

Quantum Jumps and Measurement Backaction in a Superconducting Qubit

by

Daniel Huber Slichter

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Irfan Siddiqi, Chair

Professor John Clarke

Professor Dmitry Budker

Professor Birgitta Whaley

Fall 2011

Abstract

Quantum Jumps and Measurement Backaction in a Superconducting Qubit

by

Daniel Huber Slichter

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Irfan Siddiqi, Chair

Real-time monitoring of a quantum state provides powerful tools for studying the backaction of quantum measurement and performing quantum feedback. Historically, this monitoring capability has been the exclusive province of atomic and optical physics. This thesis describes the implementation of the first such high-fidelity readout scheme in a solid state circuit, a superconducting quantum bit (qubit) coupled to a microwave cavity in the circuit quantum electrodynamics (circuit QED) architecture. The qubit-state-dependent resonance frequency of the cavity is probed with a microwave drive tone, and the resulting signal amplified using a fast, ultralow-noise superconducting parametric amplifier. This arrangement enables the observation of quantum jumps between the qubit states in real time.

The ability to monitor the qubit continuously with high fidelity and resolve quantum jumps can be used to investigate the backaction of the measurement process on the qubit. This thesis examines the quantum Zeno effect—where strong measurement inhibits the evolution of a quantum system—as well as the transition to non-ideal measurement with increasing measurement strength in the circuit QED architecture, a phenomenon shown to be due to the upconversion of low-frequency dephasing noise. These data allow probes of “universal” flux noise in previously inaccessible frequency ranges. The work presented here opens the door for quantum feedback and error correction in solid-state quantum systems using continuous weak measurement.

Quantum Jumps and Measurement Backaction in a Superconducting Qubit

Copyright 2011

by

Daniel Huber Slichter

For my family

Contents

List of Figures	v
List of Symbols and Abbreviations	viii
1 Introduction	1
1.1 Superconducting qubits	2
1.2 Quantum jumps	3
1.3 Superconducting parametric amplifiers	5
1.4 Thesis overview	6
1.5 Summary of key results	7
2 Superconducting qubits and circuit quantum electrodynamics	8
2.1 Qubits and quantum information	8
2.1.1 Quantization of an electrical circuit	10
2.1.2 Superconducting qubits	11
2.1.3 The transmon qubit	14
2.2 Quantum non-demolition measurement	18
2.3 Circuit quantum electrodynamics	19
2.3.1 The Jaynes-Cummings Hamiltonian	21
2.3.2 The dispersive approximation	23
2.3.3 Transmon dispersive shift	25
2.3.4 Measurement rate and readout signal-to-noise ratio	27
3 Amplification and the quantum limit	30
3.1 Amplification and noise	30
3.2 Quantum limits on amplifiers	34
3.2.1 Phase-sensitive and phase-preserving amplifiers	36
3.3 Parametric amplification	37
3.3.1 Brief history of parametric amplification	38
3.4 The Lumped Josephson Parametric Amplifier (LJPA)	39
3.4.1 Mathematical description	41
3.4.2 Theoretical gain and bandwidth	43
3.4.3 Physical description of operation	45
3.4.4 Saturated regime operation	51

4 Sample fabrication	54
4.1 Electron-beam lithography	54
4.1.1 Resist selection	56
4.1.2 Cold development	58
4.2 Thin-film deposition	59
4.2.1 Oxygen plasma cleaning	62
4.2.2 Ground plane fabrication	63
4.3 Sample design and parameters	65
4.3.1 Transmons	65
4.3.2 Paramaps	67
5 Experimental apparatus	68
5.1 Sample boxes and launches	68
5.2 Fridge wiring	70
5.2.1 Low-frequency wiring	70
5.2.2 Microwave wiring and switches	72
5.2.3 Microwave “roach motel” filters	76
5.2.4 Hot/cold load setup	78
5.3 Room temperature electronics	80
5.3.1 Pulse and tone generation	80
5.3.2 Variable attenuators	81
5.3.3 Demodulation and digitization	82
6 Calibration experiments	84
6.1 Paramamp sample packaging	84
6.2 Paramamp biasing procedure	85
6.3 Paramamp performance	89
6.3.1 Paramamp gain	89
6.3.2 Noise performance	91
6.3.3 Saturated regime operation	93
6.4 Calibration of the qubit/cavity system	94
6.4.1 Qubit and cavity spectroscopy	94
6.4.2 Rabi, Ramsey, T_1 , and tomography	97
6.4.3 Photon number and ac Stark shift	101
7 Quantum jumps	104
7.1 What are quantum jumps?	104
7.2 Historical background	105
7.3 Experimental design	106
7.4 Observation of quantum jumps	109
7.4.1 Jumps without a paramamp	115
7.4.2 Three-level jumps	116
7.5 Signal-to-noise ratio and fidelity	118
7.5.1 Theoretical SNR	119
7.5.2 Experimental measurements of SNR	122

7.5.3	Measurement fidelity	123
8	Measurement backaction	126
8.1	Automated qubit state extraction	126
8.1.1	Determining the qubit state	127
8.1.2	Jump time distribution	131
8.1.3	Maximum likelihood estimation	134
8.1.4	Tests of the extraction algorithm	141
8.2	Measurement-induced state mixing	144
8.2.1	Dressed dephasing theory	144
8.2.2	Experimental scheme	146
8.2.3	State mixing with added flux tones	148
8.2.4	Flux noise measurements	152
8.3	Quantum Zeno effect	154
9	Conclusions and outlook	161
9.1	Future work	161
9.1.1	Quantum feedback and error correction	161
9.1.2	Single microwave photon source and detector	162
9.1.3	Improved qubits	162
9.1.4	Parametric amplifier development	163
9.2	Conclusions	164
Bibliography		166
A	Microwave “roach motel” filters	182
A.1	Design and modeling of roach filters	182
A.2	Millikelvin filter performance	185
B	Fabrication recipes	189
B.1	Resist Spinning	189
B.2	E-beam lithography	189
B.3	Resist Development	191
B.4	Thin film deposition	191
B.5	Silicon nitride deposition	192
C	Experimental schematics	193

List of Figures

1.1	The quantronium qubit.	3
1.2	The first observation of quantum jumps.	4
2.1	Quantum two-level systems.	9
2.2	Energy levels of quantized oscillators.	12
2.3	Generalized superconducting qubit.	14
2.4	Schematic of transmon qubit.	15
2.5	Energy structure of the Cooper pair box and transmon qubits.	17
2.6	Schematic of cavity QED.	20
2.7	Schematic of circuit QED.	21
2.8	Avoided crossing from the Jaynes-Cummings Hamiltonian.	22
2.9	Principle of circuit QED readout.	24
2.10	Readout signal and noise in the IQ plane.	29
3.1	Amplification process.	31
3.2	Nonlinear resonance response.	40
3.3	Schematic of the LJPA.	41
3.4	Nonlinear resonance solutions.	42
3.5	Theoretical gain and bandwidth.	44
3.6	Paramp transfer function: reflected phase vs. pump amplitude.	46
3.7	Phase-sensitive amplification in the IQ plane.	47
3.8	Phase-preserving amplification in the IQ plane.	49
3.9	Amplification and noise for a detuned signal.	50
3.10	Saturated regime amplification.	52
4.1	Tools of the fab trade.	55
4.2	Resist mask troubleshooting.	57
4.3	Double-angle evaporation.	60
4.4	SlichTECH e-beam evaporator and NRC thermal evaporator.	61
4.5	O ₂ plasma cleaning to remove the “black veil of death”.	63
4.6	Measured loss tangent of microlab SiN _x	64
4.7	Transmon samples.	66
4.8	Paramp sample.	67
5.1	Sample boxes and launches.	69

5.2	VeriCold dilution refrigerator and sample boxes.	71
5.3	Flux bias noise from a Keithley sourcemeter.	72
5.4	Circulators and microwave switches.	75
5.5	Microwave “roach motel” filters.	77
5.6	Hot/cold load setup for calibrating noise temperature.	79
6.1	Paramp sample box.	85
6.2	Typical paramp tuning with flux.	86
6.3	Paramp transfer function.	87
6.4	Paramp phase response.	88
6.5	Paramp gain and bandwidth.	90
6.6	Unusual gain profile.	90
6.7	Paramp noise measurement scheme.	92
6.8	System noise temperature with paramp.	93
6.9	Saturated regime paramp response.	94
6.10	Avoided crossing of qubit and resonator.	95
6.11	Qubit “punchout” calibration.	96
6.12	Qubit coherence measurements: Rabi, Ramsey and T_1	98
6.13	Purcell effect.	100
6.14	Quantum state tomography.	101
6.15	ac Stark shift and photon number calibration.	102
7.1	Base temperature apparatus.	108
7.2	Overview of experimental setup.	109
7.3	Individual measurement traces.	110
7.4	Jumps with varied state preparation.	111
7.5	Histograms of many individual traces.	113
7.6	Jump times and population decay.	114
7.7	Jumps with simultaneous readout and excitation.	114
7.8	Quantum jumps with and without paramp.	116
7.9	Three-level jumps.	117
7.10	Paramp biasing for three-level jumps.	118
7.11	Readout signal cartoon.	120
7.12	Readout SNR.	122
8.1	Optimal smoothing of time traces.	128
8.2	Hysteretic data thresholding.	129
8.3	Automated qubit state determination.	131
8.4	State diagram for finite bandwidth detection.	132
8.5	Expected dwell time distributions.	133
8.6	Log-likelihood in parameter space.	140
8.7	Performance of automated state extraction algorithm.	143
8.8	Spurious qubit state mixing.	145
8.9	Transmon qubit and resonator with fast flux line.	147
8.10	Spurious excitation and qubit spectroscopy.	149

8.11	Spurious excitation with coherent fast flux tone.	150
8.12	Qubit population with no fast flux tone.	151
8.13	Qubit population with added noise.	153
8.14	Spectral density of flux noise vs frequency.	154
8.15	Quantum Zeno effect.	159
9.1	Preliminary TWPA performance.	164
A.1	Roach filter stripline geometry.	183
A.2	Permittivity and permeability of Eccosorb MFS-117.	184
A.3	Millikelvin S-parameters of roach filters.	185
A.4	Thermalization measurement setup.	186
A.5	Thermalization measurement.	187
C.1	Fast flux noise generation.	193
C.2	Typical room temperature wiring.	194
C.3	Fridge wiring for Chapter 7 experiments.	195
C.4	Fridge wiring for Chapter 8 experiments.	196

List of Symbols and Abbreviations

\hat{a}, \hat{a}^\dagger	photon/harmonic oscillator annihilation and creation operators
$A_{s }, A_{s\perp}$	in-phase and quadrature amplitudes of LJPA input signal
A_0, A_{pump}	unmodulated and modulated LJPA pump amplitude
B	noise bandwidth
B_{det}	detection bandwidth
B_{LJPA}	instantaneous half-width half-maximum bandwidth of LJPA
C	capacitance
C_B	transmon shunt capacitance
C_{cav}	readout cavity capacitance
C_g	transmon gate capacitance
C_J	junction capacitance
C_Σ	total transmon capacitance
e	elementary charge
E_C	charging energy $e^2/2C_\Sigma$
E_J	Josephson energy $\hbar I_0/2e$
E_m, E_i	energy of m^{th} or i^{th} transmon level
E_{01}	transmon qubit energy $E_1 - E_0$
f	dimensionless LJPA signal detuning frequency $f = 2Q(\omega_s/\omega_{p0})$
f_q	qubit frequency
f_φ	dephasing noise in control parameter

$f_{3\text{dB}}$	filter 3 dB corner frequency
F	single-shot measurement fidelity
g	qubit-cavity coupling strength (taken to be g_{01} if no subscript is specified)
g_{ij}	qubit-cavity coupling strength for transitions between qubit states $ i\rangle$ and $ j\rangle$
G	amplifier power gain
G_s, G_i	signal and idler power gain of parametric amplifier
H_{cav}	readout cavity Hamiltonian
H_{int}	interaction Hamiltonian
H_{meter}	meter Hamiltonian
H_{qubit}, H_q	qubit Hamiltonian
\hbar	reduced Planck's constant
$ i\rangle$	i^{th} eigenstate of transmon qubit
$ 0\rangle, 1\rangle, 2\rangle$	lowest three transmon eigenstates
I	current
I_c	critical current of a two-junction SQUID
I_d	paramp drive current
I_0	critical current of a Josephson junction
k	LJPA bandwidth scaling factor
k_B	Boltzmann's constant
K_i	multilevel Kerr shift
L	inductance
L	log-likelihood function
L_J	Josephson inductance
L_{J0}	zero-current Josephson inductance
\mathcal{L}	likelihood function
m	mass
\hat{n}	Cooper pair number operator

n_g	gate charge Cooper pair offset for transmon and Cooper pair box
\bar{n}	mean cavity photon occupation
n_{cens}	number of censored observations
n_{crit}	mathematical limit of dispersive approximation $n_{\text{crit}} = \Delta^2/4g^2$
\bar{n}_e	mean cavity photon occupation, excited state
\bar{n}_g	mean cavity photon occupation, ground state
\bar{n}_{res}	mean cavity photon occupation, driven on resonance
\hat{p}	momentum operator
P_c	critical power for bifurcation of the LJPA
P_{ff}	fast flux excitation power
p_i	generalized Hamiltonian momentum
P_{out}	power output from reflection geometry resonator
P_{rad}	power radiated from resonator
Q, \hat{Q}	capacitor charge, capacitor charge operator
Q	resonator quality factor
$Q_{\text{int}}, Q_{\text{ext}}$	internal and external resonator quality factors
q_i	generalized Hamiltonian coordinate
R_n	normal state resistance of tunnel junction
$S(\omega)$	noise power spectral density
$S_{\delta\omega_q}(\omega)$	power spectral density of qubit frequency fluctuations
$S_{\Phi}^{1/2}(\omega)$	flux noise amplitude spectral density
S_i	multilevel ac Stark shift
s_{hi}, s_{lo}	high and low voltage states for qubit state extraction algorithm
SNR_{meas}	measured signal-to-noise ratio
T	temperature
T_{eff}, T_s	effective temperature of source impedance

T_H, T_L	temperature of hot and cold loads for noise calibration
$T_n, T_{n,i}$	amplifier noise temperature
T_{sys}	system noise temperature of amplification chain
T_1	qubit relaxation time
T_2	qubit dephasing time
T_φ	qubit pure dephasing time
U	energy stored in a Josephson junction
V	voltage
V_C	voltage across a capacitor
V_d	drive voltage
$V_{e,t}, V_{g,t}$	output voltages from transmission mode cavity for qubit excited and ground states
V_g	gate voltage for transmon and Cooper pair box
V_h, V_l, V_{mid}	hysteretic threshold voltages for automated qubit state extraction algorithm
$V_{\text{in}}, V_{\text{out}}$	voltages of incoming and outgoing waves from LJPA
V_J	voltage across Josephson junction
V_L	voltage across an inductor
$V_{\text{out,r}}$	output voltage from reflection mode cavity
V_s	signal voltage
$V_{\text{sig},r}, V_{\text{sig},t}$	readout signal voltage in reflection and transmission geometries
V_{rms}^0	zero-point voltage of the readout cavity $V_{\text{rms}}^0 = \sqrt{\hbar\omega_{\text{cav}}/2C_r}$
\hat{x}	position operator
Y	Y-factor of hot/cold load test
Z	resonator impedance
Z_0	microwave shunt impedance for LJPA
α	qubit anharmonicity
$\alpha, \alpha_{ 0\rangle}, \alpha_{ 1\rangle}, \alpha_{ 2\rangle}$	complex cavity amplitude and its value corresponding to qubit states $ 0\rangle$, $ 1\rangle$, and $ 2\rangle$

β	transmon capacitance ratio $\beta = C_g/C_\Sigma$
β	resultant vector between readout cavity pointer states
δ	phase difference across Josephson junction
$\hat{\delta}$	phase operator (conjugate to number operator \hat{n})
$\delta_{ }, \delta_{\perp}$	quadrature amplitudes for solutions of driven LJPA
$\tan \delta$	dielectric loss tangent
$\Delta(T)$	temperature-dependent superconducting gap
Δ	qubit cavity detuning (angular frequency): $\Delta = \omega_q - \omega_{\text{cav}}$, equivalent to Δ_0
Δ_i	qubit-cavity detuning (angular frequency) for i^{th} qubit transition $\Delta_i = (\omega_{i+1} - \omega_i) - \omega_{\text{cav}}$
γ_{purc}	Purcell relaxation rate $\gamma_{\text{purc}} = \kappa g^2/\Delta^2$
γ_q	ac-Stark-broadened qubit linewidth
$\Gamma_{\uparrow}, \Gamma_{\downarrow}$	qubit state transition rates
$\Gamma_{\uparrow\downarrow, \text{DD}}$	qubit dressed dephasing transition rates
Γ_A, Γ_B	transition rates between states A and B of a Poissonian two-level system
Γ_d	measurement-induced dephasing rate
$\Gamma_{\text{det}}, \hat{\Gamma}_{\text{det}}$	detection bandwidth and its maximum likelihood estimate
$\Gamma_{\text{dn}}, \hat{\Gamma}_{\text{dn}}$	rate of downward state transitions and its maximum likelihood estimate
Γ_{jump}	total rate of qubit state transitions
Γ_m	measurement rate
Γ_{Rabi}	Rabi decay rate
$\Gamma_{\text{up}}, \hat{\Gamma}_{\text{up}}$	rate of upward state transitions and its maximum likelihood estimate
Γ_1	qubit relaxation rate
Γ_2	qubit dephasing rate
Γ_φ	qubit low-frequency pure dephasing rate
ϵ	solution for small signal perturbation on LJPA drive

λ	parameterization variable for jump time probability distribution $\lambda = \Gamma_{\text{up}} + \Gamma_{\text{dn}} + \Gamma_{\text{det}}$
λ_i	ratio of coupling to detuning for i^{th} transmon level $\lambda_i = g_{i,i+1}/\Delta_i$
κ	cavity linewidth (angular frequency)
ν	flux-to-qubit-frequency transfer coefficient
ρ	density matrix
ρ	dimensionless LJPA drive amplitude $\rho = Q\delta_0^2/8$
ρ_{max}	drive amplitude for maximum LJPA gain at zero detuning
$\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z$	Pauli operators for qubit
$\hat{\sigma}_+, \hat{\sigma}_-$	qubit raising and lowering operators
$\theta, \hat{\theta}$	parameter of statistical distribution and its maximum likelihood estimate
θ_{up}	parameterization variable for jump time probability distribution, $\theta_{\text{up}} = \sqrt{\lambda^2 - 4\Gamma_{\text{up}}\Gamma_{\text{det}}}$
θ_{dn}	parameterization variable for jump time probability distribution $\theta_{\text{dn}} = \sqrt{\lambda^2 - 4\Gamma_{\text{dn}}\Gamma_{\text{det}}}$
ζ	cavity pull
$\Phi, \hat{\Phi}$	branch flux, branch flux operator
Φ_{app}	applied magnetic flux
Φ_0	flux quantum $h/2e$
χ	1/2 of linear dispersive shift (angular frequency)
χ_{ij}	transmon partial dispersive shift (angular frequency)
χ_{01}	transmon lamb shift (angular frequency)
$ \Psi\rangle$	general qubit state
ω_c	critical frequency for bifurcation of the LJPA
ω_{cav}	bare cavity resonant frequency
$\tilde{\omega}_{\text{cav}}$	qubit-state-dependent cavity resonant frequency
ω_d	drive frequency
ω_{ff}	fast flux excitation frequency

ω_i	frequency of i^{th} transmon state E_i/\hbar
ω_{ij}	frequency of splitting between i^{th} and j^{th} transmon states
$\tilde{\omega}_{ij}$	Lamb-shifted frequency of splitting between i^{th} and j^{th} transmon states
ω_p	pump frequency
ω_{p0}	LJPA linear resonant frequency
ω_q	qubit frequency
ω_{ro}	readout frequency
ω_s	signal detuning frequency from pump
$\omega_{\text{sig}}, \omega_{\text{id}}$	signal and idler frequencies
Ω	dimensionless drive detuning for LJPA $\Omega = 2Q(1 - \omega_d/\omega_{p0})$
$\Omega_{\text{Rabi}}, \Omega$	Rabi frequency, Rabi frequency at zero detuning
CPB	Cooper pair box
DD	dressed dephasing
JC	Jaynes-Cummings
LJPA	lumped Josephson parametric amplifier
QED	quantum electrodynamics
SNR	signal-to-noise ratio
SQUID	superconducting quantum interference device
TWPA	traveling wave parametric amplifier

Acknowledgments

There are many people who played a role in the success of this thesis, and I would like to acknowledge and thank them.

I would first like to thank my advisor, Prof. Irfan Siddiqi. Irfan has taught me an enormous amount, on topics such as the inner workings of dilution refrigerators, the nuances of microwave electronics and measurements, how to give a successful talk, and the virtues of patience and perseverance when dealing with complex experiments. Irfan created a warm and collegial atmosphere in QNL which made it a pleasure to work there, and I always felt I could come to him with scientific questions and have an interesting discussion. He personally taught me to fabricate samples and understand microwave techniques, back when the lab was very first starting. I am very grateful for the support Irfan has given me over my graduate career. I would like to thank him as well as the rest of my thesis committee for their extensive and thoughtful comments on this document.

I have been fortunate to work with a number of excellent graduate students and postdocs in my time at Berkeley, and they have enriched my experience enormously. I would like to especially thank Dr. R. Vijay and Dr. Ofer Naaman, the two postdocs with whom I worked during my time in QNL. The work in this thesis was carried out in very close collaboration with Vijay, who taught me a great deal about the subtleties of microwave experiments and the fine points of quantum measurement theory. I have enjoyed our many sessions together discussing outlandish experimental ideas and doping out the meaning of confusing or unexpected results. He has been a superb collaborator as well as a dear friend. I worked with Dr. Ofer Naaman during the earlier stages of my career at QNL. Ofer did the wiring in our dilution fridge and taught me about all of the various pitfalls associated with this difficult activity. Ofer helped me a great deal developing and measuring the roach filters. In addition to Vijay and Ofer, my work at UC Berkeley has been done in the company of many other hardworking and lively physicists, from whom I learned a great deal about physics and life. I would like to especially thank Kater Murch, Andy Schmidt, Eli Levenson-Falk, Ned Henry, Natania Antler, Steve Weber, Chris Macklin, Michael Hatridge, Sarah Busch, Jed Johnson, and Emile Hoskinson. The work in this thesis was done with the help of numerous collaborators at other institutions as well. I would like to thank Dr. Joe Aumentado at NIST for his enthusiastic assistance with the design of the SlichTECH evaporator, a project I could not have completed by myself. Dr. Jay Gambetta, Prof. Rob Schoelkopf, Prof. Michel Devoret, and Prof. Steve Girvin have all provided insightful comments and lively discussion on issues related to understanding quantum jumps. Dr. Gambetta, Prof. Alexandre Blais and Maxime Boissonneault have also been very helpful in collaborating with us on the dressed dephasing work.

