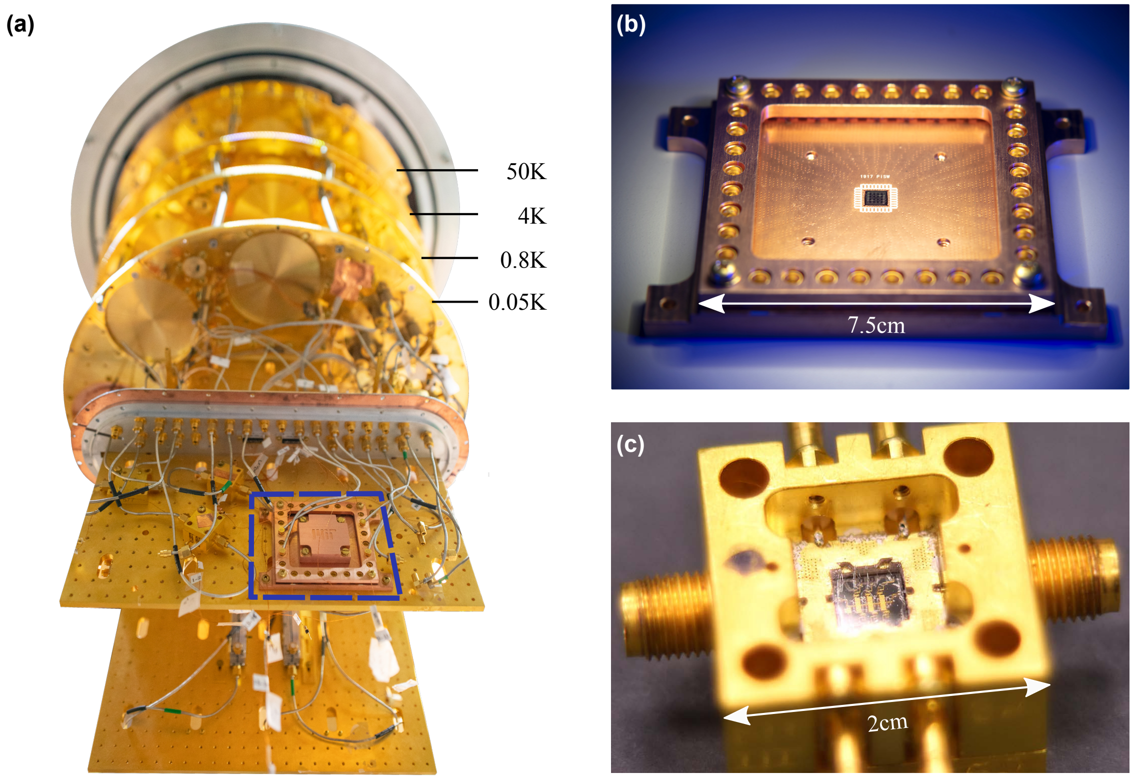


Figure 3-3: Dilution Refrigerator and Microwave Package. (a) Dilution refrigerator with multiple temperature stages holding the qubit chip enclosed in a microwave package. The microwave package interfaced with through microwave lines is mounted on a cold finger in the mixing chamber reaching a base temperature of approximately $10\,\mathrm{mK}$. (b) The next microwave package generation is highlighted with a blue rectangle and discussed in the subsequent Chapter 4. The package consists of a copper casing, a multilayer interposer to perform signal fanout, a shielding cavity in the package center, and 24 microwave connectors. (c) Previous microwave package generation with a gold-plated copper casing and six control lines.

tromagnetic environment of the qubits and connects the quantum circuit to the coaxial control lines. While the microwave electronics for qubit control are mostly commercially available, the fabrication of superconducting quantum circuits and surrounding packages have not yet reached the stage of commercialization.

A microwave package's primary purpose is to simultaneously shield the quantum circuit from the environment while enabling efficient control and thermalization. There are two kinds of general package architectures. The package can either act as a 3D cavity with an engineered mode spectrum and a high-quality resonance mode used for qubit readout [25, 26], or merely provide an electromagnetic environment with suppressed spurious modes in the frequency spectrum of interest [27, 28]. This thesis focuses on the second type, further discussed in the following Chapter 4 and shown in Figure 3-3.

3.4 Summary

Over the past two decades, superconducting qubits have emerged as a leading quantum computing platform [21, 29]. Superconducting circuits are operated in the microwave regime, enabling a high degree of control and providing a rich toolbox of experimental techniques. Today, individual qubits with coherence times exceeding $100\,\mu\mathrm{s}$ [10], gate times of a few tens of nanoseconds [7], and single- and two-qubit gate operation fidelities above the most lenient thresholds for quantum error correction have been demonstrated for devices with up to 50 qubits [1, 7]. However, considerable work is still needed to retain and even further improve these fidelities as systems increase in size and complexity [30]. Among the imminent challenges are scalable control hardware and software tools. In this thesis, microwave package hardware for multi-qubit systems and software tools to improve multi-qubit readout are developed.

References

[1] 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**, (2019).

[2] M. Steffen, M. Ansmann, R. C., N. Bialczak, E. L. Katz, R. McDermott, M. Neeley, E. M. Weig, A. N. Cleland, and J. M. Martinis, "Measurement of the entanglement of two superconducting qubits via state tomography," Science **313**, 1423–1425 (2006).

[3] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. OMalley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, and J. M. Martinis, "Superconducting quantum circuits at the surface code threshold for fault tolerance," Nature **508**, 500–503 (2014).

[4] S. Gustavsson, O. Zwier, J. Bylander, F. Yan, F. Yoshihara, Y. Nakamura, T. P. Orlando, and W. D. Oliver, "Improving quantum gate fidelities by using a qubit to measure microwave pulse distortions," Physical Review Letters **110**, (2013).

[5] 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," Physical Review A **69**, 1–14 (2004).

[6] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, R. S. Huang, J. Majer, S. Kumar, S. M. Girvin, and R. J. Schoelkopf, "Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics," Nature **431**, 162–167 (2004).

[7] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hart-

mann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, "Quantum supremacy using a programmable superconducting processor," Nature **574**, 505–510 (2019).

[8] J. Clarke and F. K. Wilhelm, "Superconducting quantum bits," Nature **453**, (2008).

[9] 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**, 042319 (2007).

[10] X. Y. Jin, A. Kamal, A. P. Sears, T. Gudmundsen, D. Hover, J. Miloshi, R. Slattery, F. Yan, J. Yoder, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "Thermal and residual excited-state population in a 3d transmon qubit," Physical Review Letters **114**, 240501 (2015).

[11] R. J. Schoelkopf and S. M. Girvin, "Wiring up quantum systems," Nature **451**, 664–669 (2008).

[12] W. D. Oliver and P. B. Welander, "Materials in Superconducting Qubits," MRS Bulletin **38**, 816–825 (2013).

[13] L. N. Cooper, "Bound electron pairs in a degenerate fermi gas," Phys. Rev. **104**, 1189–1190 (1956).

[14] B. Josephson, "Possible new effects in superconductive tunnelling," Physics Letters **1**, 251–253 (1962).

[15] V. Bouchiat, D. Vion, P. Joyez, D. Esteve, and M. H. Devoret, “Quantum coherence with a single Cooper pair,” Physica Scripta **1998**, 165 (1998).

[16] G. Ithier, E. Collin, P. Joyez, P. J. Meeson, D. Vion, D. Esteve, F. Chiarello, A. Shnirman, Y. Makhlin, J. Schriefl, and G. Schön, “Decoherence in a superconducting quantum bit circuit,” Phys. Rev. B **72**, 134519 (2005).

[17] D. Vion, A. Aassime, A. Cottet, P. Joyez, H. Pothier, C. Urbina, D. Esteve, and M. H. Devoret, “Manipulating the quantum state of an electrical circuit,” Science **296**, 886–889 (2002).

[18] M. Hutchings, J. Hertzberg, Y. Liu, N. Bronn, G. Keefe, M. Brink, J. M. Chow, and B. Plourde, “Tunable superconducting qubits with flux-independent coherence,” Physical Review Applied **8**, (2017).

[19] M. H. Devoret and R. J. Schoelkopf, “Superconducting ciruits for quantum information: an outlook,” Science **339**, 1169–1174 (2013).

[20] G. Wendin and V. Shumeiko, “Handbook of Theoretical and Computational Nanotechnology,” 3, edited by M. Rieth and W. Schommers (American Scientific publishers, Los Angeles, 2006) Chap. Quantum Circuits, Qubits and Computing, .

[21] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, “A quantum engineer’s guide to superconducting qubits,” Applied Physical Review **6**, (2019).

[22] K. Serniak, M. Hays, G. de Lange, S. Diamond, S. Shankar, L. D. Burkhart, L. Frunzio, M. Houzet, and M. H. Devoret, “Hot nonequilibrium quasiparticles in transmon qubits,” Physical Review Letters **121**, 157701 (2018).

[23] C. Macklin, K. O’Brien, D. Hover, M. E. Schwartz, V. Bolkhovsky, X. Zhang, W. D. Oliver, and I. Siddiqi, “A near–quantum-limited josephson traveling-wave parametric amplifier,” Science **350**, 307–310 (2015).

[24] D. Ristè, C. C. Bultink, M. J. Tiggelman, R. N. Schouten, K. W. Lehnert, and L. DiCarlo, "Millisecond charge-parity fluctuations and induced decoherence in a superconducting transmon qubit," Nature Communications **4**, 1913 (2013).

[25] H. Paik, D. I. Schuster, L. S. Bishop, G. Kirchmair, G. Catelani, A. P. Sears, B. R. Johnson, M. J. Reagor, L. Frunzio, L. I. Glazman, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf, "Observation of high coherence in Josephson junction qubits measured in a three-dimensional circuit QED architecture," Physical Review Letters **107**, (2011).

[26] T. Brecht, Y. Chu, C. Axline, W. Pfaff, J. Z. Blumoff, K. Chou, L. Krayzman, L. Frunzio, and R. J. Schoelkopf, "Micromachined Integrated Quantum Circuit Containing a Superconducting Qubit," Physical Review Applied **7**, (2017).

[27] B. Lienhard, J. Braumüller, W. Woods, D. Rosenberg, G. Calusine, S. Weber, A. Vepsäläinen, K. O'Brien, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "Microwave packaging for superconducting qubits," 2019 IEEE MTT-S International Microwave Symposium (2019).

[28] S. Huang, B. Lienhard, G. Calusine, A. Vepsäläinen, J. Braumüller, D. K. Kim, A. J. Melville, B. M. Niedzielski, J. L. Yoder, B. Kannan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "Microwave package design for superconducting quantum processors," (2021).

[29] X. Gu, A. F. Kockum, A. Miranowicz, Y.-X. Liu, and F. Nori, "Microwave photonics with superconducting quantum circuits," Physics Reports **718-719**, 1–102 (2017).

[30] J. M. Gambetta, J. M. Chow, and M. Steffen, "Building logical qubits in a superconducting quantum computing system," npj Quantum Information **3**, (2017).

Chapter 4

Package Design for Superconducting Qubits

Solid-state qubits that rely on microwave control to operate are among the leading candidates for realizing practical near-term quantum processors. However, significant engineering challenges constrain these devices from scaling up further. In particular, qubits require a precisely engineered microwave environment to suppress energy decay and corresponding information loss. For instance, the corruption of information can occur due to lossy package modes interacting with the qubit. Furthermore, as the number of qubits increases, quantum processor packages must be adapted to support an increasing number of control lines without creating additional loss channels.

This chapter discusses a ground-up approach to package design that addresses these challenges in the context of a multi-qubit quantum processor. First, Section 4.2 discusses the geometry and materials of microwave packages and their impact on qubit coherence times. Next, Section 4.3 focuses on maintaining high simultaneous control fidelity and coherence times for a system with many control lines. Finally, in Section 4.4, a comprehensive evaluation of the package-related loss channels is performed, demonstrating that lossy package modes can lead to coherence limits within an order of magnitude of today's state-of-the-art qubit coherence times, underscoring the importance of high-performance package engineering.

4.1 Microwave Packaging

Despite the rapid progress towards building practical quantum processors using superconducting qubits [1], the susceptibility of these artificial atoms to noise remains a significant engineering challenge to system scaling. Microwave packaging is a part of this challenge. In particular, package designs, such as displayed in Figure 4-1, must support increasing qubit numbers while also preserving qubit coherence and high-fidelity quantum operations.

Qubits interacting with their environment can lead to energy decay and dephasing. Qubit coherence generally improves as the qubit is decoupled from the environment, but it becomes more difficult to control and read out. Ideally, the qubit would couple exclusively to the control and readout environment.

Qubit energy decay is mediated by various loss channels, such as quasiparticles, vortices, surface dielectric dissipation, conductivity losses, or dissipation into spurious package modes near the qubit transition frequency. The quality factor Q_i expresses the inverse 'lossiness' of an individual loss channel i. The participation ratio p_i, a unitless factor, associates each loss channel with a normalized interaction strength between itself and the qubit [2, 3] so that the participation ratios of all loss channels sum to 1. Modeling the qubit as a harmonic oscillator—a reasonable approximation for weakly anharmonic qubits like the transmon [4] and capacitively shunted flux qubits [5]—the energy exchange rate can then be expressed as $\Gamma_1 = \omega/Q = \omega \sum_i p_i/Q_i$.

The transition frequency between the ground and excited state of the qubit can be affected by its electromagnetic (EM) environment. Fluctuating EM fields detuned from the qubit transition frequency—if coupled to the qubit—can induce qubit energy-level shifts that cause a change in the phase accumulation rate, resulting in pure dephasing of a qubit superposition state.

A microwave package's primary purpose is to simultaneously shield the quantum circuit from the environment while enabling efficient control and thermalization, as illustrated in Figure 4-1(a). There are two kinds of general package architectures.

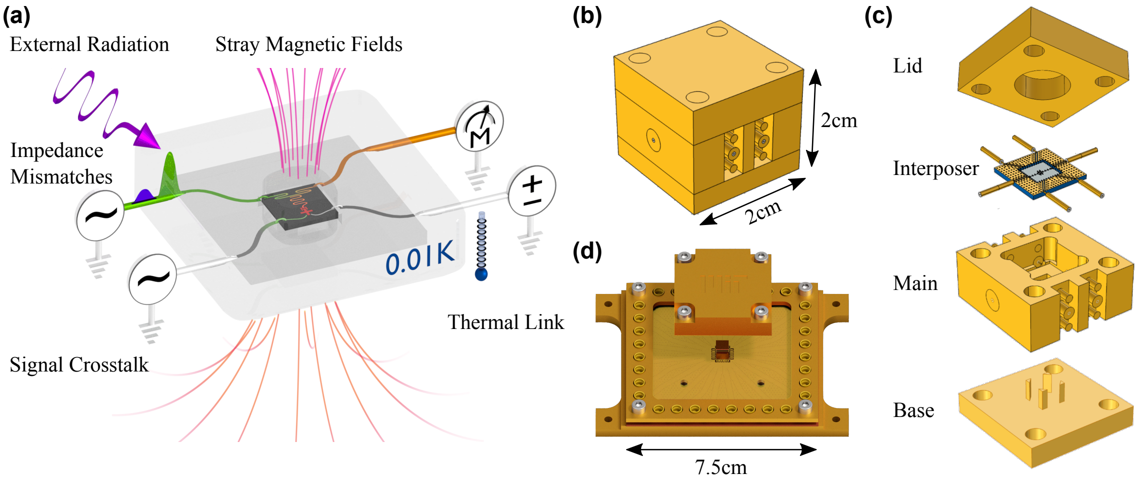


Figure 4-1: Microwave packages. (a) The purpose of a microwave package is to shield the enclosed qubit chip from external radiation (purple oscillating arrow) and stray magnetic fields (rose lines). Meanwhile, impedance-matched (transmitted green pulse and reflected blue pulse at the input and output), minimal-crosstalk communication channels (crosstalk in green at input), and a thermal link to the dilution refrigerator need to be provided. (b) Small sample package with six control channels. (c) The individual components of the package shown in panel (b). Inserted between the gold-plated copper base and the lid is the main package part. Metallic waveguides for control of the quantum circuit are imprinted on a dielectric circuit board (interposer) and $50\,\Omega$-impedance matched. The interposer routes the coaxial control line signal to the quantum circuit via wirebonds. The connectors to the control elements are coaxial connectors with $50\,\Omega$ impedance enabling the transmission of microwave signals up to about $18\,\mathrm{GHz}$ to $26\,\mathrm{GHz}$. (d) Novel sample package with 24 control lines used to study and establish the reported microwave design principles.

The package can either act as a 3D cavity with an engineered mode spectrum and a high-quality resonance mode used for qubit readout [6] or merely provide an electromagnetic environment with suppressed spurious modes in the frequency spectrum of interest. These 3D cavities—typically formed out of aluminum or copper—enable coherent interactions between microwave photons and the qubits. The 3D cavity architecture ensures a qubit environment free of spurious electromagnetic modes and therefore yields high coherence at the cost of control and scalability. On the other hand, and the focus here, the package can have the sole purpose of providing a clean electromagnetic environment.

4.2 Microwave Package Materials & Geometry

Material-dependent losses can be of magnetic ($1/Q_m$), conductive ($1/Q_c$), or dielectric ($1/Q_d$) origin [7]. Energy loss channels couple to the qubit through its electric or magnetic dipole moment. For transmon qubits, the electric dipole moment presently dominates, as shown in Figure 4-3 [4]. Qubits are fabricated using high-Q materials and substrates to reduce loss. In addition, the device geometry is designed to reduce the electric field density in lossy regions, such as surfaces and interfaces [8].

Many qubit architectures are sensitive to magnetic fields, in particular those with a tunable transition frequency. Consequently, magnetic metals or materials with magnetic compounds are generally avoided. However, to shield qubits from magnetic field fluctuations, materials with high magnetic permeability, such as mu-metal, can be used either as part of the dilution refrigerator infrastructure or package casing. An alternative approach incorporates type-I superconductors such as aluminum, tin, or lead in the package body. Once such a material turns superconducting, it expels the magnetic field from its core due to the Meissner effect, so long as the magnetic field does not exceed a specific material- and temperature-dependent threshold.

Conductivity losses arise when the electric field of the qubit induces a current in nearby normal conductors with finite conductivity. The loss depends on the conductivity σ of the material and scales as $1/Q_c \propto 1/\sqrt{\sigma}$. Losses also arise due to atomic defects in bulk dielectrics hosted on those interfaces absorbing EM energy. Dielectric losses are proportional to the imaginary dielectric coefficient $\mathrm{Im}(\epsilon)$ of the material $1/Q_d \propto \mathrm{Im}(\epsilon)$.

Commonly employed package materials include superconducting aluminum, copper, and gold-plated copper. Superconducting aluminum forms a thin oxide layer of approximately $2\,\mathrm{nm}$ [10], inducing some dielectric losses while keeping the conductivity losses at a minimum. Like aluminum, copper forms an oxide layer [11] leading to dielectric as well as conductivity loss due to its non-zero resistance [12].

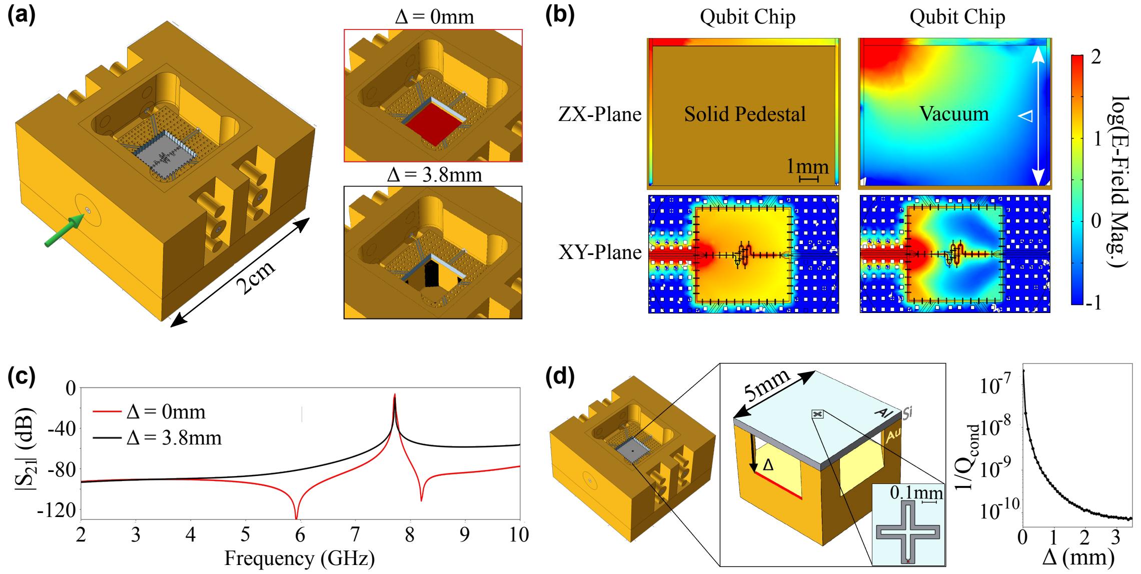


Figure 4-2: Microwave Simulations of a small Sample Package shown in Figure 4-1(b). (a) Gold-plated copper package (without lid) with wire-bonded superconducting circuit chip. The green arrow indicates the microwave input port. (b) Comparison of the transmission spectra and conductivity loss with a solid pedestal ($\Delta = 0\,\mathrm{mm}$, colored in red), and a drilled out pedestal with four corner posts ($\Delta = 3.8\,\mathrm{mm}$, black). (c) Simulated transmission magnitude spectrum $|S_{21}|$ of the chip holding an interrupted transmission line resonator with a resonance frequency at about $7.7\,\mathrm{GHz}$ mounted inside the package with pedestal (red) and a drilled out pedestal (black). The package with the drilled-out pedestal is free of any undesired package modes in the band of interest. (d) E-field magnitude ($\mathrm{V/m}$) plots in the ZX- and XY-plane at $5.9\,\mathrm{GHz}$ (indicated with an arrow in (c)). The presence of the pedestal enhances the E-field magnitude in the chip and direct vicinity. (e) Simulated conductivity loss $1/Q_{\mathrm{cond}}$ due to the normal conducting ($\sigma = 4.5 \times 10^{9}\,\mathrm{S\,m^{-1}}$ [9]) package, extracted from a transmon qubit (approximated as linear resonator) in the center. The conductivity loss depends on the gap size Δ between the chip and the pedestal (the four corner posts remain in place).

Gold-plating limits the oxide formation at the cost of an increase in conductivity losses by up to one order of magnitude [13].

The 3D finite element simulation software COMSOL Multiphysics® is used to analyze the microwave properties of the package shown in Figure 4-1(b). The package model is schematically depicted in Figure 4-2(a), holding the interposer and a chip with an interrupted transmission line resonator. Ideally, one expects a pronounced peak at the resonance frequency of this resonator at about $7.7\,\mathrm{GHz}$ and no transmission away from resonance. However, when measured inside the sample package, the transmission spectrum of the resonator chip is convoluted with a broad package mode centered at about $19\,\mathrm{GHz}$, mediating a non-zero baseline transmission. Between $2\,\mathrm{GHz}$ to $10\,\mathrm{GHz}$, the frequency range of interest, the transmission through the sample package is non-zero and gradually increases with frequency due to a package mode at about $19\,\mathrm{GHz}$. The packaged resonator chip transmission spectrum is composed of the resonator resonance peak and the package-dependent baseline transmission. Package modes depend on the package cavity geometry and increase in frequency with decreasing cavity dimensions lower bounded by the chip dimensions. The simulated transmission magnitude spectrum $|S_{21}|$ of the superconducting quantum circuit is shown in Figure 4-2(c).

Figure 4-2(b) shows schematic drawings of two sample package versions, one with the chip sitting on a solid pedestal (red) and one where the chip is supported by four corner posts (black) with the pedestal drilled out to a depth of $\Delta = 3.8\,\mathrm{mm}$ [14, 15]. The response of the on-chip resonator is visible in both simulated transmission magnitude spectra $|S_{21}|$ (red and black lines in Figure 4-2(c)). The simulation of the package with pedestal (red) reveals pronounced package modes in the relevant frequency range, which provides a potential qubit loss channel. The modes are suppressed in the version (black) with the pedestal drilled out, indicating an electromagnetic sample environment without spurious modes in the frequency range of interest. Figure 4-2(d) shows the electric field magnitude distribution at one of the box modes at $5.9\,\mathrm{GHz}$. For the package with a solid pedestal, the electric field is strongly enhanced inside and in the vicinity of the chip.

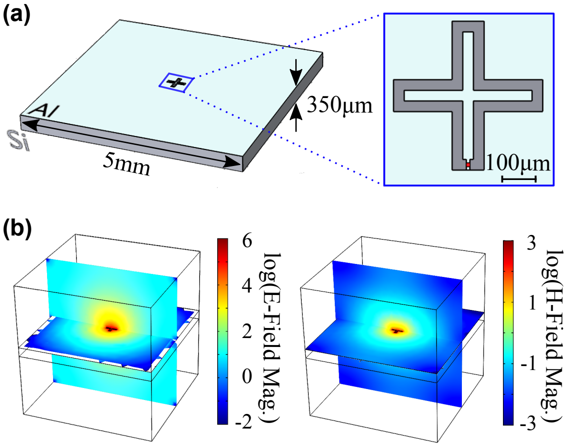


Figure 4-3: Qubit Full-Wave EM Simulations. (a) Lithographically defined superconducting qubit using aluminum and aluminum oxide on a pristine silicon chip. Blue inset shows the superconducting qubit with the Josephson junction in red, aluminum in turquoise, and silicon in grey. (b) Illustrated are the qubit's electric (E-field, left) and magnetic (H-field, right) fields. The electric field is more prominent than the magnetic field

The overall transmission magnitude in the qubit target frequency range of $2\,\mathrm{GHz}$ to $6\,\mathrm{GHz}$ is decreased for the variant with the drilled out pedestal, showing that fewer energy couples into the sample package. Consequently, the unwanted coupling of the qubit to the package mode is decreased for the version with the drilled-out pedestal.

Normal metal near the sample forms a loss channel due to finite conductivity. The conductivity loss Q_{cond}^{-1} is simulated using a surface participation model. The simulation estimates the conductivity loss of a typically sized transmon qubit with a '+'-shaped capacitor [16], defined on a chip mounted in the presented package, as a function of the distance Δ between sample and pedestal, see Figure 4-2(e). Electric and magnetic fields for such a qubit are shown in Figure 4-3. While the conductivity loss assumes a maximum for $\Delta = 0\,\mathrm{mm}$, it saturates at $\Delta \approx 3\,\mathrm{mm}$ and is suppressed by about three orders of magnitude. The extracted conductivity loss for a package with a solid pedestal is $Q_{\mathrm{cond}} = 4.5 \times 10^6$. Some of the highest experimentally achieved qubit lifetimes of $T_1 \approx 150\,\mu\mathrm{s}$ [17] correspond to a quality factor of $Q \approx 4.5 \times 10^6$ implying that conductivity loss cannot be neglected for long-lived qubits.

The device package demonstrated in this thesis, shown in Figure 4-1(d), is composed of a base and lid, both milled from oxygen-free high conductivity (OFHC) copper. A layer of aluminum with a target thickness of $500\,\mathrm{nm}$ is evaporated on the lid center cavity surface to reduce conductivity losses. The qubit chip is suspended

Table 4.1: Comparison of qubit losses induced by different casing materials for the presented package geometry. The values are obtained with EM simulations of a $\nu = 5\,\mathrm{GHz}$ transmon qubit. The qubit is located in the corner of a $5\,\mathrm{mm} \times 5\,\mathrm{mm}$ chip to estimate the maximum impact of package materials-induced losses. The T_1-limit is estimated as $1/(1/Q_c + 1/Q_d) \cdot 1/(2\pi\nu)$. The following material-dependent characteristics are assumed: Copper (Cu) has a conductivity of $5 \times 10^{10}\,\mathrm{S\,m}$ [12] and an oxide thickness of $10\,\mathrm{nm}$ [11]; Aluminum (Al) has a conductivity of infinity and an oxide thickness of $2\,\mathrm{nm}$ [10]; Gold (Au) has a conductivity of $5 \times 10^{9}\,\mathrm{S\,m}$ [13] and no oxide layer.

Center Cavity	Conductivity Loss $1/Q_c$	Dielectric Loss $1/Q_d$	T_1-limit (s)
Bare Cu	2×10^{-9}	1×10^{-12}	0.020
Al-evaporated Cu	–	5×10^{-12}	5.830
Au-plated Cu	5×10^{-9}	–	0.006

by at least $3\,\mathrm{mm}$ to form a cavity above and below it following the results to increase the package fundamental mode frequency and suppress material-induced losses, as shown in Figure 4-2(e) and Reference [8].

Full-wave EM simulations (COMSOL Multiphysics®) indicate the layer of aluminum on the center cavity surface to reduce the material-induced loss channels by three orders of magnitude. As such, the material-induced losses of the presented package are negligible, enabling it to support qubits with lifetimes up to seconds, see Table 4.1.

A device package needs to provide an efficient thermal link to the dilution refrigerator to ensure the qubits reach and remain close to the base temperature. The accumulated thermal energy due to all sources, including the material loss channels, should not significantly heat the qubits. Furthermore, as the number of qubits increases, the efficiency of the thermal link has to increase accordingly. Electrical connections remove heat to the wiring and attenuation at cryogenic temperatures—where non-equilibrium electrons can thermalize—and by mechanical (phonon) connections to the package.

The loss channel contributions of the material have to be balanced with its thermal conductivity. For example, the thermal conductivity of copper decreases linearly with temperature and reaches a value of approximately $0.5\,\mathrm{W\,cm^{-1}\,K^{-1}}$ at $20\,\mathrm{mK}$ [18]. On the other hand, the thermal conductivity of superconducting aluminum decreases exponentially faster at similar temperatures [19] and can be estimated to be around $0.025\,\mathrm{W\,cm^{-1}\,K^{-1}}$ at $20\,\mathrm{mK}$. However, at cryogenic temperatures, heat flow from the chip is almost entirely dominated by Kapitza boundary resistance interfaces [20].

In the presented package, the qubit chip is kept in place with aluminum wirebonds and pressed down on copper posts located in the corners of the qubit chip. A more efficient thermal link could be formed using conductive adhesives such as silver paste or polymeric adhesives. However, adhesives are avoided as their conductivity is typically more than three orders of magnitude smaller than copper, resulting in a measurable increase in conductivity losses [21].

4.3 Control and Readout Signals

A package's signal paths introduce various challenging factors, including a distorted step response and insertion loss (Section 4.3) and crosstalk (Section 4.3), which can diminish a qubit's controllability. Subsequently, the engineering considerations for waveguides imprinted within the interposer (Section 4.3.1) and wirebonds (Section 4.3.2) connected to the qubit chip are described.

Step Response and Insertion Loss

Good impedance matching leads to lower insertion losses and improved signal integrity, critical for high-fidelity control and readout. For a linear time-invariant system, the ideal temporal response for a step-like input is instantaneous and step-like as well. There are two related measures. First, the rise-time is the duration until the signal reached the desired amplitude. Second, the settling time is the time that elapses for the signal to stabilize at the output, typically after ringing or repeated oscillations. For one-qubit gates, these distortions can lead to under and over-rotations and reduce gate fidelity. In two-qubit gates, such as the controlled-phase gate, deviations from the carefully shaped flux pulses can lead to leakage away from the computational subspace [22]. Furthermore, a long settling time will make the action of the gate depending on the history of pulses applied [23, 24]. While a nonideal step response can be straightforwardly compensated for by an arbitrary waveform generator using predistortion, prolonged distortions require greater memory depth. Therefore, near-ideal impedance matching is vital.

The frequency response of the system determines the step response. In general, the characteristic impedance of a transmission line is $Z = \sqrt{(R + i\omega L) / (G + i\omega C)}$, where R is the series resistance of the line, G the parallel conductance of the dielectric, C the parallel capacitance, L the inductance, and ω is the angular frequency of the signal. Thus, for example, frequency dependence can arise in the case of nonzero resistance, skin effects—which reduce the effective inductance at higher frequencies [25]—and the frequency-dependent nature of the permittivity of the dielectric, which can cause C to vary.

Decreasing the resistance in the signal lines and return path, either by employing superconductors or high conductivity materials, lowers this frequency dependence and can effectively reduce the settling and rise times of the system [26]. Furthermore, in the equivalent time-domain picture, poor forward and backward matching cause repeated signal reflections in the system that lead to frequency-dependent standing waves, leading to ringing in combination with unmatched reactance in the system. As a result, geometric transitions and waveguides within the signal chain must be designed to ensure a consistent impedance.

Crosstalk

Second, crosstalk—the undesired transfer of a signal between separate communication lines—needs to be suppressed in a package. Control signals that leak to nearby qubits can induce unwanted gate operations, resulting in reduced computational performance. Furthermore, these errors may be frequency-dependent and challenging to compensate for if the qubits add significant non-linearity to the system. Such errors are particularly harmful to standard quantum error correction codes, which presume local and uncorrelated errors [27].

Crosstalk decreases as the physical distance between signal lines increases. Interactions between adjacent lines are classified into two distinct categories: near-end crosstalk (NEXT)—measured at the same end as the interfering transmitter—and far-end crosstalk (FEXT)—measured at the opposite end [illustrated in Figure 4-5 (a)]. Coupling between adjacent lines occurs through their mutual capacitance and inductance. Capacitive coupling induces a positive current on both ends of the disturbed line, while inductive coupling leads to a current moving parallel to the instigating line. As a result, the crosstalk can interfere with the desired control signal on both ends of the signal line depending on the ratio of capacitive and inductive couplings.

4.3.1 Interposer Design

There are three basic transmission line designs: coplanar waveguides, microstrips, and striplines [7]. Compared to coplanar waveguides and microstrips, symmetrical buried striplines are surrounded by a homogeneous dielectric. This environment leads to the same capacitive and inductive coupling to the top and bottom ground plane, canceling the forward-propagating current and thereby suppressing far-end crosstalk at the cost of higher fabrication complexity. Furthermore, the top and bottom ground planes shield the fully buried signal line from the far-field environment [28].

The presented interposer is constructed using a three-layer, low-loss Rogers 4350™laminate composed of glass-woven hydrocarbon and ceramics. The laminate is composed of two Rogers 4350™cores, each with a thickness of $0.338\,\mathrm{mm}$, and bonded using a layer of thermoplastic Fluorinated Ethylene Propylene (FEP) film. The loss tangent of Rogers 4350™laminates is relatively low (specified to be 0.0037 at $10\,\mathrm{GHz}$ and room temperature). The stack consists of three copper layers, with signal routing being performed in the center layer and both connector and wirebond launches patterned on the top. Blind vias are used to route signals between the top and middle layers to prevent parasitic stub resonances, as shown in Figure 4-4(b).

The package is equipped with Rosenberger non-magnetic, gold-plated SMP-type connectors. The employed connectors have a manufacturer-specified insertion loss of $\leq 0.1\sqrt{f(\mathrm{GHz})}\mathrm{dB}$ and offer a frequency range of DC-$40\,\mathrm{GHz}$, covering the operating range of most superconducting qubits.

The interposer uses symmetric copper striplines embedded in the printed circuit board (PCB) dielectric protected by via-fences. EM simulations and time-domain modeling is employed to minimize impedance mismatches. The simulations ensure that microwave connector transitions, composed of a grounding cage and a signal via, as well as the wirebond launches, are properly impedance matched [29–31].

The connector transitions are characterized using time-domain reflectometry (TDR) on a $14\,\mathrm{GHz}$ bandwidth Keysight E5063 network analyzer. As a result, the connector launch is well matched to the waveguide, as shown in Figure 4-4(d). The interposer waveguide has a measured impedance of $53.5\,\Omega$, leading to a voltage standing wave ratio of 1.07 and a mismatch loss of $0.05\,\mathrm{dB}$ ($0.1\,\%$ of the incoming power is reflected).

The crosstalk is further suppressed using via-fences—rows of metalized holes drilled through the substrate material to shield in-plane EM-field coupling between pairs of signal lines. As opposed to guarding structures, which are grounded microstrips between signal lines that provide limited isolation beyond a few hundred MHz [32], fences also work at higher frequencies. The shielding effectiveness at a particular frequency ν depends on the via spacing. The spacing between vias should remain small compared to the wavelength $\lambda = v_{\mathrm{m}}/\nu$ with the material-dependent wave velocity of the waveguide, v_{m}. As a rule of thumb, the spacing should not exceed $\lambda/20$ to ensure that the via-fence appears solid to an impinging wave [33] and to minimize loading on the signal-carrying line, which can affect signal integrity [34].

The self-resonance frequency of the vias must also be taken into consideration. A via's intrinsic shunt inductance [35] is approximately $L \approx c_1 h\,(1 + \ln(4h/d))$, with constant $c_1 = 1.95 \times 10^{-6}\,\mathrm{H\,m^{-1}}$, height h and diameter d, and a parasitic capacitance $C \approx c_2 \epsilon_\mathrm{r} h d_1/(d_1 - d_2)$, where $c_2 = 5.6 \times 10^{-11}\,\mathrm{F\,m^{-1}}$, d_1 is the diameter of the antipad (gap opening in the surrounding ground plane), d_2 is the diameter of the via pad, and ϵ_r is the relative permittivity of the dielectric. These lumped elements lead to a self-resonance for a single via typically in the range of a few hundred MHz to a few GHz. Furthermore, the combination of the vias' conductivity and the capacitance formed between the two large ground planes can result in a resonant mode. In both cases, many vias can ameliorate these issues and increase the resonance frequency beyond the qubit frequency operation range.

Employing the introduced via-fence, the crosstalk between directly neighboring control and readout lines is suppressed to below $-40\,\mathrm{dB}$ and next-nearest neighbors to $-60\,\mathrm{dB}$ up to $10\,\mathrm{GHz}$, presented in Figure 4-4(e) and (f).