I would like to thank several of my past advisors and mentors, who have been instrumental in bringing me to this point. Prof. Jene Golovchenko at Harvard took me in to his freshman seminar on experimental physics, where I first learned about the joys and frustrations of trying to do an experiment nobody else has done before. He later hired me on as an undergraduate research assistant in his lab, where I learned a great deal about carbon nanotubes, DNA sequencing, and electron microscopy. Prof. J.C. Campuzano at the University of Illinois-Chicago hired me to work in his lab when I was between college and graduate school, trying to decide if I wanted to pursue physics further. He gave me an

enormous boost of confidence, and taught me a lot about ultra-high-vacuum techniques. Prof. John Clarke at UC Berkeley first introduced me to the world of superconducting devices, and sparked my interest in high-sensitivity measurements. I will always look back fondly on my excellent first-year graduate quantum mechanics class, taught with great *élan* by Prof. Robert Littlejohn.

I am deeply grateful to the Fannie and John Hertz Foundation for the graduate fellowship they awarded me, and to Ray Sidney for endowing that fellowship. The Hertz Fellowship was a crucial aspect of my success in graduate school; beyond the simple monetary considerations, the Hertz community provided me with encouragement and support, while opening doors to a wide range of experiences and fields of scientific endeavor. I have made many lasting friendships with Hertz fellows and alumni, and I am very grateful to the Foundation for giving me the opportunities it has.

I doubt I would have made it this far without the help of the physics staff, who work tirelessly to keep everything running smoothly in the department. Special thanks go to Anne Takizawa, Donna Sakima, and Claudia Trujillo in the student services office, as well as Anthony Vitan, Eleanor Crump, and Katalin Markus in support services. These people happily helped me with problems large and small, and I am very thankful to them.

I have been lucky to have a great group of friends here in Berkeley, both from my incoming physics cohort and elsewhere. I want to thank them for making me laugh and keeping me sane for the past six and a half years. I will always have fond memories of Wednesday night dinners, Bay to Breakers, skiing in Tahoe, Big Sur trips, and many other fun adventures we shared.

Finally, I would like to thank my family, without whom I never would have made it this far. Mom, Dad, Sumner, Bill, Jacob, Ann, David, and Yolanda—I have always felt your unwavering love and support, and it has given me great strength and confidence. Thank you for your advice, your encouragement, and your enthusiasm for my work. I love you, and dedicate my thesis to all of you.

Chapter 1

Introduction

In the past 30 years, it has become possible to isolate and interrogate individual quantum systems, allowing a wide variety of experiments on fundamental aspects of quantum mechanics. Many of these experiments were initially conceived of as *Gedanken-experiments* in the early days of quantum mechanics, when the technology did not exist to carry them out in a laboratory. Questions of entanglement, “spooky action at a distance”, and the nature of quantum measurement have since transitioned from the realm of theory into the realm of experiment.

Experiments on the fundamentals of quantum mechanics require controllable quantum systems to use as test beds. Perhaps the simplest quantum system, familiar to any student of elementary quantum mechanics, is the quantum two-state system, called a quantum bit or qubit for short. The terminology comes from the word “bit”, used to describe a classical two-state system. A bit could be a flipped coin (heads or tails), a switch (on or off), or a small magnetic domain on a hard disk (aligned up or down). Crucially, a classical bit can be in only one of its two possible states at any given time.

In contrast, a qubit can exist in either of its two eigenstates, or in a superposition of those two eigenstates. In addition, if one has several coupled qubits, it is possible to form an entangled state of the qubits which cannot be written as a simple product of individual qubit states. Using n classical bits, one can represent just a single n -bit configuration out of the 2^n possible configurations. By contrast, n qubits can be placed in an entangled superposition state of all 2^n possible configurations with varying complex amplitudes, a state which can contain exponentially more information than the classical n -bit state. The enormous dimensionality of the Hilbert space for entangled qubits could theoretically be harnessed to build a quantum computer capable of solving certain classes of problems exponentially faster than any classical computer [1]. Such a system could also be used as a large-scale quantum simulator to study problems in many-body physics or to probe the quantum-classical boundary [2, 3].

There are a number of physical realizations of quantum two state systems: a spin-1/2 particle in a magnetic field, two electronic levels of an atom, or two polarizations of a photon, for example. In this thesis, we are study the quantum properties of engineered qubits made from superconducting electrical circuits.

1.1 Superconducting qubits

Superconducting qubits [4, 5] are quantum two-level systems realized in an electrical circuit by exploiting macroscopic quantum degrees of freedom. The notion that macroscopic variables could behave according to the rules of quantum mechanics was first put forward in the 1980s by Leggett [6], and subsequent experimental tests demonstrated the quantum behavior of the macroscopic phase degree of freedom of a Josephson junction [7]. This laid the groundwork for constructing qubits from superconducting Josephson circuits. In 1999, the research group at NEC in Japan observed coherent oscillations between two states of definite Cooper pair number on a submicron superconducting island [8], a design called the Cooper pair box [9]. The oscillations displayed coherence times on the order of 1 ns.

In subsequent years, many different superconducting qubit designs have been developed. Qubits based on macroscopic states of the phase of a Josephson junction [10] or the circulating currents in a superconducting loop [11], called “phase” and “flux” qubits respectively, came on the scene. An improved Cooper pair box design called the quantronium qubit (shown in Figure 1.1) exhibited reduced sensitivity to environmental noise and dramatically increased coherence times [12]. A further modified version of the Cooper pair box, called the transmon qubit, ushered in a new era of reliable qubits with greatly reduced sensitivity to charge noise [13]. Other new designs promising reduced sensitivity to various types of noise, rich energy structures, and other exotic features have arrived on the scene in the past few years, among them the fluxonium qubit [14], the capacitively shunted flux qubit [15], the tunable coupling qubit [16], and the 3D transmon qubit [17]. In the twelve years since the original NEC experiments, coherence times have improved by more than four orders of magnitude; the 3D transmon design currently boasts typical coherence times approaching 50 μ s.

As the coherence times of superconducting qubits improved, it became possible to entangle them and use them to perform quantum algorithms. Several groups have performed experiments demonstrating entanglement among multiple superconducting qubits [18, 19, 20] and violations of Bell’s inequality [21, 22, 23], providing strong evidence of the inherently quantum nature of superconducting qubits. Superconducting qubits are easy to manipulate, tunable, and can be mass produced, opening avenues for creating large-scale engineered quantum systems. Work is underway to build scalable architectures for implementing a quantum computer using superconducting circuits [22, 24, 25].

Although superconducting qubit technology has advanced by leaps and bounds in the past decade, it still lags many competing model quantum systems, such as trapped ions, on issues of measurement. Until the work presented in this thesis, there had been no measurement system for superconducting qubits which allowed continuous high-fidelity monitoring of the qubit state. Without this capability, it had been difficult or impossible to study the dynamics of superconducting qubits under measurement. In particular, it was impossible to observe one of the most basic manifestations of the effects of quantum measurement, namely quantum jumps.

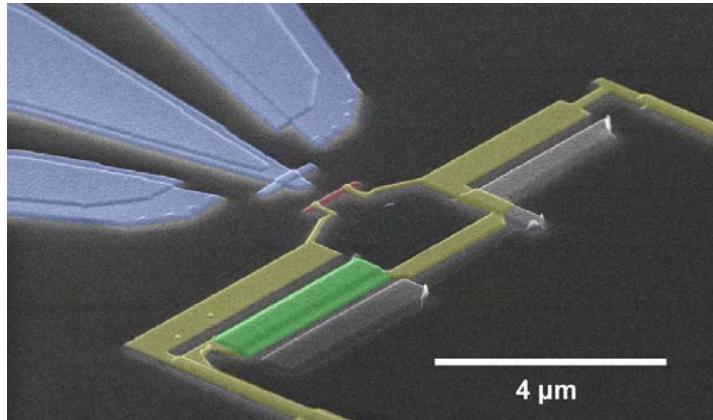


Figure 1.1: The quantronium qubit.

This false color image shows a quantronium qubit, a modified Cooper pair box whose qubit energy can be tuned to a noise-insensitive “sweet spot”. The Cooper pair box is shown in red, with qubit control implemented by applying voltages to the gate electrode (blue) and magnetic fluxes to the qubit loop (yellow). A large Josephson junction used for readout is shown in green.

1.2 Quantum jumps

The notion that quantum systems can evolve by “jumping” abruptly between eigenstates was first proposed by Bohr almost a century ago in 1913 [26]. For three quarters of a century, the concept of quantum jumps remained a purely theoretical curiosity, and a subject of substantial debate. By the early 1980s, though, advances in atomic physics allowed for the trapping and cooling of single ions [27, 28], perhaps giving a chance to settle the question of quantum jumps experimentally. Only a few years later, in 1986, three groups simultaneously reported the observation of quantum jumps between the electronic states of individual trapped ions [29, 30, 31]. Data from [29] are shown in Figure 1.2.

Since that time, quantum jumps have been observed in a variety of other systems, starting with the electronic states of single molecules embedded in a crystal [32]. It was found that single electrons in cyclotron orbits undergo quantum jumps between Landau levels [33], and that single microwave photons are suddenly created and annihilated by thermal processes inside a Fabry-Perot cavity [34]. In solid state systems, quantum jumps have been observed in a microscopic defect in a Josephson junction [35], while more recently, work showed that the state of a single nuclear spin in a diamond NV center undergoes quantum jumps [36], as does the electronic state of an electron in an indium gallium arsenide quantum dot [37]. Quantum jumps from spin flips of a single trapped proton in a magnetic field were also reported recently, paving the way for precision tests of matter-antimatter asymmetry [38].

The observation of quantum jumps requires a quantum non-demolition (QND) measurement scheme, that is, one which leaves the system in an eigenstate of the measured

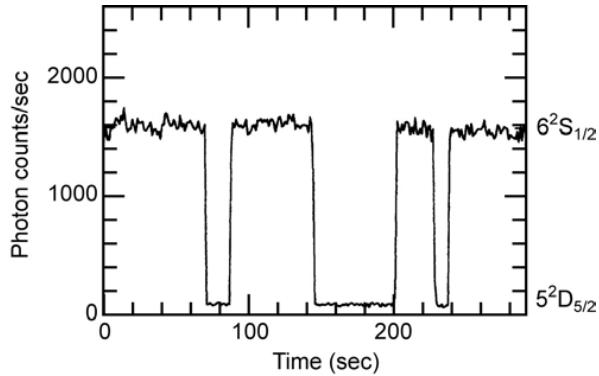


Figure 1.2: The first observation of quantum jumps.

This figure shows quantum jumps between two electronic states of a single Ba^+ ion as determined from the fluorescence of a fast cycling transition between one of the two jump states and a third state. Note the horizontal time axis. Figure adapted from ref. [29].

observable [39], thus allowing repeated measurements. One must also be able to perform the measurements on a timescale much faster than that of the qubit dynamics in order to resolve the jumps. All of the previous quantum jump experiments mentioned above have used microscopic quantum degrees of freedom with long relaxation times ($\sim \text{ms}$ to s), enabling the observation of quantum jumps even with low-bandwidth measurement techniques.

Superconducting qubits typically have much shorter relaxation times ($\sim \mu\text{s}$) on account of strong coupling to their environment, so the observation of quantum jumps requires substantially higher measurement bandwidth than in previous quantum jump experiments. Fortunately, a technique capable of continuous, high-bandwidth readout of a superconducting qubit has been developed. The technique is known as circuit quantum electrodynamics (circuit QED), and involves dispersively coupling the qubit to a microwave cavity [40]. Circuit QED was developed in direct analogy to the techniques of cavity quantum electrodynamics (cavity QED), a technique used in optics to explore the coupling between atoms and the electromagnetic field in a Fabry-Perot cavity. In the dispersive regime, where the frequency of the qubit and the cavity are far detuned relative to the strength of their coupling, the resonant frequency of the cavity depends on the state of the qubit. Probing the cavity frequency implements a continuous, high visibility QND measurement whose bandwidth is only limited by the cavity linewidth [41].

Despite successfully demonstrating QND measurement with several kinds of superconducting qubits [14, 42, 43], circuit QED implementations with linear cavities have typically suffered from low single-shot fidelity¹, precluding the observation of quantum jumps. This is primarily due to inefficient amplification of the readout photons leaving the cavity carrying information about the qubit state. The noise added by state-of-the-art cryogenic semiconductor microwave amplifiers is considerably larger than the signal from the cavity, necessitating repeated state preparation/measurement experiments to resolve the

¹Single-shot fidelity describes the ability to resolve the qubit state faithfully in a single experiment.

qubit state [44]. Using a stronger readout tone to detect the cavity frequency can induce qubit state mixing [45], thus limiting the fidelity.

Other high fidelity readout schemes exist for superconducting qubits, but they are not suitable for observing quantum jumps. Latching readouts based on the dynamical bifurcation of a nonlinear oscillator can provide the desired QND measurement [46, 47], but the averaging time required to distinguish the qubit state is sufficiently long compared to the qubit relaxation time that one cannot resolve quantum jump events. Hysteretic switching readouts involving current-biased Josephson junctions transitioning to the voltage state can have fidelities better than 90%, but they either destroy the qubit state [48] or again require too long a reset time for use in measuring quantum jumps [49].

The single-shot fidelity problems of linear cavity circuit QED could be alleviated if the added noise of the following amplifiers were considerably lower, but no commercial solution to this problem currently exists. We instead turn to a class of amplifiers called superconducting parametric amplifiers, long a topic of research but recently the subject of redoubled efforts driven by the need for low noise amplification of quantum signals.

1.3 Superconducting parametric amplifiers

Parametric amplifiers achieve gain by varying a parameter of the amplifier system harmonically in time. The energy used to vary the parameter is called the pump. The modulation of the system parameters at the pump frequency causes some of the pump energy to be transferred into other frequency modes, amplifying signals present in those modes. Because parametric amplification can occur without any dissipation, it has the potential to have quantum-limited noise performance [50].

Superconducting parametric amplifiers based on the nonlinear Josephson inductance were first demonstrated in 1975 [51]. A number of superconducting parametric amplifier designs followed in the 1980s and 1990s [52, 53, 54, 55, 56], but they tended to be plagued by spurious “noise rise” and were not used broadly in applications. However, the rapid progress in solid-state quantum measurement and quantum information of the past decade has renewed interest in superconducting parametric amplifiers. A number of recent results [57, 58, 59, 60] have demonstrated quantum-limited noise performance and high gain using new amplifier designs.

One of the main limiting factors on the use of parametric amplifiers has been their signal bandwidth, which was typically of order 1 MHz or less. The most recent designs, including the one discussed in this thesis, have instantaneous bandwidth on the order of 10 MHz or greater, making them suitable for amplifying qubit readout signals [59, 60, 61]. The key is to lower the quality factor Q of the amplifier resonator, which allows larger bandwidth without sacrificing gain.

With this large bandwidth, quantum-limited noise, and high gain, we finally have the device we need to amplify the circuit QED readout signal. The stage is set for continuous monitoring of the state of a superconducting qubit and the observation of quantum jumps.

1.4 Thesis overview

The thesis begins with a brief introduction to superconducting qubits in Chapter 2, giving some information on the quantization of electrical circuits and the Hamiltonian for our qubit of choice (the transmon) and its solutions. It then describes the circuit QED system, how this system can be used to realize a QND measurement of the qubit state, and how the readout signal manifests itself.

Chapter 3 discusses basic ideas of amplification and noise, including a derivation of the quantum limit for the noise performance of an amplifier. The chapter goes on to describe some basic features of parametric amplification, followed by some detailed theory of our particular realization of superconducting parametric amplifier, called the Lumped Josephson Parametric Amplifier or LJPA. We provide mathematical expressions for the performance as well as a physical description of the amplification mechanism.

To close the first half of the thesis, Chapters 4 and 5 give details on the fabrication of the qubit and amplifier samples, as well as the experimental methods used to interrogate the samples.

The second half of the thesis, Chapters 6 to 8, consists of experimental results. First, we detail the experimental performance of the LJPA in Chapter 6, describing the procedures for correctly biasing the amplifier and giving measurements of its gain and noise. We also present a series of tests and calibrations used to determine important properties of the qubit/cavity system such as the qubit energy structure and coherence times, the qubit-cavity coupling, and the measurement strength.

In Chapter 7, we present some theoretical and historical background on quantum jumps before detailing the experiments carried out to observe them in a superconducting qubit. The chapter contains information on a variety of test experiments used to demonstrate that the jumps seen are in fact those of the superconducting qubit, and closes with a discussion of the signal-to-noise ratio and measurement fidelity achieved with our readout technique.

The penultimate chapter, Chapter 8, describes the experiments done to examine measurement backaction. It begins with a description of the automated qubit state extraction algorithm used to process individual time trace data into useful ensemble metrics. The phenomenon of spurious qubit state mixing during measurement is detailed and a theoretical framework is presented to explain this behavior. We then show experimental data supporting the theory, finding that flux noise well below the qubit Larmor frequency is the root cause of the observed spurious state mixing. The experiment allows us to measure the spectral density of flux noise at frequencies around 1 GHz. Finally, we show data illustrating the quantum Zeno effect, where the presence of strong measurement inhibits the evolution of the qubit.

After discussions of potential future work and extensions of the thesis, including quantum error correction and feedback, single-photon sources and detectors, and improved qubits and parametric amplifiers, we give some brief conclusions in Chapter 9.

1.5 Summary of key results

The work presented in this thesis represents the first observation of quantum jumps in a macroscopic quantum degree of freedom, a superconducting qubit [62]. It provides further evidence of the quantum nature of superconducting qubits, which had been shown previously through entanglement and violation of Bell's inequality. A variety of tests show that the observed jumps are indeed quantum jumps of the qubit. The experiment relies crucially on the high bandwidth and quantum-limited noise performance of the LJPA.

With the capability to perform continuous qubit state monitoring, we look at qubit dynamics during measurement using state extraction from many single-shot time traces. We develop mathematical methods to estimate the qubit state accurately from individual measurement traces and retrieve both single-shot and ensemble dynamics in presence of noise and bandwidth limitations. Using these methods, we examine the backaction of measurement in two phenomena. The first phenomenon, spurious qubit state mixing, is found to result from the upconversion of low-frequency flux noise to noise at the qubit frequency due to the presence of measurement photons. We use the measurements of spurious excitation to estimate the flux noise power spectral density at frequencies around 1 GHz, a region previously inaccessible to experiments, and show that the spectrum of flux noise appears to obey a $1/f^\alpha$ power law over 11 decades in frequency. The second phenomenon displaying measurement backaction is the quantum Zeno effect, where the presence of strong measurement inhibits qubit state evolution. We show that our qubit system exhibits quantum Zeno behavior with scalings roughly commensurate with those predicted by theory.

Chapter 2

Superconducting qubits and circuit quantum electrodynamics

This chapter is intended to give an introduction to superconducting qubits, in particular the transmon qubit, and to describe basic principles and phenomena of circuit quantum electrodynamics (circuit QED). We will demonstrate how circuit QED can be used to measure the state of a qubit, and how to extract the theoretical parameters of a qubit-cavity system from experimental measurements.

Superconducting qubits and circuit QED have been the subjects of a great deal of work in recent years, and there are a number of excellent sources for finding information in greater detail than presented here. For more on superconducting qubits, the reader is directed to several recent review articles by Siddiqi [4], Clarke and Wilhelm [5], and You and Nori [63]. A nice introduction to circuit QED is presented in Dave Schuster's thesis [64], and subsequent theses from the Yale group detail various extensions to multiple qubits and microwave quantum optics.

2.1 Qubits and quantum information

A qubit, short for “quantum bit”, is a quantum two-state system. This terminology comes from the word “bit”, used to described a classical two-state system and commonly used in information theory and computer science. A bit could be a flipped coin (heads or tails), a switch (on or off), or a small magnetic domain on a hard disk (aligned up or down). A bit can be in only one of its two possible states at any given time. A series of n classical bits has 2^n possible configurations, but can only be in one of those configurations at a given time.

Qubits, on the other hand, can exist in either of its two eigenstates, or in a superposition of those two eigenstates. In addition, if one has multiple coupled qubits, it is possible to form states which cannot be written as product states of the individual qubits. For example, given two qubits labeled a and b , the two-qubit state

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle_a|0\rangle_b + |1\rangle_a|1\rangle_b) \quad (2.1)$$

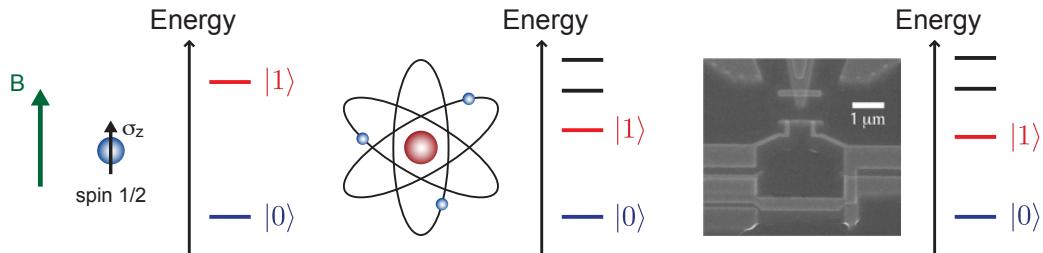


Figure 2.1: Quantum two-level systems.

The canonical qubit is a spin-1/2 particle in a magnetic field, with the energy level spacing given by the Zeeman splitting. Any quantum two-state system can be mapped on to this model. Systems with more than two quantum states can act as qubits if they are restricted to be in a two-level subspace. In this way, qubits can be realized using two electronic energy levels of a single atom or two energy levels of a quantized electrical circuit.

cannot be written as a product of states of a and b in isolation. This type of non-separable state is called an entangled state¹. Given n qubits, one can create entangled states consisting of linear combinations of all 2^n possible n -bit configurations, with a different complex amplitude for each of the 2^n terms. Such a state can contain exponentially more information than a classical n -bit state. By realizing a computer relying on qubits rather than classical bits, it may be possible to harness this exponentially large state space to carry out certain computations at speeds far exceeding that of even the most powerful classical computer [1]. Quantum computers have been theoretically shown to provide an exponential speedup for the prime factorization of large numbers [65], as well as a quadratic speedup for search on unsorted data [66].

Qubits are also interesting because they are one of the simplest test beds for studying the behavior of quantum systems. The classic example of a quantum two-state system, a spin-1/2 particle in a magnetic field, is a ubiquitous pedagogical tool in quantum mechanics textbooks. Many of the unexpected features of quantum mechanics are first laid out for students in *Gedankenexperiments* involving a Stern-Gerlach apparatus and a beam of spin-1/2 particles. However, for decades such experiments were destined to remain purely theoretical.

In the past 30 years, it has become possible to isolate and interrogate individual quantum systems, allowing a wide variety of experiments on fundamental aspects of quantum mechanics. A great variety of quantum two-level systems have been used as test beds in these efforts. Three examples are shown in Figure 2.1: a spin-1/2 particle in a magnetic field, two electronic levels of an atom, and two macroscopic quantum states of an electrical circuit. It is on this third type of qubit that this thesis will focus. In the next sections, we describe how an electrical circuit can be thought of as a qubit.

¹The particular entangled state $|\Phi^+\rangle$ given here is a type of maximally entangled two-qubit state called a Bell state.