4.3.2 Wirebonding Considerations

Thermosonically bonded $25\,\mu\mathrm{m}$-diameter aluminum wirebonds are employed to allow reconfigurable connections and enable rapid prototyping (gold wirebonds can similarly be used as an alternative). This also ensures compatibility with 3D-integrated multi-chip modules that use a silicon interposer to fan-out signals to pads for wirebonding [36]. Alternatively, vertical spring-loaded contacts can be directly fanned out into coaxial cables [37] or connected to a multi-layer PCB [14]. The device can also be directly clamped onto the PCB [38] like flip-chip ball grid array packaging used in conventional room-temperature electronics [39, 40]. However, dielectric losses induced by the proximity of the PCB substrate and surface roughness may reduce qubit coherence [36]. Inserting a pristine silicon or sapphire interposer between the PCB interposer and qubit chip [1, 36] can reduce such performance-limiting effects. These interposer stacks can be constructed using superconducting indium bump-bonds and superconducting through-silicon vias [41], enabling the construction of multi-chip modules and off-chip resonators [36]. As in the demonstrated package, a combination of these interconnect techniques can be

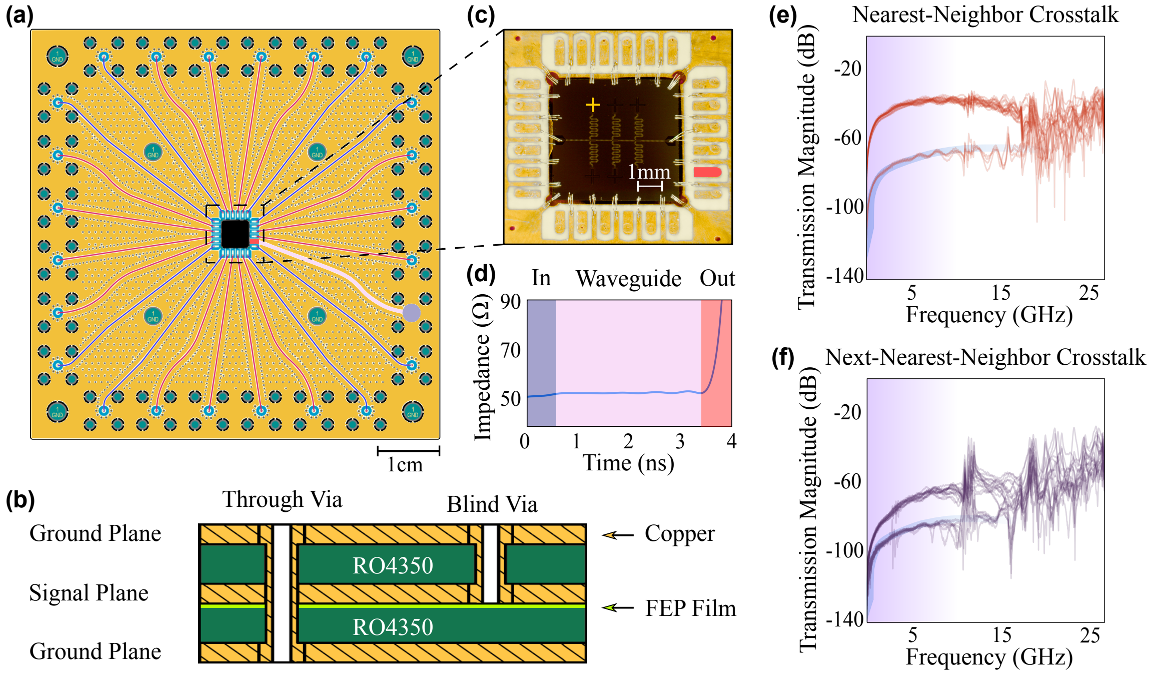


Figure 4-4: Interposer Design. (a) Interposer layout for a 24-line package fabricated out of a three-layer Rogers 4350™controlled impedance glass-reinforced ceramic laminate. Stripline-based waveguides with dense via shielding are utilized to reduce signal crosstalk. (b) Schematic diagram of the interposer stack configuration (not to scale). The board consists of two Rogers 4350™cores bonded by a layer of FEP film, indicated in purple. Through vias are used for grounding and shielding, whereas blind vias are utilized for signal routing to minimize parasitic resonances. (c) Picture of a 5 mm by 5 mm qubit chip mount with signal launches in the periphery. Wirebonds are used to provide signal connections and ground the device. (d) Single-ended TDR measurement of the SMP-to-interposer connector transition. The reference plane of the connector is located at 0 ns. The wirebond launch is left open (no chip connected), resulting in a steep increase in impedance. (e) and (f) Measured nearest neighbor and second nearest signal crosstalk. A high-port count network analyzer using Keysight M9374A PXIe modules was used to obtain the full scattering matrix. All transmission parameters corresponding to signal crosstalk information are overlaid on the plots. The purple background fading out with increasing frequency indicates the decreasing relevance of modes as their frequency separation to the qubit transition frequency increases. Note that the separate group of traces with increased isolation, as highlighted in (e) and (f), correspond to crosstalk that is reduced at corners of the package, as indicated by the pairs of blue lines in (a).

utilized, such as using an interposer with fixed-length wirebonding to the PCB.

Its parasitic inductance dominates the wirebond impedance. For a wirebond with a $25\,\mu\mathrm{m}$-wide diameter, the inductance approximately scales as $1\,\mathrm{nH}$ per $1\,\mathrm{mm}$ in length [15]. Taking into account the effect of a ground plane at a distance h, the inductance of a wirebond with a diameter d and length ℓ is approximately $L \approx \mu_0 \ell\, \mathrm{arcosh(2h/d)}/(2\pi)$ with the vacuum permeability μ_0 [42]. Similarly, the parasitic capacitance of the wirebond can modeled by a wire and a uniform metal sheet as $C = 2\pi\ell\epsilon_0/(\mathrm{arcosh(h/d)})$ with ϵ_0 being the permittivity of vacuum [43, 44].

For wirebonds $1\,\mathrm{mm}$ in length, the inductance is around $1\,\mathrm{nH}$ and the capacitance is in the range of $20\,\mathrm{fF}$. A $1\,\mathrm{mm}$-long wirebond—modeled as a transmission line—yields a characteristic impedance of $Z = \sqrt{L/C} = 223\,\Omega$ and a reflection coefficient of $\Gamma = (Z - Z_0)/(Z + Z_0) \approx 0.63$. Reducing the impedance mismatch requires shorter wirebonds, several wirebonds in parallel, or measures to decrease the parasitic capacitance.

Mutual coupling between parallel wirebonds hinders the combined inductance from decreasing as quickly as the inverse of the number of wirebonds: about five parallel $1\,\mathrm{mm}$ long wirebonds are required to reduce the impedance to the range of $50\,\Omega$. The parasitic inductance can be reduced by positioning signal launches close to the edge of the chip, thus shortening the wirebond, and by using several parallel wirebonds for each signal connection, spreading them in a V-shape to minimize mutual inductance.

In applications that need a low insertion loss, or when the length of the wirebond cannot be minimized, an impedance matching structure can be used [45]. The capacitance is tuned using capacitive structures, such as flares or ‘matching dots,’ which are empty metal pads on the interposer that can be galvanically connected to change the capacitance [46].

For basic applications without a bandwidth-limit, an inductor-capacitor structure can be utilized with an impedance of $Z = \sqrt{L_{\mathrm{parasitic}}/(C_{\mathrm{parasitic}} + C_{\mathrm{tuning}})}$. The passband can be increased using a third-order, capacitor-inductor-capacitor low pass filter. For a given design cutoff frequency (typically set above the qubit frequency operating range), there is a maximum inductance that can be accommodated in this way due to the broadband requirement. As shown in Figure 4-5(b), a Butterworth filter can compensate for a wirebond with an inductance $1.6\,\mathrm{nH}$ at a cutoff frequency of $10\,\mathrm{GHz}$ (corresponding to a length of approximately $1.6\,\mathrm{mm}$). In contrast, a Chebyshev configuration can compensate for around $0.6\,\mathrm{nH}$.

The physical distance and the exposure of wirebonds contribute significantly to signal crosstalk. Combining the parasitic inductance and capacitance of the wirebond and the capacitance formed between the chip and the package ground form a resonance mode. An impedance ladder model predicts that such crosstalk falls off exponentially with distance at low frequencies but reaches unity at the resonance frequency of the aforementioned mode [15]. As a result, pulling back the ground plane below the chip to decrease the capacitance (for example, designing a cavity underneath the qubit chip) and increasing wirebond density to lessen the inductance reduce signal crosstalk.

EM simulations confirm the formation of a coupling channel due to a resonance mode between the chip and the chip mount, as visualized in Figure 4-5 (c) as a high electric field density surrounding the chip. Without grounding wirebonds, the demonstrated particular model, which uses a square $5\,\mathrm{mm} \times 5\,\mathrm{mm}$ chip, predicts a maximum crosstalk of $-8\,\mathrm{dB}$ that peaks around $5.5\,\mathrm{GHz}$. Increasing the grounding wirebond density drives up the resonance frequency of this mode. It exponentially suppresses the crosstalk, with both far- and near-end coupling dropping to below $-40\,\mathrm{dB}$ up to $8\,\mathrm{GHz}$ when four wirebonds are employed between signal lines, as shown in Figure 4-5 (d). In the demonstrated package, two wirebonds between each pair of signal lines suppress crosstalk to less than $-30\,\mathrm{dB}$, at which the performance limitation is negligible relative to other system limitations imposed by on-chip and interposer crosstalk.

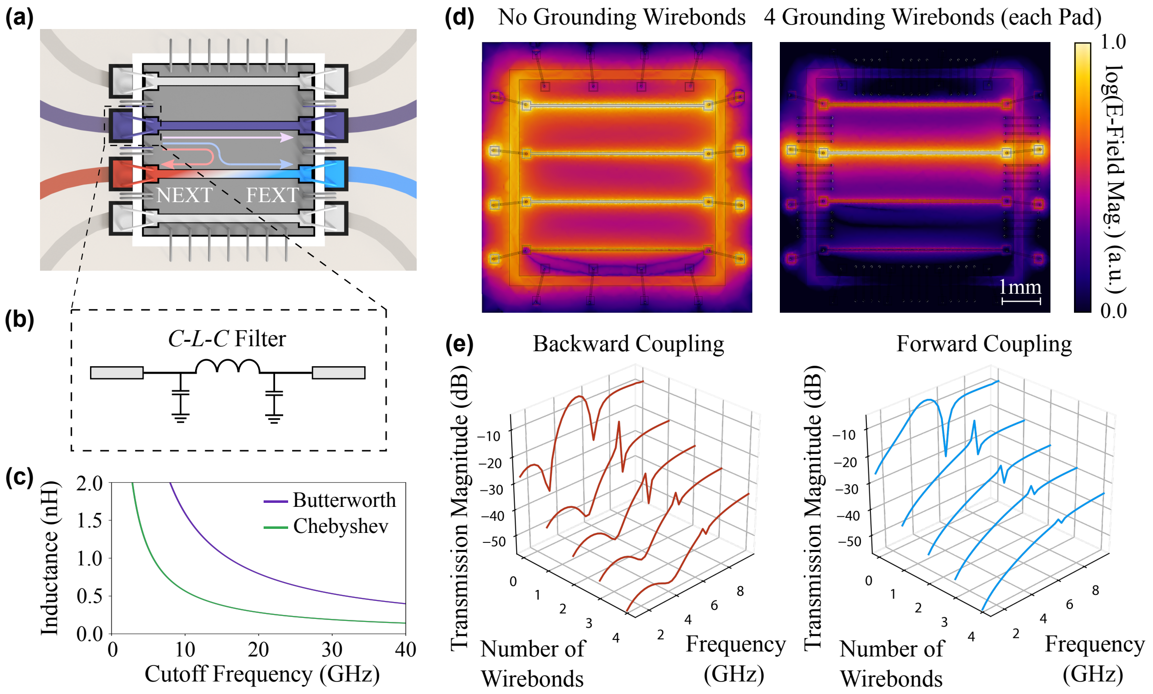


Figure 4-5: Wirebond Design. (a) Diagram of a chip (gray) with four transmission lines wirebonded to an interposer (beige) surrounding it. The instigating line, indicated in purple, can lead to near-end (NEXT) and far-end (FEXT) crosstalk in the disturbed line, shaded red and blue. (b) Schematic representation of a wirebond interconnect and its lumped element model. The flares located on the ends of the interposer and chip's transmission lines, as highlighted by the dashed box in (a), correspond to the two tuning capacitors to the ground, while the wirebond forms the series inductance. Note that the wirebonds used to connect the signal lines are spread out in a V-shape to minimize mutual inductance. (c) Plot of maximum inductance compensation versus cutoff frequency for Butterworth and Chebyshev filter designs, with $1\,\mathrm{nH}$ roughly corresponding to $1\,\mathrm{mm}$ in wirebond length. (d) and (e) The effect of wirebond configuration on signal crosstalk. A simplified model with the exact spatial dimensions as the 24-pin design is employed to simulate the relative reduction of signal wirebond crosstalk with more grounding wirebonds. As the number of grounding wirebonds increases, the electric field strength between the chip ground and the package cavity rapidly decreases, resulting in the suppressed coupling between adjacent signal wirebonds.

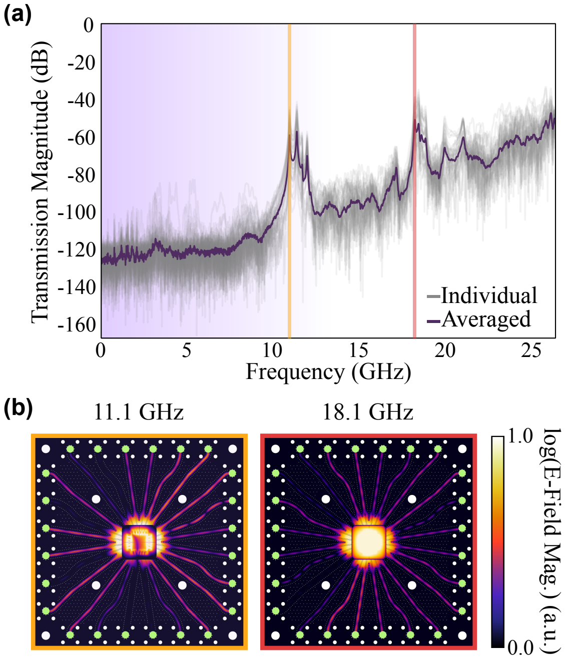


Figure 4-6: Package Mode Measurements and Simulation. (a) Package modes are probed at liquid nitrogen temperature using a multi-port vector network analyzer. The scattering matrix elements corresponding to transmission across the package cavity are measured via the ports marked green in (b). The results are overlaid (gray) and averaged (purple), with the relevant frequency range for the qubit indicated by the fading purple background shading. (b) EM simulations reveal eigenmodes at $11.1\,\mathrm{GHz}$ (orange) and $18.1\,\mathrm{GHz}$ (red) respectively, which correspond well with the measured peaks as marked by the correspondingly colored vertical lines in (a).

4.4 Package Modes

In addition to crosstalk, suppressing package modes is key to a successful design. These resonance modes can reduce the qubit lifetime and induce decoherence. A two-level quantum system can model the interaction —the qubit—coupled to an EM cavity—the resonant mode—with a rate g. For a small detuning $\Delta = |\omega_\mathrm{q} - \omega_\mathrm{m}| \ll g$ between the qubit frequency ω_q and the package mode frequency ω_m, their energy levels hybridize, and excitations are coherently swapped between the qubit and the mode. However, since the package modes are often lossy—i.e., have a low-quality factor Q_m—they lead to a reduction in the qubit lifetime.

The Jaynes-Cummings model can describe the coupling between a far-detuned mode and a qubit in the dispersive approximation [47]. A mode coupling to a qubit ac-Stark shifts the qubit transition frequency by an amount proportional to the average number of photons $\bar{n}$ present in the mode. Fluctuating photon numbers fluctuate the qubit frequency and induce pure dephasing of the qubit at a rate

$$\Gamma_\phi = \frac{\kappa}{1 + \frac{\kappa^2\Delta^2}{4} g^4}\bar{n} \tag{4.1}$$

with $\kappa = \omega_\mathrm{m}/Q_\mathrm{m}$, the decay rate of the mode [48].

Wide-band fluctuations within several GHz of the qubit transition frequency lead to pure dephasing. Moreover, modes on the order of MHz detuned from the qubit transition frequency can also lead to qubit energy decay due to the Purcell effect, thus reducing the qubit lifetime. If the coupling of the resonant mode to the qubit is small compared to the frequency separation between them, the qubit energy decay rate is

$$\gamma_{\mathrm{pkg\ mode}}^{\mathrm{Purcell}} = \frac{g^2\kappa}{(\omega_\mathrm{m} - \omega_\mathrm{q})^2}. \tag{4.2}$$

The cavity and interposer geometries often necessary in the design of a package can support resonant modes. However, to ensure high-fidelity qubit performance, either the coupling of these package modes to the qubit needs to be suppressed, or their resonance frequency must be far detuned from the qubit operational frequency spectrum.

4.4.1 Box Modes

The first class of these modes, referred to as box modes, arises in the enclosing metal cavity and interacts directly with the qubit. A box-like cavity is often used to reduce radiative losses from the qubit and significantly cut down on the number of environmental modes at the expense of Q-enhancement of the modes that remain. Furthermore, the walls of this cavity should be offset from the qubits to reduce material-induced losses, as introduced in Section 4.2, including the 'floor' below the chip. As a result, the space above and below the chip forms resonant cavities. The frequencies of the transverse electric and magnetic mode (TE_{nml} and TM_{nml}) in a rectangular cavity are given as

$$f_{nml} = \frac{c}{2\pi\sqrt{\mu_\mathrm{r}\epsilon_\mathrm{r}}}\sqrt{\left(\frac{n\pi}{b}\right)^2 + \left(\frac{m\pi}{a}\right)^2 + \left(\frac{l\pi}{d}\right)^2} \tag{4.3}$$

where TE_{101} and TM_{110} are the lowest frequency modes, μ_r is the relative permeability, ϵ_r is the relative permittivity, c is the speed of light, and a, b, d are the three dimensions of the cavity. The components μ_r, ϵ_r, and c depend on the mode carrying medium, typically the dielectric. The inclusion of any media with an ϵ_r above the vacuum permittivity shifts the resonance frequency down. Higher electric field densities around these dielectrics, caused by structures such as wirebonds, can further reduce the mode frequency by increasing the effective dielectric constant [7].

4.4.2 Interposer-Based Modes

Modes also arise within the package interposer. For example, slotlines, composed of two metallic planes separated by a dielectric gap, support a quasi-transverse electric mode propagating along it. Slotline modes arise when there is a high impedance between two ground planes, for example, due to poor galvanic contact between cavity components (such as the lid and the body of the package) and the gap separating the chip and the interposer. Furthermore, the gaps between the signal trace and the ground planes on coplanar waveguides support slotline modes that can be excited by discontinuities or asymmetries [49], such as T-junctions and sharp bends on the order of the signal wavelength [50].

Similarly, modes can form due to impedance mismatches or open transmission lines with frequencies $f = c/(\ell\sqrt{\epsilon_\mathrm{r}})$ on the order of a few GHz for waveguides with length ℓ in the range of centimeters. In the presented package, transmission lines that are not terminated have quarter-wave modes ranging from $1.15\,\mathrm{GHz}$ to $1.38\,\mathrm{GHz}$ and can couple to the qubits through unused bond pads. This coupling can be suppressed by grounding unused bond pads using wirebonds.

4.4.3 Chip-Based Modes

Spurious modes also arise from the device substrate, which has a higher ϵ_r than vacuum. For a silicon qubit chip, the first eigenmode (TM_{110}) is expected at $12.41\,\mathrm{GHz}$ for a $5\,\mathrm{mm}\times5\,\mathrm{mm}$ chip, while for dimensions of $10\,\mathrm{mm}\times10\,\mathrm{mm}$, the first eigenmode drops to $6.20\,\mathrm{GHz}$. These modes can couple very strongly to the qubit due to their physical proximity. One approach to suppress the chip modes is through a change of geometry from a square to a rectangular shape. For example, a $20\,\mathrm{mm}\times5\,\mathrm{mm}$ device, which has the same area as the $10\,\mathrm{mm}\times10\,\mathrm{mm}$ layout, has its lowest eigenmode at $9.04\,\mathrm{GHz}$. Alternatively, through-silicon vias in the qubit chip can be used to pin the resonating modes. Metallization on the chip edge can also enforce the cavity boundary conditions regardless of the surrounding environment to provide a predictable mode environment.

4.4.4 Mode Characterization

A combination of two techniques can be used to identify package modes. First, using a multi-port network analyzer, the average transmission parameters for geometrically opposing ports—different sides relative to the interposer aperture—result in a transmission spectrum as displayed in Figure 4-6(a). This method relies on the coupling between signal launches in an unpopulated chip cavity (package and interposer without a chip) and the resonances to resolve spurious modes in the system. A package eigenmode can be resolved in this way because crosstalk across the interposer aperture—typically less than $-100\,\mathrm{dB}$—is significantly weaker than the transmission induced by the package eigenmode itself. The signal-to-noise ratio is further improved by measuring in liquid nitrogen at about $77\,\mathrm{K}$, which increases the quality factor of the resonance modes. Measuring across a single pair of connectors does not paint a complete picture. Still, by taking a large number of scattering parameters, modes around $11\,\mathrm{GHz}$ and $18\,\mathrm{GHz}$ in the presented package can be resolved. In the underlying case, the mode profile is mapped out by constructing the full scattering matrix of the package using repeated measurements on a

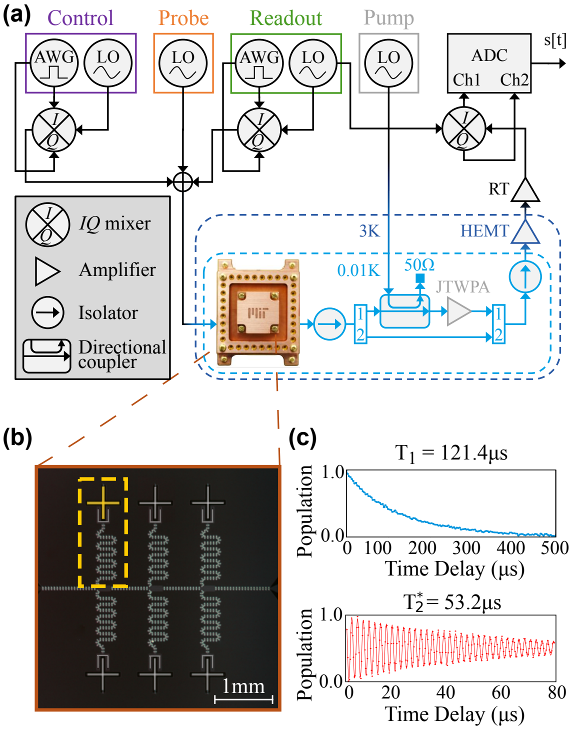

(a)
Control
Probe
Readout
Pump
AWG
LO
LO
AWG
LO
LO
ADC
Ch1
Ch2
s[t]
I
Q
RT
IQ mixer
Amplifier
Isolator
Directional coupler
3K
HEMT
0.01K
50Ω
JTWPA
1
2
1
2
(b)
1mm
(c)
T1 = 121.4μs
Population
1.0
0.0
0
100
200
300
400
500
Time Delay (μs)
T2* = 53.2μs
Population
1.0
0.0
0
20
40
60
80
Time Delay (μs)

Figure 4-7: Setup to Identify Package Modes. (a) Measurement setup used to obtain the package mode profile. The control, readout, and probe signals are combined and sent down the dilution refrigerator. The readout signal is amplified by a Josephson traveling-wave parametric amplifier (JTWPA), a high-electron-mobility transistor (HEMT) amplifier, and an amplifier at room temperature before being down-converted and subsequently digitized. Note, the JTWPA can be bypassed if necessary. (b) Superconducting qubit chip with six fixed-frequency transmon qubits with individual readout resonators coupling to a common transmission line. In the following panels, the results of the qubit and resonator indicated with a yellow dashed rectangle are explicated. (c) The qubit has an average qubit lifetime of $T_1 \approx 121.4\,\mu\text{s}$ and coherence time of $T_2^* \approx 53.2\,\mu\text{s}$ measured in intervals across a period of 12 hours.

20-port network analyzer and averaging the 144 traces that correspond to transmission across the package cavity. The two dominant peaks in the experimental results are in agreement with the 3D EM simulations of the entire design, shown in Figure 4-6(b), which revealed two high-Q eigenmodes at $11.1\,\text{GHz}$ and $18.1\,\text{GHz}$ respectively.

Detecting modes using a transmission spectrum has two main restrictions. First, the qubit may couple to resonance modes not visible to signal launches on the cavity periphery, such as chip modes and resonances resulting from device wirebonding. Conversely, the method may detect modes that do not affect qubit operation, such as resonances localized within the interposer. Second, due to the indirect nature of the transmission measurements, the plot in Figure 4-6(a) can only be used as a qualitative tool. The relative amplitudes of the peaks and the mode's coupling strength to the qubits cannot be accurately established.

A second qubit-based technique for probing package modes is the hidden-mode experiment [51], where a fixed-frequency qubit can be used as a mode sensor. It is preferable to use qubits with a long coherence time and stable baseline Ramsey oscillations to resolve subtler mode structures. A continuous-wave probe tone is injected into the package, either through the readout line or a dedicated port, and swept through the frequency range of interest. For each probe frequency, a T_2 measurement is performed using Ramsey interferometry on a fixed-frequency qubit. As the probe frequency sweeps in resonance with a package mode, the package mode will be populated by coupling the transmission line and the mode itself. Depending on the mode photon number fluctuations, the coupling to the qubit, and its detuning, the qubit will dephase. The degree of induced dephasing can be inferred with a T_2 measurement. Due to the wide-band nature of the mode-induced qubit energy level shift, the ac-Stark effect, this technique provides the advantage of facilitating mode measurements across a broad frequency range, typically several tens of GHz.

4.4.5 Measurement Setup

The measurement setup used to perform the hidden mode experiment is depicted in Figure 4-7(a). First, control pulses for the qubits are created using two separate Keysight M3202A PXI arbitrary waveform generators with sampling rates of $1\,\mathrm{GSa/s}$. Next, the in-phase and quadrature signals are upconverted to the qubit transition frequency using an IQ-mixer, which acts as a single-sideband mixer. Next, the probe tone is created using a separate signal generator. Finally, the control, readout, and probe tones are combined and sent to the dilution refrigerator via a single microwave line.

There is a total of $60\,\mathrm{dB}$ of attenuation distributed within the dilution refrigerator wiring to reduce thermal noise from room temperature and the higher temperature stages of the refrigerator. The signal has to pass through the readout resonator to reach the qubit, which acts as a filter. A control pulse length ranging from $100\,\mathrm{ns}$ to $150\,\mathrm{ns}$ is used to excite the various qubits.

The state of the qubit is determined via dispersive readout. The frequency of the resonators coupled to each qubit change slightly depending on the qubit state. This difference can be measured by sending a measurement tone near the corresponding resonator frequency down the central transmission line and recording the transmitted signal.

The signal can be first boosted using a traveling-wave parametric amplifier (TWPA) which has a gain of up to $30\,\mathrm{dB}$, a very low noise temperature, and a wide bandwidth that enables multiplexed readout [52]. The TWPA requires a pump tone, which is sourced from a signal generator at room temperature. The microwave line carrying the pump tone is attenuated by $50\,\mathrm{dB}$ and fed into the TWPA via a set of a directional coupler and isolator located at the $10\,\mathrm{mK}$ stage of the refrigerator. Next, the signal is further amplified by a Low Noise Factory high-electron-mobility transistor (HEMT) amplifier thermally anchored to the $3\,\mathrm{K}$ stage.

At room temperature, the readout signal is fed into a heterodyne detector. Next, the down-converted in-phase and quadrature signals are digitized with a Keysight M3102A PXI Analog to Digital Converter with a $500\,\mathrm{MSa/s}$ sampling rate. Finally, the signal is integrated into the internal field-programmable gate array (FPGA) of the digitizer to extract the occupation probability of the qubit in a given state.

4.4.6 Chip Design

Figure 4-7(b) depicts a device composed of six superconducting high-coherence transmon qubits used to characterize the package modes. A simple two-port device is chosen to limit the interference of the device with the hidden-mode survey. These fixed-frequency qubits are weakly coupled to a readout resonator coupled to a shared transmission line. The readout and control signals for all six qubits are frequency-multiplexed and combined with a probe tone from a tunable coherent source. While the probe tone can be injected into the package through control lines, unused signal launches, or a bandpass filter designed to limit qubit energy decays due to the Purcell effect [53], the underlying experiment utilizes a transmission line. A transmission line enables a homogeneous coupling to potential modes across

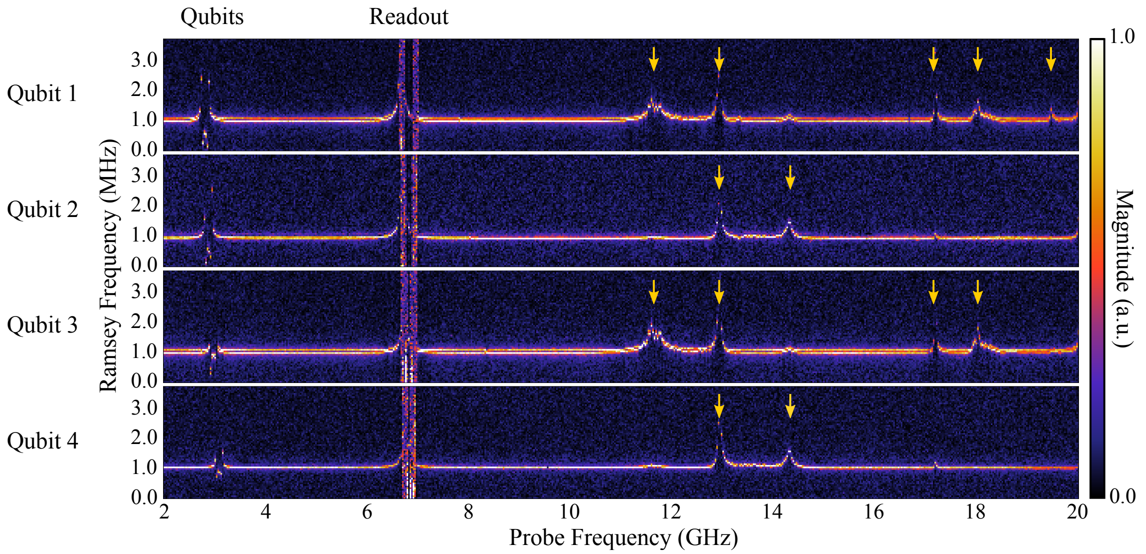


Figure 4-8: Experimental Package Mode Evaluation using four Qubits. The Ramsey spectra are taken simultaneously from four different qubits while a probe tone is injected through the central transmission line. Note the similarity between the spurious modes (indicated by yellow arrows) measured by qubits 1 and 3, as well as qubits 2 and 4. The orange lines indicate qubit-related features, which shift depending on the sensor qubit used.

the frequency spectrum. Furthermore, as the magnitude as the measured Ramsey frequency shift is roughly proportional to $1/\Delta$, the probe power is kept proportional to Δ—the frequency difference between the probe tone and the qubit—to resolve modes evenly across the spectrum of interest.

The device reported in the main text has a geometry of $5\,\mathrm{mm}$ by $5\,\mathrm{mm}$ and was fabricated following a similar process as described in [5] on a high resistivity $275\,\mu\mathrm{m}$ (001) Si wafer ($>3500\,\Omega\,\mathrm{cm}$). The chip consists of aluminum superconducting coplanar waveguides (CPW) and six superconducting fixed-frequency transmon qubits around $3\,\mathrm{GHz}$. The qubits are capacitively coupled to individual quarter-wave resonators that couple again inductively to a $50\,\Omega$ feedline in the chip's center. The resonator lengths are varied to frequency multiplex the resonances in the range of $6.69\,\mathrm{GHz}$ to $6.81\,\mathrm{GHz}$ with a spacing of approximately $25\,\mathrm{MHz}$.

4.4.7 Spatial Mode Characterization

An extension of the mode characterization performed in the main text is using a multi-qubit device to resolve spatially dependent modes. For example, different resonances in the package may couple to the qubits with different strengths depending on their electric field distribution. Such a technique is beneficial in identifying spurious modes that may be hidden from any given sensor qubit, particularly when the device area and field variability increase.

Qubit 2 is selected for Figure 4-9 as it had the longest coherence time to provide the best resolution in performing the Ramsey sweep. Here, the experiment is extended by performing simultaneous readout on four qubits (the remaining two qubits were not operational due to issues unrelated to the package). Conducting this multiplexed experiment came at the cost of a beating effect in the Ramsey oscillations of qubit 1 and 3, as seen by the two lines persistent throughout the probe frequency sweep. The beating is a consequence of spurious tones arising from instrumentation in the underlying particular multiplexed readout configuration. However, the relevant frequency shifts can still be identified.

The multi-qubit hidden mode survey in Figure 8 reveals several spurious modes positioned at the same frequency across all qubits. For example, the mode at $11.65\,\mathrm{GHz}$ more strongly affects qubits 1 and 3, while the mode at $14.3\,\mathrm{GHz}$ interferes with qubits 2 and 4, suggesting spatially-dependent coupling strengths. On the other hand, the mode at $12.94\,\mathrm{GHz}$ and the fundamental package mode at $17.18\,\mathrm{GHz}$ couple to all qubits and do not reveal a spatially-dependent coupling strength. Finally, several features between $2\,\mathrm{GHz}$ to $9\,\mathrm{GHz}$ (highlighted with orange vertical lines) vary in frequency depending on the sensor qubit used. These are qubit-dependent; namely, they are associated with the qubit $|0\rangle$ to $|1\rangle$ and $|0\rangle$ to $|3\rangle$ transitions, as well as the respective readout resonators located around $7\,\mathrm{GHz}$.

We attribute this to the chip layout: each set is located on one side of the device, separated by a transmission line. The different coupling strengths are likely due to asymmetries in chip placement. For example, an offset that causes stronger

Table 4.2: Identified Package Modes. The four identified package modes between $2\,\mathrm{GHz}$ and $20\,\mathrm{GHz}$. The table contains the mode's resonance frequency $\omega_m/(2\pi)$, the mode linewidth $\kappa/(2\pi)$, coupling strength to the qubit $g/(2\pi)$ as extrapolated using Equation 4.5, and the resulting qubit energy relaxation time $T_{\text{pkg mode}}^{\text{Purcell}} = 1/\gamma_{\text{pkg mode}}^{\text{Purcell}}$ (Equation 4.2).

Mode	$\omega_m/(2\pi)$ (GHz)	$\kappa/(2\pi)$ (MHz)	$g/(2\pi)$ (MHz)	$T_{\text{pkg mode}}^{\text{Purcell}}$ (ms)
I	11.65	25	13.05	2.77
II	12.94	53	14.15	1.48
III	14.30	81	18.23	0.73
IV	17.18	20	17.73	5.08

capacitive coupling between the chip and the interposer aperture on one side will affect the two sets of qubits differently.

This measurement only revealed two distinct mode profiles due to the symmetric arrangement of the chip. However, a measurement of this type may be beneficial in probing the mode structure of more complex devices. In larger chips, spurious modes can arise locally within the chip substrate, and qubits at different locations within the package may couple to these modes with varying strengths. Understanding the spatial distribution of modes is needed to ensure the consistent performance of all qubits on the device.

Four package modes are identified between $2\,\mathrm{GHz}$ and $20\,\mathrm{GHz}$. The linewidth κ of each spurious mode—a measure of its lossiness—can be determined directly by performing a fine frequency sweep of the hidden-mode experiment. Thus, the coupling strength can be calculated, as shown in Figure 4-9(c) for the fourth mode. The ac-Stark shift is proportional to the average number of photons $\Delta_{\text{Stark}} = \alpha\bar{n}$ with a factor α. Similarly, the mode induced dephasing is proportional to $\beta\bar{n}$ with a proportionality factor β. The measured Γ_2^* is given by

$$\Gamma_2^* = \Gamma_{\text{mode}} + \frac{1}{T_{2,\text{intrinsic}}^*} = \beta\bar{n} + \frac{1}{T_{2,\text{intrinsic}}^*}. \tag{4.4}$$

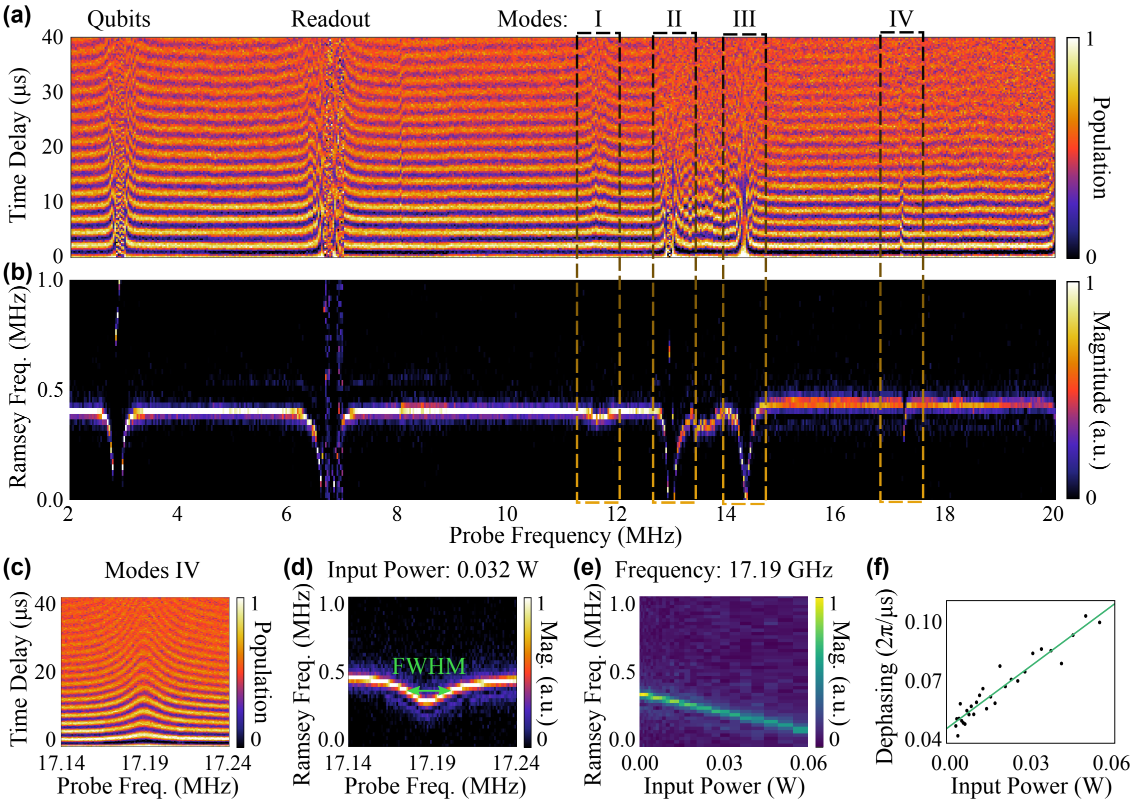


Figure 4-9: Experimental Package Mode Characterization. (a) and (b) The microwave environment between $2\,\mathrm{GHz}$ and $20\,\mathrm{GHz}$ registered by the qubit is mapped out using the qubit itself as a sensor. A continuous-wave probe tone—added to the qubit readout and control signal—excites the qubit microwave environment. Probe tone-dependent indirect and direct effects on the qubit are recorded using Ramsey spectroscopy. In the time-domain (a) and frequency-domain (b) panels, the qubits, readout resonators, and the ground to third excited state transition are identifiable as shifts in the Ramsey frequency. Furthermore, four features between $11\,\mathrm{GHz}$ and $18\,\mathrm{GHz}$ can be identified. The following mode characterization procedure is exemplified on the fourth mode, the expected package cavity mode shown augmented again in (c) at $17.18\,\mathrm{GHz}$. (d) The linewidth of a mode $\kappa/(2\pi)$ is the full width at half maximum (FWHM) of the Fourier transformed frequency-dependent Ramsey scan, here $20\,\mathrm{MHz}$. (e) The Ramsey frequency change as the power of the probe tone—parked at the mode's resonance frequency of $17.189\,\mathrm{GHz}$—is varied. The power is measured at the signal generator. (f) The qubit dephasing resulting from the varying probe tone power is extrapolated by performing fits to T_2^* experiments. The presented analysis yields a coupling rate of $g/2\pi \approx 17.73\,\mathrm{MHz}$ for Mode IV. Note, the Josephson traveling-wave parametric amplifier (JTWPA) was bypassed for this experiment to prevent interference with the probe tone.

The linear change in the Ramsey frequency against the input power feeding the mode enables the extrapolating of the absolute value of the slope, α, displayed in Figure 4-9(d). The slope's sign depends on whether the qubit drive tone is set above or below the qubit frequency for the measurement. Similarly, β is the slope of the Ramsey decay rate as the input power is increased. Combined, one can calculate the qubit-mode coupling strength

$$g = \sqrt{\frac{\beta\kappa(\omega_{\mathrm{m}} - \omega_{\mathrm{q}})}{4|\alpha|}}, \tag{4.5}$$

This method is demonstrated by performing a power sweep for Mode IV [Figure 4-9(b)] and yields a coupling strength $g/2\pi$ of $17.73\,\mathrm{MHz}$. Furthermore, the Purcell limit caused by the mode can be calculated using Equation 4.2 to be $5.09\,\mathrm{ms}$. Table 4.2 summarizes the identified modes and their characteristics. Extrapolating the limit these package modes (pkg mode) impose on the lifetime of the presented qubits, using

$$T^{\mathrm{Purcell}}_{\mathrm{pkg\ mode}} = 1 \Big/ \sum_i \gamma^{\mathrm{Purcell}}_{\mathrm{pkg\ mode_i}}, \tag{4.6}$$

$T^{\mathrm{Purcell}}_{\mathrm{pkg\ mode}} = 384\,\mu\mathrm{s}$ is obtained for the modes I-IV. While this is sufficient for current devices with coherence times in the range of $10\,\mu\mathrm{s}$ to $100\,\mu\mathrm{s}$, this result needs to be highlighted. Despite the absence of strong spurious modes up to $11\,\mathrm{GHz}$ in the discussed package, a comprehensive survey reveals that the higher frequency modes can still have a diminishing effect on qubit lifetime. These limits will likely be saturated soon as qubit lifetimes increase, underscoring the need for further package design improvements.