2.1.1 Quantization of an electrical circuit

If we would like to make a qubit from an electrical circuit, we must describe the circuit in quantum mechanical terms. In this section, we will demonstrate how to quantize an electrical circuit using the example of an LC resonator, one of the simplest electrical circuits. The classical variables used to describe the classical LC resonator will be transformed into operators in the quantum version. This derivation and its attendant methods are described in thorough and rigorous detail in refs. [67, 68].

In a lossless parallel LC resonator, the voltages across the inductor and capacitor are equal, by Kirchoff's laws. The voltage can be expressed either in terms of the inductance or the capacitance (we adopt the sign convention that the positive current direction and positive voltage direction across a component are opposite, and choose relative voltage orientations such that $V_L = V_C$):

$$V_C = Q/C, \quad V_L = LI. \quad (2.2)$$

Here Q is the charge on the capacitor and I is the current through the inductor, which because of our choice of sign convention is given by $I = -\dot{Q}$, where the dot denotes differentiation with respect to time. For reasons that will become clear, we choose to define a quantity for the inductor called the branch flux Φ , which is given by the time integral of the voltage across the inductor:

$$\Phi = \int_{-\infty}^t V(t') dt' = LI. \quad (2.3)$$

This voltage integral uses the boundary condition that all voltages and currents are zero at time $t = -\infty$. We can write the energy stored in the LC resonant circuit as:

$$E = \frac{1}{2}LI^2 + \frac{1}{2}CV_C^2, \quad (2.4)$$

which we can express as a Hamiltonian using the branch flux defined above, giving the form

$$H = \frac{\Phi^2}{2L} + \frac{Q^2}{2C}. \quad (2.5)$$

A Hamiltonian is parameterized by generalized coordinates q_i and momenta p_i , which are said to be “canonical” if they obey the relation given by the Hamilton equations:

$$\frac{\partial H}{\partial q_i} = -\dot{p}_i \quad (2.6)$$

$$\frac{\partial H}{\partial p_i} = \dot{q}_i. \quad (2.7)$$

Applying these expressions to (2.5), we can see that the branch flux Φ and charge Q act as canonical coordinate and momentum variables respectively:

$$\frac{\partial H}{\partial \Phi} = \Phi/L = I = -\dot{Q} \quad (2.8)$$

$$\frac{\partial H}{\partial Q} = Q/C = L\dot{I} = \dot{\Phi}, \quad (2.9)$$

where we have used the relations from (2.2) and (2.3). Since Φ and Q are canonical variables, we can quantize them by converting them to quantum mechanical operators $\hat{\Phi}$ and \hat{Q} , which will obey the commutation relation

$$[\hat{\Phi}, \hat{Q}] = i\hbar. \quad (2.10)$$

Having quantized the Hamiltonian of this LC resonator, we can map it to the Hamiltonian of a particle moving in a harmonic potential, which is given by

$$H = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2. \quad (2.11)$$

If we take equation (2.5) and make the correspondences $\hat{Q} \rightarrow \hat{p}$, $\hat{\Phi} \rightarrow \hat{x}$, $m \rightarrow C$, and $1/\sqrt{LC} \rightarrow \omega$, we find that it maps directly on to equation (2.11), showing that an LC circuit can be quantized as a harmonic oscillator. It is usual to express the Hamiltonian of a quantum harmonic oscillator in terms of dimensionless annihilation and creation operators:

$$H = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2}), \quad (2.12)$$

where the annihilation operator is given by

$$\hat{a} = \frac{1}{\sqrt{2\hbar Z}}(\hat{\Phi} + iZ\hat{Q}), \quad (2.13)$$

using the definition of the resonator impedance $Z = \sqrt{L/C}$. The expected value of the LC circuit's energy is just equal to the number of photons at the resonant frequency multiplied by the energy per photon, plus a half-photon zero-point term. Our model is for a lossless LC circuit in isolation, whereas a real physical system will be coupled to other degrees of freedom and sources of dissipation. We can model dissipation in a natural way using input-output theory by coupling the LC circuit to an external continuum bath of harmonic oscillator modes. We will not discuss this here, but instead direct the reader to references [69] and [70], both of which give clear and detailed accounts on the subject.

2.1.2 Superconducting qubits

Having shown how to quantize an electrical circuit, we can now examine how to make an electrical circuit into a qubit². A harmonic oscillator at low temperatures such that $\hbar\omega \gg k_B T$ will be primarily in the ground state, with little to no population in its excited states³. If we could change the state back and forth between the ground state and first excited state by applying radiation at frequency ω , we would have a qubit. Unfortunately, the states of a harmonic oscillator are all evenly spaced in energy, so any radiation that

²A somewhat more detailed description of the following arguments can be found in ref. [71].

³Technically, we mean that the density matrix of the harmonic oscillator is such that the probability of measuring it to be in a state other than the ground state is very small.

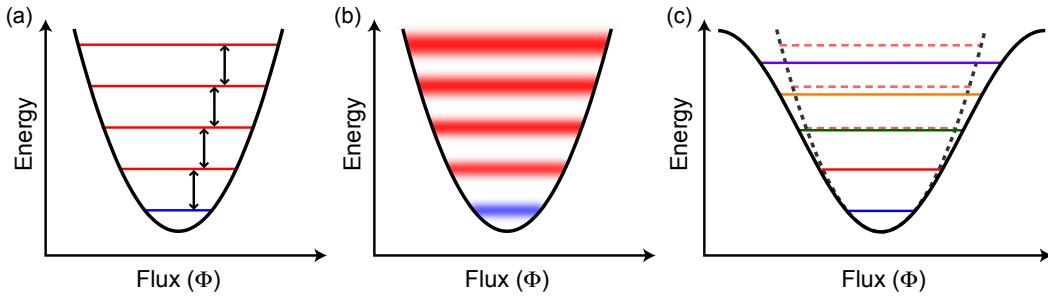


Figure 2.2: Energy levels of quantized oscillators.

Panel (a) shows the potential (black) as a function of branch flux for a lossless harmonic oscillator. The energy levels (red and blue) are equally spaced. If we add loss to the harmonic oscillator, the levels broaden as shown in (b). An anharmonic oscillator with a non-parabolic potential, such as the softening potential in (c), has unequally spaced energy levels. The potential and energy levels of a harmonic oscillator with the same lowest-level energy spacing are shown as dashed lines to highlight the uneven level spacing of the anharmonic oscillator.

drives a transition between the ground and first excited states would also drive transitions between all of the other levels as well, as shown in Figure 2.2(a). This means that the harmonic oscillator is unsuitable for use as a qubit; we require instead an electrical circuit with uneven spacing between energy levels, an anharmonic oscillator.

We should also consider the presence of dissipation. Real electrical circuits have loss, which we have so far ignored in our calculations. Loss causes decay of the quantum state of the circuit, so we would like to be able to minimize it. In circuit language, loss occurs in the presence of a real (rather than an imaginary) shunt impedance, such as a resistance. One way to create circuits with low loss is to make them from superconducting elements, which have zero resistance in the limit of zero frequency.

We have determined that in order to make a qubit out of an electrical circuit, we should create an anharmonic (nonlinear) oscillator which operates at low temperature ($\hbar\omega \gg k_B T$) and with low loss. One would ideally make such a circuit using superconductors. Fortunately, a lossless nonlinear circuit element exists for superconductors: the Josephson junction. A Josephson junction is composed of two superconductors coupled by a weak link, in our case a thin insulating barrier⁴. The junction can support a dissipationless supercurrent of any magnitude up to a certain limit called the critical current, which depends on the geometry of the junction and on the superconducting material. We can describe the behavior of a Josephson junction with a pair of constitutive relations known as the Josephson relations [72]:

$$I = I_0 \sin \delta \quad (2.14)$$

⁴A Josephson weak link can also be made using a superconducting constriction or a layer of non-superconducting metal [72].

$$V = \frac{\Phi_0}{2\pi} \dot{\delta}. \quad (2.15)$$

Here I is the supercurrent through the Josephson junction, I_0 is the critical current, V is the voltage across the junction, Φ_0 is the flux quantum $h/2e$, and δ is the difference between the phases of the superconducting order parameter on each side of the junction. Even though δ depends on the behavior of many individual electrons, it has been shown experimentally to behave as a single macroscopic quantum degree of freedom [7].

If we combine the two Josephson relations, we find the junction has a relation between the voltage and the time derivative of the current, in other words, an inductance. The Josephson inductance is given by

$$L_J = \frac{\Phi_0}{2\pi I_0 \cos \delta} \equiv \frac{L_{J0}}{\cos \delta}, \quad (2.16)$$

defining $L_{J0} \equiv \Phi_0/2\pi I_0$ as the Josephson inductance in the absence of supercurrent flow. We can use (2.14) and trigonometric identities to express the Josephson inductance as

$$L_J = \frac{L_{J0}}{\sqrt{1 - (I/I_0)^2}}. \quad (2.17)$$

This form highlights the dependence of the Josephson inductance on the current in the junction. This nonlinear inductance can be harnessed to make an anharmonic oscillator.

One can construct a tunable Josephson inductance using a superconducting loop interrupted by two junctions. This configuration is called a dc Superconducting QUantum Interference Device, or dc SQUID [73]. This nomenclature exists for historical reasons; there is nothing intrinsically low-frequency about a dc SQUID, which can happily exhibit dynamics at tens of GHz. In the presence of a magnetic flux inside the SQUID loop, a circulating current is set up which reduces the effective critical current of the SQUID. In the limit where the loop inductance L is much smaller than the Josephson inductances L_{J0} of the individual junctions, the critical current I_c of the SQUID can be expressed in terms of the applied magnetic flux bias Φ_{app} as [73]

$$I_c(\Phi_{app}) = 2I_0 \left| \cos \left(\frac{\pi \Phi_{app}}{\Phi_0} \right) \right|, \quad (2.18)$$

where I_0 is the critical current of one junction (the two junctions are assumed to be identical). The SQUID thus behaves as though it were a single junction with a flux-tunable critical current, allowing us to tune the Josephson inductance *in situ* by applying an external magnetic field.

A tunnel junction also has some capacitance in parallel with its inductance (one can think of a junction as a parallel plate capacitor, with the insulating tunnel barrier serving as the dielectric). As a result, a Josephson junction by itself is already a nonlinear oscillator. We can choose to shunt the junction with additional linear capacitance and/or inductance, as shown in Figure 2.3, to change the properties of this nonlinear oscillator to some desired regime. All superconducting qubits can be modeled in this way, as a junction with capacitance and nonlinear inductance shunted by external capacitance and/or

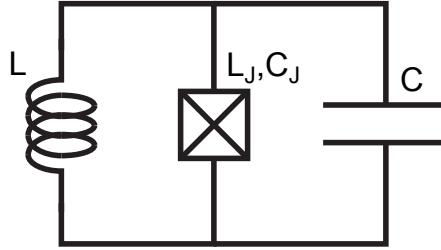


Figure 2.3: Generalized superconducting qubit.

The generalized equivalent circuit for a superconducting qubit consists of a Josephson junction, with its capacitance C_J and nonlinear Josephson inductance L_J , shunted by additional linear inductance L and capacitance C . The various types of superconducting qubits are

inductance. Charge, phase, and flux qubits, as well as variants such as the transmon, fluxonium, and capacitively shunted flux qubit, can all be seen as different limits of this protean circuit.

To find the energy levels of a qubit, we need to write down its Hamiltonian. We can write down an energy associated with the junction, assuming $\delta = 0$ at time $t = -\infty$ and integrating the power in the junction over time until we are at phase δ at time t . The power is just IV , so our energy is

$$U = \int_{-\infty}^t I(t')V(t')dt', \quad (2.19)$$

or substituting in the Josephson relations,

$$U = \int_{-\infty}^t \frac{I_0\Phi_0}{2\pi} \sin \delta \frac{\partial \delta}{\partial t'} dt'. \quad (2.20)$$

If we define the Josephson energy scale $E_J = \Phi_0 I_0 / 2\pi = \hbar I_0 / 2e$, we can rewrite the energy in the junction as:

$$U = E_J(1 - \cos \delta). \quad (2.21)$$

Since δ is a macroscopic quantum degree of freedom, we can turn this expression into a quantum Hamiltonian by replacing δ with the operator $\hat{\delta}$.

2.1.3 The transmon qubit

We now examine the specific case of the transmon qubit, which is a nonlinear LC resonator consisting of a Josephson junction shunted by an external capacitor (this is the “universal qubit circuit” shown above with the linear inductance removed, or taken to be $L \rightarrow \infty$). A full, detailed account of the transmon qubit is found in ref. [13]. The transmon is depicted in Figure 2.4. We can write down its Hamiltonian in terms of the

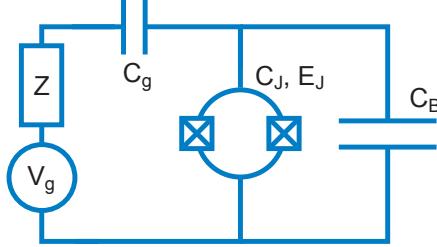


Figure 2.4: Schematic of transmon qubit.

The transmon qubit consists of a SQUID loop, characterized by a Josephson energy E_J and a parallel junction capacitance C_J , shunted by an external capacitance C_B and coupled to the external circuit voltage V_g and shunt impedance Z through a coupling capacitance C_g . The value of E_J can be tuned by applying a magnetic flux bias to the SQUID loop.

energy scales $E_J = \hbar I_c / 2e$, where I_c is the flux-dependent critical current of the SQUID given in equation (2.18), and $E_C = e^2 / 2C_\Sigma$, where C_Σ is the total capacitance shunting the Josephson inductance. In both cases, we use the convention that e is the elementary charge; the reader should be aware that some references use the convention that e is the Cooper pair charge, twice the elementary charge. The shunt capacitance C_Σ is given by $C_\Sigma = C_J + C_B + C_g$, where C_J is the junction capacitance, C_B is the external capacitance shunting the junction and C_g is the coupling capacitance to the external circuit. Using these definitions, the Hamiltonian can be written as

$$H = 4E_C(\hat{n} - n_g)^2 - E_J \cos \hat{\delta}, \quad (2.22)$$

where we have left out the constant offset term from (2.21). The two operators here are \hat{n} , the number of Cooper pairs transferred through the junction, and $\hat{\delta}$, the junction phase, while n_g is a parameter representing a voltage bias V_g from the external circuit in units of Cooper pairs ($V_g = 2en_g/C_g$). Note that the flux bias in the SQUID loop is implicit in the definition of E_J .

The operators \hat{n} and $\hat{\delta}$ are canonically conjugate, analogous to \hat{Q} and $\hat{\Phi}$ for the harmonic oscillator circuit described above, with the commutation relation

$$[\hat{\delta}, \hat{n}] = i \quad (2.23)$$

The relationship between $\hat{\delta}$ and \hat{n} is very interesting; each operator can be expanded as a Fourier series in the basis of the other operator. The fact that \hat{n} , representing the number of Cooper pairs tunneling through the junction, is constrained to have integer values causes $\hat{\delta}$ to be a periodic function; it is therefore the discreteness of the Cooper pairs tunneling through the junction which gives rise to the nonlinearity of the Josephson junction. These ideas are presented in more detail in [68].

The Hamiltonian (2.22) can be solved exactly in the phase basis representation

(where $\hat{n} = -i \frac{\partial}{\partial \delta}$) in terms of Mathieu functions, as detailed in [74]⁵. The eigenenergies are then given by [13]

$$E_m(n_g) = E_C a_{2[n_g+k(m,n_g)]}(-E_J/2E_C), \quad (2.24)$$

where $a_r(q)$ is Mathieu's characteristic value, $m \in \{0, 1, 2, \dots\}$ is the index of the eigenstates, and $k(m, n_g)$ is a function for appropriately sorting eigenvalues:

$$k(m, n_g) = \sum_{\ell=\pm 1} [\text{int}(2n_g + \ell/2) \bmod 2] \times \left(\text{int}(n_g) + \ell(-1)^m [(m+1) \bmod 2] \right). \quad (2.25)$$

In this expression, $\text{int}(x)$ rounds to the integer closest to x , $a \bmod b$ is the modulo operation, and $a \bmod b$ is the integer quotient of a and b . Mathieu's characteristic value $a_r(q)$ is implemented as a built-in Mathematica function, making the eigenenergies straightforward to evaluate.

These solutions are general for this circuit, regardless of the values of E_J and E_C ; it is the particular choice of E_J and E_C that determines the type of qubit which the circuit represents. We plot the energy spectrum versus n_g in Figure 2.5 for several values of the ratio E_J/E_C . The qubit transition energy is the simply the energy spacing between the two lowest-lying levels. In the limit $E_J \approx E_C$, this circuit is called a split Cooper pair box qubit [74], and the eigenenergies of the Hamiltonian, shown in Figure 2.5(a), are essentially those of the charge eigenstates (parabolic with n_g), with the tunnel coupling E_J lifting the degeneracy where the levels intersect. The energy levels of the Cooper pair box depend strongly on n_g , so noise in n_g caused by the motion of charged defects in the qubit substrate (known as charge noise) is problematic for qubit coherence [64]. At special charge bias values where n_g is a half-integer, the qubit energy splitting $E_{01} = E_1 - E_0$ is insensitive to n_g (and thus charge noise) to first order. This bias point is called the “charge sweet spot”.

The transmon qubit sidesteps the problem of charge noise by expanding this “charge sweet spot”. In the transmon limit $E_J \gg E_C$, usually reached by making C_B large, the energy levels of (2.22) become very flat with respect to n_g , as seen in Figure 2.5(d). The charge dispersion of the transmon levels actually decreases exponentially in the ratio E_J/E_C , so that by the time one reaches $E_J/E_C = 100$ the ground state energy varies by less than a part in 10^{-8} over the whole range of n_g , giving us a “charge sweet spot” everywhere. The remarkable feature of this reduced sensitivity to charge noise with increasing E_J/E_C is that it is not accompanied by a decrease in coupling to charge-based external control fields (this phenomenon is described extensively in [13]), affording us a qubit which we can control with capacitively coupled excitations but which does not suffer from charge-noise-related decoherence. Another nice feature of the transmon is that the external shunting capacitor C_B can be designed to have very low loss, improving the qubit relaxation time. Transmon qubits hold the current record for the longest typical qubit relaxation times ($T_1 = 20 - 60 \mu\text{s}$), with correspondingly long dephasing times ($T_2 = 6 - 20 \mu\text{s}$) [17]; only one other qubit sample in history, a Cooper pair box, has shown a longer reported relaxation time of $200 \mu\text{s}$ [75].

⁵The reader should be aware that ref. [74] uses a different convention for E_C , which differs from ours by a factor of 4.

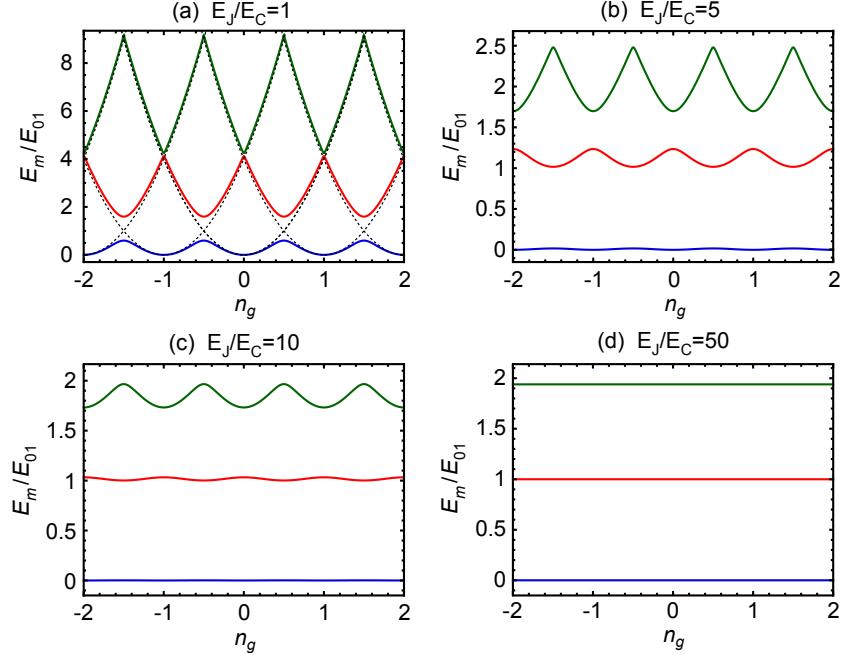


Figure 2.5: Energy structure of the Cooper pair box and transmon qubits.

This figure shows the lowest three energy levels of equation (2.24) as a function of the ratio E_J/E_C and the charge bias n_g . The energies are normalized to the value of $E_{01} = E_1 - E_0$ at the degeneracy point $n_g = 1/2$. For $E_J/E_C = 1$, the levels are close approximations to those of the uncoupled charge eigenstates (black dotted lines), with degeneracies lifted by the tunnel coupling E_J . This is the energy structure of the Cooper pair box. As E_J/E_C becomes larger, the energy levels become flatter with respect to n_g and exhibit reduced anharmonicity. By $E_J/E_C = 50$, we are in the transmon regime, where the energy levels are essentially independent of n_g . The eigenstates here are a linear combination of many uncoupled charge states. Adapted from ref. [13].

The price we pay for insensitivity to charge noise is reduced qubit anharmonicity. The Cooper pair box levels in Figure 2.5(a) are very anharmonic, but the transmon levels in (d) are almost equally spaced. As a rough rule of thumb, the qubit frequency of a transmon is given by $\sqrt{8E_JE_C} - E_C$, while the anharmonicity is about equal to E_C [13]. The reduced anharmonicity of the transmon has two major impacts. First, it means that we cannot perform qubit state manipulations too quickly, lest we accidentally excite higher energy levels. This problem can be ameliorated somewhat through the use of specially shaped qubit pulses [76]. The second impact is that the higher levels of the transmon must be accounted for when calculating readout parameters such as dispersive shifts. We will discuss and quantify this in detail in section 2.3.3.

2.2 Quantum non-demolition measurement

Having described our quantum systems of interest, superconducting qubits, we need to develop a method to measure their quantum states. For this we will use circuit quantum electrodynamics, or circuit QED, described in the next section. Before we delve into circuit QED, though, we need to examine the notion of quantum measurement and understand what it is we seek in a measurement technique.

Quantum measurement is a process by which the information in a quantum state is mapped to the state of a macroscopic meter which can then be read out classically. We can describe this in terms of a Hamiltonian for a qubit, meter, and their interaction:

$$H = H_{\text{qubit}} + H_{\text{meter}} + H_{\text{int}}. \quad (2.26)$$

Ideally one would like to be able to turn the interaction on and off (i.e. be able to make $H_{\text{int}} = 0$ when desired). When the interaction is off, the qubit can undergo free quantum evolution, and then when the interaction is turned on at the desired point the information in the qubit will be mapped into the meter where we can access it.

The textbook quantum measurement simply projects the state of the system under measurement into one of the eigenstates of the measured observable with probabilities given by the state of the system before measurement. This sort of idealized measurement is known as a projective quantum non-demolition or QND measurement [39, 77]. A projective QND measurement leaves the system in the measured eigenstate after it has completed, such that repeated measurements will give identical results (modulo the evolution of the quantum system between measurements). One can compare to a non-QND measurement, where the quantum system is left in a state other than the measured eigenstate upon completion of the measurement; here, the result of a subsequent measurement may or may not be correlated with the result of the first measurement⁶. QND measurement need not be projective; it is possible to make a weak QND measurement, where one receives only partial information about the measured observable and the system is only nudged somewhat towards the corresponding eigenstate. Projective QND measurement can be thought of as a limit of many sequential weak QND measurements [50].

Whenever quantum measurement occurs, it is accompanied by backaction on the quantum system under study. The very act of projecting the system into an eigenstate of the measured observable is a form of backaction, but it is one that we desire during measurement. However, the act of projecting into an eigenstate of one observable means that information in the eigenstates of non-commuting observables is necessarily lost, a different sort of backaction. If we envision the case of a repeated Stern-Gerlach experiment, a measurement of S_z on the beam of atoms will scramble the values of S_x and S_y , which do not commute with S_z . Repeated measurements of S_z will give the same result, but if we measure S_z , then S_x , then S_z again, the information about S_z will have been scrambled by the

⁶We note that there is vocal opposition to the terminology of “QND” from some quarters [78], with the contention that any quantum measurement can more straightforwardly be described as a perfect measurement combined with some specified or unknown measurement backaction. We employ the “QND” terminology in this thesis because it is prevalently used, but the reader should be aware of the controversy surrounding this terminology.

backaction of the intervening S_x measurement and the results of the two S_z measurements will be uncorrelated.

A QND measurement is one where all backaction (aside from projection) occurs in observables other than the one we are measuring. This means that the mechanism by which we couple to and measure the qubit must not disturb the states of our measured observable. We can express this mathematically in terms of commutation relations. A measurement of the observable A is QND if and only if:

$$[H_{\text{qubit}}, A] = 0 \quad (2.27)$$

and

$$[H_{\text{int}}, A] = 0. \quad (2.28)$$

We also note that, trivially,

$$[H_{\text{meter}}, A] = 0 \quad (2.29)$$

because A represents a qubit degree of freedom, which by definition commutes with all meter degrees of freedom. All terms involving both qubit and meter degrees of freedom are part of H_{int} by definition.

One can think of these criteria in terms of the time evolution of the quantum state. From the time-dependent Schrödinger equation, the quantum state $|\Psi\rangle$ of the qubit evolves in time according to

$$|\Psi(t)\rangle = e^{iHt/\hbar}|\Psi(0)\rangle. \quad (2.30)$$

If $|\Psi(0)\rangle$ is an eigenstate of H , the time evolution operator $e^{iHt/\hbar}$ just becomes a c-number phase factor and does not cause any state transitions. If A commutes with H , any eigenstate of A will also be unaffected by the time evolution operator (modulo a phase factor) and so repeated measurements of A will give the same result. The commutation relations above simply state that neither the time evolution dynamics of the qubit nor those of the qubit-meter interaction will couple different eigenstates of A , so once the qubit state has collapsed it will remain there. Because the time evolution operator does not mix eigenstates of A , we say that A is a “constant of the motion”.

2.3 Circuit quantum electrodynamics

The method we choose to couple our superconducting qubits to the outside world is called circuit quantum electrodynamics, or circuit QED. Circuit QED is an extension to microwave circuits of the techniques of cavity QED used in atomic physics. In cavity QED, an atom is coupled to the electromagnetic field inside a Fabry-Perot cavity. One of the mirrors of the cavity is made slightly transparent so that photons may occasionally escape. As a simplification, we look at just one atomic transition and treat the atom as a quantum two-level system. Because the interaction between light and atoms is typically weak, the cavity has the effect of amplifying the interaction strength by allowing a photon bouncing back and forth inside many chances to interact with the atom before the photon leaves the cavity. The cavity QED system is characterized by the coupling strength g between the atomic transition and the electromagnetic field in the cavity, the rate κ at which photons

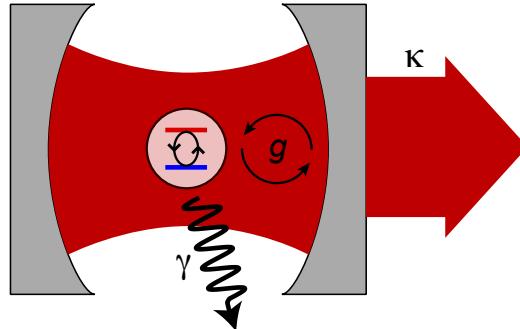


Figure 2.6: Schematic of cavity QED.

A two-level atom (circle) interacts with the electromagnetic field inside a Fabry-Perot cavity with a strength g . Photons leave the cavity at rate κ , and the atom decays into unobserved channels with a rate γ .

escape the cavity, and the rate γ at which the atom decays into modes other than the cavity mode. A typical cavity QED system is shown schematically in Figure 2.6.

The Yale group has pioneered the application of these ideas to electrical circuits [40, 41]. In circuit QED, the atom is replaced by a superconducting qubit, while the cavity is replaced by a superconducting resonator (a harmonic oscillator of the type discussed in section 2.1.1). The coupling capacitance between the qubit and the resonator sets g , while the coupling capacitance between the resonator and the microwave environment sets κ . Figure 2.7 shows a schematic of the circuit QED system with a transmon qubit.

Circuit QED has proven to be a very successful qubit measurement technique. It has been shown experimentally to give a unit visibility readout of the qubit state (that is, one where the qubit state is mapped into the readout signal perfectly) [44]. The circuit QED architecture has been used to couple qubits together and perform quantum algorithms [79, 80, 19, 22, 81, 24, 25], to prepare and detect quantum states of the microwave photon field [82, 83, 84], and to perform multi-qubit error correction protocols [24, 85, 86].

The following sections give an outline of the salient features of circuit QED relevant to the quantum jump experiments to be presented later, but are by no means an exhaustive survey of the rich and varied topics which fall under the aegis of circuit QED. For those seeking additional information, Dave Schuster's thesis gives a very thorough description of circuit QED, including for the transmon qubit [64], and Lev Bishop's thesis adds more theoretical detail [87]. Other Yale theses contain further details of circuit QED relevant to control of photons [88] and coupling and manipulation of multiple qubits [89], topics which are not directly addressed in this thesis.

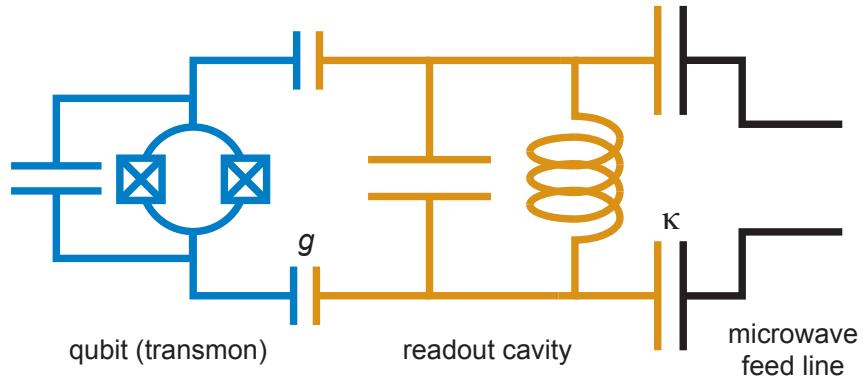


Figure 2.7: Schematic of circuit QED.

In circuit QED, we use a qubit (here a transmon) in place of an atom and a linear superconducting resonator as the readout cavity, in analogy to the cavity QED system in Figure 2.6. The qubit/cavity coupling g is set by one pair of coupling capacitors, while the cavity linewidth κ is set by another pair of capacitors. In this schematic, the cavity is probed in reflection geometry with differential excitation, but it is also possible to use transmission geometry and/or single-ended excitation.

2.3.1 The Jaynes-Cummings Hamiltonian

The circuit QED Hamiltonian consists of three terms: a qubit term, a cavity term, and a coupling or interaction term. Schematically, we can write it in the form

$$H = H_q + H_{\text{cav}} + H_{\text{int}}. \quad (2.31)$$

The H_q term is a classic Zeeman term of the form $\frac{1}{2}\hbar\omega_q\hat{\sigma}_z$, where $\hbar\omega_q$ is the qubit energy splitting and $\hat{\sigma}_z$ is one of the Pauli matrices. The H_{cav} term expresses the energy of the electromagnetic field in the cavity in harmonic oscillator form using annihilation and creation operators \hat{a} and \hat{a}^\dagger . The uncoupled (or bare) cavity resonant frequency is $\omega_{\text{cav}}/2\pi$. The H_{int} term represents a dipole coupling between the qubit and the electromagnetic field in the cavity, characterized by the strength g . Combining these expressions, we have the total Hamiltonian:

$$H = \frac{1}{2}\hbar\omega_q\hat{\sigma}_z + \hbar\omega_{\text{cav}}(\hat{a}^\dagger\hat{a} + \frac{1}{2}) + \hbar g(\hat{a} + \hat{a}^\dagger)(\hat{\sigma}_+ + \hat{\sigma}_-), \quad (2.32)$$

where $\hat{\sigma}_+$ and $\hat{\sigma}_-$ are qubit raising and lowering operators given by $\frac{1}{2}(\hat{\sigma}_x \pm i\hat{\sigma}_y)$. If we multiply out the products in parenthesis in the H_{int} term, we get four terms with \hat{a} or \hat{a}^\dagger times σ_+ or σ_- , the qubit lowering and raising operators. Since we are in the limit where the qubit and cavity energies are large compared to the available thermal energy, we invoke the rotating wave approximation, meaning that we discard any terms which do not conserve energy⁷. For example, the term $\hat{a}^\dagger\sigma_+$ does not conserve energy, because it adds a quantum

⁷In some other systems, notably trapped ions, the energy scales are such that one cannot make the rotating wave approximation and all four terms must be kept. The energy non-conserving terms of this

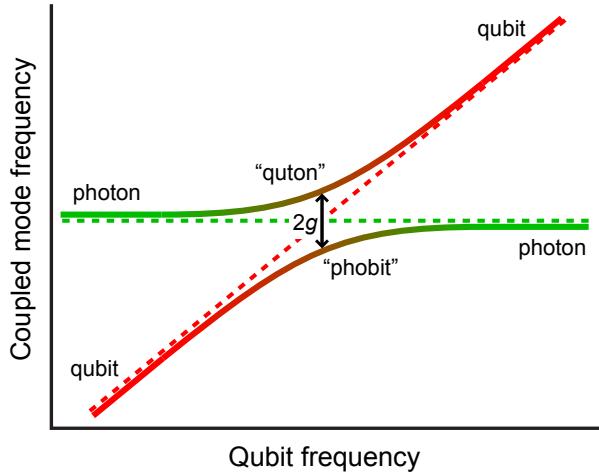


Figure 2.8: Avoided crossing from the Jaynes-Cummings Hamiltonian.

The eigenstates of the Jaynes-Cummings Hamiltonian change as the qubit frequency is swept through the cavity frequency. When qubit and cavity are far detuned, the eigenenergies (solid lines) are similar to the uncoupled eigenenergies of the qubit (red dotted line) and cavity (green dotted line), and the corresponding eigenstates are primarily qubit-like or photon-like, respectively (the color of the solid lines indicates the qubit/photon nature of the eigenstates). For $\omega_q = \omega_{\text{cav}}$, the qubit-cavity coupling g lifts the degeneracy and gives rise to an avoided level crossing. The eigenstates here are equal superpositions of qubit and photon states.

of energy to both the cavity and the qubit. After making the rotating wave approximation, our Hamiltonian becomes:

$$H_{\text{JC}} = \frac{1}{2}\hbar\omega_q\hat{\sigma}_z + \hbar\omega_{\text{cav}}(\hat{a}^\dagger\hat{a} + \frac{1}{2}) + \hbar g(\hat{a}\hat{\sigma}_+ + \hat{a}^\dagger\hat{\sigma}_-). \quad (2.33)$$

The expression in (2.33) is known as the Jaynes-Cummings (JC) Hamiltonian, and has been used extensively in quantum optics since its development in the 1960s [70]. The interaction term couples the qubit and cavity states by allowing them to exchange quanta of energy. Figure 2.8 shows a cartoon of the energy spectrum of (2.33) as one tunes ω_q from below ω_{cav} to above ω_{cav} . The dotted lines indicate the values of the qubit and cavity (photon) eigenenergies in the absence of coupling, while the solid lines give the eigenvalues of the JC Hamiltonian. The qubit and cavity maintain their individual character when far detuned but become mixed as their frequencies become closer.

In the resonant limit, where $\omega_q = \omega_{\text{cav}}$, the degeneracy of the uncoupled spectra of the qubit and cavity is lifted by the presence of the coupling term in (2.33). The

Hamiltonian are the basis for sideband cooling techniques used to bring individual trapped ions to their motional ground states [90].

eigenstates of the system are then equal-weighted linear combinations of qubit and cavity eigenstates, sometimes referred to whimsically as “phobit” and “quton” states because of their hybridized nature, which differ in energy by $2g$. A quantum of energy will be resonantly swapped back and forth between qubit and cavity continuously at a rate $2g$. This effect is known as vacuum Rabi oscillation, and the splitting between the eigenstates of the JC Hamiltonian is known as the vacuum Rabi splitting.

When ω_q and ω_{cav} are detuned from each other by an amount much larger than the coupling g , the eigenstates of the JC Hamiltonian are approximately product states of qubit and cavity. This is known as the dispersive regime, and is the subject of the next section.

2.3.2 The dispersive approximation

When the qubit and cavity are detuned from each other by an amount $\Delta \equiv \omega_q - \omega_{\text{cav}} \gg g$, we are in the dispersive regime. Unlike the resonant limit shown above, in the dispersive limit no energy is exchanged between the qubit and the cavity, and the eigenstates of the system are well approximated by product states of the qubit and cavity. If we examine (2.33) in the interaction picture, we can perform a perturbation expansion in the small parameter g/Δ to yield the dispersive approximation to the Jaynes-Cummings Hamiltonian:

$$H_{\text{disp}} = \frac{1}{2}\hbar \left(\omega_q + \frac{g^2}{\Delta} \right) \hat{\sigma}_z + \hbar\omega_{\text{cav}}(\hat{a}^\dagger \hat{a} + \frac{1}{2}) + \frac{\hbar g^2}{\Delta} \hat{a}^\dagger \hat{a} \hat{\sigma}_z. \quad (2.34)$$

There are again three terms, corresponding to the qubit, the cavity, and their interaction respectively. We note that the dispersive approximation has caused a shift in the qubit frequency of g^2/Δ ; this is the Lamb shift, caused by the interaction of the qubit with the zero-point energy of the cavity field. In addition, the interaction term has now taken on a new form which allows us to make a QND measurement of the qubit state $\hat{\sigma}_z$ using the cavity as the meter. This can be seen because our desired observable $\hat{\sigma}_z$ and the Hamiltonian obey the necessary commutation relations from (2.27) and (2.28):

$$[H_q, \hat{\sigma}_z] = 0 \quad (2.35)$$

$$[H_{\text{int}}, \hat{\sigma}_z] = 0. \quad (2.36)$$

We have established that (2.34) satisfies the conditions to give a QND measurement of the qubit, but it is not immediately evident how the measurement manifests itself. However, a simple rearrangement of terms highlights the effect of the qubit on the meter state:

$$H_{\text{disp}} = \frac{1}{2}\hbar\omega_q \hat{\sigma}_z + \hbar \left(\omega_{\text{cav}} + \frac{g^2}{\Delta} \hat{\sigma}_z \right) (\hat{a}^\dagger \hat{a} + \frac{1}{2}). \quad (2.37)$$

We see that the cavity resonant frequency now depends on $\hat{\sigma}_z$, so probing the resonant frequency of the cavity will allow us to determine the state of the qubit. The shift between the two effective cavity frequencies $\tilde{\omega}_{\text{cav}}(|0\rangle)$ and $\tilde{\omega}_{\text{cav}}(|1\rangle)$ corresponding to the two qubit

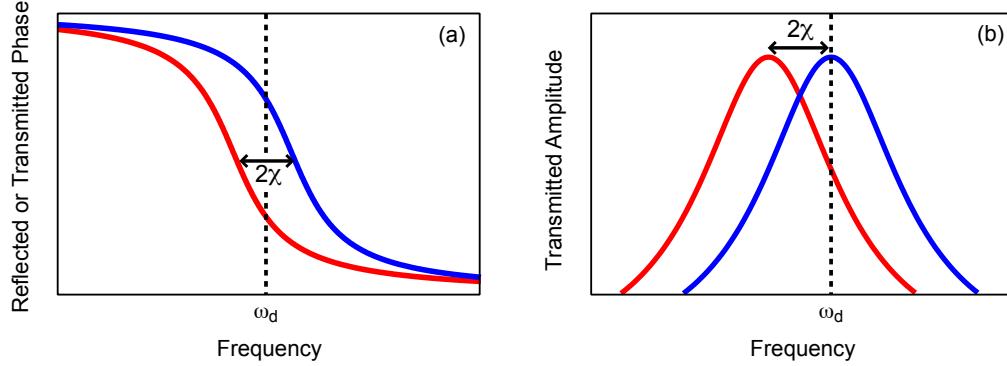


Figure 2.9: Principle of circuit QED readout.

In the dispersive approximation, the cavity resonance frequency depends on the state of the qubit. Depending on whether the qubit is in the ground or excited state, the cavity frequency is shifted by an amount 2χ . If we probe the cavity with a drive tone at an appropriate frequency, this qubit-induced cavity shift manifests itself as a change in the cavity phase response (left, for either reflection or transmission geometry) or amplitude response (right, only for transmission geometry).

states $|0\rangle$ and $|1\rangle$ is called the dispersive shift and denoted 2χ . For the dispersive JC Hamiltonian with a two-state qubit as in (2.37), we have

$$\chi = \frac{g^2}{\Delta}. \quad (2.38)$$

Figure 2.9 shows how to use the dispersive shift to read out the state of the qubit. In either reflection or transmission geometry one can look at the phase shift of the cavity response, which changes rapidly as one goes through resonance as shown in part (a). If the dispersive shift 2χ is of the order of the cavity linewidth κ , there will be an appreciable difference in the phase shift of the cavity output signal between the two qubit states if we drive at an appropriate frequency, labeled ω_d . In transmission geometry, the cavity resonant frequency can also be probed by looking at the transmitted amplitude, as shown in part (b). If we look at the complex amplitude of the output signal in the IQ plane, we find that the optimal choice of ω_d to maximize the output signal is halfway between $\tilde{\omega}_{\text{cav}}(|0\rangle)$ and $\tilde{\omega}_{\text{cav}}(|1\rangle)$, with $2\chi = \kappa$ [91].

If we group the terms in (2.34) in yet another way, we can look at the interaction term not as a shift of the cavity frequency but as a shift of the qubit frequency:

$$H_{\text{disp}} = \frac{1}{2}\hbar \left(\omega_q + \frac{2g^2}{\Delta} \hat{a}^\dagger \hat{a} + \frac{g^2}{\Delta} \right) \hat{\sigma}_z + \hbar \omega_{\text{cav}} (\hat{a}^\dagger \hat{a} + \frac{1}{2}) \quad (2.39)$$

This arrangement shows that the qubit frequency will be shifted by the presence of photons in the cavity, a phenomenon known as the ac Stark effect. We can use the ac Stark effect to calibrate the average cavity photon occupation $\bar{n} = \langle \hat{a}^\dagger \hat{a} \rangle$ for a given value of the readout

drive power [42, 92]. This is a crucial part of the calibration of a circuit QED experiment, and will be discussed in further detail in sections 2.3.3 and 6.4.3.

The dispersive approximation breaks down mathematically when the average cavity photon occupation \bar{n} becomes larger than a critical value $n_{\text{crit}} = \Delta^2/4g^2$, but numerical simulations of the full Jaynes-Cummings Hamiltonian show that many of the predictions of the dispersive approximation continue to be reasonable for $\bar{n} > n_{\text{crit}}$ [92].

The eigenstates of the qubit in the dispersive JC Hamiltonian are primarily uncoupled qubit states, but they do have a small photonic component. This photonic component gives rise to the Purcell effect, discussed experimentally in section 6.4.2. The Purcell effect gives a qubit relaxation rate $\gamma_{\text{purc}} = \kappa g^2/\Delta^2$, which can be thought of as the photonic part of the qubit eigenstate escaping from the cavity. A detailed derivation of the eigenstates of the dispersive JC Hamiltonian and the resulting Purcell effect can be found in [64]. For a multilevel system such as the transmon, the problem of diagonalizing H_{disp} exactly becomes more difficult and one must resort to numerical methods [13].

2.3.3 Transmon dispersive shift

Because the transmon has many levels with relatively low anharmonicity, we need to account for their effects in the dispersive approximation as well. For the transmon qubit, the expression for the dispersive shift χ given in (2.38) is not accurate. We must instead define the dispersive shift in terms of partial dispersive shifts between neighboring transmon states, which tend to cancel each other and reduce χ from the value for a true two-level system.

We will denote the transmon eigenenergies are given by (2.24) with a single subscript, as $\hbar\omega_i$. The energy difference between any two transmon levels $|i\rangle$ and $|j\rangle$ is denoted with two subscripts, as $\hbar\omega_{ij}$. The dipole coupling strength g between the cavity and the transmon levels now depends on which two transmon states are being coupled, so it receives two subscripts as well and becomes g_{ij} . These couplings are given by:

$$\hbar g_{ij} = 2\beta e V_{\text{rms}}^0 \langle i | \hat{n} | j \rangle \quad (2.40)$$

where $\beta = C_g/C_\Sigma$ is the ratio of the gate capacitance to the total capacitance of the transmon, $V_{\text{rms}}^0 = \sqrt{\hbar\omega_{\text{cav}}/2C_{\text{cav}}}$ is the RMS zero-point voltage of the cavity, and \hat{n} is the charge operator from the transmon Hamiltonian in equation (2.22). To a reasonable approximation, the values of the coupling strengths are given by [13]:

$$\hbar g_{i,i+1} \approx 2\beta e V_{\text{rms}}^0 \sqrt{\frac{i+1}{2}} \left(\frac{E_J}{8E_C} \right)^{1/4}, \quad (2.41)$$

The coupling elements between non-neighboring states tend to zero at large E_J/E_C :

$$g_{ij} \rightarrow 0 \text{ for } |i - j| > 1 \text{ as } E_J/E_c \rightarrow \infty. \quad (2.42)$$

The expressions above are only approximate, and any precise calculations should use values of g_{ij} computed using the exact numerical values of the matrix elements $\langle i | \hat{n} | j \rangle$. For notational simplicity, subsequent experimental parts of this thesis drop the subscripts on g , using g to mean g_{01} .