4.5 Summary & Conclusion

The approach to package design was validated by using a newly engineered package to examine various elements that can affect superconducting qubit coherence. A comprehensive characterization of the effect of package modes on superconducting transmon qubits corroborated with results from simulation tools and room-temperature measurements is presented. The main findings are summarized in Table 4.3. For the presented particular qubit design and configuration, the package limits the qubit lifetime to approximately $T_{\mathrm{limit}}^{\mathrm{pkg}} = 1/(\gamma_{\mathrm{pkg\ mode}}^{\mathrm{Purcell}} + \gamma_{\mathrm{material}}) = 384\,\mu\mathrm{s}$. This lifetime is due almost entirely to qubit loss to hidden package modes via the Purcell effect ($1/\gamma_{\mathrm{pkg\ mode}}^{\mathrm{Purcell}} = 384\,\mu\mathrm{s}$). Package material losses ($1/\gamma_{\mathrm{material}} = 9.87\,\mathrm{s}$) contribute only at the $15\,\mathrm{ns}$ level. While the package does not limit the lifetime of the measured qubits, the estimated lifetime limit is within the same order of magnitude as other loss channels. Constructing packages with larger devices and qubit lifetimes that are likely achievable soon will require a thorough engineering approach that focuses on mode and signal line engineering.

Looking forward, package design will become increasingly critical for larger quantum devices due to their increased complexity. As the number of qubits in today's noisy intermediate-scale quantum (NISQ) devices [54] increases, the precise characterization and suppression of electromagnetic modes and signal crosstalk become even more critical. These established principles for superconducting qubit packages are similarly pertinent for future work as packaging techniques are being advanced for systems in the range of 100 to 1000 qubits.

Here, a modular package design is proposed. In particular, a 24-line package, which is designed for chips of dimensions $5\,\mathrm{mm}$ by $5\,\mathrm{mm}$, can be up-scaled in a modular fashion for larger devices by subdividing the enclosing microwave cavity or pinning antinodes of the lowest mode via out-of-plane wires [55, 56]. Figure 4-10 depicts the modular approach for four subcells and 48-lines. The subdivision of the central cavity with fins as well as of the $10\,\mathrm{mm}$ by $10\,\mathrm{mm}$ qubit chip using through-silicon vias [41] prevent the formation of modes in the qubit frequency

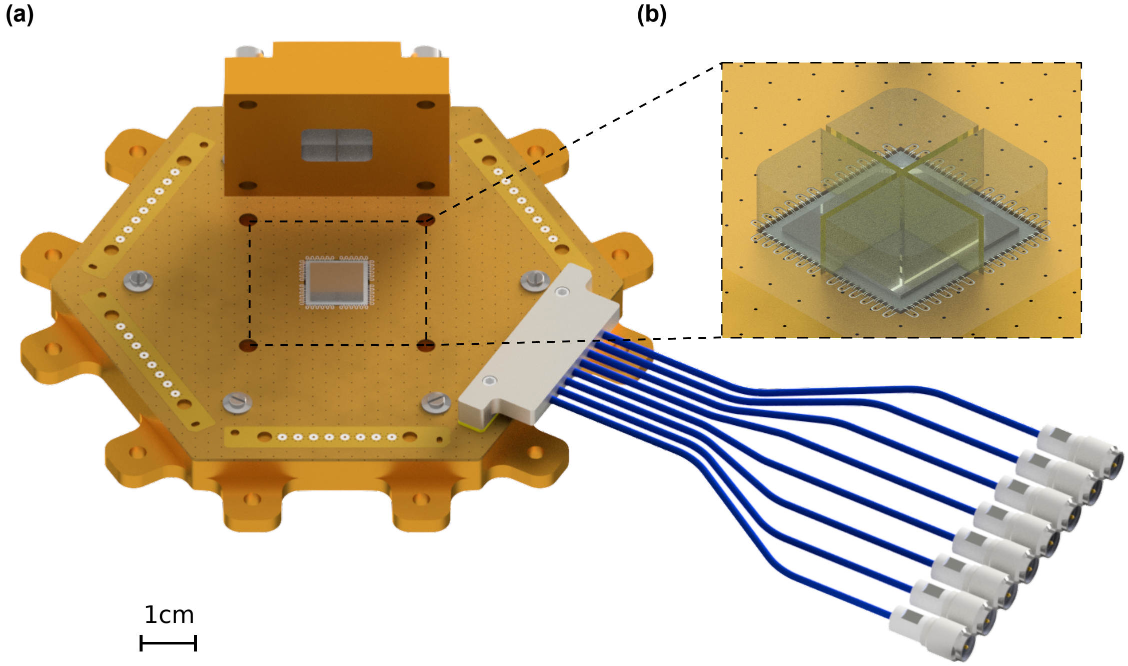


Figure 4-10: Next-Generation Packaging. (a) Blueprint for scaling up the package design to support NISQ devices. By dividing the cavity into subcells corresponding to $5\,\mathrm{mm}\times5\,\mathrm{mm}$ chip areas, each with the same number of signal launches, the crosstalk, and mode performance can be preserved while increasing the number of pins. Qubit control and readout are conducted via multi-coax Ardent connectors with eight channels each. (b) Fins or spring-loaded pins above the qubit device suppress spurious modes.

operation spectrum. The center aperture of the interposer is extended to provide 48 signal ports to the larger qubit chip. The signal transmission performance of the interposer is maintained as the distance between neighboring signal lines and their specific length is not altered. To efficiently connect the microwave package to the control signals, high-density, spring-loaded multi-coax connectors or waveguides built into flexible PCB ribbon cables [57] can be employed.

While current state-of-the-art packages still employ wirebonds to connect a multi-chip stack and the device package [1], several promising candidates such as pogo pins [37], direct chip-to-interposer connections, 3D-integrated packaging [58], and the aforementioned out-of-plane wiring [55] may provide a larger-scale interconnect solution. Such wiring techniques are directly compatible with the proposed modular package design. However, with a higher density, these techniques will face even more significant challenges in signal crosstalk, requiring precise impedance matching and the use of shielding structures. Combining these factors will necessitate thorough simulation and design characterization, building on those presented in this thesis.

Table 4.3: Package Design Considerations. Summary of package design considerations with a focus on qubit-energy-loss channels, impedance matching, signal crosstalk, and package-mode profile.

Design considerations	Associated design parameter
Material-induced qubit-energy losses	• The distance between the qubit and the package surfaces should be at least $2\,\mathrm{mm}$ [8]. • In the qubit vicinity (about $2\,\mathrm{mm}$), polymeric materials and silver paste with an electrical conductivity below that of copper should be avoided [8].
Impedance matching	• For connector and waveguide impedances within $10\,\%$ of each other, the reflected power remains below $0.25\,\%$. • For $1\,\mathrm{mm}$-long wirebonds, the capacitance is in the range of $20\,\mathrm{fF}$ and the inductance is around $1\,\mathrm{nH}$ [15] which corresponds to an impedance of $223\,\Omega$. For impedance matching, shorter wirebonds, several wirebonds in a V-shape, or impedance matching structures [45] can be employed.
Signal crosstalk	• Symmetrical buried striplines limit signal crosstalk [7, 28]. • Via-fences reduce crosstalk between waveguides. The spacing between vias should not exceed $\lambda/20$ to ensure that the via-fence appears solid to an impinging wave at wavelength λ [33]. • For a qubit chip mounted with wirebonds, three grounding wirebonds between signal wirebonds reduce the wirebond crosstalk by about $20\,\mathrm{dB}$.
Package-mode profile	• The fundamental package-mode frequency should be greater than twice the maximum qubit transition frequency. The fundamental mode of the package cavity can be increased by breaking the package cavity into subcavities [56]. • The chip mode frequency should be greater than twice the maximum qubit transition frequency. The chip mode depends on the chip dimensions and the via positioning [41].

References

[1] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, "Quantum supremacy using a programmable superconducting processor," Nature **574**, 505–510 (2019).

[2] G. Calusine, A. Melville, W. Woods, R. Das, C. Stull, V. Bolkhovsky, D. Braje, D. Hover, D. Kim, X. Miloshi, D. Rosenberg, A. Sevi, J. Yoder, E. Dauler, and W. Oliver, "Analysis and mitigation of interface losses in trenched superconducting coplanar waveguide resonators," Applied Physics Letters **112**, 062601 (2018).

[3] W. Woods, G. Calusine, A. Melville, A. Sevi, E. Golden, D. K. Kim, D. Rosenberg, J. L. Yoder, and W. D. Oliver, "Determining Interface Dielectric Losses in Superconducting Coplanar-Waveguide Resonators," Physical Review Applied **12**, 014012 (2019).

[4] 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**, 042319 (2007).

[5] F. Yan, S. Gustavsson, A. Kamal, J. Birenbaum, A. P. Sears, D. Hover, T. J. Gudmundsen, D. Rosenberg, G. Samach, S. Weber, J. L. Yoder, T. P. Orlando, J. Clarke, A. J. Kerman, and W. D. Oliver, “The flux qubit revisited to enhance coherence and reproducibility,” Nature Communications **7**, 12964 (2016).

[6] H. Paik, D. I. Schuster, L. S. Bishop, G. Kirchmair, G. Catelani, A. P. Sears, B. R. Johnson, M. J. Reagor, L. Frunzio, L. I. Glazman, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf, “Observation of high coherence in Josephson junction qubits measured in a three-dimensional circuit QED architecture,” Physical Review Letters **107**, (2011).

[7] D. M. Pozar, *Microwave engineering, 4th edition* (Wiley, 2012).

[8] B. Lienhard, J. Braumüller, W. Woods, D. Rosenberg, G. Calusine, S. Weber, A. Vepsäläinen, K. O’Brien, T. P. Orlando, S. Gustavsson, and W. D. Oliver, “Microwave packaging for superconducting qubits,” 2019 IEEE MTT-S International Microwave Symposium (2019).

[9] D. R. Lide, “Crc handbook of chemistry and physics: a ready-reference of chemical and physical data, 85th ed,” Journal of the American Chemical Society **127**, 4542–4542 (2005).

[10] C. M. Quintana, A. Megrant, Z. Chen, A. Dunsworth, B. Chiaro, R. Barends, B. Campbell, Y. Chen, I.-C. Hoi, E. Jeffrey, J. Kelly, J. Y. Mutus, P. J. J. O’Malley, C. Neill, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Cleland, and J. M. Martinisa, “Characterization and reduction of microfabrication-induced decoherence in superconducting quantum circuits,” Applied Physics Letters **105**, 062601 (2014).

[11] N. Fredj and T. D. Burleigh, “Transpassive Dissolution of Copper and Rapid Formation of Brilliant Colored Copper Oxide Films,” Journal of the Electrochemical Society **158**, C104–C110 (2011).

[12] J. E. Campbell, E. A. Eldridge, and J. K. Thompson, *Handbook on Materials for Superconducting Machinery* (Metals and Ceramics Information Center, 1974).

[13] D. R. Lide, *Handbook of Chemistry and Physics* (CRC Press, 1996), 11–41.

[14] N. T. Bronn, V. P. Adiga, S. B. Olivadese, X. Wu, J. M. Chow, and D. P. Pappas, "High coherence plane breaking packaging for superconducting qubits," Quantum Science and Technology **3**, 024007 (2018).

[15] J. Wenner, M. Neeley, R. C. Bialczak, M. Lenander, E. Lucero, A. D. O'Connell, D. Sank, H. Wang, M. Weides, A. N. Cleland, and J. M. Martinis, "Wirebond crosstalk and cavity modes in large chip mounts for superconducting qubits," Superconductor Science and Technology **24**, 065001 (2011).

[16] R. Barends, J. Kelly, A. Megrant, D. Sank, E. Jeffrey, Y. Chen, Y. Yin, B. Chiaro, J. Mutus, C. Neill, P. O'Malley, P. Roushan, J. Wenner, T. White, A. N. Cleland, and J. M. Martinis, "Coherent Josephson qubit suitable for scalable quantum integrated circuits," Physical Review Letters **111**, 080502 (2013).

[17] W. D. Oliver and P. B. Welander, "Materials in Superconducting Qubits," MRS Bulletin **38**, 816–825 (2013).

[18] A. Dupré, A. V. Itterbeek, and L. Michiels, "Heat conductivity of copper below 1K," Physics Letters **8**, 99–100 (1964).

[19] J. Sharma, "Heat conductivities below 1K," I. Cryogenics **7**, 141–156 (1967).

[20] E. T. Swartz and R. O. Pohl, "Thermal boundary resistance," Reviews of Modern Physics **61**, 605–668 (1989).

[21] J. Goetz, F. Deppe, M. Haeberlein, F. Wulschner, C. W. Zollitsch, S. Meier, M. Fischer, P. Eder, E. Xie, K. G. Fedorov, E. P. Menzel, A. Marx, and R. Gross, "Loss mechanisms in superconducting thin film microwave resonators," Journal of Applied Physics **119**, (2016).

[22] L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, "Demonstration of two-qubit algorithms with a superconducting quantum processor," Nature **460**, 240–244 (2009).

[23] M. A. Rol, L. Ciorciaro, F. K. Malinowski, B. M. Tarasinski, R. E. Sagastizabal, C. C. Bultink, Y. Salathe, N. Haandbaek, J. Sedivy, and L. DiCarlo, "Time-domain characterization and correction of on-chip distortion of control pulses in a quantum processor," Applied Physics Letters **116**, 054001 (2020).

[24] N. K. Langford, R. Sagastizabal, M. Kounalakis, C. Dickel, A. Bruno, F. Luthi, D. J. Thoen, A. Endo, and L. DiCarlo, "Experimentally simulating the dynamics of quantum light and matter at deep-strong coupling," Nature Communications **8**, 1715 (2017).

[25] W. H. Hayt and J. A. Buck, "Transmission lines," *Engineering electromagnetics* (McGraw Hill, 2001), 434–448.

[26] B. Foxen et al., "High speed flux sampling for tunable superconducting qubits with an embedded cryogenic transducer," Superconductor Science and Technology **32**, 015012 (2018).

[27] M. Sarovar, T. Proctor, K. Rudinger, K. Young, E. Nielsen, and R. Blume-Kohout, "Detecting crosstalk errors in quantum information processors," (2019).

[28] J. M. Sage, V. Bolkhovsky, W. D. Oliver, B. Turek, and P. B. Welander, "Study of loss in superconducting coplanar waveguide resonatorss," Journal of Applied Physics **109**, 06391 (2011).

[29] A. Isidoro-Munoz, R. Torres-Torres, M. A. Tlaxcalteco-Matus, and G. Hernandez-Sosa, "Scalable models to represent the via-pad capacitance and via-traces inductance in multilayer pcb high-speed interconnects," 2017 International Caribbean Conference on Devices, Circuits and Systems (ICCDCS), (2017).

[30] S. Haroche and J.-M. Raimond, *Exploring the quantum - atoms, cavities, and photons* (Oxford University Press, 2006).

[31] *Demystifying Vias in High-Speed PCB Design*, Keysight HSD Seminar, Mastering SI & PI Design, 2017.

[32] F.-J. Pajares, M. Ribo, J.-R. Regue, P. Rodriguez-Cepeda, and L. Pradell, “A multimodal analysis of the effects of guard traces over near wideband signal paths,” 2005 International Symposium on Electromagnetic Compatibility, (2005).

[33] I. J. Bahl, “Via holes,” *Lumped elements for rf and microwave circuits* (Artech House, 2003), 290–296.

[34] A. Suntives, A. Khajooeizadeh, and R. Abhari, “Using via fences for crosstalk reduction in pcb circuits,” 2006 IEEE International Symposium on Electromagnetic Compatibility, 2006. EMC 2006., (2006).

[35] H. W. Johnson and M. Graham, “Vias,” *High-speed digital design: a handbook of black magic* (Prentice Hall, 1993), 249–262.

[36] D. Rosenberg, D. Kim, R. Das, D. Yost, S. Gustavsson, D. Hover, P. Krantz, A. Melville, L. Racz, G. O. Samach, S. J. Weber, F. Yan, J. L. Yoder, A. J. Kerman, and W. D. Oliver, “3D integrated superconducting qubits,” npj Quantum Information **3**, (2017).

[37] J. H. Béjanin, T. G. McConkey, J. R. Rinehart, C. T. Earnest, C. R. H. McRae, D. Shiri, J. D. Bateman, Y. Rohanizadegan, B. Penava, P. Breul, S. Royak, M. Zapatka, A. G. Fowler, and M. Mariantoni, “Three-Dimensional Wiring for Extensible Quantum Computing: The Quantum Socket,” Physical Review Applied **6**, (2016).

[38] Q. Liu, M. Li, K. Dai, K. Zhang, G. Xue, X. Tan, H. Yu, and Y. Yu, “Extensible 3d architecture for superconducting quantum computing,” Applied Physics Letters **110**, 232602 (2017).

[39] R. N. Master, R. Jackson, S. K. Ray, and A. Ingraham, “Ceramic mini-ball grid array package for high speed device,” 1995 proceedings. 45th electronic components and technology conference (1995), 46–50.

[40] K. Yamanaka, H. Mori, and Y. Tsukada, "High performance ball grid array utilizing flip chip bonding on buildup printed circuit board," Twenty first ieee/cpmt international electronics manufacturing technology symposium proceedings 1997 iemt symposium (1997), 369–375.

[41] D. R. W. Yost, M. E. Schwartz, J. Mallek, D. Rosenberg, C. Stull, J. L. Yoder, G. Calusine, M. Cook, R. Das, A. L. Day, E. B. Golden, D. K. Kim, A. Melville, B. M. Niedzielski, W. Woods, A. J. Kerman, and W. D. Oliver, "Solid-state qubits integrated with superconducting through-silicon vias," npj Quantum Information **6**, 59 (2020).

[42] C. R. Paul, "Use of filamentary currents and current redistribution," *Inductance: loop and partial* (Wiley, 2010), 162–163.

[43] C. Snow, *Formulas for computing capacitance and inductance* (U.S. Government Publishing Office, 1954), 1–69.

[44] Y. Iossel, E. Kochanov, and M. Strunskiy, *The calculation of electrical capacitance* (Foreign Technology Div Wright-Patterson Afb Oh, 1969).

[45] S. Beer, B. Ripka, S. Diebold, H. Gulan, C. Rusch, P. Pahl, and T. Zwick, "Design and measurement of matched wire bond and flip chip interconnects for d-band system-in-package applications," 2011 IEEE MTT-S International Microwave Symposium, (2011).

[46] J. C. Whitaker, *The electronics handbook* (Taylor & Francis, 2005), 335.

[47] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "A quantum engineer's guide to superconducting qubits," Applied Physical Review **6**, (2019).

[48] D. I. Schuster, A. Wallraff, A. Blais, L. Frunzio, R.-S. Huang, J. Majer, S. M. Girvin, and R. J. Schoelkopf, "Ac stark shift and dephasing of a superconducting qubit strongly coupled to a cavity field," Physical Review Letters **94**, 123602 (2005).

[49] G. E. Ponchak, J. Papapolymerou, and M. M. Tentzeris, “Excitation of coupled slotline mode in finite-ground cpw with unequal ground-plane widths,” IEEE Transactions on Microwave Theory and Techniques **53**, 713 (2005).

[50] C. Schuster and W. Fichtner, “Parasitic modes on printed circuit boards and their effects on emc and signal integrity,” IEEE Transactions on Electromagnetic Compatibility **43**, 416–425 (2001).

[51] S. Sheldon, M. Sandberg, H. Paik, B. Abdo, J. M. Chow, M. Steffen, and J. M. Gambetta, “Characterization of hidden modes in networks of superconducting qubits,” Applied Physics Letters **111**, 222601 (2017).

[52] C. Macklin, K. O’Brien, D. Hover, M. E. Schwartz, V. Bolkhovsky, X. Zhang, W. D. Oliver, and I. Siddiqi, “A near–quantum-limited josephson traveling-wave parametric amplifier,” Science **350**, 307–310 (2015).

[53] E. A. Sete, J. M. Martinis, and A. N. Korotkov, “Quantum theory of a bandpass purcell filter for qubit readout,” Physical Review A **92**, 012325 (2015).

[54] J. Preskill, “Quantum Computing in the NISQ era and beyond,” Quantum **2**, (2018).

[55] J. Rahamim, T. Behrle, M. J. Peterer, A. Patterson, P. A. Spring, T. Tsunoda, R. Manenti, G. Tancredi, and P. J. Leek, “Double-sided coaxial circuit qed with out-of-plane wiring,” Applied Physics Letters **110**, 222602 (2017).

[56] P. Spring, T. Tsunoda, B. Vlastakis, and P. Leek, “Modeling enclosures for large-scale superconducting quantum circuits,” Physical Review Applied **14**, 024061 (2020).

[57] S. Deshpande, J.-P. Paquette, M. Vahidpour, M. Selvanayagam, R. Lion, M. Pelstring, S. Caldwell, M. Reagor, and D. Russell, “Integrating high-density microwave signalling and packaging with superconducting qubits,” (2019).

[58] D. Rosenberg, S. J. Weber, D. Conway, D. W. Yost, J. Mallek, G. Calusine, R. Das, D. Kim, M. E. Schwartz, W. Woods, J. L. Yoder, and W. D. Oliver, “Solid-

state qubits: 3d integration and packaging," IEEE Microwave Magazine **21**, 72–85 (2020).

Chapter 5

Machine Learning

Algorithms tasked to learn from data are attributed to the field of machine learning. The algorithm aims to identify a pattern in the data to optimize a decision process. The decision process can range from simple classification to complex optimal behavior in an unknown environment. In this thesis, the focus is on machine learning algorithms to improve superconducting qubit readout.

The following chapter is a basic introduction to machine learning and the tools used in this thesis. Of the three categories generally constituting machine learning—unsupervised, supervised, and reinforcement learning—the chapter focuses on supervised and reinforcement learning. In Sections 5.2-5.4, supervised learning methods ranging from matched filters to neural networks are presented to improve superconducting qubit-state discrimination. Section 5.5 focuses on reinforcement learning. Finally, deep reinforcement learning is discussed to improve superconducting qubit readout pulse shaping in Section 5.6.

5.1 Machine Learning Categories

The study of computational algorithms with the ability to improve through experience is typically referred to as machine learning [1]. These algorithms strive to identify patterns in sample data and create an approximate model of an underlying decision process without explicit instructions. While many machine learning ideas are several decades old, they only recently became widely applicable due to the development of sufficient computational resources and are applied today in many areas such as image processing [2], natural language processing [3], or playing advanced games such as chess [4].

Machine learning can be broadly divided into three categories: unsupervised, supervised, and reinforcement learning [1].

5.1.1 Unsupervised Learning

The class of unsupervised learning algorithms comprises approaches to identify patterns in unlabelled data. Such methods typically relate new data points with the previously processed ones to recognize underlying patterns. However, the resulting interpretation and performance evaluation of unsupervised learning techniques can at times be cumbersome.

5.1.2 Supervised Learning

In supervised learning, trusted data-label pairs, the training set, are utilized to make predictions on the labels of an unseen dataset. Typically, the input-output pairs for training are acquired by the 'supervisor,' hence the terminology. For complex predictions, the training of a supervised learning model can require large amounts of training instances. The quality of the learned mapping function can be probed utilizing an additional set of trusted input-output pairs, the test set. Comparing the performance of a supervised learning method on the training set compared to the test set is referred to as generalization.

5.1.3 Reinforcement Learning

Reinforcement learning algorithms aim to have a so-called agent learn appropriate actions given a specific situation of the particular environment. The agent improves the behavior by maximizing a reward. An essential aspect of the algorithm's success rate is the balance between exploration and exploitation. Unlike in unsupervised or supervised learning techniques, the agent learns from interacting with the environment instead of a prepared dataset.

In this thesis, the focus is first on supervised learning methods to optimize the classification of qubit-state readout data. Secondly, reinforcement learning algorithms are discussed and how such methods can be used to optimize qubit-readout pulses. The developed tools to optimize qubit-readout pulse shaping and qubit-state discrimination replace the current manual qubit-readout calibration process with a process governed by machine learning.

5.1.4 Supervised Learning for Qubit-State Discrimination

A qubit-state measurement is a multi-dimensional complex vector. Each dimension represents a time-bin. The dimensionality can be reduced with filters. A frequently-used filter is a rectangular window filter (boxcar filter) that projects the complex vector onto a real and imaginary part. Matched filters project the complex vector to a single real scalar value. Generally, discriminators using filtered data are computationally more efficient at the cost of a reduced discrimination accuracy relative to discriminators using unfiltered multi-dimensional data. In the following section, discriminators using filtered data are discussed first. Next, a focus is put on support vector machines before neural networks are discussed.

5.2 Matched Filter

A matched filter is the optimal linear filter in terms of the signal-to-noise ratio (SNR) in the presence of additive stochastic noise [5]. The matched filter aims to reduce the noisy measurement data to a single scalar to differentiate a ground-state and excited-state measurement. In the absence of state transitions during the qubit measurements, a linear matched filter is optimal in SNR and measurement fidelity [6]. The scalar values of the output of the matched filter can be discriminated by employing a simple threshold as the discriminator boundary. While such classifiers are not classical machine learning algorithms, the filter tune-up and threshold optimization require a 'training' step using a set of labeled measurement data.

A discrete qubit measurement signal $s_{|i\rangle}[n]$ with qubit either in the ground $(i = 0)$ or excited-state $(i = 1)$ can be modeled as $s_{|i\rangle}[n] \propto \alpha_{|i\rangle}[n] + \xi[n]$. The resonator response for the qubit in the ground-state is $\alpha_{|0\rangle}[n]$ and for the qubit in the excited-state $\alpha_{|1\rangle}[n]$. The measurement signal is a linear combination of the resonator response and a zero-mean stochastic noise term $\xi[n]$ [6, 7]. The filtered measurement signal with filter kernel $k[n]$ is described as

$$S_{|i\rangle} = \sum_n k_n s_{|i\rangle\, n}. \tag{5.1}$$

The mean difference $\langle \Delta S \rangle$ between the ground-state signal $S_{|0\rangle}$ and excited-state signal $S_{|1\rangle}$ follows as

$$\langle \Delta S \rangle = \langle S_{|0\rangle} - S_{|1\rangle} \rangle = \sum_n k_n \langle \alpha_{|0\rangle\, n} - \alpha_{|1\rangle\, n} \rangle. \tag{5.2}$$

The variance of the average difference $\langle \Delta S \rangle$ is consequently

$$N^2 = \text{var}(\Delta S) = \sum_n k_n^2 \left[\text{var}\left(\alpha_{|0\rangle\, n} - \alpha_{|1\rangle\, n} \right) + \text{var}(\xi_n) \right]. \tag{5.3}$$

The optimal filter kernel elements k_n can be derived as follows,

$$\frac{\partial}{\partial k_n} \frac{|\Delta S|}{N} = 0 \quad \rightarrow \quad \text{solve for } k_n \quad \Rightarrow \quad k_n = \frac{\langle \alpha_{|0\rangle\, n} - \alpha_{|1\rangle\, n} \rangle}{\text{var}\left(\alpha_{|0\rangle\, n} - \alpha_{|1\rangle\, n} \right) + \text{var}(\xi_n)} \tag{5.4}$$

A more thorough derivation of the optimal kernel element is presented in Appendix C.

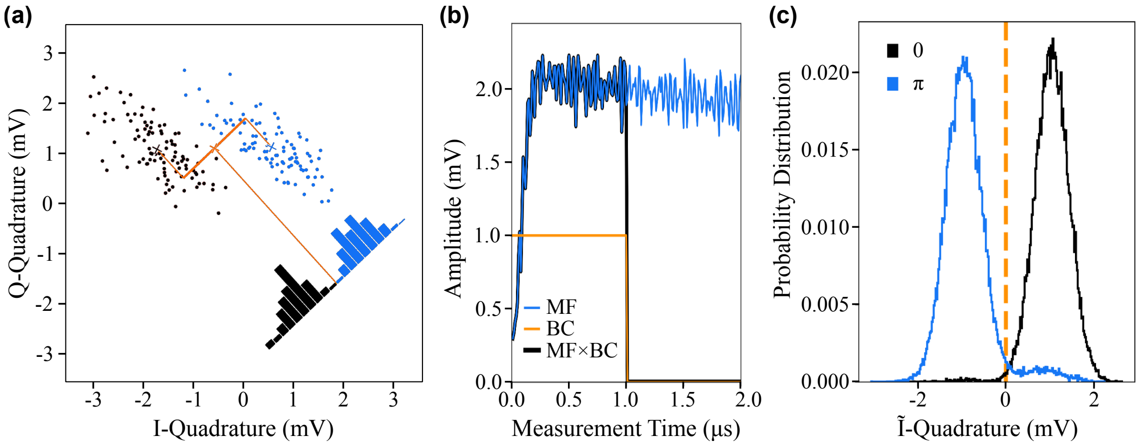


Figure 5-1: Matched Filter Threshold Discriminator. (a) A matched filter projects the input data to a single dimension. The projection is optimized in terms of signal-to-noise ratio. (b) The matched filter trace for a qubit measurement. (b) Windowed matched filter (MF) kernel using a boxcar filter (BC) to reduce the impact of qubit-state transitions. (c) Histogram of matched-filtered qubit-state measurements. The discrimination boundary is situated at the origin for a two-class system complying with the assumed noise model for the described linear matched filter.

For superconducting qubit measurements, the assumption of negligible state transitions is often violated. Therefore, the derived kernel needs to be filtered by an additional window function. The rectangular window function balances the maximum qubit-state assignment fidelity by reducing the considered measurement data to limit the increasing probability of state transitions, as depicted in Figure 5-1(b). Therefore, a matched filter is only approximately optimal when the measurement time is considerably shorter than the qubit-state lifetime. As a consequence, more advanced discriminators with fewer underlying assumptions may provide improved discrimination fidelities.

5.3 Support Vector Machine

Support vector machines (SVMs)—known for their robustness and good generalization—are fundamental two-class discriminators that draw a single decision boundary, called a hyperplane, in a supervised learning scheme [8, 9]. The margin between the classes and the hyperplane can be maximized by penalizing misclassified data

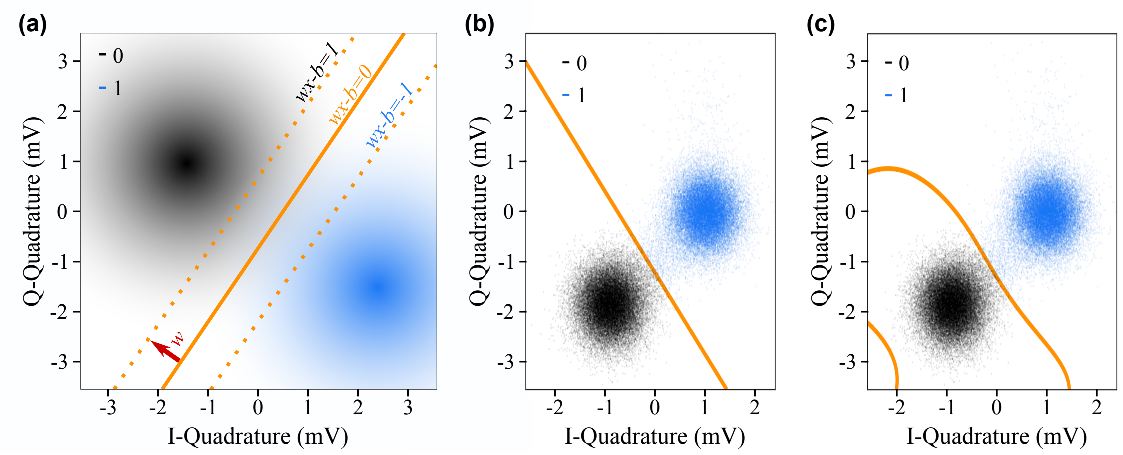


Figure 5-2: Support Vector Machines (SVM). (a) The linear decision boundary can be mathematically described as $wx - b = 0$. The margin width—the distance between the two dashed lines is equal to $2/||w||$. Depending on which side of the discrimination boundary, the label is either 1 or -1. The Linear SVM. (b) Linear SVM and (c) Nonlinear SVM are used to discriminate the states of a single qubit.

points and data points within the margin boundaries. The penalty for data points within the margin boundaries can be varied using a regularization term. A lenient penalty results in a so-called soft-margin SVM, which can better cope with not linearly separable problems. The generalized equation to optimize a soft-margin SVM with output values $y = 1$ for the first class and $y = -1$ for the second is

$$\min_{w,b,\xi} \frac{||w||^2}{2} + C\sum_{n=1}^{N} \xi_n \quad \text{subject to} \quad y^{(i)}\left(w^T x^{(i)} + b\right) \geq 1 - \xi_i, \;\; i \in 1, \ldots, N, \tag{5.5}$$

where w is a vector normal to the decision boundary, b is a parameter to determine the offset of the decision boundary from the origin, ξ is a positive slack variable (slack variables are introduced to allow specific constraints to be violated), and C is a parameter to indicate the 'softness' of the SVM. For a soft-margin SVM, C>0; and if C=0, then the SVM is referred to a hard-margin SVM. The SVM and the optimization parameters are graphically illustrated in Figure 5-2.

The hyperplane dimension equals one less than the number of features–the input data's dimensions. The location of a new data point relative to the hyperplane decides the associated label. Unfortunately, this deterministic decision process is not

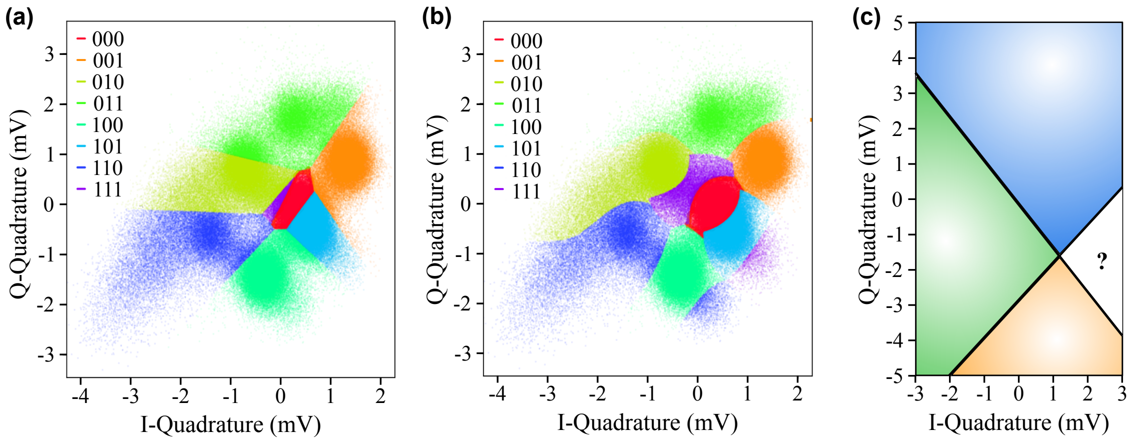


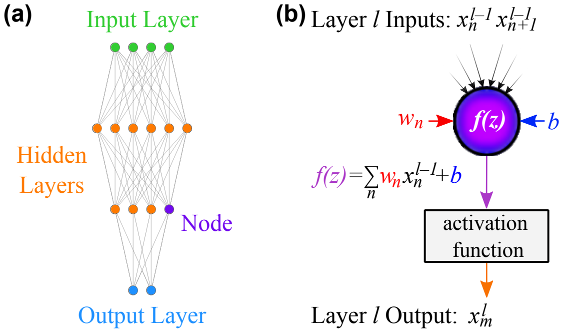
Figure 5-3: Multi-Class Support Vector Machines (SVM). (a) A linear and (b) nonlinear SVM are used to discriminate the states of three qubits (eight states). (c) A SVM tasked to discriminate more than two classes can show a diminished discrimination accuracy due to ambiguous areas. For three classes, a plane should show three distinct areas. However, dividing a plane with two lines generally results in four areas, meaning one area is not assigned to a single class.

probabilistic, and the information on the probability of label association is thus not directly accessible. While hyperplane separations only work for linearly-separable data, nonlinear SVMs map the data points to higher dimensions via a nonlinear transformation and find a hyperplane in that higher-order feature space.

5.3.1 Multi-Class Support Vector Machine

Several SVMs can be trained in concert for multi-class discrimination to divide the feature space into areas associated with distinct classes [10]. For an N-class ($N > 2$) classification task, the number of necessary hyperplanes is at least $N-1$ if each class is discriminated against the rest, referred to as 'one-versus-all.' Each class requires a hyperplane separating itself from the remaining collective of classes. However, separating space in more than two classes results in ambiguous areas that cannot be associated with a single class [1], as illustrated in Figure 5-3(c).

Figure 5-4: Neural Network Architecture. (a) Basic feedforward neural network architecture with an input layer, two hidden layers, and an output layer. (b) Neural network node values, apart from input layer nodes, are determined by calculating the product between the previous layer's node values x^{l-1} and their associated weights w plus a bias value b. An activation function further processes the resulting value.


(a)
Input Layer
Hidden Layers
Node
Output Layer
(b) Layer l Inputs: x_n^{l-1} x_{n+1}^{l-1}
w_n
$f(z)$
b
$f(z)=\sum_n w_n x_n^{l-1}+b$
activation function
Layer l Output: x_m^l

5.4 Neural Networks

Neural networks are inspired by biology and consist of layers comprising nodes. Typically, a neural network consists of an input layer composed of several nodes—the number of nodes depends on the input data dimension—and an output layer that contains the computed output values. In between the input and output layer are layers of neurons—so-called hidden layers as their output value is not directly accessible—with unique tasks per layer. A neuron's input and output channels are called edges, illustrated in Figure 5-4(a). Each neuron can be described as a mathematical function of incoming weighted parameters—typically output values of other neurons—and external parameters. The function output generally passes through a typically nonlinear filter, often referred to as an activation function, before serving as an input to other neurons, depicted in Figure 5-4(b). Varying the connectivity, neuron functions, and the activation function at each neuron output provides a flexible toolset to engineer a broad spectrum of neural network types. Supervised training of such a network can optimize the weights for each neuron input and external parameter to arbitrarily approximate any function. Training a neural network is, in its essence, the training of the weights of all nodes' input values and their bias terms to achieve a particular objective, such as optimal qubit-state discrimination.

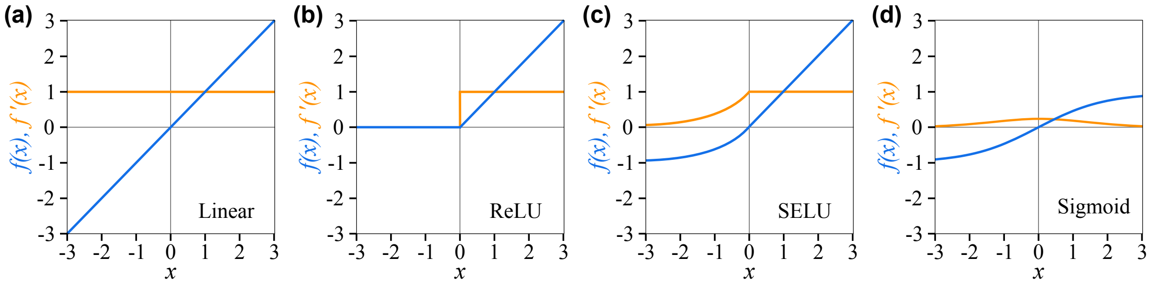


Figure 5-5: Activation Functions. Activation functions $f(x)$ are shown in blue, and their respective derivatives $f'(x)$ are shown in orange. (a) Linear, (b) Rectified Linear Unit (ReLU), (c) Scaled Exponential Linear Unit (SELU), and (d) Sigmoid are common activation functions used in machine learning.

5.4.1 Activation Functions

The linear activation function returns the activation function's input value directly to the output. The linear character of this activation function makes it impractical for most neural network applications. nonlinear activation functions are preferred.

The Rectified Linear Unit (ReLU) is a commonly used default activation function. The ReLU function outputs the input value only if the input value exceeds zero and outputs zero otherwise. ReLU functions are commonly used due to their simple structure resulting in computationally efficient optimization routines.