Given expressions for g_{ij} , we can write down a generalized Jaynes-Cummings Hamiltonian for a multilevel system with M levels denoted by $|i\rangle$. This Hamiltonian takes the form [13, 93]:

$$H = \hbar\omega_{\text{cav}} \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) + \sum_{i=0}^{M-1} \hbar\omega_i |i\rangle\langle i| + \sum_{i=0}^{M-2} \hbar g_{i,i+1} \left(\hat{a}^\dagger |i\rangle\langle i+1| + \hat{a} |i+1\rangle\langle i| \right) \quad (2.43)$$

A quick check shows that this is just a natural generalization of equation (2.33), with the interaction term swapping single quanta back and forth between qubit and cavity. If we take the dispersive limit with an expansion to second order in $\lambda_i = g_{i,i+1}/\Delta_i$, where $\Delta_i = (\omega_{i+1} - \omega_i) - \omega_{\text{cav}}$, we can derive an expression for the effective dispersive shift χ :

$$\chi = \chi_{01} - \chi_{12}/2, \quad (2.44)$$

where χ_{ij} are the partial dispersive shifts given by:

$$\chi_{ij} \equiv \frac{g_{ij}^2}{\omega_j - \omega_i - \omega_{\text{cav}}}. \quad (2.45)$$

Using the approximate expressions above for g_{ij} and defining the qubit anharmonicity $\alpha = \omega_{12} - \omega_{01} \approx -E_C/\hbar$, we calculate the dispersive shift as [13]:

$$\chi = \frac{g_{01}^2}{\Delta_0} \left(\frac{\alpha}{\Delta_0 + \alpha} \right) \quad (2.46)$$

In the regime where the qubit is detuned far from the cavity such that $|\Delta_0| \gg |\alpha|$, we find that the dispersive shift is substantially reduced from the typical value given in (2.38) by the presence of the higher transmon levels. For notational simplicity, the experimental parts of this thesis drop the subscript on Δ , using Δ to mean Δ_0 .

For improved accuracy in calculating the dispersive shift, which is a crucial calibration element for our system, we expand the generalized Jaynes-Cummings Hamiltonian (2.43) to fourth order in λ_i , which gives us [93]:

$$H_{\text{disp}} = \tilde{H}_0 + \sum_{i=0}^{M-1} \hbar S_i |i\rangle\langle i| \hat{a}^\dagger \hat{a} + \sum_{i=0}^{M-1} \hbar K_i |i\rangle\langle i| (\hat{a}^\dagger \hat{a})^2 \quad (2.47)$$

We use the shorthand \tilde{H}_0 to refer to the uncoupled qubit and cavity terms in the Hamiltonian, including the Lamb shifts. The S_i represent ac Stark shifts of the type described above, while the K_i are shifts arising from the Kerr nonlinearity (terms of the form \bar{n}^2) which appears in our fourth order dispersive expansion. The overall dispersive shift of the transmon will be given by:

$$2\chi = (S_1 - S_0) + (K_1 - K_0)\bar{n}. \quad (2.48)$$

For the ac Stark shift, the total shift in qubit frequency due to the cavity (known as the cavity pull) is given by

$$\zeta = 2\chi\bar{n} = (S_1 - S_0)\bar{n} + (K_1 - K_0)\bar{n}^2. \quad (2.49)$$

We can write out analytical expressions for the S_i and K_i , which take the form [93]:

$$\begin{aligned} S_i &= [\chi_{i-1}(1 - \lambda_i^2) - \chi_i(1 - \lambda_{i-1}^2) - 2\chi_{i-1}\lambda_{i-1}^2] \\ &\quad + \frac{1}{4} \left(9\chi_{i-2}\lambda_{i-1}^2 - 3\chi_{i-1}\lambda_{i-2}^2 - \chi_i\lambda_{i+1}^2 \right. \\ &\quad \left. + 3\chi_{i+1}\lambda_i^2 \right) - g_i^{(2)}\lambda_i^{(2)} - 3g_{i-2}^{(2)}\lambda_{i-2}^{(2)}, \end{aligned} \quad (2.50)$$

$$\begin{aligned} K_i &= \frac{1}{4} \left(3\chi_{i-2}\lambda_{i-1}^2 - \chi_{i-1}\lambda_{i-2}^2 + \chi_i\lambda_{i+1}^2 - 3\chi_{i+1}\lambda_i^2 \right) \\ &\quad + (\chi_i - \chi_{i-1})(\lambda_i^2 + \lambda_{i-1}^2) + g_i^{(2)}\lambda_i^{(2)} - g_{i-2}^{(2)}\lambda_{i-2}^{(2)}, \end{aligned} \quad (2.51)$$

where we use the shorthand notation $\lambda_i = g_{i,i+1}/\Delta_i$, $\chi_i = g_{i,i+1}^2/\Delta_i$, $g_i^{(2)} = \lambda_i\lambda_{i+1}(\Delta_{i+1} - \Delta_i)$ and $\lambda_i^{(2)} = -g_i^{(2)}/(\Delta_{i+1} + \Delta_i)$, and take $\chi_i = \lambda_i = 0$ for all $i \notin [0, M-2]$.

2.3.4 Measurement rate and readout signal-to-noise ratio

In section 2.3.2, we established that the circuit QED architecture can be used to make a QND measurement of the state of the qubit when operated in the dispersive regime, with $|\Delta| \gg g$. The information about the qubit state manifests itself as a shift in the resonant frequency of the cavity, which we can probe by driving the cavity with a microwave tone and looking at its response. We will now examine the nature and strength of the signal mathematically.

When driven with a microwave tone, the electromagnetic field inside the cavity will be characterized by a complex amplitude α which depends on the cavity linewidth, the strength of the drive, and the drive detuning. The photon occupation of the cavity \bar{n} is the square modulus of this amplitude $|\alpha|^2$, while $\arg(\alpha)$ gives the phase of the field. Because the cavity resonant frequency depends on the qubit state, in general the complex amplitude α will as well. For a given drive power and frequency, there will be two amplitudes $\alpha_{|0\rangle}$ and $\alpha_{|1\rangle}$ corresponding to the qubit being in state $|0\rangle$ or $|1\rangle$ respectively. We call these the pointer states, since they indicate the state of the qubit. The pointer states are coherent states of the microwave field, and as such are characterized by noise (uncertainty) in both amplitude and phase. We can average the readout signal in time to reduce this uncertainty, assuming the qubit does not change state.

The measurement of the state of the qubit depends on the vector difference between the pointer states $\alpha_{|0\rangle}$ and $\alpha_{|1\rangle}$, which we call β . The value of β depends on the relationship between the dispersive shift 2χ and the cavity linewidth κ , as well as the frequency and power of the cavity drive tone. If $\beta = 0$, the pointer states are independent of the qubit state; we can record the readout signal to our heart's content and never know what the state of the qubit was. As $|\beta|$ becomes larger and larger, it is easier and easier to determine the qubit state from the pointer states in a given amount of time. We can characterize this effect in terms of a measurement rate Γ_m [91]:

$$\Gamma_m(t) = \kappa|\beta(t)|^2, \quad (2.52)$$

where κ is the cavity linewidth. The measurement rate parameterizes the strength of the measurement; since the readout signal is continuous in time, Γ_m tells us how rapidly we accumulate information about the state of the qubit. As one might expect, Γ_m depends both on our ability to resolve the pointer states and on the rate at which photons leave the cavity carrying information about the qubit. Figure 2.10(a) and (b) show two possible cases, one with weak measurement (a) and another with stronger measurement (b). Note that the amplitude of the pointer states is the same for both cases, but that their relative phase shift is different (due to a different ratio of χ to κ). For a given cavity drive power, $|\beta|$ is maximized when the drive frequency is midway between $\tilde{\omega}_{\text{cav}}(|0\rangle)$ and $\tilde{\omega}_{\text{cav}}(|1\rangle)$, with $2\chi = \kappa$ [91].

As we saw in our two alternate associative groupings of the dispersive JC Hamiltonian, the qubit-cavity interaction can be thought of as shifting either the cavity frequency or the qubit frequency. If the qubit state affects the cavity frequency (i.e. $|\beta| \neq 0$), then the photons in the cavity will affect the qubit frequency through the ac Stark shift. Because the readout cavity photon state is a coherent state, there are fluctuations in the cavity photon occupation \bar{n} . As a result, there are fluctuations in the ac-Stark-shifted qubit frequency, which gives rise to dephasing. This is the measurement induced dephasing which is a necessary consequence of any quantum measurement; as one projects to an eigenstate of the measured observable, the phase information of a superposition state must be lost. The measurement-induced dephasing rate Γ_d is related to the measurement rate by:

$$\Gamma_d \geq \Gamma_m \quad (2.53)$$

This relationship means that we cannot acquire information about the qubit state faster than we cause it to dephase. A detailed theoretical treatment of measurement-induced dephasing is given in refs. [50] and [92]. The reader is warned that these references use conventions for Γ_d that differ by a factor of 2. The convention in the above expression is that of [50].

Given a certain measurement rate Γ_m at which we accumulate information about the qubit state, we can ask what our readout signal-to-noise ratio (SNR) will be. This topic will be discussed in more detail in section 7.5.1, and is covered extensively in ref. [94] (although the reader should be careful as some of the conventions used in that paper are different from this thesis).

In general, the signal to noise ratio will be proportional to $\eta\Gamma_m$, where η is a measure of our efficiency in detecting the readout signal with our amplification chain. For a state-of-the-art commercial low-noise microwave amplifier, $\eta \ll 1$, and we find that the noise added by our amplification chain severely limits our ability to resolve the pointer states. Figure 2.10 (c) and (d) show this effect schematically. When the pointer states exit the cavity, as seen in part (c), they each have some quantum noise but are still readily distinguishable from each other. If one includes the effects of added noise from the post-amplification chain, shown in part (d), the noise has become so large that it drowns out the signal from the pointer states and we cannot distinguish them anymore. The circuit QED readout signal is giving us a projective QND measurement of the qubit state, but the noise added by post-amplification prevents us from distinguishing the readout pointer states reliably. In order to observe the qubit dynamics using the circuit QED architecture, we will have to use amplifiers which add less noise to the readout signal.

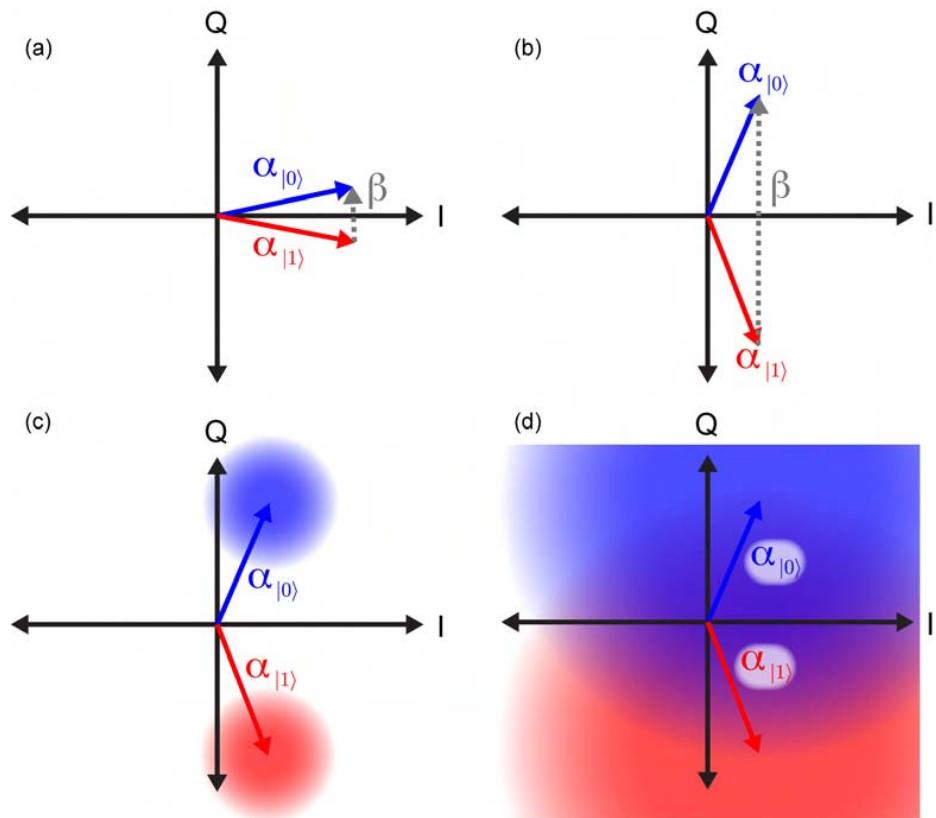


Figure 2.10: Readout signal and noise in the IQ plane.

Parts (a) and (b) show the readout pointer states for weak and strong measurements, respectively. Part (c) shows the effects of noise on our ability to resolve the pointer states given high efficiency $\eta \approx 1$ of our amplification chain. The colored circles correspond to the uncertainty in the output signal amplitude. Part (d) shows the pointer states in the presence of the noise typically added by state of the art amplifiers, where $\eta \ll 1$. It is not possible to resolve the pointer states cleanly in this instance.

Chapter 3

Amplification and the quantum limit

The circuit QED formulation and techniques mentioned in the previous chapter have been widely used for the readout of superconducting qubits. However, typical circuit QED implementations have poor single-shot readout fidelity because of noise added by following amplifiers. These amplifiers are necessary to make the signal large enough to measure at room temperature. In this chapter, we delve into the theory of amplification and noise and show that the ultimate noise performance of an amplifier is limited by quantum mechanics. We will then describe the theory and operating principles of the parametric amplifier we developed to improve the noise performance of circuit QED readout.

For a thorough, mathematically detailed treatment of general ideas of quantum noise and amplification, the reader is directed to the excellent review by Clerk and co-workers [50]. Some of the derivations in this chapter follow those presented in that reference. For an excellent general reference on classical noise and amplification, we recommend the book by Motchenbacher and Connelly [95].

3.1 Amplification and noise

Experiments with quantum mechanical systems inevitably deal with very low-level probe and readout signals. Such signals cannot be too strong, lest they destroy the quantum effects we are trying to observe. However, the typical data recording or measurement apparatus of experimental physics, sitting at room temperature, adds thermal noise to all incoming signals, so much so that signals of the level used to probe quantum effects are lost in the noise. What we need is a method for increasing the power of the low-level quantum signals until they are large enough to “lay our grubby, classical hands on”, in the famous words of Carlton Caves [96]. This is what we mean by amplification: it is a method to increase the power of a signal. Unfortunately, this power gain comes with a price, namely the addition of noise beyond that of the original signal, as shown in Figure 3.1. This added noise is an inevitable consequence of amplification which arises from quantum mechanical constraints; this will be discussed in detail in the next section. The type of amplification one typically thinks of in physics is linear amplification, which means that the

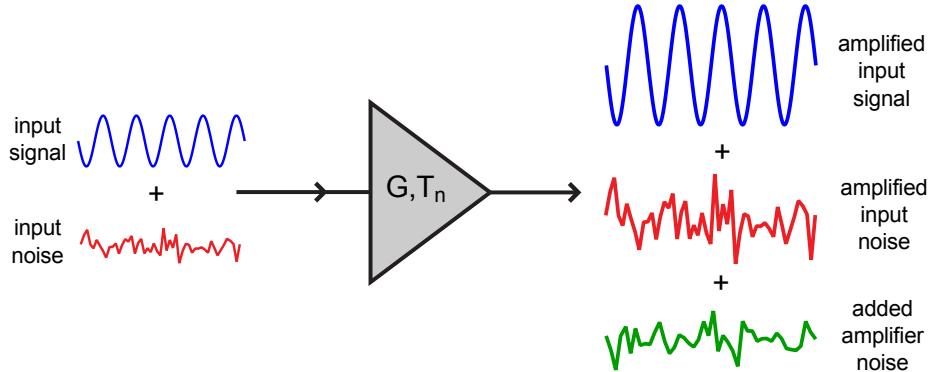


Figure 3.1: Amplification process.

A linear amplifier takes an input signal and its associated noise and outputs copies of these inputs with an amplitude increased by \sqrt{G} , where G is the power gain of the amplifier. The amplifier also adds additional noise to the output signal, the amount of which is characterized by the amplifier noise temperature T_n .

output signal is linearly related to the input signal (for example, multiplied by some fixed amplitude gain \sqrt{G}). We will deal primarily with linear amplification in this chapter, since it is straightforward to treat mathematically¹.

As discussed above, noise is intimately tied to amplification. Noise is typically characterized by a power spectral density $S(\omega)$, the intensity of noise at a given frequency ω , found by taking the Fourier transform of the time-domain noise trace². The power spectral density $S(\omega)$ can also be equivalently parameterized in terms of the amplitude spectral density $S^{1/2}(\omega)$ of noise in quantities such as voltage, current, or flux.

Classical electronic noise occurs due to the thermal fluctuations of charge carriers, and is often referred to as Johnson or Nyquist noise, after the men who discovered and explained it, respectively [97, 98]. Johnson/Nyquist noise has a spectral density which is independent of frequency (noise which is frequency-independent in this way is referred to as “white” noise). A resistor at temperature T with resistance R will exhibit voltage noise with a spectral density of $\sqrt{4k_BTR}$ V/ $\sqrt{\text{Hz}}$ and current noise with spectral density $\sqrt{4k_BT/R}$ A/ $\sqrt{\text{Hz}}$ [95]. If we multiply these expressions, we find that the total noise power in a resistor in a given bandwidth B is given by $P = 4k_BTB$. In the case where the resistor is connected to an impedance-matched load, the voltage noise occurs across the parallel resistance $R/2$ and the current noise through the series resistance $2R$. Assigning half of the dissipated power to the source impedance and half to the load impedance, and we recover the famous expression for noise power dissipated in a matched load [95]:

$$P = k_BTB. \quad (3.1)$$

¹Amplification need not be linear to be useful, and indeed the experimental work presented in later chapters involves amplification which is not necessarily linear.

²More technically, one can define the power spectral density in terms of the autocorrelation function of the noise signal. As an example, for a voltage signal $V(t)$, we find $S(\omega) = \int_{-\infty}^{\infty} \langle V(t)V(0) \rangle e^{i\omega t} dt$ [50, 95].

We can rewrite this equation in terms of the power spectral density (the noise power per unit bandwidth) as

$$S(\omega) = k_B T. \quad (3.2)$$

This expression shows that the thermal noise dissipated in a matched load is white and is linearly dependent on temperature. At very high frequencies and/or very low temperatures, however, quantum mechanical effects change this simple relationship. Accounting for quantum effects using the fluctuation-dissipation theorem [99], the result for the noise power spectral density³ becomes [50]:

$$S(\omega) = \frac{\hbar\omega}{2} \coth \left[\frac{\hbar\omega}{2k_B T} \right]. \quad (3.3)$$

Let's examine a few limits of equation (3.3). First, in the high-temperature classical limit where $\hbar\omega \ll k_B T$, we use the approximation $\coth x \approx 1/x$ for small x and find $S(\omega) \approx k_B T$, reproducing the original classical result. In the low temperature limit $\hbar\omega \gg k_B T$, the coth term goes to 1 and we find that

$$S(\omega) = \frac{\hbar\omega}{2}. \quad (3.4)$$

This term is recognizable as the zero-point energy of the electromagnetic field; we have rediscovered the result that even at zero temperature there will still be quantum fluctuations of the electromagnetic field at the level of half a photon per unit bandwidth. Since our qubits and readout operate in this low-temperature limit where $\hbar\omega \gg k_B T$, the noise on a circuit QED output signal will be made up of quantum fluctuations.

Given the nice relationship in (3.2), it is convenient to parameterize the noise spectral density in terms of an effective temperature T_{eff} , which is defined as

$$T_{\text{eff}} = S(\omega)/k_B. \quad (3.5)$$

In the classical (high temperature) regime, we find that $T_{\text{eff}} = T$, the true physical temperature. However, at sufficiently low temperatures we must use the full quantum mechanical expression (3.3) and the relationship becomes more complicated. At temperatures low enough that (3.4) applies, the effective temperature saturates to the value

$$T_{\text{eff}} = \frac{\hbar\omega}{2k_B}. \quad (3.6)$$

For a qubit experiment with readout frequencies of 6 GHz we have $T_{\text{eff}} = 144$ mK, even though the physical temperature might approach 30 mK.

³In general, the spectral density of quantum noise is not symmetric in frequency: $S(\omega) \neq S(-\omega)$. The expression in (3.3) is the symmetrized or classical noise spectral density, which is the average of the quantum noise spectral densities at frequencies ω and $-\omega$. The asymmetry of quantum noise spectra arises mathematically from the fact that the two quadratures of a signal do not commute with each other and thus their cross-correlation is non-zero. We can also understand it heuristically by considering noise as coming from absorption or emission of photons coming from a thermal bath, corresponding to noise at positive or negative frequencies respectively. The density of final states for absorption and emission processes are different when $|\omega| \gtrsim k_B T/\hbar$, so the emission and absorption rates (and thus the amount of noise at ω versus $-\omega$) will be different. Further details are provided in ref. [50].

The noise on the input signal is only part of the full story, however, since we must also account for the noise added by the amplifier itself. Let's take a microwave amplifier connected to a matched source impedance of effective temperature T_s . We can express the noise added by the amplifier in terms of an amplifier noise temperature T_n , which is defined as follows. For an amplifier with gain G and an effective source temperature T_s , the noise power at the output is given by

$$P_{n,\text{out}} = Gk_B(T_s + T_n)B. \quad (3.7)$$

By measuring the output noise power and knowing T_s , we can extract the value of T_n . The noise temperature is not a real temperature (although it may be correlated with the physical temperature of the amplifier for many devices), but rather is a convenient shorthand way of expressing the amount of noise added by the amplifier. Looking at equation (3.7), we can see that using an amplifier with noise temperature T_n is equivalent to raising the effective source temperature by T_n and then amplifying using a fictitious noiseless amplifier with the same gain G .

Often the signal of interest is so small that we must use several amplifiers in series to achieve sufficient gain; this is the case with circuit QED readout signals. In such a setup, it is essential to know how the noise temperature of each amplifier will affect the noise properties of the final output signal. We can write this down mathematically if we assume that each of the amplifiers in series (let's say there are k of them) has a gain G_i and noise temperature $T_{n,i}$. The noise power after the first amplifier stage is given by equation (3.7), and this serves as the input noise of the second amplifier stage. We add the contribution of the second amplifier and find that its output noise is

$$P_{n,\text{out}} = G_2[G_1k_B(T_s + T_{n1})B + k_BT_{n2}B]. \quad (3.8)$$

After the third stage, we have

$$P_{n,\text{out}} = G_3(G_2[G_1k_B(T_s + T_{n1})B + k_BT_{n2}B] + k_BT_{n3}B), \quad (3.9)$$

and so forth through all the stages. At the final output, we would like to calculate the equivalent input noise, which we do by dividing out the total gain $G_1G_2\cdots G_k$. We then divide by $k_B B$ and subtract the original effective source temperature T_s to give the noise temperature T_{sys} for the whole amplifier system :

$$T_{\text{sys}} = T_{n1} + \frac{T_{n2}}{G_1} + \frac{T_{n3}}{G_1G_2} + \cdots + \frac{T_{nk}}{G_1G_2\cdots G_{k-1}} \quad (3.10)$$

For large gains $G_i \gg 1$ and for noise temperatures such that $T_{n,i} \gg T_{n,i+1}/G_i$, we find that the overall system noise temperature depends primarily on the noise temperature of the first-stage amplifier. In other words, as long as the output noise from any given stage is much larger than the input noise of the next stage, the noise of this next stage will contribute little to the overall T_{sys} .