Scaled Exponential Linear Unit (SELU) [11] are similar to ReLU activation functions. SELU's differ from ReLUs if the input to the activation function is negative. While ReLU's output is zero for $x < 0$, a SELU outputs $e^{-x} - 1$. This seemingly minor change reduces the risk of a node dying due to the activation function outputting zero. With a SELU, negative inputs can be processed and remain distinct. Moreover, SELU's preserve the mean and variance from the previous layer and are thus self-normalizing.

Sigmoid activation functions are common in the output layer of neural networks used for binary classification. Typically, sigmoid functions are used for single node output layers. **Softmax activation functions** are used for multi-node outputs and thus multi-class classification and prediction tasks. Mathematically, the sigmoid function is expressed as $f(x) = 1/(1 + e^x)$ whereas the softmax function is equal to $f(x) = e^x / \sum_i e^{x_i}$.

5.4.2 Neural Network Training

The training phase is a critical element of any machine learning algorithm. During the training phase, the neural network is altered such that a defined loss function is minimized. The loss function maps the neural network's output value to a 'cost' value. The lower the cost value, the better the neural network is calibrated. Multiple training cycles, referred to as epochs, are required to ensure the neural network's output converges to the desired objective and minimizes the loss function. The number of epochs to reach a convergence plateau depends on the correction factor per cycle, the learning rate. A common strategy is to have a learning rate of 0.001—a typical value for neural networks—and gradually decrease it as the performance starts plateauing.

Similar to most optimization routines, neural networks are susceptible to overfitting, and thus, the generalization is reduced. The more parameters to be optimized, the more critical it is to choose an appropriate number of training instances to prevent under- or overfitting. Several potential regularization techniques can be used to prevent overfitting. One example is to include a regularizer in the loss function. A regularizer penalizes the model for updating too large individual weights. Randomly dividing the entire training set into normalized sub-training units, termed batches [12], is an alternative approach to regularization. Instead of updating the neural network after each training instance, the neural network is updated after each batch is processed. The choice of batch size affects the wall-clock-training time and generalization, or in other words, how well the discriminator performs on unseen data compared to the training set.

5.4.3 Neural Network Archetypes

Several neural network archetypes are analyzed to identify the most useful one in improving the qubit-state assignment fidelity and measurement time of multi-qubit devices. Fully-connected feedforward neural networks (FNN), also called multi-layered perceptron—among the most elementary neural networks—convolutional neural networks (CNN)—among the most successful image classification methods in use today—and long short-term memory (LSTM) recurrent neural networks—among the most successful architectures in language processing are explored here.

Feed forward neural networks (FFN) [13] are among straightforward neural networks. Information is fed into an input layer and transferred through hidden layers to an output layer. Generally, two adjacent layers are fully connected, meaning every neuron of one layer is connected to every neuron of the adjacent layers. FNNs are typically trained through back-propagation, giving the network trusted input-output pairs.

Convolutional neural networks (CNN) [2] are primarily used for image processing and thus quite different from most other networks. Generally, CNNs start with a rasterized scan of the input data. The resulting scanned cells can be overlapping and are then fed through a convolutional layer. Unlike nodes in FNNs that are often connected to all nodes of the previous layer, the nodes of the convolutional layer focus only on nearby cells. The raster scan and convolution can be repeated. Typically, subsequent convolutional layers decrease in size. In addition to convolutional layers, CNNs employ so-called pooling layers. Pooling layers filter out details. For example, the input of a pooling layer may output only the maximum value of the input cell. The convolutional and pooling layer take a local approach to data processing. To globally process the convoluted and pooled data, CNNs frequently finish with an FNN.

Long short-term memory (LSTM) networks [14] are composed of nodes comprising a memory unit and three gates (input, output, and forget). The gates control the flow of information. The input gate determines how much information from the previous layer is stored in the current cell. Consequently, the output gate controls the outflow of information. Finally, the forget gate controls the loss of information. At times, the network performance can be improved if some information is ignored. In general, LSTMs are challenging to train as there are three gates and associated weights to be trained for each node.

In this thesis, the three presented neural network archetypes are used for qubit-state discrimination. As discussed in the next Chapter 6, the feedforward neural network outperforms its alternatives.

5.5 Reinforcement Learning

Reinforcement learning describes algorithms with an agent aiming to maximize a reward for an action taken in a specific environment. The agent optimizes the actions by interacting with the environment rather than analyzing labeled datasets. A key aspect of reinforcement learning is that the agent does not need to know the underlying system's dynamics, and thus, is independent of a specific model. Instead, the agent generates the necessary model of the environment through the reward for specific actions. The action alters the state of the agent in the environment. Typically, the agent attempts to create a target state through actions optimized via a reward depending on the last created state, as illustrated in Figure 5-6. The effect an action has on the environment and the resulting state is referred to as state-action pairs. Reinforcement learning has gained attraction due to its success in complicated environments such as Chess and Go [15, 16] and robot automation and planning [17].

Reinforcement learning tasks are generally composed of four components. First, the environment defines the playground for an agent. It defines the states an agent can be in and what action is necessary to reach particular states. The agent is often only familiar with the direct vicinity and thus has to explore the environment to evaluate appropriate actions to optimize its state. Second, a so-called policy describes a set of guidelines to pick the following action optimally. The policy is

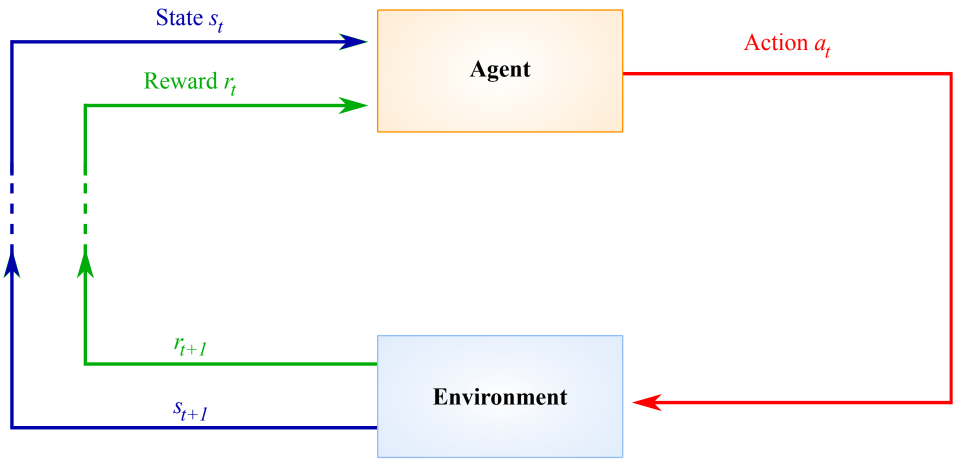


Figure 5-6: Markov Decision Process. Given a current state s_t and reward r_t at time t, an agent decided on a next action a_t. Based upon the action the state changes to s_{t+1} and a reward r_{t+1} can be associated with the new state. The cycle continues until a specific number of cycles or a particular reward is reached. To minimize computational effort, the cycle can also be stopped once the agent's progress has plateaued.

learned indirectly or directly through the interaction with the environment. Third, a reward is generated depending on the agent's goal and current state in the environment. The agent aims to maximize the reward and updates the policy accordingly. The type of reward depends on the environment, can be negative or positive, and may not be readily available after each action. Lastly, future values and potential rewards are estimated through a value function. The value function expresses the total reward an agent can expect depending on its current state and policy. The quality of a value function determines the optimization efficiency. A good value function can circumvent local optima and more efficiently reach a global optimum. The quality of a value function can be improved through the exploration of the environment.

5.5.1 Markov Decision Processes

Reinforcement learning can be mathematically described by finite Markov decision processes (MDP). For each time step t, the environment provides the agent with a current state s_t and reward r_t. The reward values the prior action a_{t-1} taken by the agent. The agent learns to understand the environment through the state-action pairs and improves its internal mapping based on the received reward. The next action is taken with that information and learned policy in mind. This routine continues until the desired state, or a specific number of cycles, is reached. For MDP models, the transitions are probabilistic, and each action only depends on the previous step.

In the generalized MDP framework, the value function is often encoded in the form of the Bellman equation [18], defined as

$$V(s) = \max_a(r(s,a) + \gamma \sum_{s'} P(s'|s,a)V(s')), \tag{5.6}$$

with the value s determined by the possible action a maximizing both the reward r and future value discounted by factor γ. The Bellman equation can be used to either optimize the value function indirectly through value iteration to infer an optimal policy or directly by using policy iteration.

5.5.2 Value Iteration

Optimal policies for the MDP problem can be found with value iteration. Value iteration aims to determine the optimal value function $V(s)$. The agent improves at optimizing the value function as experience is being gained. Consequently, the agent needs to explore many states and some even multiple times. While this is time-consuming and computationally expensive, knowing an optimal value function enables one to infer an optimal policy simply by tapping into that value function. With the optimal value function at hand, a deterministic policy Π can be generated using the algorithm presented in 1.

Algorithm 1 Value Iteration [19]

```
Randomly initialize $V(s)$ for all $s \in S$.
Initialize $\Delta > 0$ as a small threshold to determine if the value function has converged.
loop
    $\delta \leftarrow 0$
    for $s \in S$ do
        $v \leftarrow V(s)$
        $V(s) \leftarrow \max_a(r(s,a) + \gamma \sum_{s'} P(s'|s,a)V(s'))$
        $\delta \leftarrow max(\delta, |v - V(s)|$
until $\Delta > \delta$
$\Pi'(s) = \mathrm{argmax}_a(r(s,a) + \gamma \sum_{s'} P(s'|s,a)V(s'))$
```

5.5.3 Policy Iteration

Alternatively, policy iteration can determine the optimal policy directly with the algorithm presented in 2. Policy iteration differs from value iteration. In policy iteration, actions are selected from a current policy compared to computing the value of every available action and post-selecting an action. Moreover, the optimal policy is continually optimized in policy iteration. Value iteration presents an optimal policy as its very last step. Policy iteration excels in environments with a continuous action space. Determining a value for each state in a continuous environment is impossible. For discrete action spaces, however, value iteration techniques are typically more efficient [19].

5.5.4 Exploration and Exploitation

To guarantee the success of any reinforcement learning algorithm, exploration and exploitation need to be balanced. Depending on the algorithm, a larger action-state space needs to be explored. Furthermore, an algorithm needs to be sensitive enough to realize when a state is worth exploiting. However, exploiting a state is risky as the particular state might be simply a local optimum. Therefore, sufficient exploration and conscious exploitation are fundamental to the algorithm's success.

A common strategy to balance exploration and exploitation is ϵ-greedy action selection. An action is randomly taken with probability ϵ or following the generated

Algorithm 2 Policy Iteration [19]

```
Randomly initialize $V(s)$ and $\Pi(s) \in A$ for all $s \in S$.
Initialize $\Delta > 0$ as a small threshold to determine value function has convergence.

(a) Policy Evaluation
loop
    $\delta \leftarrow 0$
    for $s \in S$ do
        $v \leftarrow V(s)$
        $V(s) \leftarrow \max_a(r(s, \Pi(s)) + \gamma \sum_{s'} P(s'|s, \Pi(s))V(s'))$
        $\delta \leftarrow max(\delta, |v - V(s)|$
until $\Delta > \delta$

(b) Policy Improvement
policy-stable ← true
for $s \in S$ do
    old-action ← $\Pi(s)$
    $\Pi(s) \leftarrow \text{argmax}_a(r(s, a) + \gamma \sum_{s'} P(s'|s, a)V(s'))$
    if old-action $\neq \Pi(s)$ then
        policy-stable ← false

if policy-stable $=$ false then
    Go to (a)
```

policy with probability $1 - \epsilon$. Over time, ϵ is reduced to account for the exploration progress. With a decreasing ϵ, the algorithm starts to exploit the state increasingly.

5.6 Deep Reinforcement Learning Algorithms

The environment and state-action-space are often too large to explore. Therefore, value or policy iteration approaches can fail. Deep neural networks can be used to approximate the value function or policy distribution to remedy this issue. Reinforcement learning algorithms using deep neural networks as approximators are classified as deep reinforcement learning algorithms.

5.6.1 AlphaZero

AlphaZero is a famous deep reinforcement learning algorithm due to its success in outperforming human master players in games such as Chess or Go [15, 16]. Astonishing about AlphaZero is that its training did not include human gameplay samples for training. In quantum computing, AlphaZero has been employed in combination with quantum optimal control to calibrate two-qubit gates in simulation [20].

AlphaZero is centered around deep neural networks in conjunction with a deep look-ahead in a guided tree search. Such a setup enables a predictive hidden-variable approximation of the state-action pairs. Monte Carlo Tree Search (MCTS) is used to look ahead and evaluate different strategies. The capability to look ahead is limited for many reinforcement learning algorithms to a single step. At the same time, AlphaZero can evaluate nearly an arbitrary number of steps into the future and is only limited by computational resources. The computational effort remains manageable because MCTS can efficiently explore state-action pairs due to its probabilistic approach. Many other algorithms randomly pick initial actions to explore their environment and can thus be less efficient. The efficiency of MCTS due to its probabilistic approach comes at the price of a more sophisticated model requiring more frequent updates and computational resources.

5.6.2 Deep-Q Network

Deep-Q Network (DQN) algorithms [21] are an extension of the popular Q-Learning algorithm that allows one to explore more significant state and action spaces and take advantage of modern deep neural network techniques and hardware optimizations. In tabular Q-Learning settings, the aim is to find an optimal policy by creating a Q-table that holds a value for each state-action pair. Each table entry signifies the current value of a pair by acting in that given state and the agent's future value by making decisions from that new state-action pair. With enough iterations, all optimal values for the Q-table are computed. When the table is complete, an agent can simply find an optimal solution by performing table look-ups and selecting the action that produces the largest Q-value at each state.

The primary issue with tabular Q-Learning approaches is the reliance on storing the Q-table in memory and the algorithm's necessity to explore many state-action pairs. This becomes infeasible for environments with extensive and precise states and action spaces. Deep Q-Networks replace the Q-table with a deep neural network that takes in a state input and outputs the approximated Q-values for each action. The past network predictions are stored in a replay buffer for sampling or to prioritize samples with higher rewards periodically [22]. This algorithm requires a discrete action-space as the network produces a predicted Q-value for each possible action when provided a state. The hardware output and input channels are typically discrete and can be directly mapped to the action space.

5.6.3 Deep Deterministic Policy Gradient

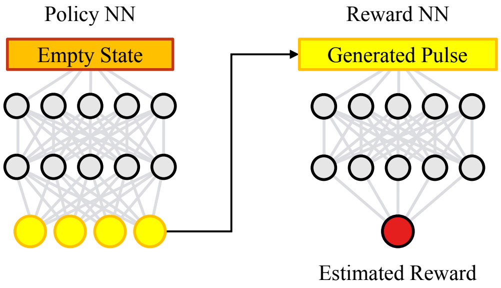
While Deep-Q Learning is a robust deep reinforcement learning tool, it is potentially limited in performance because it requires a discrete action space which can be challenging to train. In contrast, the Deep Deterministic Policy Gradient (DDPG) algorithm, a model-free algorithm that can learn optimal policies in high dimensional and uses continuous action-spaces [23]. DDPG utilizes an actor-critic method, in which an 'Actor' agent uses the current state to approximate an optimal continuous action. In contrast, the ‘Critic’ agent approximates the expected reward for the new state-action pair. Actor-critic methods work well compared to single action-value methods like DQN because they require less computation to operate in vastly larger action spaces. The separation into two processes increases overall stability [19]. Similar to DQN, deep neural network approximators can be used for the Actor and Critic.

5.6.4 Proximal Policy Optimization

Proximal policy optimization (PPO) algorithms [24] intend to identify the maximum possible improvement without unnecessary risks through an evaluation of the acquired data. Two neural networks underlay this algorithm; one serves as a policy and one as a reward estimator. The search for an optimal improvement step is an iterative process. First, a single iteration determines a maximum step size to explore the area around the current state, the trust region, before an optimal next step is defined inside this region. For conservative implementations, the following action clips the optimal steps if necessary. The area is explored using the two networks without direct interactions with the environment. The policy neural network creates a new action while the second neural network estimates the associated reward. This process enables computationally efficient exploration of the surrounding area to determine the next action-state pair. In contrast to a similar type of algorithms—trust region policy optimization [25]—PPO algorithms computationally simplified this optimization task to first-order methods such as gradient descent.

PPO belongs to a class of reinforcement learning algorithms called policy gradient methods. Meaning, the algorithm aims to learn optimal policies directly instead of inferring them from optimal value functions. PPO trains a stochastic policy in an on-policy way. This means that it explores by sampling actions according to the latest version of its stochastic policy. The amount of randomness in action selection depends on both initial conditions and the training procedure. The policy typically becomes progressively less random throughout training, as the update rule encour-

Figure 5-7: Proximal Policy Optimization Architecture. The algorithm is composed of two deep feedforward neural networks. One neural network (the policy network) generates a new action, whereas the second neural network (value network) tries to associate an expected reward to the proposed new action.



ages it to exploit rewards that it has already found. This may cause the policy to get trapped in local optima.

Actor-Critic Approach

PPO utilizes a so-called actor-critic approach. An ‘actor’ agent approximates an optimal continuous action using the current state. A ‘critic’ agent, on the other hand, approximates an expected value to the proposed new state-action pair. The PPO algorithm with the actor and critic agent is shown in Figure 5-7. Actor-critic methods remain computationally efficient even for vast numbers of action-state pairs. Furthermore, the separation of the tasks into two agents increases the overall optimization robustness [19].

5.7 Summary

This thesis focuses first on qubit-state discrimination and then on qubit-readout pulse shaping. For qubit-state discrimination, a feedforward neural network is used to discriminate multi-qubit readout-states. As elaborated, the qubit transition dynamics often violate the noise model underlying matched filters. SVMs, on the other hand, can process single-qubit data but are underperforming on multi-qubit data. Consequently, neural networks are better multi-qubit discriminators as there are no restrictions on the underlying noise model nor the number of classes that can be distinguished.

The proximal policy optimization (PPO) algorithm [24] is used to create qubit-readout pulses. PPO algorithms identify the maximum possible improvement without unnecessary risks by evaluating the acquired data. In contrast to a similar type of algorithm, trust region policy optimization [25], PPO algorithms computationally simplified the optimization task to first-order methods such as gradient descent. As shown in Figure 5-7, two neural networks underlay this algorithm: one serves as a policy and one as a reward estimator, which creates a new action. At the same time, the second estimates the associated reward. These processes enable computationally efficient exploration of the surrounding area to determine the next action-state pair.

References

[1] C. M. Bishop, *Pattern recognition and machine learning (information science and statistics)* (Springer-Verlag, Berlin, Heidelberg, 2006).

[2] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE **86**, 2278–2324 (1998).

[3] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: pre-training of deep bidirectional transformers for language understanding," (2019).

[4] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, and D. Hassabis, "A general reinforcement learning algorithm that masters chess, shogi, and go through self-play," Science **362**, 1140–1144 (2018).

[5] G. Turin, "An introduction to matched filters," IRE Transactions on Information Theory **6**, 311–329 (1960).

[6] C. A. Ryan, B. R. Johnson, J. M. Gambetta, J. M. Chow, M. P. Da Silva, O. E. Dial, and T. A. Ohki, "Tomography via correlation of noisy measurement records," Physical Review A **91**, 1–7 (2015).

[7] C. C. Bultink, B. Tarasinski, N. Haandbæk, S. Poletto, N. Haider, D. J. Michalak, A. Bruno, and L. DiCarlo, "General method for extracting the quantum efficiency of dispersive qubit readout in circuit QED," Applied Physics Letters **112**, 092601 (2018).

[8] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," Proceedings of the fifth annual workshop on computational learning theory, COLT '92 (1992), 144–152.

[9] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning **20**, 273–297 (1995).

[10] R. O. Duda and P. E. Hart, *Pattern classification and scene analysis* (John Willey & Sons, New York, 1973).

[11] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, “Self-normalizing neural networks,” Proceedings of the 31st international conference on neural information processing systems, NIPS’17 (2017), 972–981.

[12] Y. Bengio, “Practical Recommendations for Gradient-Based Training of Deep Architectures,” *Neural networks: tricks of the trade: second edition* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012), 437–478.

[13] F. Rosenblatt, “The perceptron: a probabilistic model for information storage and organization in the brain,” Psychological Review, 65–386 (1958).

[14] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural computation **9**, 1735–1780 (1997).

[15] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, “Mastering the game of go with deep neural networks and tree search,” Nature **529**, 484–503 (2016).

[16] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, and D. Hassabis, “Mastering chess and shogi by self-play with a general reinforcement learning algorithm,” (2017).

[17] S. Gu, E. Holly, T. Lillicrap, and S. Levine, “Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates,” 2017 ieee international conference on robotics and automation (icra) (2017), 3389–3396.

[18] R. Bellman, “The theory of dynamic programming,” Bull. Amer. Math. Soc. **60**, 503–515 (1954).

[19] R. S. Sutton and A. G. Barto, *Reinforcement learning: an introduction* (A Bradford Book, Cambridge, MA, USA, 2018).

[20] M. Dalgaard, F. Motzoi, J. J. Sørensen, and J. Sherson, "Global optimization of quantum dynamics with alphazero deep exploration," npj Quantum Information **6**, 6 (2020).

[21] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," Nature **518**, 529–533 (2015).

[22] T. Schaul, J. Quan, I. Antonoglou, and D. Silver, "Prioritized experience replay," (2016).

[23] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. M. O. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," CoRR **abs/1509.02971**, (2015).

[24] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," (2017).

[25] J. Schulman, S. Levine, P. Moritz, M. I. Jordan, and P. Abbeel, "Trust region policy optimization," (2017).

Chapter 6

Superconducting Qubit-State Discrimination using Supervised Learning

For contemporary quantum processors, an increase from a few to many qubits is generally accompanied by the challenge of maintaining low control and readout error rates. In multi-qubit architectures, signal crosstalk induces computational errors. Beyond merely measuring the computational output, multi-qubit readout plays a key role, for instance, for near-term algorithms or error correction protocols. State-of-the-art qubit-state discriminators assume ideal operating conditions or require considerable computational effort, limiting their applicability for accurate multi-qubit readout. In this Chapter, a holistic approach using neural networks to read out multiple qubits simultaneously is pursued. This approach leads to a reduction of crosstalk-induced errors by up to one order of magnitude and enables multi-qubit readout performance on the level of single-qubit readout.

This chapter introduces superconducting qubit readout in Section 6.1 before different discriminators are discussed in Section 6.2. Finally, the performance of the discriminators is analyzed in Section 6.3 in the context of superconducting multi-qubit readout.

Qubit errors arise during all stages of the circuit model: initialization [1, 2], computation [3, 4], and readout [5]. In many implementations, qubit readout plays a key role beyond merely measuring the computational output. For example, quantum error correction protocols require repeated readout of syndrome qubits [6–8]. Even without error correction, many of the noisy intermediate-scale quantum (NISQ) [9] era algorithms involve an iterative optimization that generates a target quantum state based on prior trial-state measurements of qubits [10, 11]. In addition, diagnosing qubit-readout errors in post-processing requires computationally expensive statistical analyses of repeated computation and measurement [12–14]. Thus, developing accurate and resource-efficient qubit-state readout is a key to realize practical quantum information processing tasks.

In this chapter, deep neural networks (DNNs) as qubit-state discriminators are evaluated by their applicability relative to contemporary methods used for superconducting qubits. Nonlinear filters such as DNNs can better cope with system-dependent nonidealities, such as readout crosstalk. A quantum system comprising five frequency-tunable transmon qubits is used for simultaneous readout via a common feedline using a standard frequency multiplexing approach to evaluate these different qubit-state discriminator techniques. In contrast to single-qubit readout, such a multi-qubit system is subject to nonidealities, such as readout crosstalk, that may benefit from more sophisticated discriminators.

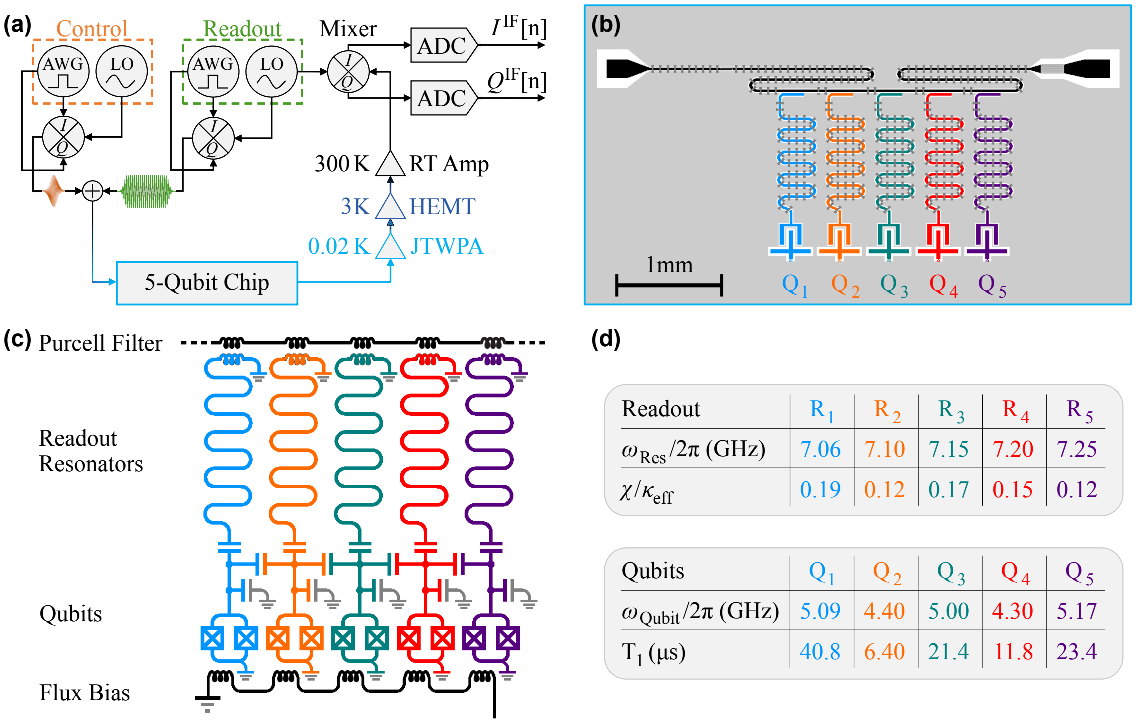
It has been shown that neural networks can learn the quantum evolution of a single superconducting qubit using merely measurement data without introducing the rules of quantum mechanics [15]. Statistical learning algorithms have been applied to superconducting qubit readout in the form of support vector machines [13], hidden Markov models [16], or a reservoir computing approach [17]. Using DNNs, improved single-qubit readout fidelity has previously been demonstrated for trapped-ions and spin qubits [18, 19]. In this chapter, the application of neural networks is extended to superconducting qubit readout and, more generally, to dispersive qubit readout. Furthermore, readout discrimination using a DNN of multiple qubits read-out on a single feedline is demonstrated.

6.1 Superconducting Qubit-State Discrimination

As described in Chapter 3.2, superconducting qubit readout is generally performed today under the paradigm of circuit quantum electrodynamics (cQED) in the dispersive regime [20]. Here, the qubit is coupled to a far-detuned resonator, such that their interaction can be treated perturbatively. The leading-order effect on the resonator is a qubit-state-dependent frequency shift $\hat{H}_{\text{disp}} = \chi \hat{a}^\dagger \hat{a} \hat{\sigma}_z$, where $\hat{a}$ is the resonator lowering operator, $\hat{\sigma}_z$ the Pauli-Z operator describing the qubit state, and χ the dispersive frequency shift. As a result, a coherent microwave signal incident on the resonator acquires a qubit-state-dependent phase shift upon transmission or reflection. The readout resonator population $\langle \hat{a}^\dagger \hat{a} \rangle$ has to remain below a critical photon number, typically tens to hundreds of photons, to remain in the dispersive readout regime. Low-noise cryogenic preamplification—a Josephson traveling-wave parametric amplifier (JTWPA) [21] at the mixing chamber ($20\,\mathrm{mK}$) and a high-electron-mobility transistor (HEMT) at $3\,\mathrm{K}$— are used to improve the signal-to-noise ratio (SNR). Subsequent heterodyne detection and digitization of the amplified signal imprints the information of the qubit state in the in-phase (*I*) and quadrature (*Q*) components of the output signal, as depicted in Figure 6-1(a).

For multi-qubit systems, there are three main qubit-state-readout approaches. First, each qubit can be measured with a separate readout resonator, feedline, and amplifier chain—a resource-intensive approach with minimal crosstalk. Alternatively, more resource-efficient readout architectures have several qubits coupled to a single readout resonator [22] or use frequency-multiplexed readout signals from multiple readout resonators [23] sharing a single feedline and amplifier chain [24]. Finally, many contemporary architectures have added Purcell filters to reduce residual off-resonant energy decay from the qubits to the resonators [25, 26].

For a qubit with static coupling to its readout resonator, energy decay and excitation during the readout are typically the primary sources of qubit measurement errors. In addition, a frequency-multiplexed readout signal contains state information on multiple qubits and is susceptible to crosstalk-induced qubit-state-readout



Readout	R_1	R_2	R_3	R_4	R_5
$\omega_{Res}/2\pi$ (GHz)	7.06	7.10	7.15	7.20	7.25
χ/κ_{eff}	0.19	0.12	0.17	0.15	0.12

Qubits	Q_1	Q_2	Q_3	Q_4	Q_5
$\omega_{Qubit}/2\pi$ (GHz)	5.09	4.40	5.00	4.30	5.17
T_1 (μs)	40.8	6.40	21.4	11.8	23.4

Figure 6-1: Measurement Setup and Chip. (a) Schematic of superconducting qubit control and readout. The control and readout pulses, generated by an arbitrary waveform generator (AWG) and up-converted to GHz frequencies using a local oscillator (LO), are sent through attenuated signal lines to the readout resonator on the five-qubit chip. The transmitted readout signal is amplified by a Josephson traveling-wave parametric amplifier (JTWPA), a high-electron-mobility transistor (HEMT), and a room-temperature amplifier. Subsequently, the signal is down-converted to MHz frequencies and digitized—in-phase $I^{\mathrm{IF}}[n]$ and quadrature $Q^{\mathrm{IF}}[n]$ sequences at intermediate frequencies (IF). Colored optical micrograph (b) and associated circuit schematic (c) comprising five superconducting transmon qubits. The qubit transition frequencies are tuned via a global flux bias. Each qubit is capacitively coupled to a quarter-wave readout resonator that couples inductively to a bandpass (Purcell) filtered feedline. (d) The resonator frequencies $\omega_{\mathrm{Res}}/2\pi$ are near 7 GHz with $\chi/\kappa_{\mathrm{eff}}$ ratios ranging from 0.12 to 0.19, where χ and κ_{eff} are respectively the dispersive shift and the effective resonator decay rate through the feedline. Table of the qubit lifetimes (T_1) and operating frequencies ($\omega_{\mathrm{Qubit}}/2\pi$). Qubit color indicate the qubit operating frequency: red (purple) → lowest (highest) operating frequency.

errors. Such crosstalk errors occur due to intrinsic interactions between the qubits themselves, qubits coupling parasitically to the readout resonators associated with other qubits, or insufficient spectral separation between readout frequencies [27].

As a result of crosstalk, state transitions due to decoherence, and other non-idealities [28], the discrimination complexity of heterodyne signals scales with the number of qubits, making state discrimination more challenging. There has been significant progress in reducing error rates and measurement times for both single- and multi-qubit devices [27, 29]. However, managing, classifying, and extracting useful information from the measured signal remains a fundamental challenge in light of the complex error mechanisms, such as crosstalk, introduced by multiplexed readout at scale.

For this thesis, the focus is on multiple frequency-tunable transmon qubits [30] arranged in a linear array with operating frequencies $\omega_{\mathrm{Qubit}}/2\pi$ between $4.3\,\mathrm{GHz}$ and $5.2\,\mathrm{GHz}$ and qubit lifetimes T_1 ranging from $7\,\mu\mathrm{s}$ to $40\,\mu\mathrm{s}$ (see Appendix B for additional details). The qubits are connected via individual co-planar waveguide resonators to the same Purcell filtered feedline, as depicted in Figure 6-1(b,c). The frequency-multiplexed readout tone comprises superimposed baseband signals at intermediate frequencies (IF) between $10\,\mathrm{MHz}$ to $150\,\mathrm{MHz}$ up-converted to the individual readout resonator frequencies ω_{Res}. After passing the feedline, the transmitted and phase-shifted tones are down-converted to IF. Up- and down-conversion is conducted with a shared local oscillator at $7.127\,\mathrm{GHz}$. Lastly, the down-converted *I*- and *Q*-components of the signal are digitized with a $2\,\mathrm{ns}$ sampling period. The resulting sequences, $I^{\mathrm{IF}}[n]$ and $Q^{\mathrm{IF}}[n]$, are subsequently digitally processed—the focus of this Chapter—to extract the individual qubit states.

6.2 Superconducting Qubit-State Discriminators

Here, supervised machine learning methods to improve superconducting qubit-state readout are employed [Chapter 5 covers the field of machine learning in more detail]. This requires a classifier capable of distinguishing the qubit-state-dependent phase shift encoded in the discrete-time $I^{\mathrm{IF}}[\mathrm{n}]$ and $Q^{\mathrm{IF}}[\mathrm{n}]$ sequences. This section will also review the current approaches to state discrimination (which are used as comparative benchmarks).

6.2.1 Boxcar Filter

Boxcar filters average the equal-weighted digitally-demodulated elements of the $I^{\mathrm{IF}}[\mathrm{n}]$ and $Q^{\mathrm{IF}}[\mathrm{n}]$ discrete-time readout signal. The digital demodulation employed here is further elaborated in Subsection 6.2.5. Each boxcar filtered digitally-demodulated sequence $\mathcal{I}[\mathrm{n}]$ and $\mathcal{Q}[\mathrm{n}]$ results in a single two-dimensional data point in the $\mathcal{IQ}$-plane [5]. Subsequently, the resulting data set can be further processed and discriminated, such as, for example, with a support vector machine, as shown in Figure 6-4.

6.2.2 Matched Filter (MF) Threshold Discriminator

Matched filter (MF) windows are generalized windowing functions with each element optimized to maximize the SNR within a given system noise model [31]. The boxcar window is the simplest example of a filter in the absence of such a noise model. For additive stationary noise independent of the qubit state and diagonal Gaussian covariance matrices, the optimal filter $k_i[\mathrm{n}]$ in terms of the SNR uses a 'window' or 'kernel,' proportional to the difference between the mean ground- and excited-state-readout signal normalized by its standard deviation, referred to as a 'matched filter' in Ref. [32], 'mode matched filter' in Ref. [27], or as 'Fisher's linear discriminant' in the context of statistics and machine learning [33] [further described in Chapter 5.2]. Applying such a matched filter reduces each readout single-shot measurement to a single one-dimensional value dependent on the qubit-state-dependent phase, allowing the qubit states to be discriminated by a simple threshold classifier. Here, a discriminator composed of a matched filter [32] and subsequently optimized threshold is referred to as MF.

While MFs are computationally efficient and provably optimal (for stationary

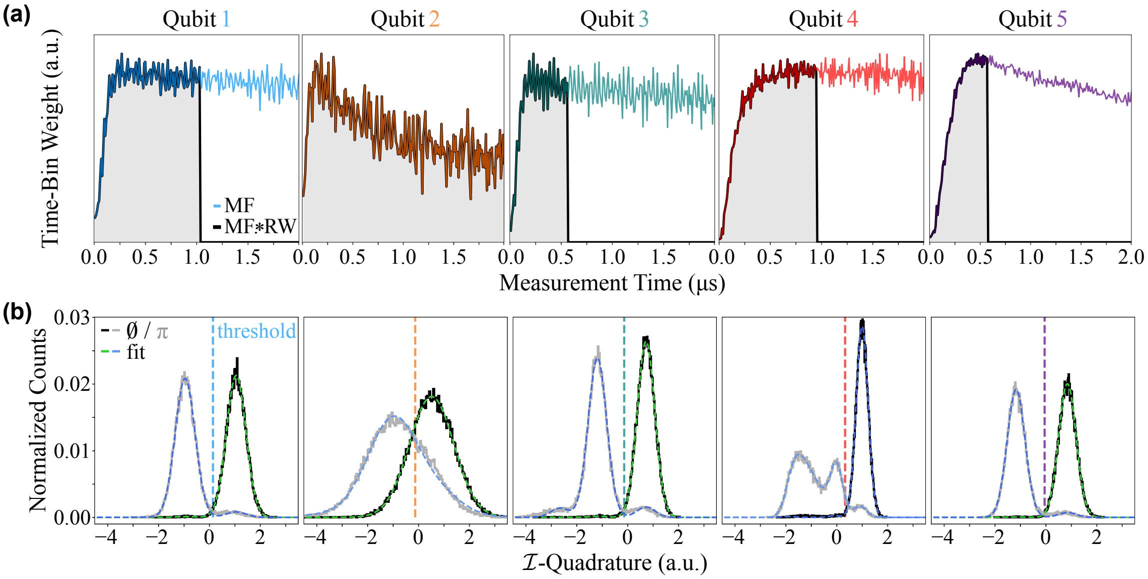


Figure 6-2: Readout Data Statistics. (a) Magnitude of the time-bin weights of the qubit-specific matched filter shapes derived using prepared ground and excited states. A rectangular window (RW) is applied to each matched filter kernel to reduce the impact of qubit-energy decays and maximize qubit-state-assignment fidelities. The resulting matched filter windows are shaded in gray. (b) Shown are the histograms of the qubit-state-readout single-shot traces after applying the optimized $1\,\mu$s-long matched filter. The dashed lines represent the optimized thresholds with the states to the right attributed to the ground state and left to the excited state. Using bimodal Gaussian fit functions for the ground state (green) and trimodal Gaussian fit functions for the excited state (blue) provides insight into the underlying dynamics such as thermal excitation or qubit-energy decays (see Table 6.1).

noise) for single qubits, as shown in Appendix C, the computational complexity to derive multi-qubit MFs scales exponentially in the number of qubits, N [34]. Consequently, in practice, multi-qubit readout is conducted per qubit with individually optimized single-qubit MFs—the approach used for many contemporary single- and multi-qubit readout schemes [14, 27, 32, 35, 36] and does not account for noise sources and nonidealities present in multi-qubit systems.

In the discussed matched filter implementation, as illustrated in Figure 6-2(a), each matched filter kernel is multiplied with a rectangular window to limit the impact of nonidealities such as qubit-energy decay. Summing up the element-wise product of the windowed matched filter kernel k_i[n] and the readout signal, $\mathcal{I}_i$[n]

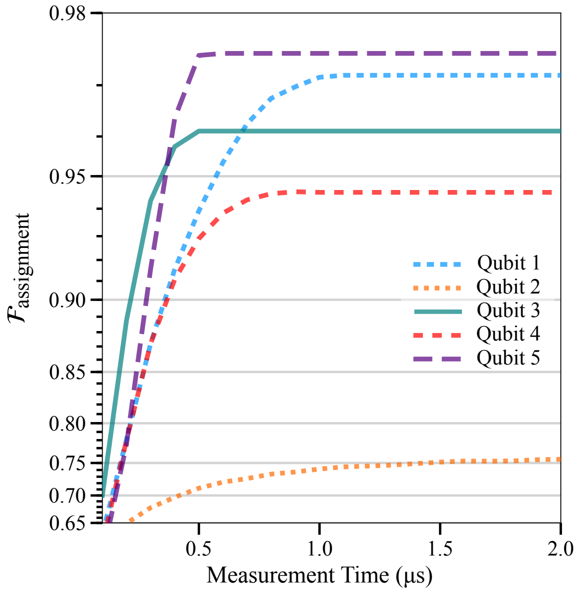


Figure 6-3: Time-dependent Qubit-State-Assignment Fidelity. Matched filter discriminator for each qubit versus measurement time. The maximum assignment fidelity $\mathcal{F}_i(t_i)$ for each qubit i is reached after $t_1 = 1\,\mu\text{s}$, $t_2 = 2\,\mu\text{s}$, $t_3 = 0.5\,\mu\text{s}$, $t_4 = 0.8\,\mu\text{s}$, and $t_5 = 0.5\,\mu\text{s}$.

and $\mathcal{Q}_i[\text{n}]$, yields a distribution along a single dimension (here, along $\mathcal{I}_i$). After matched filter summation, an optimized threshold partitions the one-dimensional projection into a ground- and excited-state class, depicted in Figure 6-2(b). The assignment fidelities for each matched filter for varying measurement times are shown in Figure 6-3. Finally, the concatenation of the one-bit labels assigned by each single-qubit discriminator yields the assigned five-qubit-state label. Note, the demodulation step at intermediate frequencies using $\mathrm{e}^{-\mathrm{j}\omega_\mathrm{i}^\mathrm{IF}\mathrm{n}}$ with ω_i^IF (as described in Chapter 3.2 and Ref. [5]) can be incorporated in the kernel tune-up.

Under the assumption of symmetric noise, the achievable assignment fidelity depends on the separation R between the ground- and excited-state-readout signals S_0 and S_1 and is referred to as the Fisher criterion [37]. The separation R is defined as

$$R = (\langle S_0 \rangle - \langle S_1 \rangle)^2 / \text{var}(S), \tag{6.1}$$

with a symmetric variance, $\text{var}(S) = \text{var}(S_0) = \text{var}(S_1)$ ($\langle f \rangle$ denotes the mean value of f). For Gaussian distributed states and diagonal covariance matrices, R can be maximized using the introduced matched filter kernel $k \propto \langle S_0 - S_1 \rangle / [\text{var}(S_0) + \text{var}(S_1)]$ [32, 33]. For a system with additive stationary noise independent of the qubit state and diagonal Gaussian covariance matrices, the maximally achievable assignment fidelity is

$$\mathcal{F}_{\mathrm{ach}} = \frac{1}{2}\left[1 + \mathrm{erf}\left(\sqrt{R/8}\right)\right], \tag{6.2}$$

with erf(z), the Gauss error function of z.

The qubit-readout-state histograms that result after the matched filters are fit with Gaussian functions are shown in Figure 6-2(b). For the fit functions, the variance for the ground and excited state are kept identical to evaluate the expected achievable discrimination fidelity $\mathcal{F}_{\mathrm{ach}}$, as presented in Table 6.1. Fitting the ground state with a bimodal and the excited state with a trimodal Gaussian fit reveals aspects of the state transition dynamics such as thermal excitations or qubit-energy decays. The product of the label and achievable fidelity provides an estimation of the upper boundary for the matched filter (MF) discriminator qubit-state-assignment fidelity $\mathcal{F}^{\overline{\mathrm{MF}}}$, as shown in the last column of Table 6.1.

In the used experimental setup [described in Appendix B.1] the highest qubit-state-assignment fidelity for MFs is achieved using time traces recorded with the other qubits (spectator qubits) initialized in their ground states, as depicted in Figure 6-6(b) for five qubits and in Figure 6-7 for a one and three qubits. This observation is a consequence of the simple noise model presumed for the MF, and thus, the MF discriminator does not capture multi-qubit readout crosstalk. In this thesis, the MF is used as a baseline to compare the following methods.

6.2.3 Support Vector Machine (SVM)

Support vector machines (SVM) are quadratic programs [38, 39] with the objective to maximize the distance between each data point and a decision boundary, a learned hyperplane separating two distinct classes. SVMs are a purely geometric approach to discrimination. For a single superconducting qubit, it has been reported that SVMs generate decision boundaries superior to that of MFs, as realistic noise deviates from the simple single-qubit noise model assumed for the MF [13].

Similar to the MF approach, multi-qubit-state discrimination can be conducted using a SVM classifier per qubit-readout signal, as exemplified using boxcar filtered

Table 6.1: Matched Filter Statistics. Numerical values extracted from Gaussian fits to readout data distribution after a $1\,\mu$s-measurement time using a matched filter, as illustrated in Figure 6-2(a,b). The peak ratio of bimodal Gaussian fits (with equal variance) to the readout-traces histograms of qubits initialized in the ground state (no pulse applied: $\emptyset$) provide insight in the thermal excitation probability $\mathcal{P}(1|\emptyset)$. Comparing the peak ratios for trimodal Gaussian fits to the readout-traces histograms of qubits initialized in the excited state (π-pulse applied: π) indicate the conditional probability for qubit-energy decays $\mathcal{P}(0|\pi)$ and second-excited state population $\mathcal{P}(2|\pi)$. $\mathcal{F}_{\text{label}} = 1 - (\mathcal{P}(1|\emptyset) + \mathcal{P}(0|\pi))/2$ denotes a lower boundary for the initialization fidelity and thus the label accuracy using the conditional state transition rates. $\mathcal{F}_\pi$ represents the fitted π-pulse fidelities resulting in the preparation fidelities $\mathcal{F}_{\text{prep}} = (1+[1-2\mathcal{P}(1|\emptyset)]\mathcal{F}_\pi)/2$. $\langle S_0 \rangle$, $\langle S_1 \rangle$, and var(S) are the mean ground state, mean excited state, and variance of both states used to derive the Fisher criterion R and achievable assignment fidelity $\mathcal{F}_{\text{ach}}$ (see Equation 6.1, 6.2). $\mathcal{F}^{\overline{\text{MF}}}$, the product of $\mathcal{F}_{\text{label}}$ and $\mathcal{F}_{\text{ach}}$, is an estimate for an upper qubit-state-assignment fidelity bound for a classifier composed of a matched filter and the subsequent optimized threshold, here referred to as MF.

Qubit	$\mathcal{P}(1_i\|\emptyset_i)$	$\mathcal{P}(2_i\|\emptyset_i)$	$\mathcal{P}(0_i\|\pi_i)$	$\mathcal{P}(2_i\|\pi_i)$	$\mathcal{F}_{\text{label}}$	$\mathcal{F}_\pi$	$\mathcal{F}_{\text{prep}}$
1	0.005	$\ll$0.001	0.038	0.001	0.979	0.999	0.995
2	0.003	$\ll$0.001	0.106	0.019	16	0.977	0.986
3	0.006	$\ll$0.001	0.057	0.052	0.968	0.965	0.977
4	0.009	0.018	0.051	0.734	0.961	0.970	0.976
5	0.003	$\ll$0.001	0.036	$\ll$0.001	0.981	0.976	0.985

Qubit	$\langle S_0 \rangle$	$\langle S_1 \rangle$	var(S)	R	$\mathcal{F}_{\text{ach}}$	$\mathcal{F}^{\overline{\text{MF}}}$
1	1.061	-17	0.388	26.817	0.995	0.974
2	0.523	-1.145	0.963	3.001	0.807	0.773
3	0.731	-1.181	0.355	28.927	0.996	0.965
4	1.003	-0.101	0.247	19.953	0.987	0.950
5	0.852	-1.164	0.348	33.614	0.998	0.979

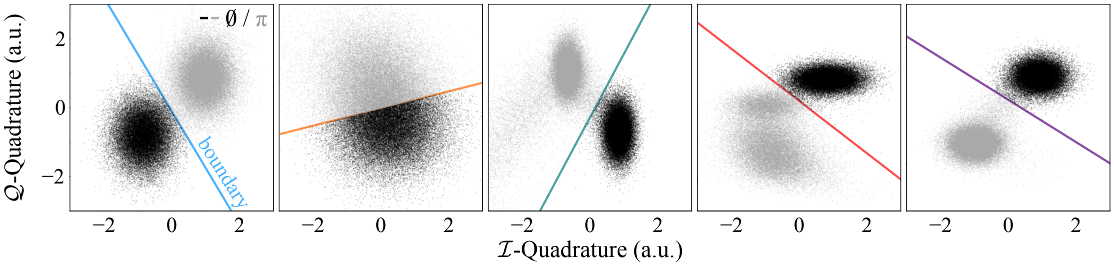


Figure 6-4: Support Vector Machine. Plotted are boxcar filtered single-shot traces of ground (black) and excited states (gray) in the $\mathcal{IQ}$-plane. A linear support vector machine trained on the two-dimensional data generates the qubit-specific colored discrimination boundary.

readout traces shown in Figure 6-4. However, in contrast to the MF tune-up, the highest assignment fidelity is achieved when the SVMs are trained using qubit-state measurement traces with the spectator qubits prepared in all combinations of ground and excited states.

Alternatively, multi-qubit states can be discriminated by a single SVM composed of several hyperplanes that partition the complete multidimensional $\mathcal{IQ}$-space, shown in Figure 6-8(c). Such a multi-qubit SVM can be tuned using a 'one-versus-all' strategy. As a result, 2^N (N, the number of qubits) two-class discrimination problems with a single qubit state as one class and the remaining qubit states as the other are solved. The analysis reveals that linear SVMs (LSVM) used as parallel single- and multi-qubit discriminators outperform their nonlinear counterparts in robustness, computational efficiency, and assignment fidelity.

The scikit-learn library is used to implement single-qubit and multi-qubit linear and nonlinear SVMs in Python [40]. The LinearSVC is employed to implement linear and SVC for nonlinear soft-margin SVMs with regularization parameters optimized per discriminator to deliver the maximally achievable qubit-state-assignment fidelity. In general, the training wall-clock-time for a SVM implemented using LinearSVC is significantly reduced relative to the training time required for SVC SVMs. However, Nonlinear SVMs can only be implemented in SVC, as LinearSVC does not offer the kernel trick (described in Chapter 5.3). In addition to the resulting unfavorable scaling of the training wall-clock-time of nonlinear SVMs, the multi-dimensional optimization problem, if tasked to discriminate multiple qubit states, mainly resulted in suboptimal hyperplanes and poor generalization. Consequently, the assignment fidelity was drastically reduced, as displayed below in the discriminator performance overview in Figure 6-7. Therefore, the study of nonlinear SVMs is limited to a basic characterization due to the lack of qubit-state-assignment fidelity robustness and the training-time requirements (for five qubits more than one day). Henceforth, the focus is on linear soft-margin SVMs as parallel single-qubit or multi-qubit discriminators (in the one-versus-all mode).

6.2.4 Neural Networks

Neural networks are mapping functions composed of arbitrarily connected nodes arranged in layers [41], as illustrated in Figure 6-5(a) and (b). Different neural network archetypes can be generated depending on the layer organization and the functions governing the connections between nodes.

In this thesis, three of the most common and successful NNs are investigated: fully-connected feedforward neural networks (FNN), which are among the most elementary neural networks, convolutional neural networks (CNN) known for their success as image classifiers, and long short-term memory recurrent neural networks (LSTM) popular in language processing. The fully-connected FNN with three hidden layers excelled in assignment fidelity compared to the other neural network types.

The fully-connected feedforward neural network (FNN) (implemented using the PyTorch library [42]) outperforms the other network architectures in qubit-state-assignment fidelity. The number of nodes composing the input layer depends on the measurement time and the size of the discrete time-bins—here $2\,\mathrm{ns}$. For a $1\,\mu\mathrm{s}$-long measurement time, the input layer contains 1,000 nodes with the in-phase and quadrature components alternating. The dimension of the first hidden layer is equal to, the second hidden layer is half of, and the third hidden layer is a quarter of the input layer dimension. Finally, the output layer consists of 2^N nodes, with

N being the number of qubits (32 for the five-qubit readout focused on here). The activation function, the nonlinear filter acting on the hidden layer nodes, is a scaled exponential linear unit (SELU) [43], instead of the common rectified linear unit (ReLU) [44] due to its improved robustness and learning rate. The output layer is filtered using a softmax function softmax$(x_i) = \exp(x_i)/\sum_j \exp(x_j)$. The network is trained (validation-training set ratio of 0.35) using the Adam optimizer [45] with categorical cross-entropy as the loss function.

Multiple training cycles, referred to as epochs, are required to ensure the discriminator output converges to the maximum qubit-state-assignment fidelity. The number of epochs to reach a convergence plateau depends on the correction factor per cycle, the so-called learning rate. It starts with a more aggressive learning rate of 0.001—a typical value for neural networks—and gradually decreases as the qubit-state-assignment fidelity starts plateauing around 250 epochs. Furthermore, the entire training set is randomly divided into normalized sub-training units, termed batches [46]. The batch size specifies after how many training samples the neural network weights are updated. The choice of batch size affects the wall-clock-training time and generalization, or in other words, how well the discriminator performs on unseen data compared to the training set. A batch size of 1,024 was found to achieve a good balance between assignment fidelity, generalization, and wall-clock-training time. An average wall-clock-training time of about half an hour for five qubits is recorded. The learning rate, generalization, and the optimal number of epochs as the number of qubits increases are shown in Figure 6-5(c).

In contrast to the MF and LSVM, the FNN can directly discriminate the frequency-multiplexed multi-qubit readout sequences I^{IF}[n] and Q^{IF}[n] without demodulation or filtering. Thus, training the network directly on the multiplexed readout signal bypasses the need for further preprocessing stages, suggesting a more efficient use of the measurement output, as illustrated in Figure 6-8(d). In addition, fewer independent operations in the readout chain may reduce the possibility of systematic errors.

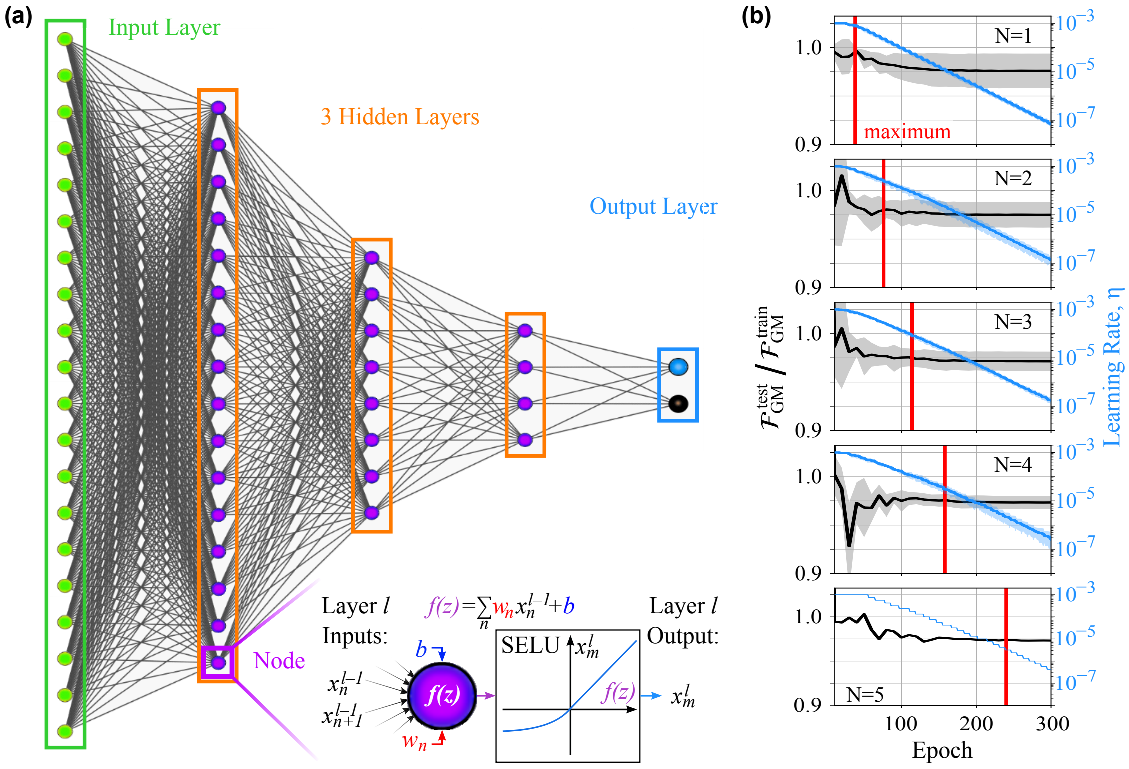

(a)
Input Layer
3 Hidden Layers
Output Layer
Node
Layer l Inputs:
$f(z)=\sum_n w_n x_n^{l-1}+b$
Layer l Output:
b
x_n^{l-1}
x_{n+1}^{l-1}
w_n
$f(z)$
SELU
x_m^l
$f(z)$
x_m^l
(b)
N=1
N=2
N=3
N=4
N=5
maximum
1.0
0.9
10^{-3}
10^{-5}
10^{-7}
$\mathcal{F}_{\mathrm{GM}}^{\mathrm{test}} / \mathcal{F}_{\mathrm{GM}}^{\mathrm{train}}$
Learning Rate, η
100
200
300
Epoch

Figure 6-5: Architecture and Training of Fully-Connected Feedforward Neural Network (FNN). (a) The FNN architecture used here comprises an input layer, three hidden layers, and an output layer. For a $1\,\mu$s-long measurement time, the input layer consists of 1,000 nodes. 1,000, 500, and 250 nodes form the first, second, and third hidden layer (the shown number of nodes per layer does not reflect the reported implementation). The output layer scales as 2^N (N denotes the number of qubits). For five qubits, the output layer encompasses 32 nodes. The inset at the bottom shows a node as a part of the hidden layer l. The node's function depends on the following parameter inputs: the output values x_n^{l-1} of the prior layer $l-1$ and a node-specific bias b. The output value x_m^l of node m corresponds to the weighted (weights w_n) sum of the inputs x_n^{l-1} and the bias b after passing through an activation function, here a scaled exponential linear unit (SELU), shown in blue. (b) Shown is the training performance for an FNN tasked to discriminate N qubits with $N = 1, 2, \ldots, 5$. The generalization—the ratio of the geometric mean test $\mathcal{F}_{\mathrm{GM}}^{\mathrm{test}}$ and training qubit-state-assignment fidelity $\mathcal{F}_{\mathrm{GM}}^{\mathrm{train}}$—as the number of epochs increases is shown in black using the left y-axis. The associated standard deviation of the generalization is indicated in gray. The number of epochs to achieve the maximum qubit-state-assignment fidelity is indicated with a red vertical bar. The learning rate η, shown in blue and using the right y-axis, is gradually reduced as the number of epochs increases.

6.2.5 Discriminator Comparison

In addition to a specific choice of discriminator, the to-be-discriminated data can be differently prepared. Typically, the discrete time readout signals at intermediate frequency, $z^{\mathrm{IF}}[\mathrm{n}] = I^{\mathrm{IF}}[\mathrm{n}] + jQ^{\mathrm{IF}}[\mathrm{n}]$, are digitally demodulated following the steps outlined in Figure 6-6(a) and Ref [5]. The signal components $\mathcal{I}_i[\mathrm{n}] = \mathrm{Re}\,(z_i[\mathrm{n}])$ and $\mathcal{Q}_i[\mathrm{n}] = \mathrm{Im}\,(z_i[\mathrm{n}])$ can be boxcar filtered [5] or kept as a sequences $\mathcal{I}_i[\mathrm{n}]$ and $\mathcal{Q}_i[\mathrm{n}]$. For digitally demodulated data and multi-qubit discrimination, $z^{\mathrm{IF}}[\mathrm{n}]$ are demodulated at each intermediate frequency. The resulting digitally demodulated time traces need to be stacked up to form a single data block before being used as the input to the multi-qubit discriminator.

Furthermore, the training data set can be composed of all the qubit states' permutations or a specific subset. Here, the focus is on either training discriminators with qubits not involved in the training process, the spectator qubits, in all combinations of the ground and excited state (indicated as $*$), or kept in the ground state (denoted by $\emptyset$).

The comparison is conducted for a measurement time of $1\,\mu\mathrm{s}$ after which four out of five qubits have reached their maximum assignment fidelity for matched filters, as shown below in Figure 6-3. For five qubits, a $1\,\mu\mathrm{s}$-long measurement time, and 10,000 training instances, a comparison of the qubit-state-assignment fidelity of the above introduced single- and multi-qubit discriminator approaches is shown in Figure 6-6(b). Optimizing the threshold of MFs and using training data with the spectator qubits in the ground state increases the qubit-state-assignment fidelity.

Single-qubit linear SVMs perform best if tasked to discriminate vectorized digitally-demodulated data and trained with a data set with all qubit-state combinations represented. Multi-qubit linear SVMs appear to perform better if tasked to discriminate digitally demodulated readout signals. On the contrary, neural networks perform the best if unprocessed data is used. The feedforward neural network outperforms its counterparts, the recurrent and convolutional neural network, in achieving qubit-state-assignment fidelity.

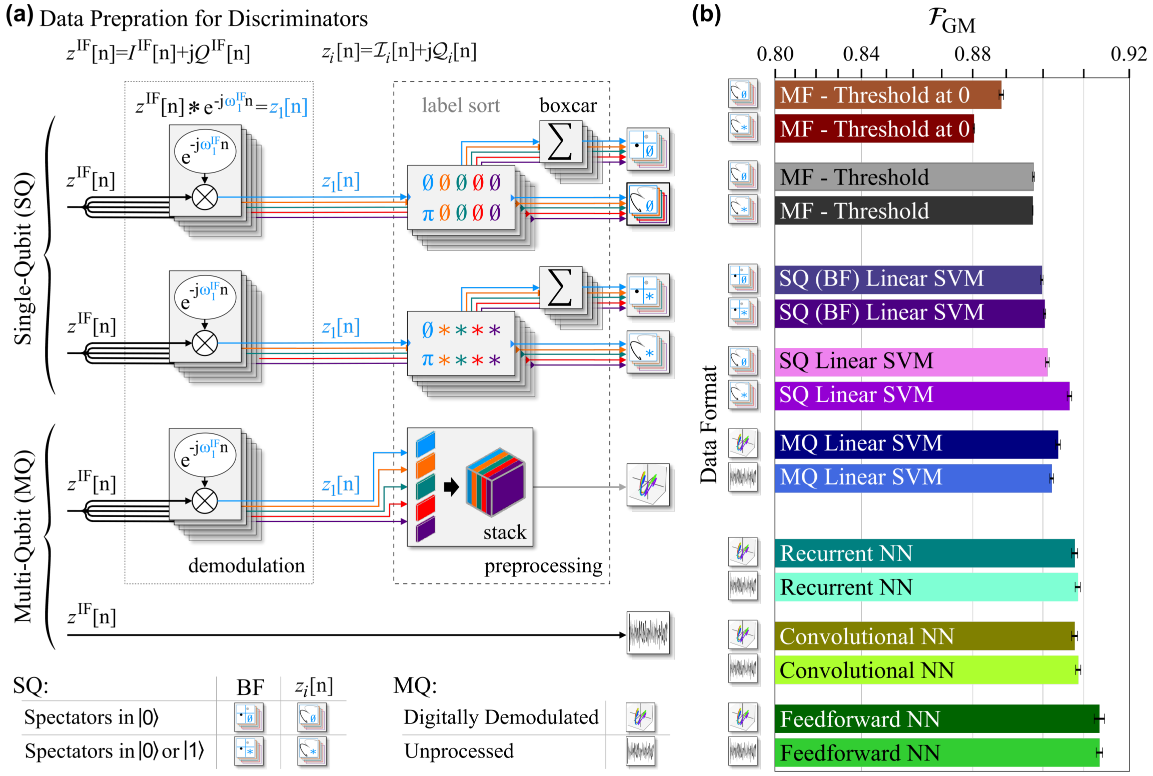


Figure 6-6: Measurement Data Processing and Discrimination. (a) M-dimensional data (z^{IF}[n]) processing for single-qubit (SQ) and multi-qubit (MQ) discrimination. For single-qubit discrimination, z^{IF}[n] is digitally demodulated at the intermediate frequency of a resonator i. Thereafter, the signal z_i[n] can either be simplified with a boxcar filter (BF) $[\frac{1}{M}\sum_n z_i[\mathrm{n}] = \bar{\mathcal{I}}_i + j\bar{\mathcal{Q}}]$ or kept as sequences $\mathcal{I}_i$[n] and $\mathcal{Q}_i$[n]. The discriminators can either be trained with the spectator qubits exclusively in their ground state (denoted by $\emptyset$) or in either their ground or excited state (denoted by $*$). For multi-qubit discriminators, the digitally demodulated signals z_i[n] at all resonator frequencies i are stacked up. The resulting data block is subsequently used for the discriminator training. Alternatively, the discriminator can be tasked to discriminate z^{IF}[n] directly without any digital preprocessing. (b) Comparison of the geometric mean qubit-state-assignment fidelity for five qubits after a $1\,\mu$s-long measurement and 10,000 training instances per qubit-state configuration. All single-qubit discriminators are evaluated using training data with the spectator qubits in the ground and all ground and excited state combinations. The matched filter (MF) threshold discriminator [the matched filter is part of the discriminator and thus not shown in (a)] is shown in two configurations; the threshold set to 0 and the threshold optimized. The linear support vector machine (SVM) is applied to boxcar-filtered (BF) and time-trace data of $\mathcal{I}_i$[n] and $\mathcal{Q}_i$[n]. The multi-qubit discriminators are evaluated utilizing digitally demodulated and unprocessed data. Shown are a multi-qubit linear SVM, a recurrent neural network (NN), a convolutional NN, and feedforward NN.

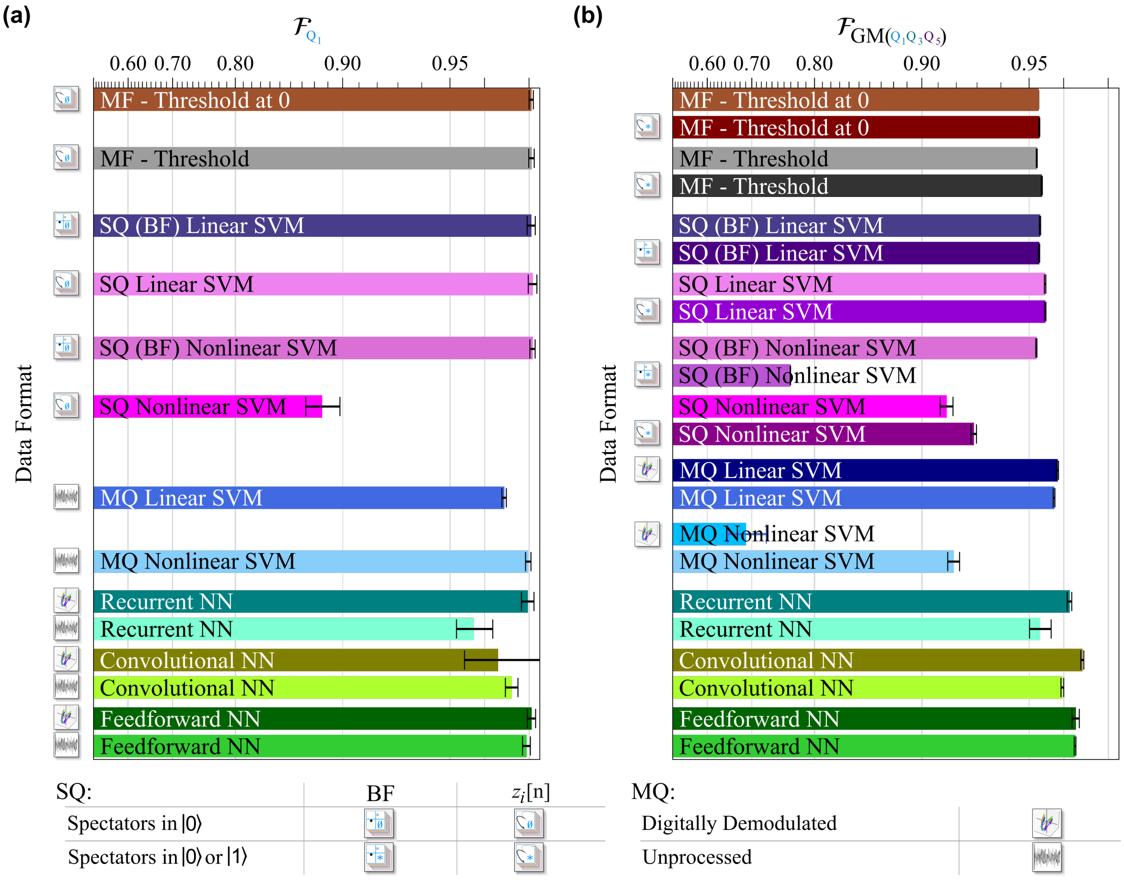


Figure 6-7: Single-Qubit and Three-Qubit Discrimination. (a) Comparison of the qubit-state-assignment fidelity for qubit 1. (b) Comparison of the geometric mean qubit-state-assignment fidelity for qubit 1, 3, and 5. The analysis in (a) and (b) is conducted for a $1\,\mu$s-long measurement and 10,000 training instances per qubit-state configuration. All single-qubit discriminators are evaluated using training data with the spectator qubits in the ground and all ground and excited state combinations. The matched filter (MF) threshold discriminator is evaluated in two configurations; the threshold set to 0 and the threshold optimized. The linear support vector machine (SVM) is applied to boxcar-filtered (BF) and time-trace data of $\mathcal{I}_i$[n] and $\mathcal{Q}_i$[n]. The multi-qubit discriminators are evaluated utilizing digitally demodulated and unprocessed data. Shown are a multi-qubit linear SVM, a recurrent neural network (NN), a convolutional NN, and feedforward NN.

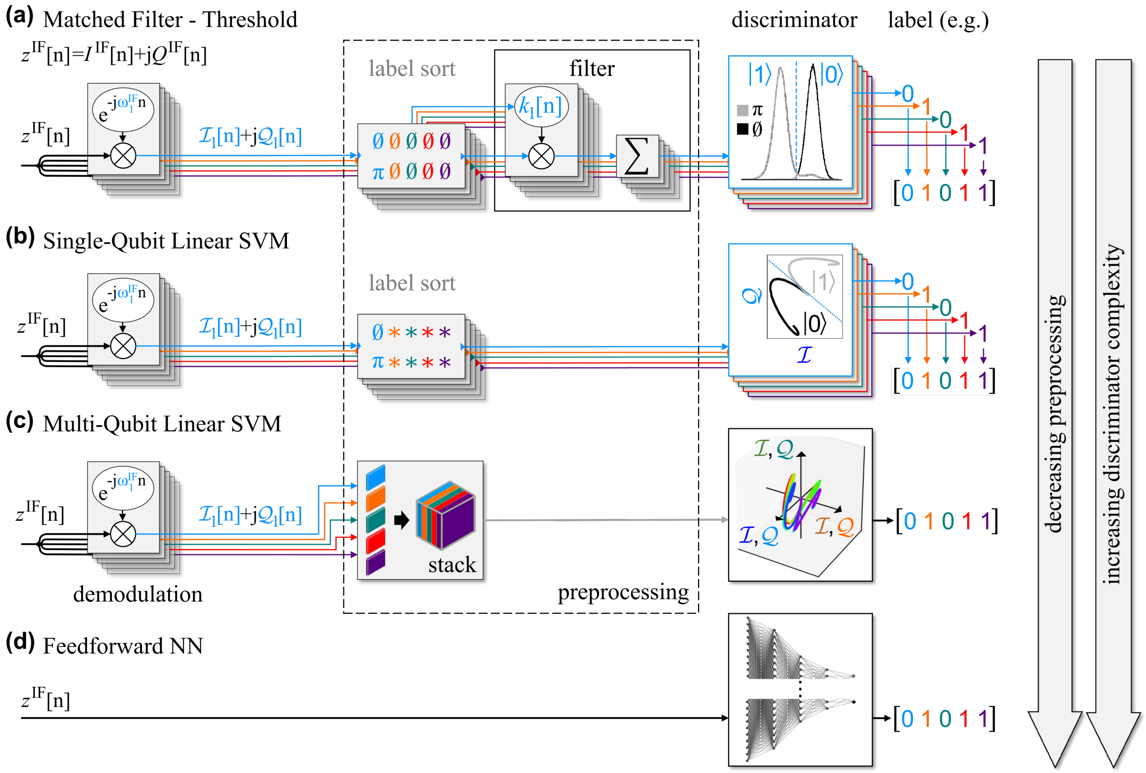

(a) Matched Filter - Threshold
$z^{IF}[n]=I^{IF}[n]+jQ^{IF}[n]$
$z^{IF}[n]$
$e^{-j\omega_1^{IF}n}$
$\mathcal{I}_1[n]+j\mathcal{Q}_1[n]$
label sort
$\emptyset\ \emptyset\ \emptyset\ \emptyset\ \emptyset$
$\pi\ \emptyset\ \emptyset\ \emptyset\ \emptyset$
filter
$k_1[n]$
$\sum$
discriminator
label (e.g.)
|1⟩
|0⟩
π
∅
0
1
0
1
1
[0 1 0 1 1]
(b) Single-Qubit Linear SVM
$\emptyset * * * *$
$\pi * * * *$
$\mathcal{Q}$
$\mathcal{I}$
(c) Multi-Qubit Linear SVM
stack
demodulation
preprocessing
$\mathcal{I},\mathcal{Q}$
(d) Feedforward NN
decreasing preprocessing
increasing discriminator complexity

Figure 6-8: Focused Measurement Data Processing and Discrimination. (a) Superconducting qubit-state discrimination can be accomplished using a single-qubit matched filter (MF) with kernel $k_i[\mathrm{n}]$ which serves as a windowing function that projects the readout signals to a single axis and subsequent discriminator threshold optimization (no pulse applied, denoted by $\emptyset$, qubit initialized in the ground state: $\emptyset \rightarrow |0\rangle$ and labeled as 0; π-pulse applied, denoted by π, qubit initialized in the excited state: $\pi \rightarrow |1\rangle$ and labeled as 1). (b) Single-qubit linear support vector machines (SQ-LSVM), (c) multi-qubit LSVMs (MQ-LSVM), and (d) fully-connected feedforward neural networks (NN) are analyzed as alternatives to MFs. The qubit-state-assignment fidelity of the MF and LSVM is maximized if the intermediate frequency signal ($z^{\mathrm{IF}}[\mathrm{n}] = I^{\mathrm{IF}}[\mathrm{n}] + \mathrm{j}Q^{\mathrm{IF}}[\mathrm{n}]$) is digitally demodulated (e.g., for resonator 1: $z^{\mathrm{IF}}[\mathrm{n}].*^{-\mathrm{j}\omega_1^{\mathrm{IF}}\mathrm{n}} = \mathcal{I}_1[\mathrm{n}] + \mathrm{j}\mathcal{Q}_1[\mathrm{n}]$ with $.*$ indicating an element-wise multiplication). The training data is relabelled to train five parallel single-qubit discriminators (MF, SQ-LSVM). The training data can either be limited to measurements during which spectator qubits are kept in their ground state (denoted by $\emptyset$) or in all combinations of the ground and excited state (symbolized by $*$. The MQ-LSVM as a single multi-qubit discriminator requires the digitally demodulated data stacked up to form a single data block. The feedforward neural networks does not require any digital demodulation or preprocessing.

6.3 Experimental Comparison of Qubit-State Discriminators

Figure 6-6 compares the performance of parallelized single-qubit MFs, parallelized single-qubit LSVMs (SQ-LSVM), multi-qubit LSVM (MQ-LSVM), and FNN discriminator using the introduced five-qubit chip. These approaches are the best performing discriminators per classifier type. The same qubit-readout sequences I^{IF}[n] and Q^{IF}[n] with varying amounts of preprocessing [Figure 6-8]—are used for all approaches. The discrimination results, a five-bit string with each bit representing the assigned state of a qubit, are compared. The qubit-state-assignment fidelity for qubit i is

$$\mathcal{F}_i = 1 - [P(0_i|\pi_i) + P(1_i|\emptyset_i)]/2, \tag{6.3}$$

where $P(0_i|\pi_i)$ is the conditional probability of assigning the ground state with label 0 to qubit i when prepared in the excited state with a π-pulse applied. $P(1_i|\emptyset_i)$ is the conditional probability of assigning the excited state with label 1 to qubit i when prepared in the ground state (no pulse applied: $\emptyset$).

The data to train and evaluate the discriminator performance was acquired using the five-qubit chip introduced in Figure 6-1(b,c). For five qubits, all 32 qubit-state permutations are sequentially initialized, and the measurement output is recorded. The generated data set contains 50,000 single-shot sequences I^{IF}[n] and Q^{IF}[n] recorded over $2\,\mu\mathrm{s}$ for each qubit-state configuration. The recorded data set is subsequently divided into a randomized training and test set (15,000 traces per qubit-state configuration for training and 35,000 for testing). All of the following results are evaluated using 35,000 single-shot measurements per qubit-state configuration.

The assignment fidelity per qubit using the geometric mean assignment fidelity is quantified as,

$$\mathcal{F}_{\mathrm{GM}} = (\mathcal{F}_1\mathcal{F}_2\mathcal{F}_3\mathcal{F}_4\mathcal{F}_5)^{1/5}, \tag{6.4}$$

with each qubit-state-assignment fidelity defined by Equation 6.3. Both SVM approaches raise the assignment fidelity relative to the MF, with the parallelized

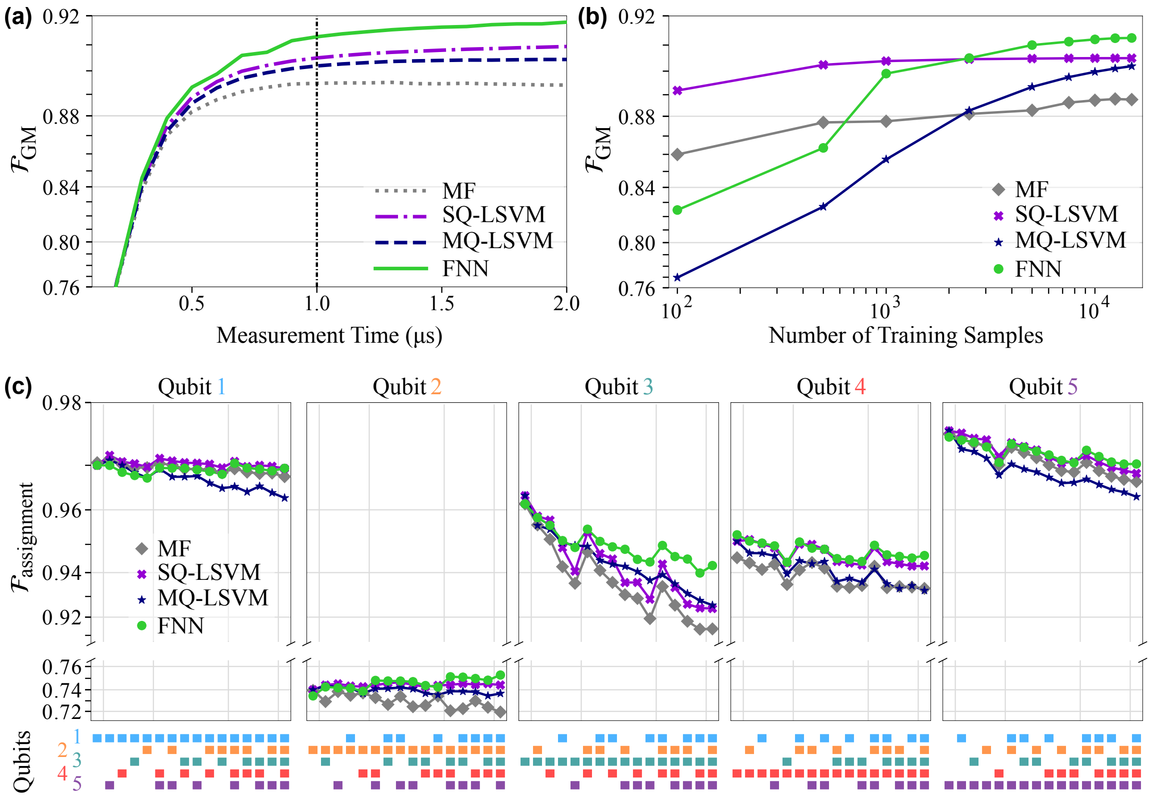


Figure 6-9: Qubit-State-Assignment Fidelity. (a) Geometric mean qubit-state-assignment fidelity $\mathcal{F}_{\mathrm{GM}}$ (Equation 6.4) for five qubits versus measurement time for the matched filter (MF), single-qubit linear support vector machine (SQ-LSVM), multi-qubit linear SVM (MQ-LSVM), and the fully-connected feedforward neural network (FNN). (b) $\mathcal{F}_{\mathrm{GM}}$ versus the number of training instances for each of the 32 qubit-state configurations evaluated after a measurement time of $1\,\mu\mathrm{s}$ [vertical dashed-dotted line in (a)]. (c) Achievable assignment fidelity $\mathcal{F}_{\mathrm{assignment}}$ per qubit when $N = \{1, 2, \ldots, 5\}$ qubits are simultaneously discriminated after a $1\,\mu\mathrm{s}$-measurement time. For each N-qubit discrimination task, the spectator qubits are initialized in their ground state. Single-qubit discrimination ($N = 1$): the first data point of each of the five panels represents the single-qubit $\mathcal{F}_{\mathrm{assignment}}$ defined by Equation 6.3. At the same time, the states of the four spectator qubits are initialized in their ground state and not discriminated. When employed as single-qubit discriminators, all methods perform similarly. Two-qubit discrimination ($N = 2$): The following four data points show $\mathcal{F}_{\mathrm{assignment}}$ when the state of each panel's qubit is simultaneously discriminated with the state of one other qubit. N-qubit discrimination ($N > 2$): the state of each panel's qubit is simultaneously discriminated with the states of $N - 1$ other qubits. For each N-qubit discrimination task, the non-spectator qubits are indicated with a colored square at the graph bottom.

single-qubit SVM outperforming the multi-qubit approach by $0.3\,\%$ after a $1\,\mu$s-measurement time. For multi-class discriminators such as the MQ-LSVM, geometric constraints result in ambiguous regions without a unique class assigned [47], which leads to poor performance relative to the other approaches. After a $1\,\mu$s-long measurement time, the FNN, compared to the MF, increases the qubit-state-assignment fidelity from 0.885 to 0.913—a reduction of the single-qubit assignment error $[1-(1-\mathcal{F}_{\mathrm{FNN}})/(1-\mathcal{F}_{\mathrm{MF}})]$ by 0.244. Compared to the SQ-LSVM, the FNN increases the qubit-state-assignment fidelity from 0.905 to 0.913 and thus reduces the single-qubit assignment error by 0.084. The FNN yields the highest qubit-state-assignment fidelity regardless of measurement time [Figure 6-9(a)].