For circuit QED experiments, we need several stages of amplification. State-of-the-art cryogenic semiconductor microwave amplifiers have noise temperatures around 3 K and power gains on the order of $10^3 - 10^4$, or 30 – 40 dB (gain for microwave amplifiers is usually

quoted in logarithmic units (dB), with the relationship between linear and logarithmic power gains given by $G_{\text{dB}} = 10 \log_{10} G_{\text{linear}}$. This means that as long as the next amplifier in the chain has $T_n \ll 3,000K$, the value of T_{sys} is set by the cryogenic amplifier. If we would like to reduce the system noise temperature further, we need an amplifier with noise temperature well below 3 K and enough gain to overwhelm the input noise of the cryogenic semiconductor amplifier. If we make an amplifier with $T_n \approx 150 \text{ mK}$ (one which adds half a photon of noise at 6 GHz), it should ideally have a power gain of at least 100 (20 dB in log units) to ensure that it is the primary contribution to the system noise temperature.

3.2 Quantum limits on amplifiers

In the previous section, we alluded to the fact that quantum mechanics sets limits on the noise temperature of a linear amplifier. We will now derive this result theoretically and examine its consequences. At the source of all quantum noise constraints is the quantum mechanical notion of commuting and non-commuting operators. Qualitatively, quantum noise limits arise when we wish to know the value of two non-commuting observables simultaneously. Our knowledge of non-commuting observables is limited by the Heisenberg uncertainty principle, which in effect makes our estimates of the observed values “noisy”.

Let’s take the canonical example of a harmonic oscillator, described by the position operator \hat{x} and the momentum operator \hat{p} , which do not commute with each other. Quantum mechanics dictates that there will exist a relationship between the position and momentum uncertainties Δx and Δp given by

$$\Delta x \Delta p \geq \hbar/2 \quad (3.11)$$

Any measurement of position will disturb the value of momentum and vice versa, with the amount of disturbance depending on the precision of the measurement of the other observable. If one would like to know both the position and the momentum simultaneously, the uncertainty relationship (3.11) limits the total precision of the combined measurement. Even when the harmonic oscillator is in its lowest energy state there is some spread in the values of position and momentum, an effect called the “zero-point motion” of the oscillator. In a heuristic sense, one can think of the zero-point motion as a fundamental quantum-mechanical “noise” obscuring our knowledge of the exact position and momentum of the oscillator [39].

The Hamiltonian for the electromagnetic field at a given frequency is isomorphic to that of a harmonic oscillator, with the roles of position and momentum mapped into the two quadrature amplitudes, or alternatively the amplitude and phase. We usually characterize the electromagnetic field in terms of the creation and annihilation operators \hat{a}^\dagger and \hat{a} , which obey the commutation relation $[\hat{a}, \hat{a}^\dagger] = 1$.

Since these operators do not commute, it will be impossible to know their values simultaneously with arbitrary precision, just as in the case of position and momentum for the harmonic oscillator. However, if we want to amplify an electromagnetic wave faithfully, in general we need to “know” the value of both operators simultaneously (equivalent to saying we need to know both the amplitude and phase of a signal simultaneously) so we can

make an enlarged copy at the amplifier output. Heuristically, we can say that the amplifier adds noise because it cannot copy and enlarge both quadratures of the input signal at the same time without being a bit uncertain about what the true value of the input signal is. As a result, the output signal is never a perfect copy of the input, but rather a slightly noisier copy. This way of looking at added noise suggests that we could potentially amplify without adding noise if we only care about one quadrature; this is in fact the case and will be discussed in section 3.2.1.

We can examine this in a more mathematically rigorous way in the following derivation, which is due to Haus and Mullen [100] and was expanded by Caves [96]. We take an amplifier whose input and output can be described by a set of bosonic modes, photons for example. A classical input signal $E(t)$ at frequency ω takes the general form

$$E(t) \propto (ae^{-i\omega t} + a^* e^{i\omega t}), \quad (3.12)$$

where a and its complex conjugate a^* characterize the two quadrature amplitudes (or the amplitude and phase) of the signal. For quantum mechanical signals, we must convert these complex amplitudes to operators $a \rightarrow \hat{a}$ and $a^* \rightarrow \hat{a}^\dagger$, the annihilation and creation operators for the electromagnetic field at the input. We would like to map this input signal to an output signal which is also a state of the electromagnetic field. We described the output field in terms of output annihilation and creation operators \hat{b} and \hat{b}^\dagger . These operators all obey the standard commutation relations:

$$[\hat{a}, \hat{a}^\dagger] = 1, \quad [\hat{b}, \hat{b}^\dagger] = 1. \quad (3.13)$$

We would like the output of our amplifier to be a copy of the input with some power gain G . We would like both quadratures of the input signal to be amplified with the same gain, so we have the relationship

$$\hat{b} = \sqrt{G}\hat{a}, \quad \hat{b}^\dagger = \sqrt{G}\hat{a}^\dagger. \quad (3.14)$$

Unfortunately, these expressions clearly violate the commutation relations in (3.13). We have to introduce a second term $\hat{\mathcal{F}}$ to satisfy the commutation relations, changing (3.14) to

$$\hat{b} = \sqrt{G}\hat{a} + \hat{\mathcal{F}}, \quad \hat{b}^\dagger = \sqrt{G}\hat{a}^\dagger + \hat{\mathcal{F}}^\dagger. \quad (3.15)$$

This operator $\hat{\mathcal{F}}$ represents noise added by the amplifier, which should be uncorrelated with the input signal \hat{a} . As a result, we can take $[\hat{\mathcal{F}}, \hat{a}] = 0$ and $[\hat{\mathcal{F}}, \hat{a}^\dagger] = 0$. Since $\hat{\mathcal{F}}$ is a random noise, we take the expectation values $\langle \hat{\mathcal{F}} \rangle = \langle \hat{\mathcal{F}}\hat{a} \rangle = \langle \hat{\mathcal{F}}\hat{a}^\dagger \rangle = 0$ as well. If we now insist that \hat{b} satisfy the commutation relations in (3.13), we find that:

$$[\hat{\mathcal{F}}, \hat{\mathcal{F}}^\dagger] = 1 - G. \quad (3.16)$$

We would now like to determine the mean-square fluctuations of the output signal, given by $(\Delta b)^2$. This can be expressed as

$$(\Delta b)^2 = \frac{1}{2}\langle\{\hat{b}, \hat{b}^\dagger\}\rangle - |\langle\hat{b}\rangle|^2, \quad (3.17)$$

where $\{\hat{b}, \hat{b}^\dagger\} = \hat{b}\hat{b}^\dagger + \hat{b}^\dagger\hat{b}$. An identical formula holds for $(\Delta a)^2$. Using a series of useful commutator and anticommutator relations presented in [96], as well as the relations between $\hat{\mathcal{F}}$ and \hat{a} presented above, we can write the expression for $(\Delta b)^2$ as:

$$(\Delta b)^2 = G(\Delta a)^2 + \frac{1}{2}\langle\{\hat{\mathcal{F}}, \hat{\mathcal{F}}^\dagger\}\rangle. \quad (3.18)$$

Using another identity from [96] and recalling that $\langle\hat{\mathcal{F}}\rangle = 0$, we can rewrite this as an inequality:

$$\begin{aligned} (\Delta b)^2 &\geq G(\Delta a)^2 + \frac{1}{2}|\langle[\hat{\mathcal{F}}, \hat{\mathcal{F}}^\dagger]\rangle| \\ &\geq G(\Delta a)^2 + \frac{|G - 1|}{2}. \end{aligned} \quad (3.19)$$

In the limit of large gain, we have $G \approx G - 1$. Dividing the output noise by the gain G to get the noise referred to the input, we find

$$(\Delta b)^2/G \geq (\Delta a)^2 + \frac{1}{2}. \quad (3.20)$$

We have recovered the result that phase-preserving amplification must add at least half a quantum of noise, referred to the input. In the no-amplification limit $G = 1$, we find that the amplifier is not required to add noise. Looking at (3.16), we can see that for $G > 1$ it is possible to write a model for the added noise as

$$\hat{\mathcal{F}} = \sqrt{G - 1}\hat{d}^\dagger, \quad \hat{\mathcal{F}}^\dagger = \sqrt{G - 1}\hat{d}. \quad (3.21)$$

This formulation represents a single additional input mode \hat{d} which is amplified by $G - 1$. One can readily check that the noise inequality in (3.20) becomes an equality in this case, and the amplifier's added noise represents just the vacuum fluctuations of the input mode \hat{d} , which is required to maintain the commutation relations and still achieve gain. This additional mode represents the “idler” mode in a parametric amplifier, to be discussed in section 3.3.

3.2.1 Phase-sensitive and phase-preserving amplifiers

The above derivation is dependent on the fact that we wish to amplify both \hat{a} and \hat{a}^\dagger in the same way, in other words, that we wish to amplify both quadratures of the input signal with the same gain. This type of amplification is often referred to as “phase-preserving” amplification⁴, since the amplified signal contains the information from both quadratures of the input signal and maintains phase and amplitude relationships.

There are certain situations in which one might only be interested in one quadrature of the signal, without caring what happens to the other quadrature. Amplification of just one quadrature of a signal is referred to as “phase-sensitive” amplification, since the output signal depends on the relative phase of the input signal. We might suspect that this

⁴It is also occasionally called “phase-insensitive” amplification, a nomenclature we avoid here because it is so similar to “phase-sensitive”.

allows us to amplify without adding noise, since the single quadrature we amplify commutes with itself.

Phase-sensitive amplification is equivalent to amplifying \hat{a} without regard to what happens to information in \hat{a}^\dagger . Given this flexibility, we can write down an analogue to equation (3.14) where we amplify one quadrature but deamplify the other⁵:

$$\hat{b} = \sqrt{G}\hat{a}, \quad \hat{b}^\dagger = \frac{1}{\sqrt{G}}\hat{a}^\dagger. \quad (3.22)$$

We can see by inspection that these definitions of \hat{b} and \hat{b}^\dagger satisfy the commutation relations in (3.13), and so we have no need to introduce extra modes and their associated noise into the output signal. As a result, we find that the output noise referred to the input is $(\Delta b)^2/G = (\Delta a)^2$; we have accomplished noiseless amplification! For general signals, phase-sensitive amplification is not appropriate because we typically care about the information in the deamplified quadrature. However, for some signals, including qubit readout signals as will be detailed in later sections, we can profitably employ phase-sensitive amplification and benefit from its superior noise performance.

3.3 Parametric amplification

Having talked about amplification in general, we narrow our focus to a specific and basic type known as parametric amplification. The principle of parametric amplification is simple: gain is achieved by varying a parameter of the amplifier system harmonically in time. The energy used to vary the parameter is called the pump. The modulation of the system parameters at the pump frequency causes some of the pump energy to be transferred into another frequency mode, chosen to be the signal frequency (additional pump power is sent to another mode, called the idler). In this way, the signal achieves power gain and amplification. Detailed mathematical derivations of parametric amplification can be found in refs. [101, 102], among many others.

The classic example of parametric amplification is a child on a swing. Even when nobody is around to push him or her, a child can still swing to great heights just by “pumping” his or her legs. Whenever the child reaches the maximum swing excursion, he or she changes leg position from bent to outstretched or vice-versa. The direction is chosen to raise the child’s center of gravity, which changes the effective length of the swing. This is the parameter of the oscillator (the swing) which is modulated. The energy from the child’s “pumping” is, appropriately, the pump, which is at twice the natural frequency of the swing. The energy from this pump is transferred into the two degenerate normal modes of the swing ($e^{\pm i\omega t}$) and causes the amplitude of the swing’s oscillations to increase.

In a parametric amplifier, often called a “paramp” for short, the pump couples two frequency modes together and transfers energy to both of them. These modes are called the signal and the idler. The frequencies of the signal, idler, and pump are related

⁵This operation is also called squeezing, because it takes a coherent state input, which has equal width in both quadratures in the IQ plane, and “squeezes” it into a more elliptical state with different widths along its two principal axes but the same phase space area as the initial coherent state. Squeezed states are of interest in quantum optics [70], and phase-sensitive amplification is a typical way of generating them.

in a specific way depending on the exact nature of the system. The two most common types of parametric amplifiers are called three-wave or four-wave amplifiers. In a three-wave amplifier, the relationship between the pump frequency ω_p and the signal and idler frequencies ω_{sig} and ω_{id} is given by

$$\omega_p = \omega_{\text{sig}} + \omega_{\text{id}}, \quad (3.23)$$

while for a four-wave amplifier the relationship is

$$\omega_p + \omega_p = \omega_{\text{sig}} + \omega_{\text{id}}. \quad (3.24)$$

Thinking in the context of quantum mechanics, one need only multiply these equations by \hbar to realize that they represent expressions of energy conservation⁶. This also suggests the reason for the names “three-wave” and “four-wave”; a three-wave amplifier turns a pump photon into a signal photon plus an idler photon, while a four-wave amplifier turns two pump photons into a signal photon plus an idler photon. This fact is codified mathematically by the so-called Manley-Rowe relations, which can be derived from the completely classical theory of coupled oscillator modes and state that, in a parametric amplifier, the power in a given mode divided by the frequency of that mode is constant across all modes [101].

It is possible to have a parametric amplifier where $\omega_{\text{sig}} = \omega_{\text{id}}$; such an amplifier is called a degenerate parametric amplifier, because the frequencies of the signal and idler modes are degenerate. A degenerate parametric amplifier is necessarily phase-sensitive; the pump serves as a “reference clock” for incoming signals, and signals which are not in phase with the pump will not be amplified. For a parametric amplifier to be phase-preserving, we require that $\omega_{\text{sig}} \neq \omega_{\text{id}}$. This dovetails nicely with the description of quantum noise limits in the previous section, where we showed that a phase-preserving amplifier must have at least one additional input mode (which we recognize to be the idler) in order to satisfy the constraints of photon commutation relations, while a phase-sensitive amplifier does not require an additional input mode (which we recognize as the signal and idler being at the same frequency).

3.3.1 Brief history of parametric amplification

Parametric amplification has been known and used for a long time [103]. The idea that a parametrically driven system exhibits energy transfer between coupled modes was noted by Faraday in 1831 and Lord Rayleigh in the 1880s, who considered oscillations of mechanical systems in their calculations. Almost a century ago, the first circuit-based parametric amplifiers were developed to amplify and transmit radio signals [104]. These amplifiers relied on parametric modulation of the inductance of a resonant circuit by periodic saturation of the inductor’s iron core. Just a few years later, though, high power vacuum tubes became available and took over the market for power amplification of radio signals. It was not until the late 1940s and early 1950s, after the development of radar and microwave electronics during World War II led to a need for high-frequency, low-noise amplifiers, that parametric amplification returned to the scene. Transistors were still a new technology, and

⁶It also gives an idea of why the idler must exist; if not for the idler, the only way energy conservation can be satisfied is to have ω_p be a multiple of ω_{sig} , which as we shall see represents phase-sensitive amplification.

early transistors did not function well as amplifiers at microwave frequencies. Parametric amplifiers using varactor diodes, which operate as voltage-dependent capacitors in certain bias regimes, became a topic of substantial research interest because they had superior gain and noise performance at microwave frequencies. In the 1970s, the technology of heterostructure transistors became sufficiently advanced that transistor amplifiers could once again compete with parametric amplifiers at microwave frequencies, and in the ensuing years transistor amplifiers have become the dominant technology for amplifying microwave signals.

Work on microwave-frequency parametric amplifiers has continued despite the dominance of the transistor. For example, the traveling-wave tube amplifier (TWTA) uses parametric modulation of a high-voltage beam of electrons to amplify microwave signals to very large powers, up to 70 dBm (10 kW). TWTAAs are widely used for satellite communications [105]. The first superconducting parametric amplifiers, based on the nonlinear Josephson inductance, were demonstrated in 1975 [51]. A variety of superconducting parametric amplifier designs followed in the 1980s and 1990s [52, 53, 54, 55, 56], but they tended to be plagued by spurious “noise rise” and were not used broadly in applications.

Optical-frequency parametric amplifiers employing nonlinear optical fibers were also demonstrated in the mid-1970s, and development work continues to the present day. The performance of these fiber amplifiers has improved in recent years and may someday be competitive for use in fiber-optic repeaters [106]. Optical parametric amplifiers have also been used extensively in quantum squeezing experiments [70].

The rapid progress in solid-state quantum measurement and quantum information of the past decade has renewed interest in superconducting parametric amplifiers. A number of recent results [57, 58, 59, 60, 107, 108, 109, 110, 111] have demonstrated quantum-limited noise performance, gains in excess of 30 dB, and quantum squeezing of the microwave vacuum. These properties make superconducting parametric amplifiers attractive candidates for improving qubit readout.

3.4 The Lumped Josephson Parametric Amplifier (LJPA)

The work in this thesis relies on a particular implementation of superconducting parametric amplifier called the Lumped Josephson Parametric Amplifier (LJPA)⁷. The LJPA is a nonlinear microwave resonator consisting of two Josephson junctions in a small SQUID loop, shunted by a capacitance (the equivalent circuit diagram is shown in Figure 3.3). This is the same basic circuit as the transmon qubit described in Chapter 2, with the exception that the critical current of the LJPA junctions is much higher than that of the transmon junctions. In particular, the critical current is large enough that there are many energy levels in the potential well defined by E_J , so we can treat the LJPA as a classical nonlinear oscillator rather than a quantum one. In the zero-voltage state, the two-junction SQUID loop acts like a single Josephson junction with a flux-tunable critical current, allowing us to tune the resonant frequency of the LJPA over an octave in frequency, from 4-8

⁷The “lumped” in the name comes from the fact that the circuit is sufficiently small to be treated as a lumped element at microwave frequencies.

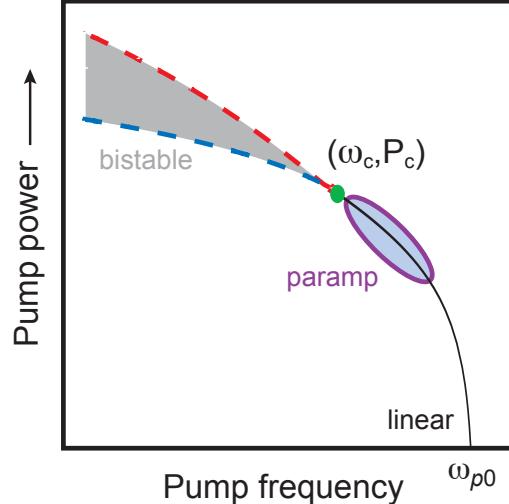


Figure 3.2: Nonlinear resonance response.

The response of the LJPA resonance is linear for low drive powers, but moves to lower frequency as the drive power increases, eventually becoming bistable for certain bias parameters. The LJPA can be used as a paramp when biased into the region labeled “paramp” by a strong drive tone.

GHz. The fact that one can use this design of lumped Josephson resonator as a paramp was first noted in R. Vijay’s thesis [108].

The response of the nonlinear LJPA resonance is shown in cartoon form in Figure 3.2. At low drive power, the resonance is linear, with a resonant frequency $\omega_{p0} = \sqrt{1/L_{J0}C}$, where L_{J0} is the Josephson inductance in the absence of driving and C is the shunt capacitance. As the drive power increases, the resonance bends to lower frequencies, a consequence of the particular type of nonlinearity exhibited by the Josephson junctions, and sharpens. The diagram shows a critical point (ω_c, P_c) , beyond which the system becomes bistable for some bias parameters. This bistable regime has been accessed in the Josephson bifurcation amplifier (JBA)—essentially the same circuit as the LJPA—for high-fidelity qubit readout [46, 108].

For parametric amplification, we bias the LJPA into the region labeled “paramp” with a strong drive tone that functions as the pump. The pump modulates the nonlinear Josephson inductance L_J . Since L_J is an even function of the current through the junctions, L_J is modulated at twice the pump frequency, and the LJPA acts like a four-wave parametric amplifier. Accordingly, the signal and idler frequencies are symmetric around the pump frequency.

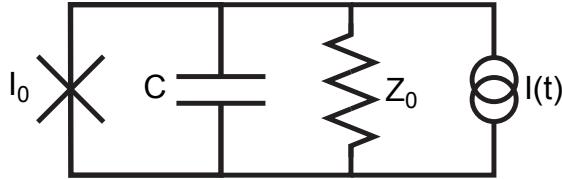


Figure 3.3: Schematic of the LJPA.

The LJPA consists of a SQUID loop, modeled as a single Josephson junction with flux-dependent critical current I_0 , shunted by a capacitance C and the impedance Z_0 of the microwave environment⁸. We can study the dynamics of the system mathematically by adding a drive current $I(t)$.

3.4.1 Mathematical description

A schematic diagram of the LJPA circuit is shown in Figure 3.3. We treat the SQUID loop as a junction with critical current I_0 , shunted by a capacitance C and the impedance Z_0 of the microwave environment⁸. We apply a current $I(t)$ to the circuit from an external source. Using the Josephson relations and Kirchoff's laws, we can describe the behavior of this circuit with the following differential equation:

$$C \frac{\Phi_0}{2\pi} \frac{d^2\delta(t)}{dt^2} + \frac{1}{Z_0} \frac{\Phi_0}{2\pi} \frac{d\delta(t)}{dt} + I_0 \sin(\delta(t)) = I(t). \quad (3.25)$$

Let's take our externally applied current to be a sinusoidal pump tone of the form $I(t) = I_d \cos(\omega_d t)$, with $I_d < I_0$, and replace the $\sin \delta$ term with its Taylor expansion, keeping only the first nonlinear term. The resulting equation, known as the Duffing equation, greatly simplifies the algebra while accurately reproducing the dynamics of the system [112], and is given by

$$\frac{d^2\delta(t)}{dt^2} + 2\Gamma \frac{d\delta(t)}{dt} + \omega_{p0}^2 \left(\delta(t) - \frac{\delta(t)^3}{6} \right) = \frac{2\pi}{\Phi_0 C} I_d \cos(\omega_d t), \quad (3.26)$$

where $\omega_{p0} = \sqrt{2\pi I_0 / \Phi_0 C}$ is the oscillator's resonant frequency (or plasma frequency) at low drive amplitude and $\Gamma = 1/2Z_0C$ is the damping rate of the oscillation amplitude. We can solve this equation using an ansatz of the form

$$\delta(t) = \delta_0 \cos(\omega_d t - \theta). \quad (3.27)$$

To make it easier to solve equation (3.26), it is useful to parameterize the ansatz as a pair of quadrature amplitudes $\delta_{||}$ and δ_{\perp} , with $\delta_0^2 = \delta_{||}^2 + \delta_{\perp}^2$ and $\tan \theta = \delta_{\perp} / \delta_{||}$. Our ansatz then takes the form

$$\delta(t) = \delta_{||} \cos(\omega_d t) + \delta_{\perp} \sin(\omega_d t), \quad (3.28)$$

⁸We use differential excitation with a 180° hybrid for our paramps as described in sections 5.1 and 6.1, so in our case $Z_0 = 100 \Omega$.