Next, the assignment fidelity for different numbers of training samples per qubit configuration, presented in Figure 6-9(b), is evaluated. The assignment fidelity of five parallel single-qubit discriminators (MF, SQ-LSVM) saturates around 1,000 training samples per qubit-state configuration. The assignment fidelity of the FNN exceeds that of parallelized single-qubit discriminators after 2,500 training samples and saturates around 10,000 training samples per qubit-state configuration. It is estimated that the multi-qubit LSVM plateaus after approximately 40,000 training samples per qubit-state configuration. The FNN architecture here is solely optimized to maximize the qubit-state-assignment fidelity, with no consideration of the size of training data required. Thus, these results should not be taken as an indication that DNN approaches will generically perform poorly for small training sets. The remaining discriminator analysis is conducted after a $1\,\mu$s-measurement time and 10,000 training samples per qubit-state configuration.

The assignment fidelity per qubit discriminated individually and in parallel with up to $N=5$ qubits is presented in Figure 6-9(c). For N-qubit discrimination tasks with $N>2$, the FNN starts outperforming its discriminator alternatives. Except for qubit 2, the per-qubit-assignment fidelity decreases with an increasing number of discriminated qubits. A more substantial assignment fidelity decrease can be observed if the resonators involved in the discrimination are closer in frequency, suggesting the occurrence of readout crosstalk. In addition to readout crosstalk, qubit

3 reveals control crosstalk with qubit 1 and 5, the qubits closest in frequency. Under the assumption of additive stationary noise independent of the qubit state and diagonal Gaussian covariance matrices, the estimated upper qubit-state-assignment fidelity bound per qubit for MFs [13] including the label confidence listed in Table 6.2. $\mathcal{F}_2^{\overline{\mathrm{MF}}}$ is primarily reduced due to T_1-events and limited qubit-state separation in the $\mathcal{IQ}$-plane. The different discriminators yield a similar assignment fidelity within a few tenths of a percent of the upper MF assignment fidelity bound—except for qubit 2, where it is off by a few percent—when tasked to discriminate a single qubit, as shown in Table 6.2. The small discrepancy between this upper bound and the achieved assignment fidelity suggests that the noise sources affecting single-qubit readout in the presented devices are reasonably well approximated by additive stationary noise independent of the qubit state and diagonal Gaussian covariance matrices. As the number of simultaneously discriminated qubits increases, the assignment fidelity increasingly deviates from $\mathcal{F}_i^{\overline{\mathrm{MF}}}$, revealing system dynamics unaccounted for by the Gaussian noise model.

Table 6.2: Qubit-Assignment Fidelity Statistics. Qubit-assignment fidelity if discriminated individually, $\mathcal{F}_i^{1\mathrm{Q}}$, and in parallel with all other qubits, $\mathcal{F}_i^{5\mathrm{Q}}$. The last five columns present the assignment fidelity for an N-qubit discrimination process with $N = \{1, 2, \ldots, 5\}$. $\langle \mathcal{F}^{N\mathrm{Q}} \rangle$ represents the mean assignment fidelity of all qubit permutations. The single-qubit assignment fidelity is similar for all discriminator approaches. For a two-qubit discrimination task, the SQ-LSVM and FNN outperform the MF and MQ-LSVM. For N-discrimination tasks with $N > 2$, the FNN outperforms all other methods.

	Qubit 1		Qubit 2		Qubit 3		Qubit 4		Qubit 5	
	$\mathcal{F}_1^{1\mathrm{Q}}$	$\mathcal{F}_1^{5\mathrm{Q}}$	$\mathcal{F}_2^{1\mathrm{Q}}$	$\mathcal{F}_2^{5\mathrm{Q}}$	$\mathcal{F}_3^{1\mathrm{Q}}$	$\mathcal{F}_3^{5\mathrm{Q}}$	$\mathcal{F}_4^{1\mathrm{Q}}$	$\mathcal{F}_4^{5\mathrm{Q}}$	$\mathcal{F}_5^{1\mathrm{Q}}$	$\mathcal{F}_5^{5\mathrm{Q}}$
MF	0.971	0.968	0.740	0.719	0.962	0.914	0.946	0.934	0.976	0.967
SQ-LSVM	0.970	0.969	0.740	0.744	0.963	0.924	0.951	0.943	0.976	0.968
MQ-LSVM	0.970	0.963	0.740	0.737	0.963	0.926	0.951	0.933	0.976	0.963
FNN	0.970	**0.969**	0.735	**0.753**	0.962	**0.943**	0.953	**0.946**	0.975	**0.970**

	$\langle \mathcal{F}^{1\mathrm{Q}} \rangle$	$\langle \mathcal{F}^{2\mathrm{Q}} \rangle$	$\langle \mathcal{F}^{3\mathrm{Q}} \rangle$	$\langle \mathcal{F}^{4\mathrm{Q}} \rangle$	$\langle \mathcal{F}^{5\mathrm{Q}} \rangle$
MF	0.9185	0.9100	0.9042	0.8993	0.8946
SQ-LSVM	**0.9201**	**0.9148**	0.9112	0.9083	0.9053
MQ-LSVM	0.9201	0.9130	0.9078	0.9033	0.8997
FNN	0.9188	0.9141	**0.9129**	**0.9126**	**0.9122**

6.3.1 Confusion Matrix

The confusion matrix, a matrix $\mathcal{P}_{\text{assign}}$ with the qubit-state-assignment probability distribution for each prepared qubit-state configuration as rows, provides further insight into the underlying error mechanisms. The confusion matrix is the identity matrix if each prepared state is correctly labeled and assigned. In practice, in addition to misclassification, the preparation of states can be imperfect. The estimated mean state preparation fidelities for each qubit are (see Table 6.1): $\mathcal{F}_1^{\text{prep}} \approx 0.995$, $\mathcal{F}_2^{\text{prep}} \approx 0.986$, $\mathcal{F}_3^{\text{prep}} \approx 0.977$, $\mathcal{F}_4^{\text{prep}} \approx 0.976$, and $\mathcal{F}_5^{\text{prep}} \approx 0.985$.

To gain a deeper understanding of the qubit-state-assignment probabilities, the metric of confusion matrices are analyzed. Figure 6-10(a) and (b) illustrate the confusion matrix for the FNN and MF discriminator. For an ideal confusion matrix with all prepared states agreeing with the assigned state, the confusion matrix is an identity matrix. To evaluate the overlap between an identity matrix (entries represented as a Kronecker delta δ_{ij} with i and j representing the indices of the matrix row and column) and a confusion matrix (with entries c_{ij}), the following metric based on the Frobenius norm is proposed

$$||\mathrm{A}||_{\mathrm{F}} = \sqrt{\sum_i \sum_j |c_{ij} - \delta_{ij}|^2}. \tag{6.5}$$

To bound the Frobenius norm between 1 and 0, the Frobenius norm is normalized by the maximum value of Equation 6.5 ($\sqrt{2^{N+1}}$). The normalized Frobenius norm is equal to 0 if the confusion matrix is exactly an identity matrix. An alternative representation more closely related to the fidelity metric can be expressed as

$$\mathcal{F}_{\mathrm{N}} = 1 - \frac{||\mathrm{A}||_{\mathrm{F}}}{\sqrt{2^{N+1}}}. \tag{6.6}$$

The MF achieves $\mathcal{F}_{\mathrm{N}} = 0.644$, whereas the FNN yields a value of $\mathcal{F}_{\mathrm{N}} = 0.691$, a relative improvement of $7.3\,\%$.

Next, the qubit-state-dependent assignment probability of the FNN relative to the MF is expressed as the difference between their respective confusion matrices,

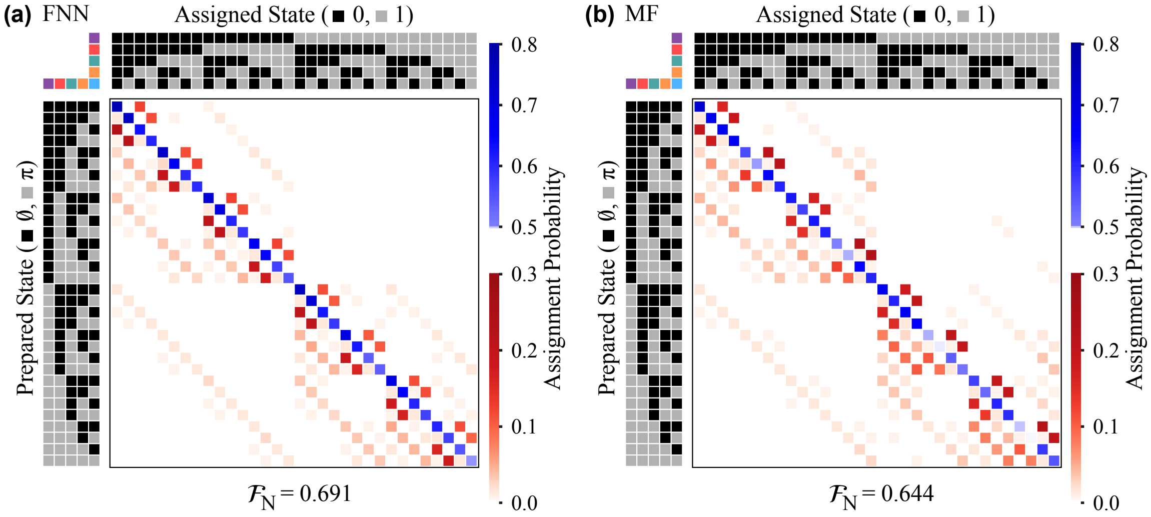


Figure 6-10: Confusion Matrix. Confusion (assignment probability) matrix of the feedforward neural network (FNN) (a) and matched filter (MF) (b). The rows of the confusion matrix encompass the probability distribution of the discriminator to assign each of the 32 qubit-state configurations to the row's prepared qubit-state configuration (no pulse applied, qubit initialized in the ground state: $\emptyset \rightarrow 0$; π-pulse applied, qubit initialized in the excited state: $\pi \rightarrow 1$). The probabilities of correctly classified states–on the diagonal–are shown in blue, whereas the misclassification probabilities–the off-diagonals–are shown in red. $\mathcal{F}_{\mathrm{N}}$, introduced in Equation 6.5, represents a metric to indicate the overlap between the confusion matrix and an identity matrix (the ideal confusion matrix). $\mathcal{F}_{\mathrm{N}} = 1$ if the confusion matrix is an identity matrix.

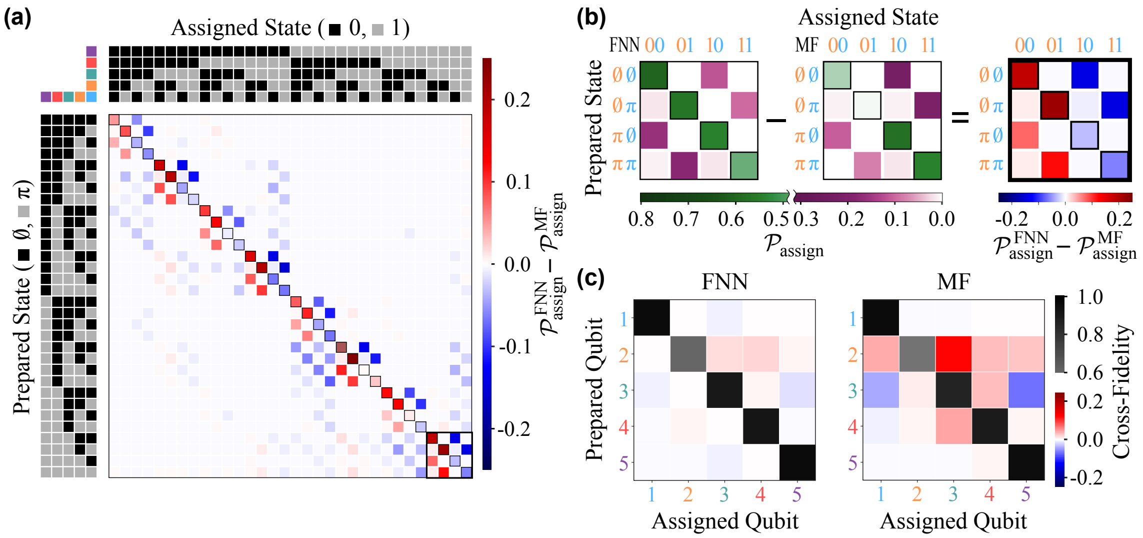


Figure 6-11: Qubit-State-Assignment Fidelity Analysis. (a) Difference between the confusion (assignment probability) matrix of the feedforward neural network (FNN) $\mathcal{P}^{\mathrm{FNN}}_{\mathrm{assign}}$ and of the matched filter (MF) $\mathcal{P}^{\mathrm{MF}}_{\mathrm{assign}}$ [shown in Figure 6-10]. The rows of the confusion matrix encompass the discriminator's probability distribution to assign each of the 32 qubit-state configurations to the row's prepared qubit-state configuration (no pulse applied, qubit initialized in the ground state: $\emptyset \rightarrow 0$; π-pulse applied, qubit initialized in the excited state: $\pi \rightarrow 1$). An increase (decrease) in the relative state-assignment probability is marked in red (blue). Red diagonal and blue off-diagonal elements indicate an improvement of the FNN over MF discrimination performance. (b) The cutouts [bold frame in the lower right corner of (a)] of the FNN, MF, and resulting relative confusion matrix display the most prominent pattern that arises due to the discrimination of qubit 2. (c) The FNN and MF cross-fidelity matrices, as defined in Equation 6.7, indicate the discrimination correlation. The off-diagonals are ideally 0. A positive (negative) matrix off-diagonal entry indicates qubit-state assignment to be correlated (anti-correlated).

$\mathcal{P}^{\mathrm{FNN}}_{\mathrm{assign}}$ and $\mathcal{P}^{\mathrm{MF}}_{\mathrm{assign}}$, shown in Figure 6-11(a). The FNN generally reduces the erroneous off-diagonal assignment probabilities relative to the MF. The most significant exception being the lower off-diagonal elements corresponding to decay of qubit 2, as presented in Figure 6-11(c).

Deviations from the ideal confusion matrix occur due to initialization errors, state transitions during the measurement, or readout crosstalk. Typically, the qubit-state misclassifications in the lower off-diagonal block outweigh those of the upper off-diagonal due to the greater likelihood of decay events at cryogenic temperatures. Here, for a $1\,\mu$s-long measurement, qubit 2—the qubit with the shortest lifetime—

Table 6.3: Readout Crosstalk. Mean absolute value, $\langle | \cdot | \rangle$, of the qubit-state-assignment correlations between readout resonators i and j ($i \neq j$) extracted from the cross-fidelity matrix $\mathcal{F}^{\mathrm{CF}}$ when using a MF or FNN discriminator.

	$\langle \lvert \mathcal{F}^{\mathrm{CF}}_{j=i\pm1} \rvert \rangle$	$\langle \lvert \mathcal{F}^{\mathrm{CF}}_{j=i\pm2} \rvert \rangle$	$\langle \lvert \mathcal{F}^{\mathrm{CF}}_{j=i\pm3} \rvert \rangle$	$\langle \lvert \mathcal{F}^{\mathrm{CF}}_{j=i\pm4} \rvert \rangle$
MF	0.020	0.015	0.006	$\sim$0
FNN	0.002	0.005	0.002	$\sim$0

has a $15\,\%$ probability of T_1-decay, such that for a significant portion of the training measurements with qubit 2 excited, the final state of qubit 2 is the ground state.

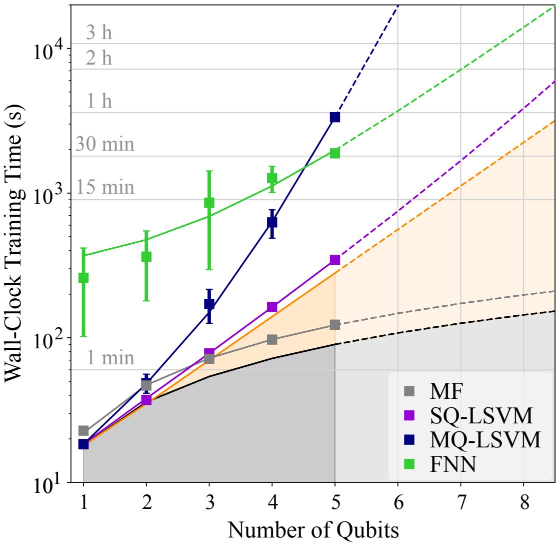
As shown in Figure 6-11(c), the FNN is more likely to assign a ground-state label to qubit 2 than an excited-state label, whereas the MF reveals the reverse trend. This suggests that the assignment probabilities of the FNN agree better with the expected error model. However, the pattern of the MF assignment probability can be attributed to a training bias. Since measurements with qubit 2 prepared in the excited state and corrupted by a T_1-decay reveal a similar integrated signal pattern as measurements with qubit 2 prepared in the ground state, the threshold optimizer overcompensates to correctly classify T_1-decay corrupted excited-state measurements at the cost of misclassification of ground-state measurements. This results in the misclassification pattern seen in Figure 6-11(c) for $\mathcal{P}^{\mathrm{MF}}_{\mathrm{assign}}$.

From the confusion matrix, the probability distribution of the non-zero Hamming distance can be extracted. The hamming distance expresses the number of misassigned qubits per qubit-state configuration. The assignment errors of the FNN (MF) are $85.8\,\%$ ($83.8\,\%$) single-qubit, $13.2\,\%$ ($15.0\,\%$) two-qubit, and $0.8\,\%$ ($1.0\,\%$) three-qubit errors. The reduction of assignment errors for the FNN compared to the MF is universal and not specific to a unique Hamming distance error, indicating a consistent reduction of crosstalk sensitivity.

The cross-fidelity matrix, which describes correlations between the assignment fidelities of individual qubits [27], is studied to study crosstalk further. The cross-fidelity $\mathcal{F}^{\mathrm{CF}}_{ij}$ is defined as

$$\mathcal{F}^{\mathrm{CF}}_{ij} = \langle 1 - \left[P\left(1_i | \emptyset_j\right) + P\left(0_i | \pi_j\right) \right] \rangle, \tag{6.7}$$

Figure 6-12: Training Wall-Clock Time. The data set acquisition for the matched filter (MF) scales linearly and exponentially for the single-qubit LSVM (SQ-LSVM), multi-qubit LSVM (MQ-LSVM), and feedforward neural network (FNN). The discriminator training remains linear for the MF while being exponential for the other three.



where $\emptyset_j$ (π_j) represent the preparation of qubit j in the ground (excited) state and 0_i (1_i) the subsequent assignment to the ground (excited) state ($\langle f \rangle$ denotes the mean value of a function f). A positive (negative) off-diagonal indicates a correlation (anti-correlation) between the two qubits. Such correlations can occur due to readout crosstalk. The off-diagonal entries for the FNN are all less than one percent, and are drastically reduced relative to the MF, as shown in Figure 6-11(c). Relative to the MF, the mean cross-fidelity, $\langle|\mathcal{F}_{ij}^{\mathrm{CF}}|\rangle$, for nearest neighbors ($j = i \pm 1$) is reduced by one order of magnitude from $\langle|\mathcal{F}_{j=i\pm1}^{\mathrm{MF\ CF}}|\rangle = 0.02$ to $\langle|\mathcal{F}_{j=i\pm1}^{\mathrm{FNN\ CF}}|\rangle = 0.002$. For neighboring readout resonators, the spectral overlap is maximized, and thus readout crosstalk most likely to occur. In general, relative to the MF, the FNN reduces the mean cross-fidelity for all $j \neq i$, as presented in Table 6.3. The FNN's reduction of assignment correlations by up to one order of magnitude corroborates the claim of the FNN's diminishing readout-crosstalk-related discrimination errors.

6.3.2 Training Effort

Frequency-multiplexed qubit readout is typical in today's most advanced NISQ devices [14, 27, 36, 48, 49]. While it is resource-efficient in operation, tuning up discriminators to compensate for system nonidealities such as readout crosstalk is computationally demanding. For the discussed five-qubit system, the train and test data set collection—a total of 50,000 $3\,\mu$s-long single-shot measurements for the 32 qubit-state permutations—required a wall-clock time of 28 minutes and generated an hdf5 file with a size of $18\,\mathrm{GB}$. The analysis reveals that 10,000 training and 3,500 validation and test instances, each $1.5\,\mu$s-long, are sufficient, which reduces the data acquisition time and file size by $1/6$.

The initial discriminator training time as the number of qubits scales in addition to the training data set acquisition is shown in Figure 6-12. While the FNN approach requires a more involved initial discrimination effort, recalibration is expected to be significantly more efficient thanks to transfer learning [46]. The FNN will require only a minimal effort for recalibration, while the other methods demand the same effort for each recalibration round.

6.4 Conclusion

An approach to multi-qubit readout using neural networks as multi-qubit state discriminators was demonstrated. The approach is more crosstalk-resilient than other contemporary approaches. A fully-connected FNN increases the readout assignment fidelity for a multi-qubit system compared to contemporary methods. The FNN compensates system-nonidealities such as readout crosstalk more effectively relative to alternatives such as matched filters (MFs) or support vector machines (SVMs). The assignment error rate is diminished by up to $25\,\%$ and crosstalk-induced discrimination errors are suppressed by up to one order of magnitude. The relative assignment fidelity improvement of the FNN over its contemporary alternatives grows as the number of simultaneously read-out and multiplexed qubits increases.

While FNNs are initially more resource-intensive in training, re-calibration can be significantly more efficient due to transfer learning [46]. Periodic re-calibration of control and readout parameters is necessary as quantum systems drift in time. For a marginal drift, neural networks can be updated at a fraction of the initial resource requirements. Furthermore, to speed up qubit readout, the techniques developed here can be transitioned to dedicated hardware such as field-programmable gate arrays (FPGA) [19].

The FNN multi-qubit-state discrimination approach was tested on a quantum system with five superconducting qubits and frequency-multiplexed readout. While the readout fidelity of Qubit 2 was relatively marginal, four qubits revealed multi-qubit readout fidelities comparable with contemporary multi-qubit systems, albeit with measurement times around $1\,\mu\mathrm{s}$, as shown in Figure 6-3, much longer than the state of the art of $100\,\mathrm{ns}$ for single-qubit systems [29]. An improvement in assignment fidelity using the FNN as a simultaneous discriminator for all qubits was demonstrated. The next step is to test the performance of FNNs on higher-fidelity multi-qubit systems with measurement times below $100\,\mathrm{ns}$ to assess if the advantage is retained on already high-performing devices.

FNNs offer a readout-state discrimination approach tailored to the underlying system. They can be readily employed to more general discrimination tasks than were considered here, such as multi-level readout in a qudit architecture [50–53]. Neural networks used for multi-qubit readout discrimination present a potential building block to scaling quantum processors while maintaining high-fidelity readout.

References

[1] D. Ristè, J. G. van Leeuwen, H.-S. Ku, K. W. Lehnert, and L. DiCarlo, "Initialization by Measurement of a Superconducting Quantum Bit Circuit," Physical Review Letters **109**, (2012).

[2] J. E. Johnson, C. Macklin, D. H. Slichter, R. Vijay, E. B. Weingarten, J. Clarke, and I. Siddiqi, "Heralded State Preparation in a Superconducting Qubit," Physical Review Letters **109**, (2012).

[3] M. D. Reed, L. DiCarlo, S. E. Nigg, L. Sun, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, "Realization of three-qubit quantum error correction with superconducting circuits," Nature **482**, 382–385 (2012).

[4] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. OMalley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, and J. M. Martinis, "Superconducting quantum circuits at the surface code threshold for fault tolerance," Nature **508**, 500–503 (2014).

[5] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "A quantum engineer's guide to superconducting qubits," Applied Physical Review **6**, (2019).

[6] D. P. DiVincenzo, "Fault tolerant architectures for superconducting qubits," Physica Scripta **T 137**, 014020 (2009).

[7] J. M. Gambetta, J. M. Chow, and M. Steffen, "Building logical qubits in a superconducting quantum computing system," npj Quantum Information **3**, (2017).

[8] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, "Surface codes: towards practical large-scale quantum computation," Physical Review A **86**, 032324 (2012).

[9] J. Preskill, "Quantum Computing in the NISQ era and beyond," Quantum **2**, (2018).

[10] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, and A. A.-G. amd Jeremy L. O'Brien, "A variational eigenvalue solver on a photonic quantum processors," Nature Communications **4213**, (2014).

[11] E. Farhi, J. Goldstone, and S. Gutmann, "A quantum approximate optimization algorithms," (2014).

[12] F. B. Maciejewski, Z. Zimborás, and M. Oszmaniec, "Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography," Quantum **4**, 257 (2020).

[13] E. Magesan, J. M. Gambetta, A. Còrcoles, and J. M. Chow, "Machine Learning for Discriminating Quantum Measurement Trajectories and Improving Readout," Physical Review Letters **114**, 200501 (2015).

[14] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, "Quantum supremacy using a programmable superconducting processor," Nature **574**, 505–510 (2019).

[15] E. Flurin, L. S. Martin, S. Hacohen-Gourgy, and I. Siddiqi, "Using a recurrent neural network to reconstruct quantum dynamics of a superconducting qubit from physical observations," Phys. Rev. X **10**, 011006 (2020).

[16] L. A. Martinez, Y. J. Rosen, and J. L. DuBois, “Improving qubit readout with hidden markov models,” Phys. Rev. A **102**, 062426 (2020).

[17] G. Angelatos, S. Khan, and H. E. Türeci, “Reservoir computing approach to quantum state measurement,” (2020).

[18] A. Seif, K. A. Landsman, N. M. Linke, C. Figgatt, C. Monroe, and M. Hafezi, “Machine learning assisted readout of trapped-ion qubits,” Journal of Physics B: Atomic, Molecular and Optical Physics **51**, 174006 (2018).

[19] Z.-H. Ding, J.-M. Cui, Y.-F. Huang, C.-F. Li, T. Tu, and G.-C. Guo, “Fast High-Fidelity Readout of a Single Trapped-Ion Qubit via Machine-Learning Methods,” Physical Review Applied **12**, 014038 (2019).

[20] 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,” Physical Review A **69**, 1–14 (2004).

[21] C. Macklin, K. O’Brien, D. Hover, M. E. Schwartz, V. Bolkhovsky, X. Zhang, W. D. Oliver, and I. Siddiqi, “A near–quantum-limited josephson traveling-wave parametric amplifier,” Science **350**, 307–310 (2015).

[22] L. DiCarlo, M. D. Reed, L. Sun, B. R. Johnson, J. M. Chow, J. M. Gambetta, L. Frunzio, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf, “Preparation and measurement of three-qubit entanglement in a superconducting circuit,” Nature **467**, (2010).

[23] M. Jerger, S. Poletto, P. Macha, U. Hübner, E. Il’ichev, and A. V. Ustinov, “Frequency division multiplexing readout and simultaneous manipulation of an array of flux qubits,” Applied Physics Letters **101**, 042604 (2012).

[24] E. Jeffrey, D. Sank, J. Y. Mutus, T. C. White, J. Kelly, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, et al., “Fast Accurate State Measurement with Superconducting Qubits,” Phys. Rev. Lett. **112**, 190504 (2014).

[25] E. A. Sete, J. M. Martinis, and A. N. Korotkov, “Quantum theory of a bandpass purcell filter for qubit readout,” Physical Review A **92**, 012325 (2015).

[26] C. Neill, P. Roushan, K. Kechedzhi, S. Boixo, S. V. Isakov, V. Smelyanskiy, A. Megrant, B. Chiaro, A. Dunsworth, K. Arya, R. Barends, B. Burkett, Y. Chen, Z. Chen, A. Fowler, B. Foxen, M. Giustina, R. Graff, E. Jeffrey, T. Huang, J. Kelly, P. Klimov, E. Lucero, J. Mutus, M. Neeley, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, H. Neven, and J. M. Martinis, "A blueprint for demonstrating quantum supremacy with superconducting qubits," Science **360**, 195–199 (2018).

[27] J. Heinsoo, C. K. Andersen, A. Remm, S. Krinner, T. Walter, Y. Salathé, S. Gasparinetti, J. C. Besse, A. Potočnik, A. Wallraff, and C. Eichler, "Rapid High-fidelity Multiplexed Readout of Superconducting Qubits," Physical Review Applied **10**, 1–14 (2018).

[28] L. C. G. Govia and F. K. Wilhelm, "Unitary-feedback-improved qubit initialization in the dispersive regime," Phys. Rev. Applied **4**, 054001 (2015).

[29] T. Walter, P. Kurpiers, S. Gasparinetti, P. Magnard, A. Potočnik, Y. Salathé, M. Pechal, M. Mondal, M. Oppliger, C. Eichler, and A. Wallraff, "Rapid High-Fidelity Single-Shot Dispersive Readout of Superconducting Qubits," Physical Review Applied **7**, 1–11 (2017).

[30] 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**, 042319 (2007).

[31] G. Turin, "An introduction to matched filters," IRE Transactions on Information Theory **6**, 311–329 (1960).

[32] C. A. Ryan, B. R. Johnson, J. M. Gambetta, J. M. Chow, M. P. Da Silva, O. E. Dial, and T. A. Ohki, "Tomography via correlation of noisy measurement records," Physical Review A **91**, 1–7 (2015).

[33] C. M. Bishop, *Pattern recognition and machine learning (information science and statistics)* (Springer-Verlag, Berlin, Heidelberg, 2006).

[34] K. Fukunaga, *Introduction to statistical pattern recognition (2nd ed.)* (Academic Press Professional, Inc., USA, 1990).

[35] N. T. Bronn, B. Abdo, K. Inoue, S. Lekuch, A. D. Còrcoles, J. B. Hertzberg, M. Takita, L. S. Bishop, J. M. Gambetta, and J. M. Chow, “Fast, high-fidelity readout of multiple qubitss,” Journal of Physics: Conference Series **834**, 012003 (2017).

[36] C. C. Bultink, B. Tarasinski, N. Haandbæk, S. Poletto, N. Haider, D. J. Michalak, A. Bruno, and L. DiCarlo, “General method for extracting the quantum efficiency of dispersive qubit readout in circuit QED,” Applied Physics Letters **112**, 092601 (2018).

[37] R. A. Fisher, “The use of multiple measurements in taxonomic problems,” Annals of Eugenics **7**, 179–188 (1936).

[38] B. E. Boser, I. M. Guyon, and V. N. Vapnik, “A training algorithm for optimal margin classifiers,” Proceedings of the fifth annual workshop on computational learning theory, COLT ’92 (1992), 144–152.

[39] C. Cortes and V. Vapnik, “Support-vector networks,” Machine Learning **20**, 273–297 (1995).

[40] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. Vanderplas, A. Joly, B. Holt, and G. Varoquaux, “Api design for machine learning software: experiences from the scikit-learn project,” (2013).

[41] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning* (The MIT Press, 2016).

[42] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, et al., “Pytorch: an imperative style, high-performance deep learning library,” *Advances in neural information processing systems 32* (Curran Associates, Inc., 2019), 8024–8035.

[43] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, "Self-normalizing neural networks," Proceedings of the 31st international conference on neural information processing systems, NIPS'17 (2017), 972–981.

[44] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," Proceedings of the 27th international conference on international conference on machine learning, ICML'10 (2010), 807–814.

[45] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," (2017).

[46] Y. Bengio, "Practical Recommendations for Gradient-Based Training of Deep Architectures," *Neural networks: tricks of the trade: second edition* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012), 437–478.

[47] R. O. Duda and P. E. Hart, *Pattern classification and scene analysis* (John Willey & Sons, New York, 1973).

[48] Z. Yan, Y.-R. Zhang, M. Gong, Y. Wu, Y. Zheng, S. Li, C. Wang, F. Liang, J. Lin, Y. Xu, C. Guo, L. Sun, C.-Z. Peng, K. Xia, H. Deng, H. Rong, J. Q. You, F. Nori, H. Fan, X. Zhu, and J.-W. Pan, "Strongly correlated quantum walks with a 12-qubit superconducting processor," Science **364**, 753–756 (2019).

[49] A. Hashim, R. K. Naik, A. Morvan, J.-L. Ville, B. Mitchell, J. M. Kreikebaum, M. Davis, E. Smith, C. Iancu, K. P. O'Brien, I. Hincks, J. J. Wallman, J. Emerson, and I. Siddiqi, "Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor," (2020).

[50] P. Kurpiers, P. Magnard, T. Walter, B. Royer, M. Pechal, J. Heinsoo, Y. Salathé, A. Akin, S. Storz, J.-C. Besse, S. Gasparinetti, A. Blais, and A. Wallraff, "Deterministic quantum state transfer and remote entanglement using microwave photons," Nature **558**, 1476–4687 (2018).

[51] S. S. Elder, C. S. Wang, P. Reinhold, C. T. Hann, K. S. Chou, B. J. Lester, S. Rosenblum, L. Frunzio, L. Jiang, and R. J. Schoelkopf, "High-fidelity measure-

ment of qubits encoded in multilevel superconducting circuits," Phys. Rev. X **10**, 011001 (2020).

[52] M. A. Yurtalan, J. Shi, G. J. K. Flatt, and A. Lupascu, "Characterization of multi-level dynamics and decoherence in a high-anharmonicity capacitively shunted flux circuit," (2020).

[53] C. Wang, M.-C. Chen, C.-Y. Lu, and J.-W. Pan, "Optimal readout of superconducting qubits exploiting high-level states," Fundamental Research **1**, 16–21 (2021).

Chapter 7

Superconducting Qubit-Readout-Pulse Shaping using Deep Reinforcement Learning

Qubit readout performance of resource-efficient quantum processors with multiple superconducting qubits is often not on par with single-qubit readout—nonidealities such as crosstalk limit the readout performance. In addition to the design layout of the qubits and resonators, qubit-state discrimination [as described in Chapter 6] and qubit-readout pulse shaping can mitigate and compensate for some of these nonidealities. One of the imminent milestones in quantum computing is the realization of error-corrected quantum processors. Fast and accurate readout is necessary to engineer error-corrected quantum processors. This chapter focuses on multi-qubit readout pulse shaping using deep reinforcement learning. Relative to conventional readout methods, the simulations reveal that deep reinforcement learning facilitates significantly shorter measurement times.

After a brief recap of superconducting qubit readout in Section 7.1, contemporary techniques and the proposed technique using deep reinforcement learning to tune up qubit-readout pulses are discussed in Section 7.2. These qubit-readout pulse shaping techniques are compared using a multi-qubit model. Finally, the results and simulation details are presented in Section 7.3.

7.1 Superconducting Qubit-Readout-Pulse Shaping

The final step of each quantum computation is qubit-readout. The accuracy of qubit readout, however, impacts the performance of quantum processors beyond the final stage. For example, quantum error correction protocols require repeated readout of syndrome qubits while the computation is still in progress [1–3]. While error-corrected quantum processors are not readily available yet, algorithms that can be implemented on contemporary noisy intermediate-scale quantum (NISQ) systems [4] require iterative readout routines to optimize a target quantum state [5, 6]. The number of computational steps is limited by the qubit coherence time. Therefore, qubit readout must be significantly faster than the quantum processor's coherence time, typically well below one microsecond for superconducting qubits. The necessary qubit-readout measurement time depends on the system design, the control setup, the pulse shape, and the readout discriminator, as discussed in Chapter 6. The focus here is on readout pulse shapes, particularly on qubit-readout-pulse shapes for multi-qubit systems using deep reinforcement learning.

As introduced in Chapter 3.2, superconducting qubit readout is generally performed in the dispersive regime [7]. In the dispersive regime, where the resonator frequency is far-detuned from the qubit, a probe signal at the resonator frequency, transmitted or reflected off the resonator, acquires a qubit-state dependent phase or amplitude shift χ. To remain in the dispersive regime, the readout resonator population has to remain below a critical photon number, typically tens to hundreds of photons. The dephasing of the qubit accelerates with rising photon numbers populating the resonator [8]. However, faster qubit-state measurements require higher photon population numbers. Therefore, a photon population number has to be carefully chosen to balance these two effects. The leakage rate of the resonator depends on the resonator linewidth κ and resonance frequency ω_r. The higher the quality factor $Q = \omega_r/\kappa$, the lower the leakage and the slower the potential measurement. However, for low-Q resonators, the qubit can suffer from dephasing through the resonator. The dephasing can be partially mitigated with bandpass filters, often referred to as Purcell filters [9]. For fast measurements, in addition to a low-Q resonator, it is critical for the leaking photons to acquire the dispersive shift sufficiently fast to enable qubit-state discrimination. It can be shown that the ratio between the dispersive shift and the resonator linewidth should be $\chi/\kappa = 1/2$ [10]. A discrimination fidelity of 99.2% after a measurement time of $88\,\mathrm{ns}$ is state-of-the-art.

Resource-efficient readout designs for multi-qubit systems are often based on frequency-multiplexed readout signals from multiple readout resonators [11] sharing a single feedline and amplifier chain [12]. Therefore, for frequency-multiplexed readout architectures, the qubit readout pulse shapes are frequency-multiplexed as well. Additionally, many contemporary architectures utilize the aforementioned Purcell filters to limit residual off-resonant energy decay from the qubits to the resonators [13, 14]. Consequently, qubit readout can suffer from crosstalk due to the shared bandpass-filtered feedline, making multi-qubit-readout pulse shaping a complicated endeavor.

For systems with multiple superconducting qubits, readout crosstalk is a combination of (1) interactions between the generated readout probe signals, (2) photon population due to a residual coupling to a probe tone or neighboring readout resonators, (3) coupling between readout resonator and neighboring qubits, and (4) interactions between reflected/transmitted readout signals in the amplifier chain, mixers, or during analog demodulation and digitization. Fast readout, such as necessary for ancilla qubits as part of a quantum error correction protocol [3, 15], requires wide resonator linewidths κ. The qubit transition frequency constrains the frequency spacing between readout resonators, the number of frequency-multiplexed probe tones, and the readout amplifier chain bandwidth. Readout crosstalk is proportional to the spectral overlap between resonators, and thus, the wider the resonator spectra, the more readout crosstalk [16]. Therefore, readout crosstalk is expected to be a particularly significant error source for fast frequency-multiplexed qubit readout.

For this thesis, the focus is on multiple frequency-tunable transmon qubits [17] arranged in a linear array with operating frequencies $\omega_{\mathrm{Qubit}}/2\pi$ between $4.3\,\mathrm{GHz}$ and $5.2\,\mathrm{GHz}$ (see Appendix B.2 for additional details). The qubits are connected via individual co-planar waveguide resonators to the same Purcell filtered feedline, as illustrated in Figure 6-1(b,c). Deep reinforcement learning (DRL) [introduced in Chapter 5.1.3] improves readout pulse shapes. Here, reinforcement learning algorithms ‘learn’ a task by interacting with the quantum system itself, and thus a model that describes the system dynamics is unnecessary [18]. In the past, DRL has been used to calibrate a two-qubit gate in simulation [19].

7.2 Superconducting Qubit-Readout-Pulse Shapes

Generally, qubit-readout pulse shaping aims to inject photons into a resonator and to deplete the photons at the end of the measurement as rapidly as possible. The section between photon injection and depletion depends on the system and the employed qubit-state discriminator. These three sections have to be optimized such that the qubit-dephasing remains limited. For single qubits, the so-called cavity level excitation and reset (CLEAR) pulse, a rectangular pulse with modified pulse flanks, can inject and deplete photons from a resonator optimally from an empirical perspective [20]. A rectangular and a CLEAR pulse, including the three pulse shaping sections, are depicted in Fig 7-1(a). The CLEAR pulse method has been shown to reduce the resonator depletion time by a factor of two compared to a passive reset. A passive reset depends on the cavity leakage rate κ while no pulse is applied. Hence the terminology ‘passive.’ Readout pulse shaping methods have not been demonstrated for multi-qubit systems. Therefore, the efficacy of CLEAR for multi-qubit systems is unclear. This chapter addresses this open question.

7.2.1 Rectangular Qubit-Readout Pulse

The most straightforward readout pulse shape is a rectangular qubit-readout pulse, as depicted in Figure 7-1(a). The rectangular pulse width defines the measurement time and does not account for the resonator reset time during which the resonator returns to the vacuum state. The second parameter, the rectangular pulse amplitude, is constant. Therefore, the optimized amplitude compromises fast injection and stabilizing the photon number such that the qubit dephasing remains limited. Rectangular readout pulses can be used for the final qubit-readout stage after all quantum computations are completed. However, for measurements during the computation, more sophisticated readout pulses are necessary to shorten the measurement time and resonator-reset time.

7.2.2 Cavity Level Excitation and Reset Qubit-Readout Pulse

Progress has been made by engineering readout pulses able to speed up the process to both populate the resonator and evacuate the photons post measurement. The cavity level excitation and reset (CLEAR) pulse can depopulate the resonator twice as fast compared to a rectangular pulse [20]. The CLEAR pulse is a rectangular pulse with modified pulse flanks, as illustrated in Figure 7-1(a). Two short segments of equal length replace both edges. The four amplitude values can be empirically optimized for a single qubit depending on the target photon population number. The pulse amplitudes corresponding to different photon target values need to be optimized separately. As the drive power increases, the resonator response becomes nonlinear, and thus the optimization more complicated.

Further improvements can be achieved using gradient ascent pulse engineering (GRAPE) to optimize the readout pulse [21]. However, these methods mainly rely on a simple rectangular pulse with modified flanks or a gradient-based optimization approach that relies on an accurate description of the system and a robust initial guess that is often difficult to identify [21].

For single-qubit readout, the CLEAR pulse suffices. However, for a system with N resonators, $4N$ segments have to be optimized at best, and 4^N segments at worst if crosstalk needs to be compensated. Therefore, the optimization routine becomes increasingly demanding as the number of resonators is increased.

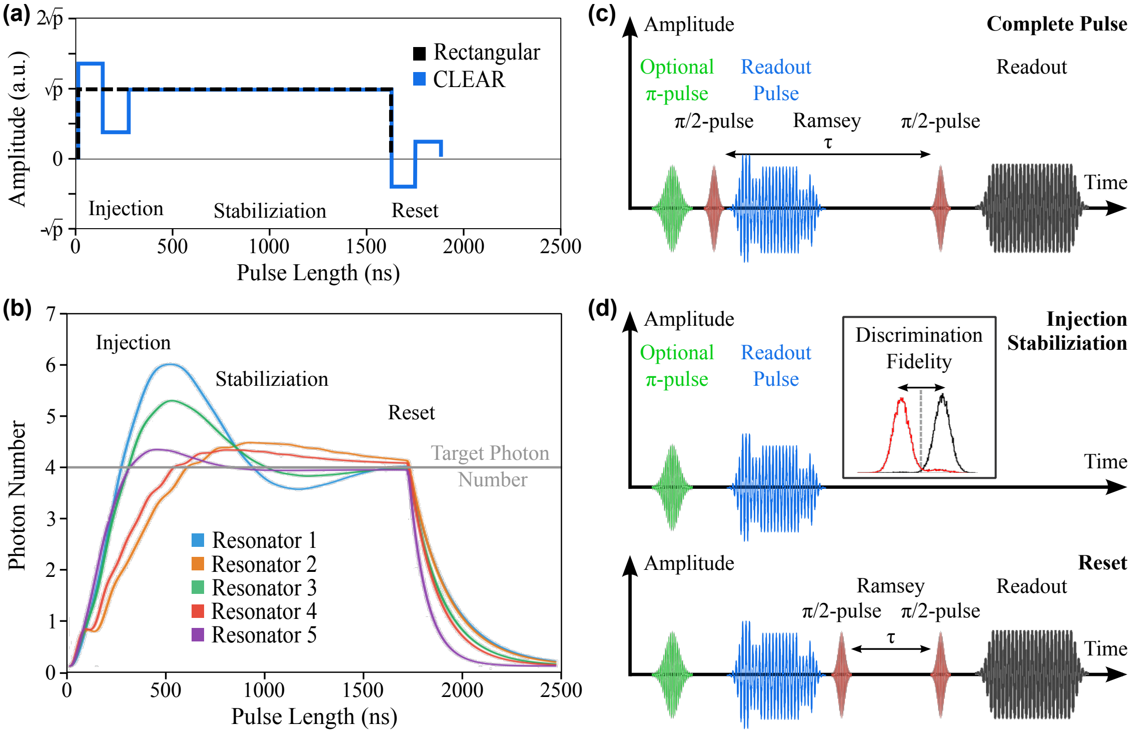


Figure 7-1: Readout Pulse Shaping Introduction. (a) A basic rectangular readout pulse (black dashed) with a normalized drive power p_{norm} contrasted by the cavity level excitation and reset (CLEAR) pulse (blue). The CLEAR pulse distinguishes itself from a rectangular pulse by the rising and falling edges. A two-amplitude segment replaces each edge to inject and deplete photons more rapidly, facilitating a shorter measurement time. (b) Shown are the resulting photon counts for the resonators in a five-qubit system for the rectangular pulse, shown in (a), frequency-multiplexed probing the quantum system. The three qubit-readout-pulse-shape sections (injection, stabilization, and reset) are indicated. The injection and reset sections are subject to an optimization routine to minimize their duration. The length of the middle part, the stabilization part, with a constant amplitude of $\sqrt{p}$ depends on how much of the signal needs to be acquired to discriminate the qubit states. As shown in (b), the photon number from a rectangular pulse oscillates at first. In (c) and (d) are two possible pathways presented to optimize readout pulse shapes. The readout pulse shape can either be optimized by enforcing to reach a specific photon number or delivering the maximally possible discrimination fidelity as quickly as possible. In (c), the readout pulse shape is optimized by targeting a specific photon number inferred by evaluating the qubit's dephasing rate with a Ramsey measurement. (d) Alternatively, the injection and stabilization part can be optimized by targeting the maximally possible discrimination rate. To optimize the reset, however, the photon number needs to be monitored.

Bayesian Optimization for CLEAR Amplitudes

Bayesian optimization is used to optimize the four amplitude segments. Bayesian optimization is an effective black-box optimization scheme attempting to generate a probabilistic model of the underlying system dynamics. The sequentially generated model enables the identification of suitable parameters more efficiently and robustly than standard optimization techniques such as example Nelder-Mead [22].

A disadvantage of Bayesian Optimization is that it scales poorly with the number of to-be-optimized segments. Specifically, each update to the posterior probability distribution makes use of all parameter-observation pairs. Therefore, more complicated environments, such as multi-qubit systems, naturally require more optimization steps to be slowed down gradually.

7.2.3 Reinforcement-Learning-Generated Qubit-Readout Pulse

For complex systems, such as multi-qubit readout, with vast numbers of different possible states (pulses), deep neural networks are often used to approximate potential actions (new pulses) and the correspondingly expected rewards (fidelity). Several reinforcement learning algorithms are investigated to efficiently tune multi-qubit readout pulses [described in Chapter]. Such deep reinforcement algorithms have offered many new avenues, such as, for example, AlphaZero, which was recently used for the optimization of two-qubit gates [19]. However, the power of this algorithm is accompanied by a significant computational effort.

The qubit-readout pulse can be mathematically modeled as an N-dimensional vector with each element representing a pulse amplitude. For qubit-readout pulse shaping, the possible number of readout pulse shapes is referred to as the action space. This action space per single dimension is bounded by the available amplitudes determined by the available hardware, particularly the arbitrary wave generator. The number of available states is nearly continuous and thus modeled as such. Finally, the reward function can depend on the number of photons in the resonator or qubit-state discrimination fidelity.

A model-free algorithm with a continuous action space can learn optimal policies in high-dimensional spaces more efficiently. The proximal policy optimization (PPO) algorithm [23] is a relatively robust DRL technique that benefits from inexpensive sampling costs of the underlying system. For instance, PPO is computationally more efficient and delivers more robust results than Deep-Q Network algorithms, an algorithm considered here [24]. A model based on PPO with its actor-critic approach comprises two separate neural networks, as shown in Figure 5-7. Two feedforward neural networks with two hidden layers comprising 64 nodes are employed as actor and critic networks in the underlying implementation. The actor generates new qubit-readout-pulse shapes where the critic network tries to associate them with an expected reward. These two networks, in concert, enable a relatively computationally efficient approach to multi-qubit readout pulse shaping.

7.3 Qubit-Readout-Pulse Shaping Simulations

Qubit-readout-pulse shaping optimization can be divided into two separate parts: photon injection and photon depletion. Photon injection can be optimized using either a target photon number or the qubit-state discrimination fidelity as the optimization metric. Photon depletion, on the other hand, is optimized using the photon number. The number of photons populating a resonator can be inferred through the dephasing rate of a qubit. The different methods are presented in Figure 7-1.

In simulations using the formalism presented in Appendix A, the discrimination fidelity is not directly accessible. Therefore, a target photon number is used for injection and reset. The photon number is directly proportional to the readout probe tone driving power in the linear dispersive regime.

For simplicity, the stabilization section of the readout pulse is kept constant for all pulse shapes and is not part of the optimization. Moreover, inspired by the idea of CLEAR [20], only half the injection section is being optimized, while the second part is kept constant, during which the photon number stabilizes.

The optimizer can pick amplitudes between $-\sqrt{p}$ and $\sqrt{p}$ with a resolution of 2^{10} for the readout-resonator reset pulse segment. The resolution depends on the employed hardware. The amplitude is mapped to the closest available discrete amplitude for methods outputting continuous amplitudes, such as CLEAR or PPO. To ensure realistic pulse shapes in simulation, the pulse shapes are filtered with a Gaussian filter with a standard deviation of $\sigma = 5$ encompassing system-dependent factors such as a nonideal step response [21, 25].

For CLEAR pulses tuned-up using Bayesian optimization, 250 random initial points and 2,500 optimization points are used. As each optimization episode is significantly faster for the PPO-generated pulses, the upper limit is set to 51,200 training episodes.

Readout-Resonator Photon Injection

The necessary qubit-readout measurement time depends on the photon injection efficiency. In general, the faster photons are injected, the faster a qubit's state can be determined. As shown in Figure 7-1(b), for a rectangular readout pulse, the photon number dynamics reveal two oscillations; one due to the pulse itself and the other due to crosstalk.

Generally, readout-pulse shapes should not depend on the qubit states: the readout pulse shape is the same whether the qubit is in the ground or excited state.

For the optimization of the injection segment of a readout-resonator pulse, the range of potential amplitudes is changed to 0 and $2\sqrt{P}$. To ensure the qubit-resonator interaction remains dispersive, amplitudes resulting in photon numbers exceeding the critical photon number [as discussed in Chapter 3.2 and calculated in Appendix B.2] are penalized by a tunable penalty term $\Phi > 100$.

To identify optimal photon-injection pulses in simulation, the reward function is set as the sum of the absolute difference between a resonator's photon count n_r and the target photon number $\hat{n}_r$, which depends on the normalized drive power. For a pulse of length L, the algorithm evaluates the stability of the photon number for the last L' nanoseconds. A photon number is deemed stable as long as it remains within ± 0.1 photons of the target photon number for L' nanoseconds. Violations of the introduced conditions reduced the reward for all resonators $r \in R = \{R_1, R_2, R_3, R_4, R_5\}$ as

$$\text{Reward} = \begin{cases} -\Phi, & \text{if } \forall r \in R, \\ & \quad \forall l \in [L - L', L], \\ & \quad \text{and } |\hat{n}_r - n_r^l| \leq 0.1, \\ -\Phi - \sum\limits_{r \in R} \left(\sum\limits_{l=L-L'}^{L} |\hat{n}_r - n_r^l| \right), & \text{otherwise.} \end{cases} \tag{7.1}$$

The stable photon count length L' is set to $100\,\mathrm{ns}$.

Readout-Resonator Photon Depletion

A passive readout-resonator reset can be time-consuming and unnecessarily dephases the qubit. While this may not matter for the final readout, it matters for periodic readout while quantum computations are in progress, such as for quantum error correction protocols [3].

For passive decay, the time-dependent photon number in a resonator r can be expressed as $n_r = e^{-\kappa t} p_{\text{norm}}$ [20], with the resonator linewidth κ and p_{norm} denoting the ratio of the applied qubit-readout-pulse power and the power required to inject a single photon in the readout resonator $p = p_{\text{norm}} P_{1\text{ photon}}$. Active readout-resonator reset aims to reduce the decay time relative to the passive exponential decay. A resonator is considered sufficiently close to its vacuum state and only thermally populated once the photon number dropped below a threshold of 0.1 photon. In addition to reducing the readout-resonator reset time, the readout-resonator-reset pulse needs to be independent of the qubit state.

To optimize the readout-resonator reset, the reward function for each readout resonator $r \in R = \{R_1, R_2, R_3, R_4, R_5\}$ is defined as the negative sum of the photons populating the readout resonator and defined as

$$\text{Reward} = \begin{cases} 0, & \text{if } \forall r \in R, \ \ n_r \leq 0.1, \\ -\sum_{r\in R} n_r, & \text{otherwise.} \end{cases} \tag{7.2}$$

If the readout-resonator reset is tuned-up separately from the photon injection, the readout resonator is populated using a $3\,\mu$s-long rectangular readout pulse with a drive power of $\sqrt{p}$ to inject $p_{\text{norm}} = 4$. Note, this injection phase is longer than needed to ensure all five resonators have reached equilibrium.

7.3.1 Single-Qubit Simulation

First, the different algorithms and pulse shapes are tested on a single-qubit system. The qubit and resonator parameters described in Appendix B.2 for Qubit 1 are used.

In Figure 7-2, the pulse shapes of the CLEAR and PPO pulse are shown for the photon injection and reset part. Furthermore, the resulting photon numbers are compared with a rectangular readout pulse. The rectangular readout pulse requires $1470\,\text{ns}$ to reach a stable photon number count and about $700\,\text{ns}$ to reset. The CLEAR and PPO pulse are nearly indistinguishable. The injection phase is reduced by $62.6\,\%$ relative to the rectangular readout pulse to $550\,\text{ns}$. The readout resonator reset is $260\,\text{ns}$ for the CLEAR pulse and $250\,\text{ns}$ for the PPO-generated pulse, a reduction of $62.9\,\%$ and $64.3\,\%$ respectively. The difference of $10\,\text{ns}$ between the CLEAR and PPO pulse is due to the defined shape with both amplitudes lasting for the same amount of time.

For single-qubit readout, the CLEAR pulse and PPO-generated pulse are effectively the same. The pulses start with a high pulse amplitude to overshoot and compensate during the injection segment. Similarly, the opposite shape behavior is observed for the reset segment. The injection and reset amplitude level tend to use the extreme amplitude levels at $\pm\sqrt{p}$. These observations agree with prior work confirming that overshooting the particular target photon number, followed by a correction, is typically faster than a rectangular pulse [12, 20, 26].

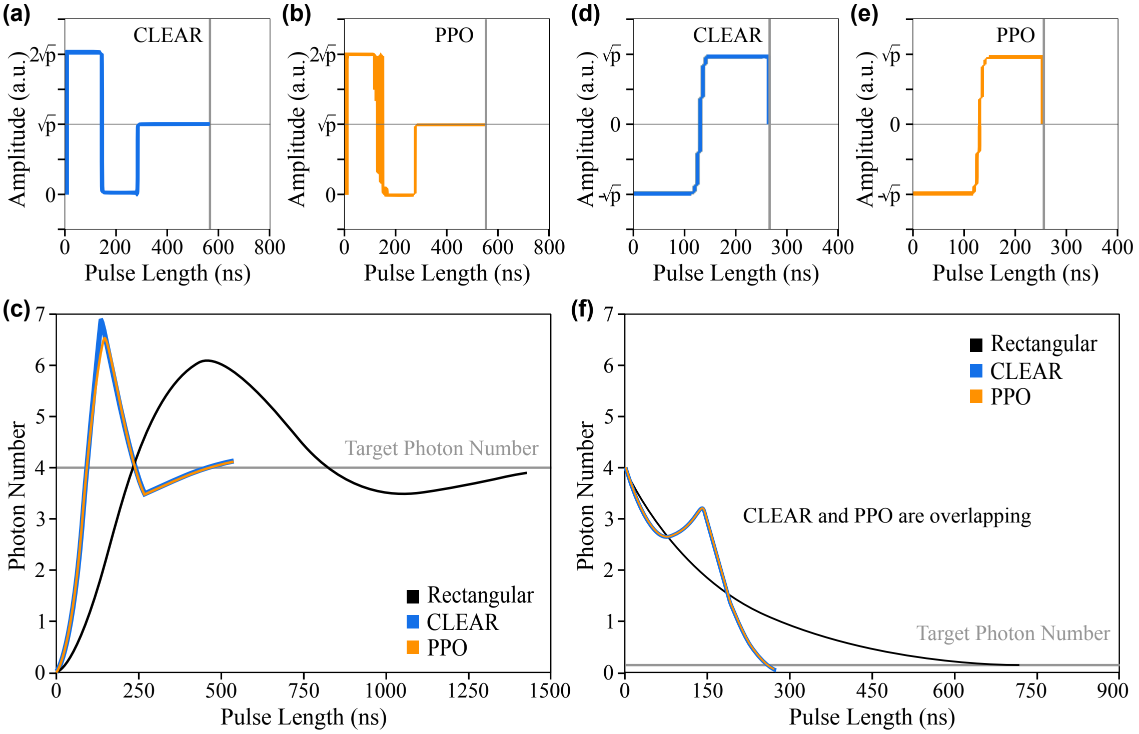


Figure 7-2: Single-Qubit Readout Pulse Shapes. (a), (b), and (c) show the pulse shapes and photon number for the photon injection part. (d), (e), and (f) present the pulse shapes and photon numbers for the photon depletion section. (a) and (d) show the CLEAR pulse. (b) and (e) the PPO pulse. (c) and (f) contains the expected photon numbers of the rectangular, CLEAR, and PPO readout-resonator pulse. For a single-qubit readout, both CLEAR and PPO outperform a rectangular readout pulse. The pulse generated by CLEAR and PPO is virtually indistinguishable.

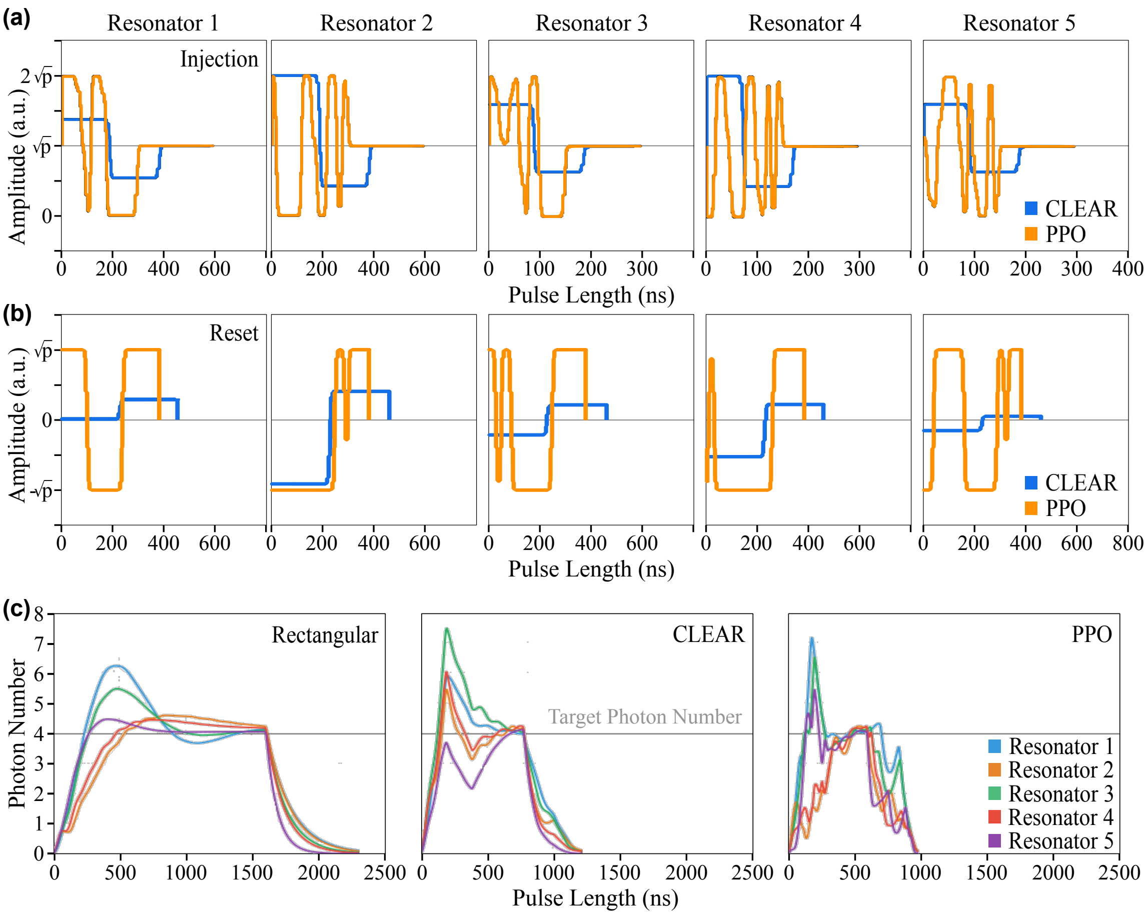


Figure 7-3: Multi-Qubit Readout Pulse Shapes. (a) Injection segment of resonator-readout-pulse shapes and (b) Reset segment of resonator-readout-pulse shapes for five qubits. (c) Photon numbers versus pulse length for a rectangular, CLEAR, and PPO-generate readout pulse.

7.3.2 Multi-Qubit Simulation

For the five-qubit simulations using frequency-multiplexed readout, the system parameters discussed in Appendix B.2 are used.

Figure 7-3 shows the pulse shapes and corresponding photon numbers for a rectangular, CLEAR, and PPO-generated pulse. A rectangular qubit-readout pulse needs $1610\,\mathrm{ns}$ to stabilize all five resonators. The CLEAR pulse in comparison, achieves the same in $780\,\mathrm{ns}$, a $51.6\,\%$ reduction. For a pulse shape generated using PPO, an additional speed-up can be generated and the photons injected in $600\,\mathrm{ns}$, reducing the injection phase by $62.7\,\%$ relative to a rectangular readout pulse. All five resonators

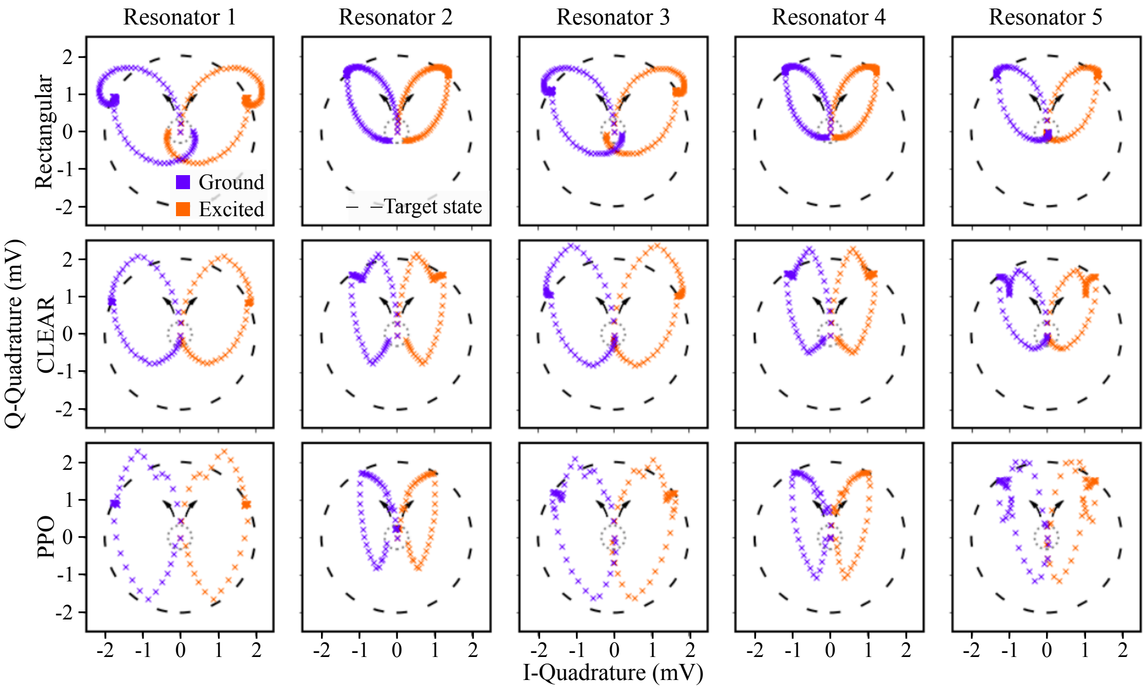


Figure 7-4: IQ-plane resonator trajectories for each resonator and each pulse shape: rectangular, CLEAR, PPO-generated. Ground-state trajectories are shown in purple and orange for the excited state. The dashed grey circle marks the target photon number of four photons. The inner dashed circle indicates 0.1 photons. The 'X'-markers are spaced $20\,\mathrm{ns}$ apart. Each trajectory starts and returns to its origin. Depending on the photon population dynamics, the trajectory changes its shape.

are reset after $700\,\mathrm{ns}$. To reset the resonator, CLEAR requires $460\,\mathrm{ns}$, a reduction of $34.3\,\%$. Similar to injection, PPO can further reduce the reset time to $380\,\mathrm{ns}$, reducing the reset time by $45.7\,\%$.

While the pulse shapes still follow the single-qubit CLEAR trend—overshooting and correcting—the pulse amplitudes no longer span the maximally allowed range. There are two explanations. First, for all resonators, the reset pulse length is the same. The resonator that requires the longest to be depleted defines that reset pulse length. Therefore, resonators that could be reset more rapidly can choose less aggressive amplitudes. Second, crosstalk from off-resonant pulses or parasitic coupling between resonators or non-paired qubits could affect the pulse shapes. The PPO-generated pulses are visibly more complicated. The pulses often fluctuate between the maximal amplitudes causing photon number fluctuations.

Table 7.1: Summary of Simulation Results. Listed is the injection and reset pulse length for the case of single-qubit and five-qubit readout. The investigated readout pulse shapes are rectangular, CLEAR, and PPO-generated.

	Single-Qubit		Multi-Qubit	
	Injection	Reset	Injection	Reset
Rectangular	1470 ns	700 ns	1610 ns	700 ns
CLEAR	550 ns	260 ns	780 ns	460 ns
vs. **Rectangular**	**−62.6 %**	**−62.9 %**	**−51.6 %**	**−34.3 %**
PPO	550 ns	250 ns	600 ns	380 ns
vs. **Rectangular**	**−62.6 %**	**−64.3 %**	**−62.7 %**	**−45.7 %**
vs. CLEAR		−62.9 %	−23.1 %	−17.4 %

As illustrated in Figure 7-3(c), the total readout pulse shapes are 2310 ns for a rectangular, 1440 ns for the CLEAR, and 980 ns for a PPO-generated pulse. The individual pulse lengths are summarized in Table 7.1.

The trajectories in the IQ-plane, shown in Figure 7-4, can provide further insight into the underlying system dynamics. The IQ-trajectory depends on the qubit-state. The path towards the target photon number is more efficient for CLEAR and PPO-generated pulses. The trajectories of resonators 2 and 4 shine light on the reasons why PPO-generated pulses outperform CLEAR pulses.

7.4 Conclusion

For a qubit dispersively coupled to a resonator, rapid resonator ring-up and ring-down ensure fast readout and limited qubit dephasing. In an efficient, frequency-multiplexed readout of multiple qubits, effects such as drive crosstalk increase the complexity of optimal readout pulse shapes. Computationally intensive methods are typically required to discover high-fidelity readout pulse shapes.

In this thesis, a pulse shaping optimization module using deep reinforcement learning (DRL) was developed. The readout measurement duration of contempo-

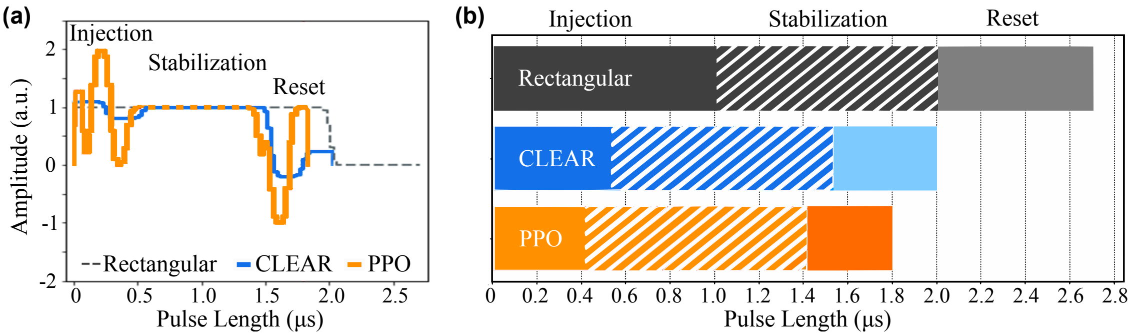


Figure 7-5: Simulation of Readout Pulse Shape versus Measurement Time for a Five-Qubit Chip. (a) Comparison of readout pulse shapes for a single qubit. The typically employed pulse shape, a rectangular pulse shape (gray, dashed), is compared with the cavity level excitation and reset (CLEAR) (in blue) and a pulse shape tune-up using the proximal policy optimization (PPO) algorithm (in orange), a deep reinforcement learning algorithm. (b) Comparison of the duration to simultaneously inject photons in 5 different readout resonators, stabilize the photon counts during a waiting period, and return the resonators close to their vacuum states. The waiting period is kept constant for all three methods. In comparison with the rectangular pulse (CLEAR), PPO reduces the injection window by 58% (22%) and reset duration by 46% (17%).

rary and the PPO pulse shape optimization module was compared on a simulated five-qubit chip. The resulting pulse shapes and measurement duration are presented in Figure 7-5. The PPO pulse optimization module yielded a reduction of the injection time by 22% and a 17% shorter reset time compared to the currently most advanced readout pulse shapes, the cavity level excitation and reset (CLEAR) pulse [20]. Interestingly, for single-qubit readout, the suggested readout pulse shape was nearly identical with the CLEAR pulse, suggesting that CLEAR pulses are optimal for single-qubit readout. However, for multi-qubit readout, better-performing readout pulse shapes exist, as demonstrated in this thesis.

The experimental verification of the simulated reduction in measurement time is currently in progress. The experiment is conducted with a system comprised of five superconducting qubits, as described in Appendix B.2. In the experiment, PPO-generated readout pulse shapes are compared with rectangular and CLEAR pulses. This proof of concept demonstration could significantly improve readout performance and facilitate practical quantum error correction schemes.

References

[1] D. P. DiVincenzo, “Fault tolerant architectures for superconducting qubits,” Physica Scripta **T 137**, 014020 (2009).

[2] J. M. Gambetta, J. M. Chow, and M. Steffen, “Building logical qubits in a superconducting quantum computing system,” npj Quantum Information **3**, (2017).

[3] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, “Surface codes: towards practical large-scale quantum computation,” Physical Review A **86**, 032324 (2012).

[4] J. Preskill, “Quantum Computing in the NISQ era and beyond,” Quantum **2**, (2018).

[5] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, and A. A.-G. amd Jeremy L. O’Brien, “A variational eigenvalue solver on a photonic quantum processors,” Nature Communications **4213**, (2014).

[6] E. Farhi, J. Goldstone, and S. Gutmann, “A quantum approximate optimization algorithms,” (2014).

[7] 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,” Physical Review A **69**, 1–14 (2004).

[8] J. Gambetta, A. Blais, D. I. Schuster, A. Wallraff, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, “Qubit-photon interactions in a cavity: measurement-induced dephasing and number splitting,” Physical Review A **74**, 042318 (2006).

[9] N. T. Bronn, B. Abdo, K. Inoue, S. Lekuch, A. D. Còrcoles, J. B. Hertzberg, M. Takita, L. S. Bishop, J. M. Gambetta, and J. M. Chow, “Fast, high-fidelity readout of multiple qubitss,” Journal of Physics: Conference Series **834**, 012003 (2017).

[10] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, “A quantum engineer’s guide to superconducting qubits,” Applied Physical Review **6**, (2019).

[11] M. Jerger, S. Poletto, P. Macha, U. Hübner, E. Il’ichev, and A. V. Ustinov, “Frequency division multiplexing readout and simultaneous manipulation of an array of flux qubits,” Applied Physics Letters **101**, 042604 (2012).

[12] E. Jeffrey, D. Sank, J. Y. Mutus, T. C. White, J. Kelly, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, et al., “Fast Accurate State Measurement with Superconducting Qubits,” Phys. Rev. Lett. **112**, 190504 (2014).

[13] E. A. Sete, J. M. Martinis, and A. N. Korotkov, “Quantum theory of a bandpass purcell filter for qubit readout,” Physical Review A **92**, 012325 (2015).

[14] C. Neill, P. Roushan, K. Kechedzhi, S. Boixo, S. V. Isakov, V. Smelyanskiy, A. Megrant, B. Chiaro, A. Dunsworth, K. Arya, R. Barends, B. Burkett, Y. Chen, Z. Chen, A. Fowler, B. Foxen, M. Giustina, R. Graff, E. Jeffrey, T. Huang, J. Kelly, P. Klimov, E. Lucero, J. Mutus, M. Neeley, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, H. Neven, and J. M. Martinis, “A blueprint for demonstrating quantum supremacy with superconducting qubits,” Science **360**, 195–199 (2018).

[15] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary edition* (Cambridge University Press, USA, 2011).

[16] J. Heinsoo, C. K. Andersen, A. Remm, S. Krinner, T. Walter, Y. Salathé, S. Gasparinetti, J. C. Besse, A. Potočnik, A. Wallraff, and C. Eichler, “Rapid High-fidelity Multiplexed Readout of Superconducting Qubits,” Physical Review Applied **10**, 1–14 (2018).

[17] 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**, 042319 (2007).

[18] R. S. Sutton and A. G. Barto, *Reinforcement learning: an introduction* (A Bradford Book, Cambridge, MA, USA, 2018).

[19] M. Dalgaard, F. Motzoi, J. J. Sørensen, and J. Sherson, “Global optimization of quantum dynamics with alphazero deep exploration,” npj Quantum Information **6**, 6 (2020).

[20] D. T. McClure, H. Paik, L. S. Bishop, M. Steffen, J. M. Chow, and J. M. Gambetta, “Rapid Driven Reset of a Qubit Readout Resonator,” Physical Review Applied **5**, 1–5 (2016).

[21] S. Boutin, C. K. Andersen, J. Venkatraman, A. J. Ferris, and A. Blais, “Resonator reset in circuit QED by optimal control for large open quantum systems,” Physical Review A **96**, 1–11 (2017).

[22] J. B. Mockus and L. J. Mockus, “Bayesian approach to global optimization and application to multiobjective and constrained problems,” J. Optim. Theory Appl. **70**, 157–172 (1991).

[23] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” (2017).

[24] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, “Human-level control through deep reinforcement learning,” Nature **518**, 529–533 (2015).

[25] F. Motzoi, J. M. Gambetta, S. T. Merkel, and F. K. Wilhelm, “Optimal control methods for rapidly time-varying hamiltonians,” Physical Review A **84**, (2011).

[26] Y. Liu, S. J. Srinivasan, D. Hover, S. Zhu, R. McDermott, and A. A. Houck, “High fidelity readout of a transmon qubit using a superconducting low-inductance undulatory galvanometer microwave amplifier,” New Journal of Physics **16**, 113008 (2014).

Chapter 8

Summary & Outlook

Before universal quantum computers [1] can become a reality, several more milestones have to be achieved and engineering challenges overcome. With the accomplishment of quantum supremacy [2, 3], error-corrected quantum computing is the next declared milestone towards practical quantum computing hardware. While small-scale quantum systems perform on a level sufficient for the most lenient quantum error correction protocols [4], the qubit performance and control fidelity generally decrease as the number of qubits increases. In addition to scaling up the quantum processor and operation protocols, control hardware and software will become a bottleneck to handle hundreds to even thousands of qubits. Several blueprints have been presented on how to realize practical quantum processors using different qubit modalities [5, 6]. It remains to be seen if industrial and academic institutions will deliver on their promise to engineer a practical quantum computer before the end of this decade.

In this thesis, three tools were established to overcome some obstacles on the path from a few to many qubits. The three tools focus on efficient readout and control of many superconducting qubits. First, a multi-qubit microwave package was engineered to investigate and establish the design principles necessary for efficient qubit control and readout with minimal impact on the qubit performance. Next, deep machine learning techniques were developed to improve superconducting qubit readout pulse shapes and discrimination of qubit states.

8.1 Package Design for Superconducting Qubit

Using a newly engineered microwave package, presented in Chapter 4 and References [7, 8], microwave package design principles were developed by systematically examining various elements that can affect superconducting qubit coherence. A comprehensive characterization of the effect of package modes on superconducting transmon qubits was presented and corroborated with results from simulations and room-temperature measurements. For the utilized qubit design and configuration, the package limits the qubit lifetime to approximately $T_{\mathrm{limit}}^{\mathrm{pkg}} = 384\,\mu\mathrm{s}$. This lifetime is due almost entirely to qubit loss to hidden package modes via the Purcell effect. Package material losses contribute only at the $15\,\mathrm{ns}$ level. While the package does not limit the lifetime of the measured qubits, the estimated lifetime limit is within the same order of magnitude as other loss channels. Constructing packages with larger devices and qubit lifetimes that are likely achievable shortly will require a thorough engineering approach that focuses on mode and signal line engineering.

While many contemporary packages still employ wirebonds to provide signal connections between a multi-chip stack and the device package [2], several promising candidates such as pogo pins [9], out-of-plane wiring [10], direct chip-to-interposer links, and 3D-integrated packaging [11] may potentially provide a larger-scale interconnect solution. With a greater wire density, these techniques will face even more challenges in signal crosstalk, requiring precise impedance matching and the use of shielding structures. Combining these factors will necessitate thorough simulation and design characterization, building on those presented in this thesis.

Package design will become increasingly critical for larger quantum devices due to their increased complexity. The precise characterization and suppression of electromagnetic modes and signal crosstalk become even more critical as the number of qubits in today's noisy intermediate-scale quantum (NISQ) devices is increased [12]. These established principles for superconducting qubit packages are similarly pertinent for future work as packaging techniques are being advanced for systems in the range of 100 to 1000 qubits.