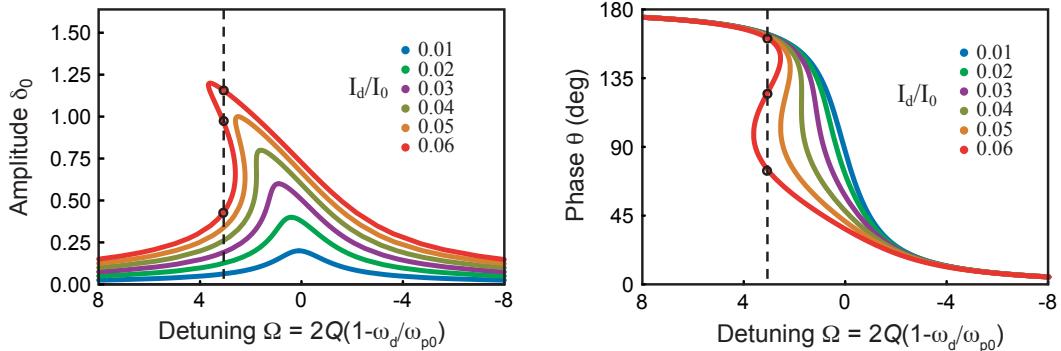


Figure 3.4: Nonlinear resonance solutions.

We plot solutions for equation (3.26) in terms of the parameters δ_0 and θ defined in equation (3.27). At low drive strengths, the response is linear. As the drive amplitude increases, the resonance moves to lower frequencies and finally becomes bistable. The vertical dashed line and three circles highlight the three solutions (of which the middle solution represents an unstable state) in the bistable regime.

and we can solve for the values of $\delta_{||}$ and δ_{\perp} by plugging into equation (3.26). We make the rotating wave approximation, replacing terms oscillating at $2\omega_d$ and $3\omega_d$ with their averages, to yield two coupled cubic equations

$$-\delta_{||} + \delta_{\perp} \left[\Omega - \frac{Q}{8} (\delta_{||}^2 + \delta_{\perp}^2) \right] = 0 \quad (3.29)$$

and

$$\delta_{\perp} + \delta_{||} \left[\Omega - \frac{Q}{8} (\delta_{||}^2 + \delta_{\perp}^2) \right] = QI_d/I_0, \quad (3.30)$$

where $\Omega = 2Q(1 - \omega_d/\omega_{p0})$ is the dimensionless detuning of the drive from the resonant frequency and $Q = \omega_{p0}Z_0C$ is the quality factor of the resonator.

Since these equations are cubic, we expect in general that there will be three solutions for $\{\delta_{||}, \delta_{\perp}\}$ and thus for $\{\delta_0, \theta\}$. For $\Omega > \sqrt{3}$, there exist some ranges of I_d for which there are three distinct real solutions; this is the bistable regime, and the largest and smallest values of δ_0 are the stable states. For $\Omega < \sqrt{3}$, where the paramp regime falls, only one real-valued solution exists for $\{\delta_0, \theta\}$.

Figure 3.4 shows a family of solutions for δ_0 and θ at different values of I_d as a function of the dimensionless detuning Ω . As shown qualitatively in Figure 3.2, the resonance moves to lower frequency with increasing drive power, finally becoming bistable at some bias values. The response is single-valued for $\Omega < \sqrt{3}$, as well as for $\Omega \geq \sqrt{3}$ at sufficiently low or high drive amplitude. The curve for $I_d/I_0 = 0.04$, with a steepened but single-valued response, represents the paramp regime. We note that the amplitude $\delta_0 \lesssim 0.8$ for the paramp regime, so our use of the Duffing approximation $\sin \delta \approx \delta - \delta^3/6$ is accurate to better than 0.4%.

3.4.2 Theoretical gain and bandwidth

Now we add a small signal slightly detuned from the pump tone to characterize the gain and bandwidth of the amplifier. A detailed derivation of these results has been presented in several theses [108, 109], so we will give only a brief sketch. We take the signal to have the form $I_s(t) = I_s \cos(\omega_d t + \omega_s t)$, where $\omega_s \ll \omega_d$ and $I_s \ll I_d$. In this case, we can treat the signal as a small perturbation on the strong pump, and we predict that the solution to (3.26) will be of the form

$$\delta(t) = \delta_0 \cos(\omega_d t - \theta) + \epsilon(t), \quad (3.31)$$

where $\epsilon(t)$ represents the response due to the presence of the weak signal. Plugging this into equation (3.26) and using the fact that $\delta(t) = \delta_0 \cos(\omega_d t - \theta)$ is a solution in the absence of a signal tone, we get a differential equation for ϵ of the form

$$\frac{d^2\epsilon(t)}{dt^2} + 2\Gamma \frac{d\epsilon(t)}{dt} + \epsilon(t)\omega_{p0}^2 \left(1 - \frac{\delta_0^2}{4} - \frac{\delta_0^2}{4} \cos(2\omega_d t - 2\theta)\right) = \frac{2\pi}{\Phi_0 C} I_s \cos(\omega_d t + \omega_s t). \quad (3.32)$$

The left-hand side of this equation is that of a parametrically driven harmonic oscillator whose resonant frequency is being modulated at $2\omega_d$ ⁹. We can use input-output theory [69] to write down the voltage amplitudes of incoming and outgoing waves in our reflection geometry in terms of the input signal amplitude $V_s = I_s Z_0$ and $V_J = (\Phi_0/2\pi)d\epsilon/dt$, the voltage across the junction. This gives us two equations:

$$V_s(t) = 2V_{\text{in}}(t) \quad (3.33)$$

$$V_{\text{out}}(t) = V_J(t) - V_{\text{in}}(t). \quad (3.34)$$

We can take the Fourier transform of equation (3.32) and solve for the relationship between $V_{\text{in}}(t)$ and $V_{\text{out}}(t)$ in the frequency domain, which makes the analysis simple because our signal is harmonic. The parametric driving gives rise to a coupling between the modes at $\omega_d + \omega_s$ and $\omega_d - \omega_s$. We recognize this second mode as the idler frequency in our four-wave parametric amplifier. The gain is defined as the ratio of the power in signal and idler output modes to that in the input mode, and depends on the signal frequency and pump parameters. We can write down closed-form expressions for this gain, which we call G_s for the signal (or direct) gain and G_i for the idler (or trans) gain¹⁰:

$$G_s(f) = 1 + \frac{4\rho^2}{[(\Omega - 2\rho)^2 - \rho^2 + 1]^2 - 2f^2[(\Omega - 2\rho)^2 - \rho^2 - 1]} \quad (3.35)$$

and

$$G_i(f) = G_s(f) - 1, \quad (3.36)$$

where we have introduced the notation $\rho = Q\delta_0^2/8$ to characterize the drive amplitude¹¹

⁹The fact that the modulation is at $2\omega_d$ while the pump frequency is at ω_d arises because the Josephson inductance is an even function of δ , so a drive at ω_d modulates the Josephson inductance and thus the resonant frequency of the oscillator at $2\omega_d$.

¹⁰We have dropped an f^4 term in the denominator of this expression, which has a negligible effect, to highlight the Lorentzian shape of the gain profile. This f^4 term is not dropped in the gain expressions given in refs. [61] and [108]. We also note that there are typographical errors in the expressions for G_s and f given in [61], which have been corrected here.

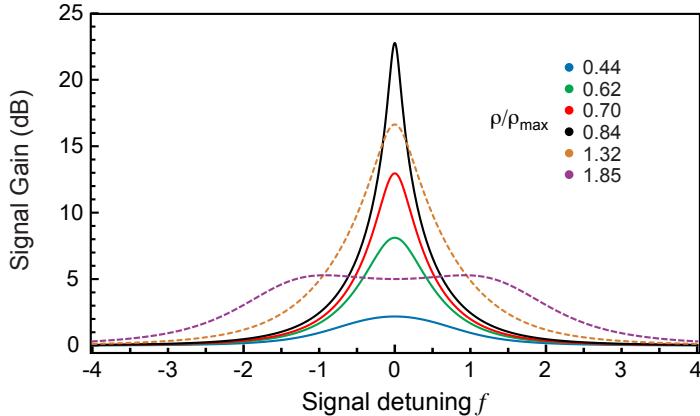


Figure 3.5: Theoretical gain and bandwidth.

These curves show the gain and bandwidth of the LJPA for varying values of the pump drive strength ρ at a pump detuning $\Omega = 0.98\sqrt{3}$. Note that the bandwidth is given in units of the reduced signal detuning f , indicating that a lower Q paramp resonator will give correspondingly broader absolute instantaneous bandwidth. We can trade gain for additional bandwidth by driving at $\rho > \rho_{\max}$.

and defined the dimensionless signal-pump detuning frequency $f = 2Q(\omega_s/\omega_{p0})$. A quick check of these equations shows that, in the limit of no pump ($\rho \rightarrow 0$), we recover $G_s = 1$ and $G_i = 0$, as we would expect for the case of no amplification. For large gain, we have $G_s \approx G_i$, again as expected. In the limit of small signal detuning $f \rightarrow 0$, we can optimize the gain with respect to the pump amplitude, yielding a value for ρ which maximizes the gain:

$$\rho_{\max} = \sqrt{\frac{1 + \Omega^2}{3}} \quad (3.37)$$

Figure 3.5 shows the value of $G_s(f)$ for several different values of the drive amplitude ρ . The bandwidth is a monotonically increasing function of ρ , while the gain reaches a peak at ρ_{\max} and then decreases at higher drive amplitudes. For signal detunings $f \neq 0$, the maximum gain may come for $\rho \neq \rho_{\max}$. We can calculate the drive current amplitude I_d at ρ_{\max} , which is given by the expression

$$I_d^2 = \frac{16I_0^2}{3\sqrt{3}Q^3}(1 + \Omega^2)(2\sqrt{1 + \Omega^2} - \sqrt{3}\Omega). \quad (3.38)$$

This gives us a scale for a typical pump power we might expect to use for amplification. For typical amplifier parameters $I_0 \approx 3 - 4 \mu\text{A}$, $Q \approx 20 - 25$, $\Omega \approx \sqrt{3}$, we find that the pump power should be around -90 to -95 dBm¹². This in turn sets a scale for the strength of the input signals, since we require that $I_s \ll I_d$ in our analysis.

¹¹Note that in ref. [108] the notation ϵ is used for what we call ρ .

¹²The units “dBm”, commonly used in microwave frequency applications, are absolute logarithmic power units, defined as $P_{\text{dBm}} = 10 \log_{10} P_{\text{mW}}$, where P_{mW} is the power in milliwatts.

In the $f \rightarrow 0$ limit, we can use the expression for ρ_{\max} to derive an expression for the maximum gain:

$$G_{\max} = 1 + \frac{1}{3 + 7\Omega^2 - 4\sqrt{3}\Omega\sqrt{1 + \Omega^2}}. \quad (3.39)$$

This expression diverges for $\Omega = \sqrt{3}$; in practice, however, higher order effects will cause the gain to remain finite. At the maximum gain point, we can substitute the value of ρ_{\max} into (3.35), re-expressing the frequency-dependent gain as:

$$G_s(f) \approx \frac{G_{\max}}{1 + f^2(kG_{\max})}, \quad (3.40)$$

where k is a constant of order unity given by the expression

$$k = \frac{-\Omega(3\Omega - 2\sqrt{3}\sqrt{1 + \Omega^2})}{1 + \Omega^2}. \quad (3.41)$$

In the limit $\Omega \rightarrow \sqrt{3}$, we find $k \rightarrow \frac{3}{4}$. Equation (3.40) highlights the Lorentzian nature of the gain profile, and gives us a relation between the gain and the bandwidth of the amplifier; using the definition of f , we can write the half-width half-maximum bandwidth B_{LJPA} as

$$B_{\text{LJPA}} = \frac{\omega_{p0}}{4\pi Q\sqrt{kG_{\max}}}. \quad (3.42)$$

This expression shows that the gain-bandwidth product¹³ is constant for the LJPA, and is given by $B_{\text{LJPA}}\sqrt{G_{\max}} = \omega_{p0}/4\pi Q\sqrt{k}$. We can see that using a low- Q resonator as our paramp increases the gain-bandwidth product.

3.4.3 Physical description of operation

It is helpful to have a physical understanding of the mechanism of amplification, beyond just the math, which is what we hope to provide in this section. All amplifiers have a transfer function which relates the input quantity to the output quantity. For a field-effect transistor, for example, the transfer function is the transconductance, which relates the gate voltage to the source-drain current. To make a useful transistor amplifier, one constructs and biases a circuit in such a way that a small change in gate voltage will give rise to a large change in source-drain current. A dc SQUID amplifier has a flux-to-voltage transfer function V_Φ , allowing small flux signals to be converted into larger voltage signals. When the flux and SQUID current are correctly biased, small changes in flux effect large changes in voltage.

For the LJPA, the transfer function is the change in the phase of the reflected pump tone with respect to the pump amplitude; note that this is not the same as the phase θ of $\delta(t)$, though the two are related. As we increase the amplitude of the pump and sweep through resonance, the phase of the reflected pump changes rapidly. The phase changes about $60^\circ - 90^\circ$ over a very small range in pump amplitude, with the overall phase shift across the resonance equal to about 180° . This phase shift is a manifestation of the

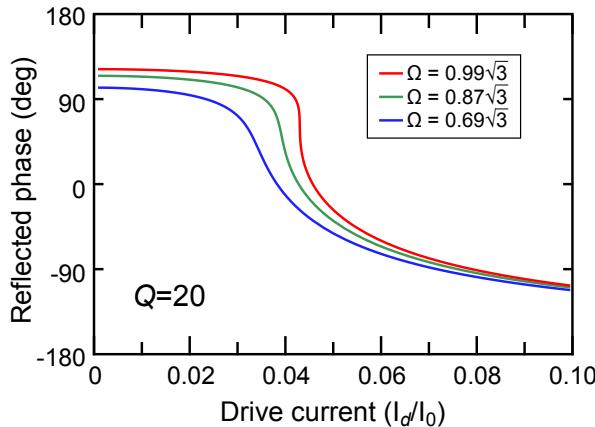


Figure 3.6: Paramp transfer function: reflected phase vs. pump amplitude.

We show three theoretical traces of the phase of the reflected pump tone as a function of the dimensionless detuning Ω for $Q = 20$. These traces can be thought of as vertical linecuts of Fig 3.2 in the paramp regime, with $\Omega = \sqrt{3}$ corresponding to the critical point.

nonlinearity of the LJPA; a linear resonator shows no phase response with changes in drive amplitude.

Figure 3.6 shows three sample theoretical traces of the reflected phase of the paramp for three different values of the dimensionless detuning Ω . As $\Omega \rightarrow \sqrt{3}$, in other words, as the pump frequency approaches the bifurcation point, the phase response steepens. The physical mechanism by which gain occurs can be understood using this transfer function. When the pump amplitude is chosen so that we are biased on the steepest part of the transfer function, any small modulations of the pump amplitude will lead to modulations of its reflected phase. The transfer function is approximately linear over a small range of pump amplitudes, so we will have a linear relation between the reflected pump phase and its amplitude. The theoretical power gain when biased at the steepest part of the transfer function is given by equation (3.39) and is 10, 70, and 10^4 for the blue, green, and red traces respectively.

We consider first the case where our signal is at the same frequency as the pump; this is the doubly-degenerate mode of operation (so named because signal and pump frequencies are degenerate, and we have four-wave mixing) and will give us phase-sensitive amplification, as described in section 3.2.1. If the signal is in phase with the pump, they will add coherently and the effective pump amplitude will be changed slightly. This is shown schematically in the IQ plane in Figure 3.7(a). Any shifts in the amplitude of the signal will cause shifts of the same magnitude in the effective pump amplitude. These amplitude shifts move us left and right along the paramp transfer function, shown in cartoon form in Figure 3.7(b), giving rise to phase shifts in the reflected pump tone. We plot the reflected

¹³The gain-bandwidth product is defined in terms of amplitude gain, not power gain; the amplitude gain is $\sqrt{G_{\max}}$.

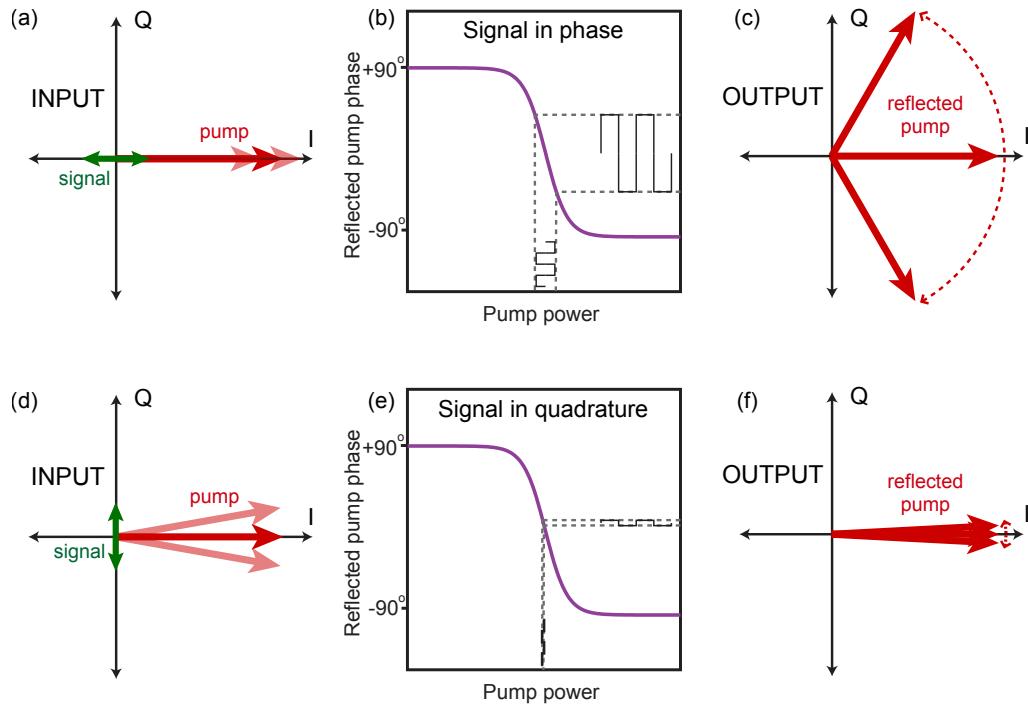


Figure 3.7: Phase-sensitive amplification in the IQ plane.

Parts (a-c) treat a signal in phase with the pump. A signal (green) in phase with the pump (red) will add or subtract coherently with it, modulating its amplitude slightly (light red arrows). Note that we exaggerate the magnitude of the signal relative to the pump in parts (a) and (d) for clarity. If the pump is biased at the steep part of the transfer function, as seen in part (b), this pump amplitude modulation will lead to a phase modulation on the reflected pump. Sketching the output signal in the IQ plane in (c), we note that the vector difference between output states (the phase-shifted reflected pump) is much larger than that between input signal states, meaning we have achieved power gain. In parts (d-f), the signal is in quadrature with the pump. As a result, the amplitude of the vector sum of the signal and pump (light red arrows) is the same to first order as the original pump amplitude. This means that the signal only makes a very small modulation of the reflected pump phase, as seen in (e), and the vector difference between output states shown in (f) is smaller than that of the input signal. This represents de-amplification of the quadrature signal.

pump in the IQ plane in part (c). For sufficiently small signal amplitudes, there will be a linear relation between the signal amplitude and the reflected pump phase.

How does this represent amplification? The output signal is the phase shift on the reflected pump tone, which is much stronger than the input signal. The information carried in the amplitude of the input signal has been mapped linearly to the phase of the much stronger reflected pump. In the IQ plane, we see the vector difference between the output states is larger than the vector difference between the input states. This represents power gain and thus is amplification by our definition. To reconstruct the original signal from the amplified one, we can simply extract the phase of the output signal and use the linear mapping between input amplitude and output phase to determine what the original signal was.

The phase-sensitive nature of this method can be seen if one considers a signal in quadrature with the pump tone. Such a signal will leave the pump amplitude unchanged to first order. This can be seen in Figure 3.7(d), where the vector sum of the signal and pump in the IQ plane has essentially the same magnitude as the original pump vector. As a result, the pump amplitude moves very little along the transfer function, seen in part (e), and thus the phase of the reflected pump changes very little. The signal ends up actually being de-amplified; the vector difference between output states in the IQ plane is less than that between the input signals.

Now let's consider the case where there is input noise in addition to the signal (we know that we will always have at least have quantum zero-point fluctuations, no matter what the signal is). Noise at the signal frequency will modulate the pump amplitude in exactly the same way as the signal does: in-phase noise will lead to noise on the reflected pump phase, while quadrature noise will be de-amplified. Since the amplifier treats the input signal and input noise on the same footing, we will have the same gain for signal and noise; in particular, no quadrature noise will be amplified. This means that the signal-to-noise ratio is preserved by the amplifier, which can only happen if the amplifier itself adds no noise (see Figure 3.1). We can see therefore from transfer function arguments why the phase-sensitive mode of the LJPA paramp should give noiseless amplification.

Phase-sensitive amplification is unfortunately only useful when one knows the phase of the incoming signal. This is not the case for arbitrary signals, but for a circuit QED readout signal one can control the phase of the readout tone relative to the paramp pump and thus ensure that the readout signal lies in the amplified quadrature. We utilize phase-sensitive amplification in the quantum jump experiments detailed in Chapters 7 and 8 because of its superior noise performance.

Phase-preserving amplification is more broadly applicable than phase-sensitive amplification, although we pay for this with decreased noise performance. We can understand the mechanism for phase-preserving amplification using the same idea of the transfer function. Here we have an input signal at frequency $\omega_d + \omega_s$. For maximum amplification to occur, we require $\omega_s \ll \omega_d/Q\sqrt{G_s}$ (the signal detuning must be small compared to the linewidth of the paramp resonance divided by the amplitude gain; compare with equation (3.42)). This criterion is roughly equivalent to $f \ll \sqrt{G_s}$, which is the bandwidth criterion for high gain seen in equation (3.40). Physically, this constraint means that ω_s is sufficiently slow that the pump amplitude can respond to modulation at that frequency.

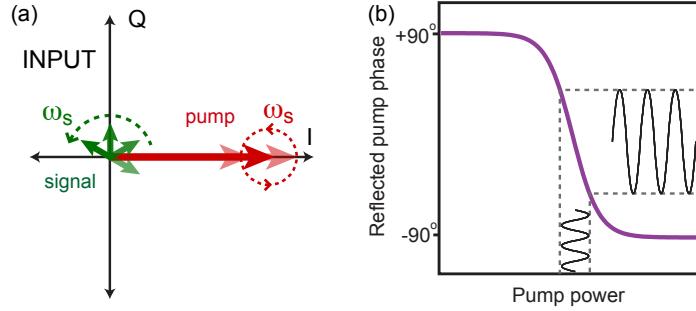


Figure 3.8: Phase-preserving amplification in the IQ plane.

In phase-preserving mode, the input appears as a small signal rotating at frequency ω_s in the IQ plane frame rotating at ω_d . This signal introduces a sinusoidal modulation on the pump amplitude, which becomes a sinusoidal phase modulation of the reflected pump as shown in (b). Since the signal amplitude is much smaller than the pump amplitude, the reflected pump exhibits a double-sideband phase modulation. These two sidebands are the signal and the idler of the paramp.