8.2 Machine Learning assisted Superconducting Qubit Readout

Qubit-state readout is a significant error source in contemporary superconducting quantum processors [13]. The fidelity of dispersive qubit-state readout depends on the readout pulse shape and resulting phase-shifted readout signal discriminator [14]. For a single qubit, fast and high-fidelity readout is achieved with minor changes to the rising and falling edge of a rectangular pulse [15] and a linear matched filter discriminator [16]. However, in resource-efficient, frequency-multiplexed readout of multiple qubits, optimizing the readout pulse shape and discriminator becomes a computationally intensive task. This thesis provides necessary components for robust and high-fidelity multi-qubit readout. Deep machine learning techniques to improve superconducting qubit readout pulse shapes and discrimination were experimentally demonstrated and compared with conventional methods. The presented readout techniques mark a significant step towards implementing quantum error correction protocols, hence realizing universal quantum computers.

8.2.1 Superconducting Qubit-State Discrimination using Supervised Learning

In Chapter 6 and Reference [17], an approach to multi-qubit readout using neural networks as multi-qubit state discriminators was presented. It was shown that neural networks are more crosstalk-resilient than other contemporary approaches. A fully-connected FNN increases the readout assignment fidelity for a multi-qubit system compared to contemporary methods. The FNN compensates system-nonidealities such as readout crosstalk more effectively relative to alternatives such as matched filters (MFs) or support vector machines (SVMs). The assignment error rate is diminished by up to $25\,\%$ and crosstalk-induced discrimination errors are suppressed by up to one order of magnitude. The relative assignment fidelity improvement of the FNN over its contemporary alternatives grows as the number of simultaneously read out and frequency-multiplexed qubits increases.

While FNNs are initially more resource-intensive in training, re-calibration can be significantly more efficient due to transfer learning [18]. Periodic re-calibration of control and readout parameters is necessary as quantum systems drift in time. For a marginal drift, neural networks can be updated at a fraction of the initial resource requirements. Furthermore, to speed up qubit readout, the techniques developed here can be transitioned to dedicated hardware such as field-programmable gate arrays (FPGA) [19].

FNNs offer a readout-state discrimination approach tailored to the underlying system. They can be readily employed to more general discrimination tasks than is considered in this thesis, such as multi-level readout in a qudit architecture [20–22]. Neural network-assisted discrimination of multiple qubit-states presents a potential flexible building block to scaling quantum processors while maintaining high-fidelity readout.

8.2.2 Superconducting Qubit-Readout-Pulse Shaping using Deep Reinforcement Learning

In an efficient, frequency-multiplexed readout of multiple qubits, effects such as drive cross-talk increase the complexity of optimal readout pulse shapes, requiring computationally intensive methods to discover high-fidelity pulse shapes. In Chapter 7 and Reference [23], existing readout optimization methods were extended to operate in multi-qubit environments. Furthermore, a new pulse shaping optimization module using deep reinforcement learning was presented. Compared to conventional readout methods in a simulated environment, the new module reduces the required readout pulse lengths by over 63% in single-qubit environments and over 57% in multi-qubit environments. The experimental verification of the simulated reduction in measurement time is currently in progress. PPO-generated readout pulse shapes are compared to rectangular and CLEAR pulses using a multi-qubit device in the experiment. The presented pulse shaping module can be readily generalized to other control tasks in quantum computing.

Using machine learning to improve multi-qubit-state readout, the presented approaches were demonstrated on a quantum system with five superconducting qubits and frequency-multiplexed readout with reasonable readout integration times. The next step will be to test the performance of these techniques on higher-fidelity multi-qubit systems with measurement times below $100\,\mathrm{ns}$ [16] to assess if the advantage will be retained on already high-performing devices. Furthermore, the temporal performance of the presented methods as system parameters drift will need to be investigated. The temporal robustness and effort to re-calibrate will provide further insight into the practicability of the presented techniques.

8.3 Global Calibration Module

Today, calibrating quantum computers is a tedious endeavor. Cross-talk between qubits and control signals is a fundamental roadblock for scaling up quantum computers from a few qubits to hundreds of qubits. The presence of multiple control signals distorts the individual control signal pulse shapes that eventually creates qubit errors. The current method of calibrating quantum computers does not address this signal cross-talk as control signals are independently calibrated for each constituent qubit. The design of control pulse shapes typically starts with a system-dependent ansatz that is subsequently fine-tuned by monitoring the response of each qubit. For a quantum computer with a few hundred qubits, a new method of controlling the qubits is necessary to perform accurate quantum computation.

The developed readout pulse shaping module using deep reinforcement learning can be expanded to a global calibration control module. Making the calibration procedure global addresses system nonidealities such as control signal crosstalk, as demonstrated for multi-qubit readout [17]. A global calibration control module comprises three components: central processing, pulse generation, and measurement unit. The previously developed pulse generation module and the readout methods will enable the central processing unit to communicate with the system in a closed-loop fashion.

The current calibration control methods prevent superconducting quantum computing from scaling up and demonstrating more complex quantum algorithms. An efficient, global, and general calibration control module using gradient-based optimization combined with deep reinforcement learning could address some of the challenges and enable more complex quantum algorithms.

References

[1] P. Shor, "Fault-tolerant quantum computation," Proceedings of 37th conference on foundations of computer science (1996), 56–65.

[2] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, "Quantum supremacy using a programmable superconducting processor," Nature **574**, 505–510 (2019).

[3] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, "Quantum computational advantage using photons," Science **370**, 1460–1463 (2020).

[4] 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**, (2019).

[5] B. Lekitsch, S. Weidt, A. G. Fowler, K. Mølmer, S. J. Devitt, C. Wunderlich, and W. K. Hensinger, "Blueprint for a microwave trapped ion quantum computer," Science Advances **3**, (2017).

[6] J. E. Bourassa, R. N. Alexander, M. Vasmer, A. Patil, I. Tzitrin, T. Matsuura, D. Su, B. Q. Baragiola, S. Guha, G. Dauphinais, K. K. Sabapathy, N. C. Menicucci, and I. Dhand, "Blueprint for a Scalable Photonic Fault-Tolerant Quantum Computer," Quantum **5**, 392 (2021).

[7] B. Lienhard, J. Braumüller, W. Woods, D. Rosenberg, G. Calusine, S. Weber, A. Vepsäläinen, K. O'Brien, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "Microwave packaging for superconducting qubits," 2019 IEEE MTT-S International Microwave Symposium (2019).

[8] S. Huang, B. Lienhard, G. Calusine, A. Vepsäläinen, J. Braumüller, D. K. Kim, A. J. Melville, B. M. Niedzielski, J. L. Yoder, B. Kannan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "Microwave package design for superconducting quantum processors," (2021).

[9] J. H. Béjanin, T. G. McConkey, J. R. Rinehart, C. T. Earnest, C. R. H. McRae, D. Shiri, J. D. Bateman, Y. Rohanizadegan, B. Penava, P. Breul, S. Royak, M. Zapatka, A. G. Fowler, and M. Mariantoni, "Three-Dimensional Wiring for Extensible Quantum Computing: The Quantum Socket," Physical Review Applied **6**, (2016).

[10] J. Rahamim, T. Behrle, M. J. Peterer, A. Patterson, P. A. Spring, T. Tsunoda, R. Manenti, G. Tancredi, and P. J. Leek, "Double-sided coaxial circuit qed with out-of-plane wiring," Applied Physics Letters **110**, 222602 (2017).

[11] D. Rosenberg, S. J. Weber, D. Conway, D. W. Yost, J. Mallek, G. Calusine, R. Das, D. Kim, M. E. Schwartz, W. Woods, J. L. Yoder, and W. D. Oliver, "Solid-state qubits: 3d integration and packaging," IEEE Microwave Magazine **21**, 72–85 (2020).

[12] J. Preskill, "Quantum Computing in the NISQ era and beyond," Quantum **2**, (2018).

[13] J. Gambetta, A. Blais, D. I. Schuster, A. Wallraff, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, "Qubit-photon interactions in

a cavity: measurement-induced dephasing and number splitting," Physical Review A **74**, 042318 (2006).

[14] 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," Physical Review A **69**, 1–14 (2004).

[15] D. T. McClure, H. Paik, L. S. Bishop, M. Steffen, J. M. Chow, and J. M. Gambetta, "Rapid Driven Reset of a Qubit Readout Resonator," Physical Review Applied **5**, 1–5 (2016).

[16] T. Walter, P. Kurpiers, S. Gasparinetti, P. Magnard, A. Potočnik, Y. Salathé, M. Pechal, M. Mondal, M. Oppliger, C. Eichler, and A. Wallraff, "Rapid High-Fidelity Single-Shot Dispersive Readout of Superconducting Qubits," Physical Review Applied **7**, 1–11 (2017).

[17] B. Lienhard, A. Vepsäläinen, L. C. G. Govia, C. R. Hoffer, J. Y. Qiu, D. Ristè, M. Ware, D. Kim, R. Winik, A. Melville, B. Niedzielski, J. Yoder, G. J. Ribeill, T. A. Ohki, H. K. Krovi, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "Deep neural network discrimination of multiplexed superconducting qubit states," (2021).

[18] Y. Bengio, "Practical Recommendations for Gradient-Based Training of Deep Architectures," *Neural networks: tricks of the trade: second edition* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012), 437–478.

[19] Z.-H. Ding, J.-M. Cui, Y.-F. Huang, C.-F. Li, T. Tu, and G.-C. Guo, "Fast High-Fidelity Readout of a Single Trapped-Ion Qubit via Machine-Learning Methods," Physical Review Applied **12**, 014038 (2019).

[20] P. Kurpiers, P. Magnard, T. Walter, B. Royer, M. Pechal, J. Heinsoo, Y. Salathé, A. Akin, S. Storz, J.-C. Besse, S. Gasparinetti, A. Blais, and A. Wallraff, "Deterministic quantum state transfer and remote entanglement using microwave photons," Nature **558**, 1476–4687 (2018).

[21] S. S. Elder, C. S. Wang, P. Reinhold, C. T. Hann, K. S. Chou, B. J. Lester, S. Rosenblum, L. Frunzio, L. Jiang, and R. J. Schoelkopf, "High-fidelity measurement of qubits encoded in multilevel superconducting circuits," Phys. Rev. X **10**, 011001 (2020).

[22] M. A. Yurtalan, J. Shi, G. J. K. Flatt, and A. Lupascu, "Characterization of multi-level dynamics and decoherence in a high-anharmonicity capacitively shunted flux circuit," (2020).

[23] B. Lienhard, C. R. Hoffer, A. Vepsäläinen, L. C. G. Govia, D. K., A. J. Melville, B. M. Niedzielski, J. L. Yoder, T. A. Ohki, H. K. Krovi, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "Readout pulse shaping using deep reinforcement learning," in preparation, (2021).

Appendix A

Cavity Bloch Equation Formulation

A.1 Dispersive Jaynes-Cummings Hamiltonian

The coherent interaction between superconducting qubits and microwave photons is described by the mathematical framework of circuit quantum electrodynamics (cQED) [1–3]. In particular, the interaction of a superconducting qubit and readout resonator can be described by the Jaynes-Cummings model in the dispersive limit. The Jaynes-Cummings Hamiltonian operating in the dispersive regime for cQED is defined as

$$H_{\text{JC, disp}} \approx \omega_r \left(\hat{a}^\dagger \hat{a} + \frac{1}{2}\right) + \left(\omega_q + 2\chi \left[\hat{a}^\dagger \hat{a} + \frac{1}{2}\right]\right) \frac{\hat{\sigma}_z}{2} \tag{A.1}$$

with the readout resonator frequency ω_r, the qubit transition frequency ω_q, the dispersive shift χ, the Pauli-Z operator $\hat{\sigma}_z$, and the creation $\hat{a}^\dagger$ and annihilation operator $\hat{a}$. Note, operators are indicated with a $\hat{\cdot}$.

A driving field at frequency ω_d can be expressed by

$$H_{\text{D}} = \epsilon(t)(\hat{a}^\dagger e^{-i\omega_d t} + \hat{a} e^{i\omega_d t}), \tag{A.2}$$

where $\epsilon(t)$ denotes the time dependent drive amplitude of the field [4]. If the resonator is driven at the bare resonator frequency ω_r, the readout resonator population can be treated as independent of the qubit state ($\Rightarrow \omega_d = \omega_r$).

While the interaction between the qubit and resonator is predominantly linear for small photon numbers, the so-called Kerr-nonlinearity can matter for moderate photon numbers. The Kerr-nonlinearity is a qubit-induced resonator shift. For a qubit in the ground state, the resonator is shifted lower in frequency. For an excited qubit, the resonator frequency is raised. The Kerr-nonlinearity K induces a shift quadratic in the number of photons ($\hat{a}^\dagger\hat{a}$ is the photon number operator) as can be seen in the following Hamiltonian

$$H_{\mathrm{K}} = K(a^\dagger a)^2. \tag{A.3}$$

The complete Hamiltonian describing a driven resonator interacting with a qubit in the dispersive regime can be expressed by

$$\begin{aligned} H_{\mathrm{JC,\ disp}} &= \omega_r \hat{a}^\dagger\hat{a} + \left(\omega_q + 2\chi\left[\hat{a}^\dagger\hat{a} + \frac{1}{2}\right]\right)\frac{\hat{\sigma}_z}{2}, \\ H_{\mathrm{D}} &= \epsilon(t)(\hat{a}^\dagger e^{-i\omega_d t} + \hat{a}e^{i\omega dt}), \\ H_{\mathrm{K}} &= K(\hat{a}^\dagger\hat{a})^2, \end{aligned} \tag{A.4}$$

$$H = H_{\text{JC, disp}} + H_{\mathrm{D}} + H_{\mathrm{K}}.$$

A.2 Master Equation

Measuring the state of a qubit requires the resonator to leak the information. Leakage is typically a non-unitary process. The resonator photon decay dynamics with decay rate κ can be included in the model using master equations. Master equations are a common mathematical framework to represent the coupling of quantum systems with their environment. A Lindblad master equation for a density matrix ρ and Hamiltonian H can be expressed as

$$\dot{\rho} = -i[H,\rho] + \kappa\mathcal{D}[\hat{a}]\rho. \tag{A.5}$$

where $\mathcal{D}[\hat{a}]\rho = \hat{a}\rho\hat{a}^\dagger - \{\hat{a}^\dagger\hat{a}, \rho\}/2 = \hat{a}\rho\hat{a}^\dagger - \hat{a}^\dagger\hat{a}\rho/2 - \rho\hat{a}^\dagger\hat{a}/2$ is the dissipator describing the leakage.

The master equation solver provided by QuTip [5] can be computationally cumbersome. For moderately-sized problems with a few qubits, the computations can already be tedious. The Bloch equation formulation, however, provides a sufficiently accurate and fast alternative.

A.3 Cavity Bloch Equation

The cavity Bloch equation formulation expresses the temporal dynamics of a quantum system [6]. To study qubit measurements, the photon number operator $\hat{a}^\dagger\hat{a}$ and resonator state operator $\hat{a}$ are sufficient.

For an arbitrary operator $\hat{A}$, the cavity Bloch equation is

$$\frac{\partial}{\partial t}\langle \hat{A} \rangle = \mathrm{Tr}\{\hat{A}\dot{\rho}\} = -i\langle [\hat{A}, H]\rho \rangle + \kappa \langle \mathcal{D}[\hat{a}]\hat{A} \rangle. \tag{A.6}$$

The cavity Bloch equation for the evolution of the resonator state operator $\hat{a}$ and photon count operator $\hat{a}^\dagger\hat{a}$ using the Hamiltonian from Equation A.4 can now be computed. For $\hat{a}$, the equation is

$$\begin{aligned} \frac{\partial}{\partial t}\langle \hat{a} \rangle &= -i\langle [\hat{a}, H]\rho \rangle + \kappa \langle \mathcal{D}[\hat{a}]\hat{a} \rangle \\ &= -i\Delta - i\chi\langle \hat{a}\hat{\sigma}_z \rangle + i\epsilon(t) - iK\langle \hat{a} \rangle - iK\langle \hat{a}^\dagger\hat{a}\hat{a} \rangle - \frac{\kappa}{2}\langle \hat{a} \rangle, \end{aligned} \tag{A.7}$$

with the difference between the resonator and drive frequency $\Delta = |\omega_r - \omega_d|$.

Consequently, the dynamics of the photon number operator $\hat{a}^\dagger\hat{a}$ are

$$\begin{aligned} \frac{\partial}{\partial t}\langle \hat{a}^\dagger\hat{a} \rangle &= -i\langle [\hat{a}^\dagger\hat{a}, H]\rho \rangle + \kappa \langle \mathcal{D}[\hat{a}]\hat{a}^\dagger\hat{a} \rangle \\ &= -2\epsilon(t)\,\mathrm{Im}\{\langle \hat{a} \rangle\} - \kappa\langle \hat{a}^\dagger\hat{a} \rangle. \end{aligned} \tag{A.8}$$

A.4 Frequency-Multiplexed Readout

The mathematical framework introduced in Section A.3 can be extended to characterize a multi-qubit system with Q qubits, R resonators, and frequency-multiplexed readout. The dynamics for each resonator state operator $\hat{a}_n$ are described by

$$\begin{aligned}\frac{\partial}{\partial t}\langle \hat{a}_n \rangle =& \sum_{n \in R} \sum_{m \in Q} -i\Delta_{n,m} + i\epsilon_n(t)) - \frac{\kappa_n}{2}\langle \hat{a}_n \rangle \\ &- i\chi_{n,m}\langle \hat{a}_n \hat{\sigma}_{z\,m} \rangle - iK_{n,m}\langle \hat{a}_n \rangle - iK_{n,m}\langle (\hat{a}^\dagger \hat{a})_n \hat{a}_n \rangle).\end{aligned} \tag{A.9}$$

The first line of Equation A.9 describes the driving and leakage of the resonator. The driving at off-resonant frequencies targeting other resonators induces control crosstalk. The qubit-induced resonator frequency shifts are expressed in the second line. In addition to the resonator frequency shifts induced by the target qubit, neighboring qubits can induce qubit-state-dependent shifts as well. While the coupling to neighboring qubits is diminished, their effect is not negligible.

The photon number operator $(\hat{a}^\dagger \hat{a})_n$ for resonator n is defined as

$$\frac{\partial}{\partial t}\langle (\hat{a}^\dagger \hat{a})_n \rangle = -2\epsilon_n(t)\,\mathrm{Im}\{\langle \hat{a}_n \rangle\} - \kappa \langle (\hat{a}^\dagger \hat{a})_n \rangle. \tag{A.10}$$

The introduced cavity Bloch equations are used to simulate a five-qubit system with experimentally evaluated parameters described in Appendix B.2. The simulator is utilied to evaluate qubit-readout pulse shapes. As described in Chapter 7, the qubit-readout pulse shapes become increasingly complex as the number of qubits is increased. The suspected reason is crosstalk that the presented simulator captures in two aspects: control crosstalk and resonators coupling to qubits in the vicinity. The simulator takes all present drive fields into account. However, the simulator is simplified only to consider resonators coupling to their target and nearest-neighbor qubit.

References

[1] 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,” Physical Review A **69**, 1–14 (2004).

[2] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, R. S. Huang, J. Majer, S. Kumar, S. M. Girvin, and R. J. Schoelkopf, “Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics,” Nature **431**, 162–167 (2004).

[3] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, “Circuit quantum electrodynamics,” Rev. Mod. Phys. **93**, 025005 (2021).

[4] S. Boutin, C. K. Andersen, J. Venkatraman, A. J. Ferris, and A. Blais, “Resonator reset in circuit QED by optimal control for large open quantum systems,” Physical Review A **96**, 1–11 (2017).

[5] J. Johansson, P. Nation, and F. Nori, “Qutip: an open-source python framework for the dynamics of open quantum systems,” Computer Physics Communications **183**, 1760–1772 (2012).

[6] F. Bloch, “Nuclear induction,” Physical Review **70**, 460–474 (1946).

Appendix B

Experimental Infrastructure

B.1 Measurement Setup

Qubit control and readout pulses—envelopes with cosine-shaped rising and falling edges encompassing a plateau—are programmed in Labber [1]. They are created using three—two for control and one for readout—Keysight M3202A PXI arbitrary waveform generators (AWG) with a sampling rate of $1\,\mathrm{GSa\,s^{-1}}$. Next, the in-phase (*I*) and quadrature (*Q*) components of the signals at MHz frequencies are up-converted to the qubit transition frequency using an *IQ*-mixer and a local oscillator (LO) (Rohde and Schwarz SGS100A) per AWG. Finally, the control and readout tones are combined and sent to the qubit chip in the dilution refrigerator via a single microwave line attenuated by $60\,\mathrm{dB}$.

The qubit chip is mounted in a microwave package following design principles described in Chapter 4. A coil—centered above the qubit chip—is mounted in the device package. A global flux bias Φ is applied through that coil to the superconducting quantum interference devices (SQUID) of the qubits using a Yokogawa GS200.

The readout signal, upon acquisition of a qubit-state-dependent phase shift, is first amplified using a Josephson traveling-wave parametric amplifier (JTWPA) with near quantum-limited performance over a bandwidth of more than $2\,\mathrm{GHz}$ and a $1\,\mathrm{dB}$ compression point of approximately $-100\,\mathrm{dBm}$ [2]. An Agilent E8267D signal generator provides the pump tone for the JTWPA. The microwave line carrying the pump tone is attenuated by $50\,\mathrm{dB}$ and fed into the JTWPA via a set of directional couplers and isolators located in the mixing chamber of the refrigerator. The signal is further amplified by a high-electron-mobility transistor (HEMT) amplifier that is thermally anchored to the $3\,\mathrm{K}$ stage.

At room temperature, the readout signal is amplified, IQ-mixed with the LO at $7.127\,\mathrm{GHz}$, and fed into a heterodyne detector. Next, the *I*- and *Q*-components of the readout signal are digitized with a Keysight M3102A PXI Analog to Digital Converter (ADC) at a sampling rate of $500\,\mathrm{MSa\,s^{-1}}$. The subsequent digital signal processing to distinguish qubit states is the focus of Chapter 6.

B.2 Superconducting Chip with Five Transmon Qubits

The quantum system comprised of five superconducting qubits is fabricated on a (001) silicon substrate ($>3500\,\Omega\,\mathrm{cm}$) by dry etching a molecular-beam epitaxy (MBE) grown aluminum film in an optical lithography process before being diced into 5×5 mm^2 chips, as described in [3].

The superconducting chip consists of coplanar waveguides and five frequency-tunable transmon qubits [4]. The target qubit transition frequencies alternate between $4.3\,\mathrm{GHz}$ and $5.2\,\mathrm{GHz}$. The qubits are detuned ($\rightarrow$ operating frequency) to limit qubit-qubit and control crosstalk. The capacitive nearest-neighbor (next-nearest-neighbor) qubit-qubit coupling rate, J_{nn} (J_{nnn}), is designed (using COMSOL Multiphysics®) to be $J_{nn}/2\pi \approx 14\,\mathrm{MHz}$ ($J_{nnn}/2\pi < 1\,\mathrm{MHz}$) and at the qubit operating frequency $< 0.3\,\mathrm{MHz}$ ($< 0.01\,\mathrm{MHz}$) [5]. Each qubit couples capacitively to a quarter-wave resonator that couples inductively to a shared bandpass (Purcell) filtered feedline. Neighboring readout resonator frequencies differ by about $50\,\mathrm{MHz}$. The qubit and resonator operation parameters are included in Table B.1 and Table B.2.

Table B.1: Superconducting Chip—Qubit Details. Chip comprising five superconducting frequency-tunable transmon qubits with alternating transition frequencies. A normalized magnetic flux bias Φ/Φ_0 (magnetic flux quantum Φ_0) detunes the qubits from their idling to their operating frequency. The qubit anharmonicities α are in the moderate transmon regime. The qubit lifetimes T_1, Ramsey coherence times $T_{2\mathrm{R}}$, and spin-echo relaxation times $T_{2\mathrm{E}}$ are measured at the qubit operating frequency.

Qubit	$\omega_{\mathrm{Qubit}}/2\pi$		Bias	$\alpha/2\pi$	T_1	$T_{2\mathrm{R}}$	$T_{2\mathrm{E}}$
	Idle (GHz)	Biased	$\left(\frac{\Phi}{\Phi_0}\right)$	(MHz)		(μs)	
1	5.249	5.092	0.124	-212	40.8	1.3	7.4
2	4.708	4.404	0.160	-216	6.4	0.6	4.1
3	5.202	5.000	0.166	-204	21.4	1.0	7.2
4	4.560	4.309	0.154	-214	11.8	0.8	5.4
5	5.196	5.165	0.085	-200	23.4	7.6	31.8

Table B.2: Superconducting Chip—Resonator Details. Chip comprising five superconducting readout resonators at bare resonance frequencies of about $7\,\mathrm{GHz}$. Signals are up-converted from MHz intermediate frequencies (IF) using a common local oscillator at $\omega_{\mathrm{LO}}/2\pi = 7.127\,\mathrm{GHz}$. Each resonator couples to a designated qubit with strength g, leading to a dispersive shift χ. The effective resonator decay rate through the Purcell filter is κ_{eff}. The qubit-resonator interaction remains in the dispersive regime for readout resonator photon populations below the critical photon number n_{crit}.

Resonator	$\omega_{\mathrm{Res}}/2\pi$ (GHz)	$\omega_{\mathrm{IF}}/2\pi$ (MHz)	$g/2\pi$	$\chi/2\pi$ (MHz)	$\kappa_{\mathrm{eff}}/2\pi$	n_{crit}
1	7.06	-65	116.3	0.83	4.29	33.8
2	7.10	-26	143.3	0.51	4.25	55.3
3	7.15	24	125.7	0.77	4.41	34.9
4	7.20	70	133.1	0.49	3.33	56.9
5	7.25	127	125.4	0.80	6.90	33.0

References

[1] Keysight, *Labber software*, 2018, `http://quantum.labber.org/`.

[2] C. Macklin, K. O'Brien, D. Hover, M. E. Schwartz, V. Bolkhovsky, X. Zhang, W. D. Oliver, and I. Siddiqi, "A near–quantum-limited josephson traveling-wave parametric amplifier," Science **350**, 307–310 (2015).

[3] F. Yan, S. Gustavsson, A. Kamal, J. Birenbaum, A. P. Sears, D. Hover, T. J. Gudmundsen, D. Rosenberg, G. Samach, S. Weber, J. L. Yoder, T. P. Orlando, J. Clarke, A. J. Kerman, and W. D. Oliver, "The flux qubit revisited to enhance coherence and reproducibility," Nature Communications **7**, 12964 (2016).

[4] 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**, 042319 (2007).

[5] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, "Circuit quantum electrodynamics," Rev. Mod. Phys. **93**, 025005 (2021).

Appendix C

Alternative Matched Filter Derivation

C.1 Derivation of Matched Filter

To reduce the computational discrimination effort, the elements of a measured single-shot readout trace are often summed up before a discriminator is applied. Filtering the readout traces before they are summed up further simplifies the discrimination process. Filtering in this context means multiplying each element of a discrete signal $s_{|i\rangle}[n]$ by a window or kernel weight k_n. If the weights are all unity over a particular range and zero otherwise, the filter is a boxcar filter.

A discrete qubit measurement signal $s_{|i\rangle}[n]$ with qubit either in the ground $(i = 0)$ or excited-state $(i = 1)$ can be modeled as $s_{|i\rangle}[n] \propto \alpha_{|i\rangle}[n] + \xi[n]$. The resonator response for the qubit in the ground-state is $\alpha_{|0\rangle}[n]$ and for the qubit in the excited-state $\alpha_{|1\rangle}[n]$. The measurement signal is a linear combination of the resonator response and a zero-mean stochastic noise term $\xi[n]$ [1, 2]. The filtered measurement signal with filter kernel $k[n]$ is described as

$$S_{|i\rangle} = \sum_{n} k_n s_{|i\rangle\, n}. \tag{C.1}$$

The average difference $\langle \Delta S \rangle$ between the ground-state signal $S_{|0\rangle}$ and excited-state signal $S_{|1\rangle}$ follows as

$$\langle \Delta S \rangle = \langle S_{|0\rangle} - S_{|1\rangle} \rangle = \sum_n k_n \langle \alpha_{|0\rangle' n} - \alpha_{|1\rangle' n} \rangle. \tag{C.2}$$

The variance of the average difference $\langle \Delta S \rangle$ is consequently

$$\text{var}(\Delta S) = \sum_n k_n^2 \left[\text{var}\left(\alpha_{|0\rangle' n} - \alpha_{|1\rangle' n} \right) + \text{var}(\xi_n) \right]. \tag{C.3}$$

To maximize the SNR, the filter kernel $k[n]$ can be derived taking the derivative of the squared mean value divided by the variance as follows (an alternative derivation is presented in Appendix C)

$$\begin{aligned}
0 &= \frac{\partial}{\partial k_i} \frac{|\langle S \rangle|^2}{\text{var}(\Delta S)} \\
&= \langle \alpha_{|0\rangle\, i} - \alpha_{|1\rangle\, i} \rangle \text{var}(\Delta S) - 2k_i \left[\text{var}\left(\alpha_{|0\rangle\, i} - \alpha_{|1\rangle\, i} \right) + \text{var}(\xi_i) \right] \langle S \rangle \\
&= \langle \alpha_{|0\rangle\, i} - \alpha_{|1\rangle\, i} \rangle \sum_n k_n^2 \left[\text{var}\left(\alpha_{|0\rangle\, n} - \alpha_{|1\rangle\, n} \right) + \text{var}(\xi_n) \right] \\
&\quad - k_i \left[\text{var}\left(\alpha_{|0\rangle\, i} - \alpha_{|1\rangle\, i} \right) + \text{var}(\xi_i) \right] \sum_n k_n \langle \alpha_{|0\rangle\, n} - \alpha_{|1\rangle\, n} \rangle \\
&= \sum_{n, i \neq n} \langle \alpha_{|0\rangle\, i} - \alpha_{|1\rangle\, i} \rangle k_n^2 \left[\text{var}\left(\alpha_{|0\rangle\, n} - \alpha_{|1\rangle\, n} \right) + \text{var}(\xi_n) \right] \\
&\quad - k_i \left[\text{var}\left(\alpha_{|0\rangle\, i} - \alpha_{|1\rangle\, i} \right) + \text{var}(\xi_i) \right] k_n \langle \alpha_{|0\rangle\, n} - \alpha_{|1\rangle\, n} \rangle
\end{aligned} \tag{C.4}$$

solving for k_i

$$k_i = \frac{\langle \alpha_{|0\rangle\, i} - \alpha_{|1\rangle\, i} \rangle}{\left[\text{var}\left(\alpha_{|0\rangle\, i} - \alpha_{|1\rangle\, i} \right) + \text{var}(\xi_i) \right]} \; \frac{\sum_{n, i \neq n} k_n^2 \left[\text{var}\left(\alpha_{|0\rangle\, n} - \alpha_{|1\rangle\, n} \right) + \text{var}(\xi_n) \right]}{\sum_{n, i \neq n} k_n \langle \alpha_{|0\rangle\, n} - \alpha_{|1\rangle\, n} \rangle}$$

Dropping the scaling factor independent of i, the kernel weights are

$$k_i = \frac{\langle \alpha_{|0\rangle\, i} - \alpha_{|1\rangle\, i} \rangle}{\left[\text{var}\left(\alpha_{|0\rangle\, i} - \alpha_{|1\rangle\, i} \right) + \text{var}(\xi_i) \right]} \tag{C.5}$$

Table C.1: Definition of used Terminology. Derivation of a matched filter to discriminate qubit states [3]. The noise source affecting the measurement is assumed to be additive and stochastic.

input signal (n dimensional):	$x[\mathrm{n}]$
output signal (1 dimensional) with projection $w[\mathrm{n}]$:	$y = w^{\mathrm{T}}[\mathrm{n}]x[n]$
Class 1 (N_1 samples), mean value:	$m_1[\mathrm{n}] = \frac{1}{N_1}\sum_{i\in C_1} x_i[\mathrm{n}]$
Class 2 (N_2 samples), mean value:	$m_2[\mathrm{n}] = \frac{1}{N_2}\sum_{i\in C_2} x_i[\mathrm{n}]$
Class 1 (N_1 samples), mean value after projection:	$m_1 = w^{\mathrm{T}}[\mathrm{n}]\frac{1}{N_1}\sum_{i\in C_1} x_i[\mathrm{n}]$
Class 2 (N_2 samples), mean value after projection:	$m_2 = w^{\mathrm{T}}[\mathrm{n}]\frac{1}{N_2}\sum_{i\in C_2} x_i[\mathrm{n}]$
Class 1 (N_1 samples), variance after projection:	$v_1 = \sum_{i\in C_1}(y_n - m_1)^2$
Class 2 (N_2 samples), variance after projection:	$v_2 = \sum_{i\in C_2}(y_n - m_2)^2$
Class distance after projection (Fisher criterion):	$F(w[\mathrm{n}]) = \frac{(m_2-m_1)^2}{(v_1+v_2)}$
Maximize Fisher criterion → maximize SNR:	

C.2 Alternative Derivation of Matched Filter

$$\text{(explicitly in } w[\mathrm{n}]\text{)} \quad F(w[\mathrm{n}]) = \frac{w^{\mathrm{T}}[\mathrm{n}]M[\mathrm{n}]w[\mathrm{n}]}{w^{\mathrm{T}}[\mathrm{n}]V_1[\mathrm{n}]w[\mathrm{n}] + w^{\mathrm{T}}[\mathrm{n}]V_2[\mathrm{n}]w[\mathrm{n}]} \tag{C.6}$$

where

$V_1[n] = \sum_{i\in C_1}(x_i[\mathrm{n}] - m_1[\mathrm{n}])(x_i[\mathrm{n}] - m_1[\mathrm{n}])^{\mathrm{T}}$,

$V_2[n] = \sum_{i\in C_2}(x_i[\mathrm{n}] - m_2[\mathrm{n}])(x_i[\mathrm{n}] - m_2[\mathrm{n}])^{\mathrm{T}}$, and

$M[\mathrm{n}] = (m_2[\mathrm{n}] - m_1[\mathrm{n}])(m_2[\mathrm{n}] - m_1[\mathrm{n}])^{\mathrm{T}}$.

Maximize explicit equation with respect to $w[\mathrm{n}]$:

$$\frac{\partial F(w[\mathrm{n}])}{\partial w[\mathrm{n}]} = \frac{\begin{array}{c}2\left(w^{\mathrm{T}}[\mathrm{n}]\left(V_1[\mathrm{n}] + V_2[\mathrm{n}]\right)w[\mathrm{n}]\right)M[\mathrm{n}]w[\mathrm{n}] \\ \hline -2\left(w^{\mathrm{T}}[\mathrm{n}]M[\mathrm{n}]w[\mathrm{n}]\right)(V_1[\mathrm{n}] + V_2[\mathrm{n}])w[\mathrm{n}]\end{array}}{\left(w^{\mathrm{T}}[\mathrm{n}]V_1[\mathrm{n}]w[\mathrm{n}] + w^{\mathrm{T}}[\mathrm{n}]V_2[\mathrm{n}]w[\mathrm{n}]\right)^2} \tag{C.7}$$

$$\begin{aligned} 0 &= \frac{\partial F(w[\mathrm{n}])}{\partial w[\mathrm{n}]} \\ &= \left(w^{\mathrm{T}}[\mathrm{n}]\left(V_1[\mathrm{n}] + V_2[\mathrm{n}]\right)w[\mathrm{n}]\right)M[\mathrm{n}]w[\mathrm{n}] \\ &= \left(w^{\mathrm{T}}[\mathrm{n}]M[\mathrm{n}]w[\mathrm{n}]\right)(V_1[\mathrm{n}] + V_2[\mathrm{n}])w[\mathrm{n}] \end{aligned} \tag{C.8}$$

$\big(w^{\mathrm{T}}[\mathrm{n}]\,(V_1[\mathrm{n}]+V_2[\mathrm{n}])\,w[\mathrm{n}]\big)$ and $\big(w^{\mathrm{T}}[\mathrm{n}]\,(M[\mathrm{n}])\,w[\mathrm{n}]\big)$ are scalar factors. The magnitude of $w[\mathrm{n}]$ does not matter for discrimination. Therefore, scalar factors can be dropped.

After dropping the scalar factors, and using

$$\frac{M[\mathrm{n}]w[\mathrm{n}]}{|M[\mathrm{n}]w[\mathrm{n}]|} = \frac{m_2[\mathrm{n}]-m_1[\mathrm{n}]}{|m_2[\mathrm{n}]-m_1[\mathrm{n}]|}, \tag{C.9}$$

replacing the vector $M[\mathrm{n}]w[\mathrm{n}]$ with $m_2[\mathrm{n}]-m_1[\mathrm{n}]$, and multiplying both sides with $(V_1[\mathrm{n}]+V_2[\mathrm{n}])^{-1}$, the optimal weight is

$$w_{\mathrm{opt}}[\mathrm{n}] \propto \frac{m_2[\mathrm{n}]-m_1[\mathrm{n}]}{V_1[\mathrm{n}]+V_2[\mathrm{n}]} \tag{C.10}$$

C.3 Matched Filter for Qubit-State Discrimination

Following the derivation of the matched filter $w_{\mathrm{opt}}[\mathrm{n}] \propto \frac{m_2[\mathrm{n}]-m_1[\mathrm{n}]}{V_1[\mathrm{n}]+V_2[\mathrm{n}]}$, the average and variance of the measured signal in a known state ($i \in N_g$ ground-state measurement $s_{g_i}[\mathrm{n}] = \mathcal{I}_{g_i}[\mathrm{n}] + j\mathcal{Q}_{g_i}[\mathrm{n}]$; $i \in N_e$ excited-state measurement $s_{e_i}[\mathrm{n}] = \mathcal{I}_{e_i}[\mathrm{n}] + j\mathcal{Q}_{e_i}[\mathrm{n}]$) can be used to generate the matched filter kernel.

$$K^{\mathrm{MF}}_{\mathcal{I},\mathcal{Q}}[2\mathrm{n}] \approx \frac{S_e[2\mathrm{n}]-S_g[2\mathrm{n}]}{V_g[2\mathrm{n}]+V_e[2\mathrm{n}]}, \tag{C.11}$$

with

$$S_g[2\mathrm{n}] = \tfrac{1}{N_g}\textstyle\sum_i^{N_g} \begin{bmatrix} \mathcal{I}_{g_i}[\mathrm{n}] \\ \mathcal{Q}_{g_i}[\mathrm{n}] \end{bmatrix} \qquad V_g[2\mathrm{n}] = \tfrac{1}{N_g}\textstyle\sum_i^{N_g} \left(\begin{bmatrix} \mathcal{I}_{g_i}[\mathrm{n}] \\ \mathcal{Q}_{g_i}[\mathrm{n}] \end{bmatrix} - S_g[2\mathrm{n}] \right)$$

$$S_e[2\mathrm{n}] = \tfrac{1}{N_e}\textstyle\sum_i^{N_e} \begin{bmatrix} \mathcal{I}_{e_i}[\mathrm{n}] \\ \mathcal{Q}_{e_i}[\mathrm{n}] \end{bmatrix} \qquad V_e[2\mathrm{n}] = \tfrac{1}{N_e}\textstyle\sum_i^{N_e} \left(\begin{bmatrix} \mathcal{I}_{e_i}[\mathrm{n}] \\ \mathcal{Q}_{e_i}[\mathrm{n}] \end{bmatrix} - S_e[2\mathrm{n}] \right)$$

References

[1] C. A. Ryan, B. R. Johnson, J. M. Gambetta, J. M. Chow, M. P. Da Silva, O. E. Dial, and T. A. Ohki, “Tomography via correlation of noisy measurement records,” Physical Review A **91**, 1–7 (2015).

[2] C. C. Bultink, B. Tarasinski, N. Haandbæk, S. Poletto, N. Haider, D. J. Michalak, A. Bruno, and L. DiCarlo, “General method for extracting the quantum efficiency of dispersive qubit readout in circuit QED,” Applied Physics Letters **112**, 092601 (2018).

[3] C. M. Bishop, *Pattern recognition and machine learning (information science and statistics)* (Springer-Verlag, Berlin, Heidelberg, 2006).