If we look in the IQ plane in the frame rotating at ω_d , the signal looks like a small vector rotating at ω_s , as shown in Figure 3.8. As is evident from the picture, this signal will modulate the pump amplitude at frequency ω_s as well. We can parameterize the input signal in the rotating frame in terms of two time-varying amplitudes, one in phase with the pump ($A_{s||}$) and one in quadrature with the pump ($A_{s\perp}$),

$$A_{s||}(t) = A_s(t) \cos(\omega_s t + \phi_s(t)) \quad (3.43)$$

$$A_{s\perp}(t) = A_s(t) \sin(\omega_s t + \phi_s(t)), \quad (3.44)$$

where $A_s(t)$ is the amplitude of the input signal and $\phi_s(t)$ is its phase¹⁴. In this way we have decomposed the signal at $\omega_d + \omega_s$ into two quadrature signals at frequency ω_d , each of which will undergo phase-sensitive amplification as described previously. The quadrature $A_{s||}(t)$ will modulate the pump amplitude (since it is in phase with the pump) and be amplified, while $A_{s\perp}(t)$ will not modulate the pump amplitude (since it is in quadrature with the pump) and thus be de-amplified. However, since $A_{s||}(t)$ contains information on both the amplitude $A_s(t)$ and phase $\phi_s(t)$ of the input signal, we will get both amplitude and phase information on the output signal, realizing phase-preserving amplification. If we assume for a moment that A_s and ϕ_s are constant, we will have a sinusoidal modulation of pump amplitude from $A_{s||}(t)$, which will result in a sinusoidal phase modulation of the reflected pump, as shown in Figure 3.8(b).

¹⁴An input signal of the form $A_s(t) \cos(\omega_d t + \omega_s t + \phi_s(t))$ can be re-expressed as $A_s(t) \cos(\omega_d t) \cos(\omega_s t + \phi_s(t)) - A_s(t) \sin(\omega_d t) \sin(\omega_s t + \phi_s(t))$ using trigonometric identities. We can rewrite this as the two quadratures of a signal at ω_d as $A_{s||}(t) \cos(\omega_d t) - A_{s\perp}(t) \sin(\omega_d t)$, where the quadrature amplitudes are as given in equations (3.43) and (3.44). The minus sign can be eliminated by the choice of reference phase.

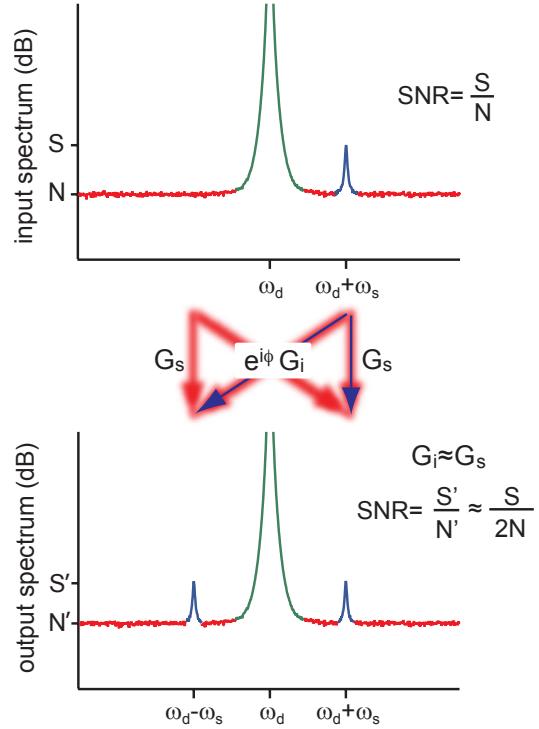


Figure 3.9: Amplification and noise for a detuned signal.

The top portion of this figure shows the input pump (green) and single-sideband input signal (blue) in addition to the input noise floor (red) composed of quantum zero-point fluctuations. The phase-preserving parametric amplification process gives both direct gain (vertical arrows) and trans gain (diagonal arrows) at the signal and idler frequencies. The output signal is a double-sideband signal, where each sideband has half the signal-to-noise ratio (SNR) of the original input signal because of the presence of additional amplified vacuum fluctuations from the idler input. This is the realization of the quantum limit on phase-preserving amplification derived in section 3.2.

The quadrature $A_{s\parallel}(t)$ and its noise will be amplified with the same gain (and noise on $A_{s\perp}(t)$ will be de-amplified), so one might guess that this is again noiseless amplification as in the phase-sensitive case. However, we have neglected to consider vacuum fluctuations at the idler frequency $\omega_d - \omega_s$. This noise will also modulate the pump amplitude at frequency ω_s (in a counter-rotating sense relative to the signal) and be amplified. Thus the output signal will have an additional $\hbar(\omega_d - \omega_s)/2$ of added noise referred to the input and the amplification process is no longer noiseless, in keeping with the quantum limits on phase-preserving amplifiers described in section 3.2.

We can see this effect more clearly by thinking in terms of the frequency spectrum. The quadrature $A_{s\parallel}(t)$ modulates the initial amplitude A_0 of the pump, and the modulated

amplitude A_{pump} is given by

$$\begin{aligned} A_{\text{pump}} &= ((A_0 + A_{s||}(t)) \cos(\omega_d t) \\ &= A_0 \cos(\omega_d t) + A_s(t) \cos(\omega_s t + \phi_s(t)) \cos(\omega_d t). \end{aligned} \quad (3.45)$$

Using trigonometric identities, we can rewrite this as:

$$A_{\text{pump}} = A_0 \cos(\omega_d t) + \frac{1}{2} A_s(t) [\cos(\omega_d t + \omega_s t + \phi_s(t)) + \cos(\omega_d t - \omega_s t - \phi_s(t))]. \quad (3.46)$$

Recalling that $A_0 \gg A_s$, we can see that this is the equation for a carrier tone at ω_d with two smaller sidebands detuned by $\pm\omega_s$. Since the pump amplitude is linearly mapped into the phase of the output signal, we should see two sidebands at $\omega_d \pm \omega_s$ on the output signal as well¹⁵. These represent the signal and the idler of the paramp, as described in section 3.3. We note that the pump amplitude contains information on both the amplitude $A_s(t)$ and phase $\phi_s(t)$ of the input signal, meaning that the output signal will also contain this information. The maintenance of both input amplitude and input phase information on the output signal demonstrates that this is phase-preserving amplification of the input signal.

Figure 3.9 draws the frequency spectrum of the input (top) and output (bottom), showing the pump, signal, and idler. The arrows in between show both the direct gain G_s (signal to signal and idler to idler) and the trans gain G_i (signal to idler and idler to signal). In the high gain regime, $G_s \approx G_i$, as seen from equation (3.36). The signal (blue arrows) and its noise (red arrows) are amplified with the same direct gain, but noise (red arrows) from the idler frequency is mixed in by the trans gain and degrades the overall signal-to-noise ratio. If the input noise at both signal and idler frequencies is just due to quantum zero-point fluctuations, this process adds half a photon of noise (referred to the input) to the amplified signal, the minimum required by quantum mechanics. In the limit of no amplification $G_s = 1$ and $G_i = 0$, so the amplifier simply reflects the signal without adding any noise or creating an idler.

3.4.4 Saturated regime operation

The above results are derived for the small-signal regime, where we assume we are always in the linear region of the transfer function. When the signal amplitude becomes large enough that we leave the linear transfer region, the output signal is no longer linearly related to the input signal, a phenomenon called saturation. In an op-amp, saturation can occur when one attempts to drive the output signal past the voltage supplies, causing it to “rail”. In the LJPA, saturation occurs when the signal modulates the pump amplitude far enough that the slope of the transfer function starts to flatten out¹⁶.

¹⁵The output signal shows phase modulation, not amplitude modulation. However, given a harmonic input signal small enough to remain in the linear regime of the transfer function, the output phase modulation has the same frequency spectrum as the pump amplitude modulation.

¹⁶The range of signal powers over which the amplifier behaves linearly is called its dynamic range; the dynamic range of the LJPA is set by the width in pump amplitude of the linear part of the transfer function. Alternatively, one can think about dynamic range and paramp saturation in terms of pump depletion. The process of parametric amplification transfers energy from the pump to the signal and the idler. If the power of the output signal becomes large enough, the transfer of power from pump to signal becomes significant enough to deplete (reduce the power of) the pump. When the pump starts to become depleted, there is less pump energy to transfer into signal and idler tones and so the gain is decreased.

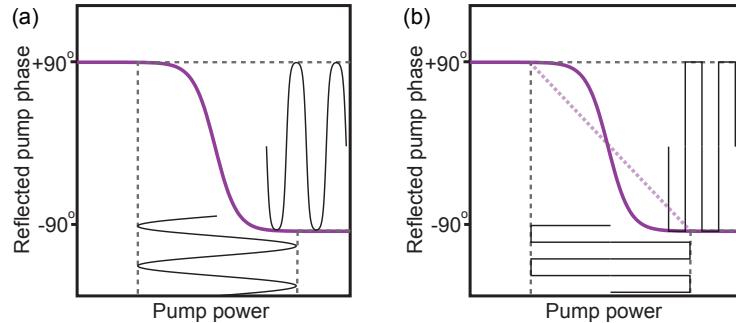


Figure 3.10: Saturated regime amplification.

A large-amplitude sinusoidal signal modulating the pump amplitude is not faithfully reproduced in the output signal. The extrema of the sinusoid are flattened and reduced in amplitude. However, a square wave signal (one which modulates the pump between two different amplitudes) will give rise to a square wave output signal. The effective square wave transfer function (dashed purple line) has a reduced slope compared to the original transfer function, indicating the reduced gain that typifies amplifier saturation.

The effects of saturation on sinusoidal and square wave modulations are shown in Figure 3.10. For signals which give a sinusoidal pump modulation, the output signal in the saturated regime will be a distorted sinusoid with flattened peaks. This distortion renders the amplifier much less useful. However, if the signal is a square wave which modulates the pump amplitude between two specific values, the output signal will also be a square wave, with a one-to-one mapping between output phase and pump amplitude. In this way, input signals which switch between two states are faithfully amplified even in the saturated regime. This is fortuitous, because we get exactly this type of signal from circuit QED qubit readout. By choosing the pump bias properly, we can use the paramp to amplify qubit readout signals even though the power from the readout signal exceeds the dynamic range of the paramp.

Let's examine the behavior of a square wave signal undergoing phase-sensitive amplification in the saturated regime. The output signal is the phase shift on the reflected pump; however, once we have saturated the paramp, this will always be $\sim 180^\circ$, no matter the input signal strength. With the output signal fixed, increases in the input signal must correspond to decreases in gain of the same amount. This reduction in gain with increasing signal amplitude is a hallmark of a saturated amplifier. If one draws an effective paramp transfer function (shown in Figure 3.10(b) as a purple dashed line), one sees that the effective transfer function is less steep than the original one, indicating reduced gain.

The noise performance is also degraded in the saturated regime. When the paramp saturates, the amplitude of the output signal no longer grows with increasing input signal amplitude. Noise on the input signal continues to be amplified, although the noise gain decreases gradually as the paramp moves farther into the saturated regime. As a result, the output signal-to-noise ratio (SNR) plateaus and grows only sub-linearly with increasing input SNR; the output signal is fixed, while the output noise is slowly decreasing. When

the paramp is deeply saturated, the output SNR will be considerably lower than the input SNR. Thus the amplification process degrades the SNR, which we can think of in terms of the amplifier adding noise. The farther we go into the saturated regime, the greater the SNR degradation, and thus the higher the noise temperature of the paramp. It is important to note, however, that the output SNR is still a monotonically increasing function of input SNR.

Chapter 4

Sample fabrication

This chapter describes the fabrication techniques used to make superconducting qubits, resonators, and parametric amplifiers. The methods for making these devices are straightforward, but achieving reliable and high-quality results requires attention to detail, a great deal of patience, and sometimes a bit of luck. We detail a few tricks used for successful sample fabrication, to save the reader the trouble of rediscovering them. Further specifics of recipes will be presented in Appendix B.

4.1 Electron-beam lithography

All of the samples described in this thesis were made with electron-beam lithography, a standard technique for fabricating very high resolution features. Electron-beam lithography allows structures to be defined by first coating the substrate of interest with some form of resist, then selectively damaging (exposing) the resist in certain desired regions using a finely focused beam of high-speed electrons. The exposed resist is then washed away by a chemical developer, leaving a patterned mask of resist on the surface. The samples described in this thesis are made by depositing thin films of aluminum or niobium on top of this mask, which serves as a sort of stencil. The mask is then chemically dissolved, with the metal on top of the mask “lifting off” to leave behind metal on the substrate in the pattern defined by the electron beam exposure.

The focused electron beam used for lithography comes from a scanning electron microscope (SEM), a modified FEI Nova NanoSEM depicted in Figure 4.1. This microscope uses a field emission tip as the electron source, so the electrons come from a smaller region than with a hot filament emitter. This allows the beam to be more finely focused than with a filament SEM, giving better resolution for both imaging and lithography. The imaging resolution has been measured to be ~ 1.2 nm, and the system is capable of reliably making 20 nm lithographic features. The SEM is turned over to remote control by a computer running Nanometer Pattern Generation System (NPGS) software during lithography. NPGS can control the position of the beam, its focus, the magnification, and even the position of the sample stage, allowing the e-beam writing process to be fully automated for samples from small test structures to entire wafers of ground planes.

For optimal resolution, the electron beam should have the highest possible accel-

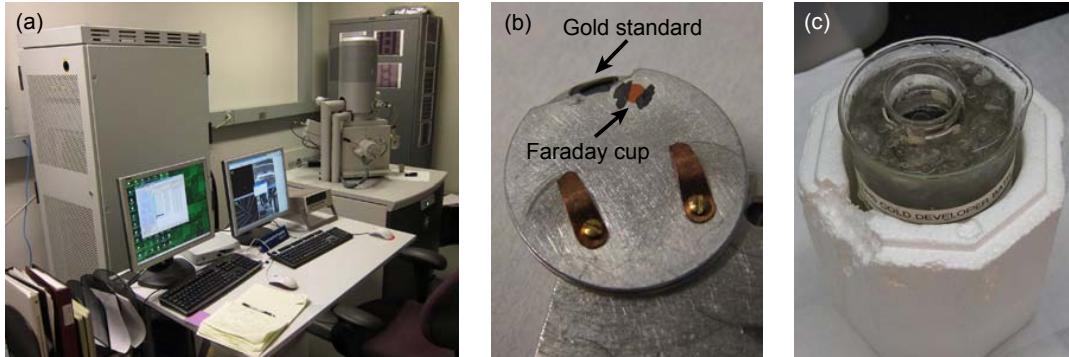


Figure 4.1: Tools of the fab trade.

The scanning electron microscope in (a) has been modified to allow remote control by NPGS pattern generation software, allowing it to be used for e-beam lithography in addition to imaging. The electron source is a field emission tip, giving very high resolution. Part (b) shows a sample holder for lithography, showing gold standard, Faraday cup, and spring clips for holding the sample. The cold developer setup is seen in (c), consisting of a beaker of ice-water slush in an insulating jacket with a smaller beaker of developer resting inside.

erating voltage; we use 30 kV, the maximum for our system. Dedicated e-beam writers in the UC Berkeley Nanolab and the Lawrence Berkeley Lab Molecular Foundry can run at 50 kV or 100 kV, but this is not necessary to achieve suitable resolution for qubit samples. Higher accelerating voltages are useful because they minimize the spread of the electron beam as it scatters off of atoms in the resist and the substrate [113]. Due to this scattering, very fine (10-20 nm) features should be made on the thinnest possible resist stack. Another essential element for good lithography is careful optimization of the electron beam profile by adjusting the focus, stigmation, and lens alignment. For this purpose, our sample holder includes a viewing standard containing gold balls on a graphite substrate, which allows very precise beam optimization before lithography. In addition, drops of a colloidal solution of 100 nm and 40 nm diameter gold balls are added on the substrate surface itself, to allow final focusing and correction for any tilt in the substrate relative to the beam axis. To ensure fine lithography, one needs to write above a certain minimum SEM magnification setting (about 1300x) so that the fine scan coils will be used instead of the coarse scan coils. This effect is shown in Figure 4.2(c) and (d). Finally, it is important to know the beam current accurately to be able to give the correct dose. To do this, the sample holder includes a Faraday cup, made by covering a small hole drilled in the holder with a single-hole TEM grid glued down with carbon paint. We measure the current by steering the beam into the Faraday cup and reading out the current on a picoammeter connected to the microscope stage.

4.1.1 Resist selection

All samples in this thesis were fabricated using a bilayer resist stack, consisting of a thick ($\sim 1 \mu\text{m}$), softer underlayer resist topped with a thinner ($\sim 300 \text{ nm}$), harder resist layer. The bilayer resist stack is used for several reasons. First, it is required to create the freestanding bridge structures necessary to fabricate junctions using double-angle evaporation (described in detail in section 4.2). Secondly, the bilayer resist stack allows for the creation of undercut, where holes patterned in the developed resist are larger next to the substrate than at the top of the resist stack; this aids in liftoff by reducing the chance that metal films will stick to the walls of the resist or form a continuous coat between the substrate and the top surface of the resist stack.

The standard resists used to fabricate superconducting qubits in the past are poly(methylmethacrylate) (PMMA) in anisole solvent as the top resist layer and methyl methacrylate/methacrylic acid copolymer (MMA) in ethyl lactate solvent as the underlayer. The resists and solvents are chosen so that the MMA is not soluble in the PMMA solvent (anisole), which allows the PMMA to be spin-coated on top of the MMA layer without disturbing or damaging it. PMMA and MMA are available in different concentrations, which yield different final thicknesses after spinning. Detailed spin and bake recipes are given in Appendix B. The PMMA/MMA resist stack is developed in a single step using a mixture of isopropanol (IPA) and methyl isobutyl ketone (MIBK) in a ratio of 1:3 MIBK:IPA (other groups have used other ratios, such as 1:7 MIBK:IPA, or even IPA/water mixtures [114]). Because MMA is more sensitive to e-beam irradiation than PMMA, it is exposed at a lower dose, so electrons scattered to the side (proximity dosing) during writing create a natural undercut in the MMA underlayer. MMA is also more easily solved in the developer solution, which further contributes to the undercut. With the ability to realize features smaller than 50 nm, and sufficient undercut to perform double-angle evaporation, PMMA/MMA is suitable for making a variety of superconducting devices such as qubits, resonators, and paramps.

There are some drawbacks to the PMMA/MMA resist stack, however, which led us to investigate alternatives. In some masks, more undercut is required than occurs naturally in the MMA layer. Sometimes it is possible to achieve additional undercut by writing “undercut boxes”, low-dose features intended to expose the MMA layer without fully exposing the PMMA above it. However, there are limitations to this method. In particular, if the size of the feature in the top PMMA layer is relatively small (a small area qubit junction, for example), the developer can only enter and leave through this small PMMA hole. If one needs a large adjacent undercut area, it is often difficult to get enough developer flow through the PMMA hole to remove all the MMA from the undercut region, and one ends up with an improperly developed sample, as seen in Figure 4.2(a) and (b). The problem can be solved by developing for a longer time, but unfortunately the PMMA will be overdeveloped as a result, resulting in poorly defined features and potentially causing collapse of small bridges.

The solution to the undercut conundrum is to use orthogonal resists, i.e. resists where different developers are required for each of the two layers, and neither one is developed by the other’s developer. In this way, one can write a pattern in the top layer and low-dose undercut boxes for the bottom layer, develop the top for a short time so it comes

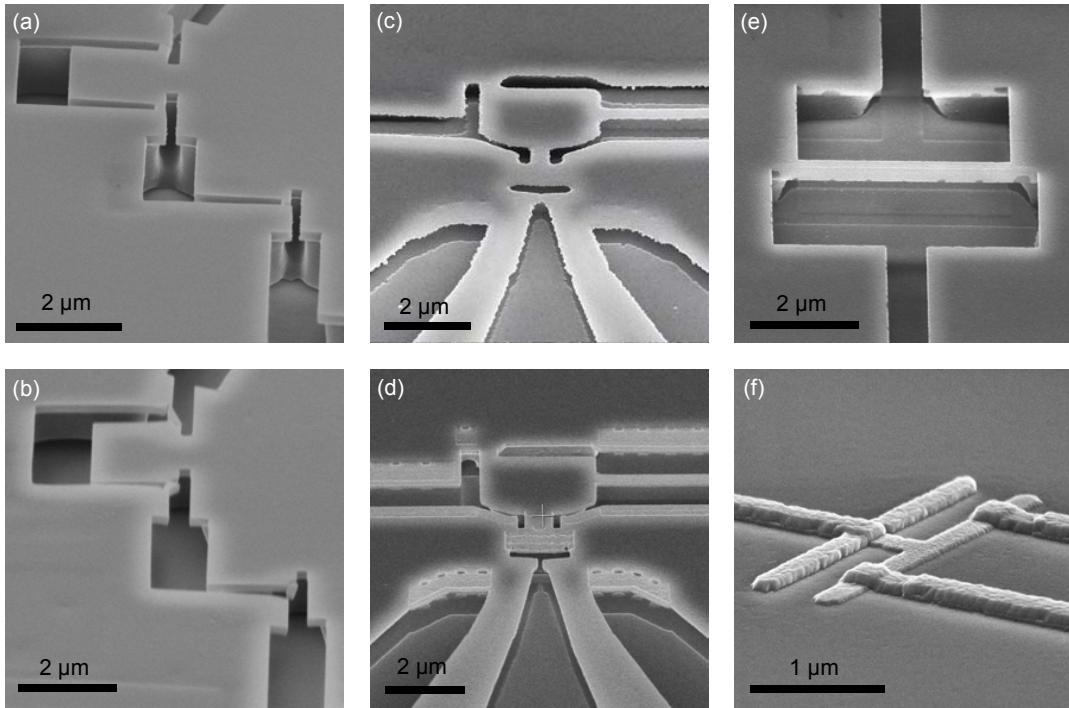


Figure 4.2: Resist mask troubleshooting.

Part (a) shows insufficient underlayer development due to small top layer features. This problem is solved by using orthogonal resists and developing the underlayer for a longer time. Part (b) gives an example of overdevelopment and “resist pulling” breaking resist bridges. Both (a) and (b) are PMMA/PMGI bilayers. In part (c), the resist appears to have a development or focus issue, but this is actually due to writing with coarse scan coils ($\text{mag} < 1300x$). The same mask is written under the same conditions in part (d), except this time at a sufficiently high magnification, and shows much better feature resolution. Part (e) shows a properly exposed mask with a double-angle evaporated junction. Note the freestanding resist bridge and the undercut, described further in section 4.2. All the masks in (a)-(e) were coated with a few nm of sputtered gold or evaporated aluminum after development to improve contrast and reduce charging while imaging. Part (f) is a detailed view of transmon qubit junctions after liftoff and post-ashing. The two shifted shadows, characteristic of double-angle evaporation, are clearly visible, the first one thin and the second one thicker.

out nicely, and then let the sample sit in the developer for the bottom layer as long as it needs to for the undercut to be fully developed. This technique was reported by Cord and coworkers [115] using poly(methylglutarimide) (PMGI) as an underlayer and PMMA as the top layer. We tried this combination initially for making flux qubit samples, but the desired features did not come out reliably because the PMGI underlayer appeared to be too soft, or

else too prone to swelling and/or contracting during development. Because the flux qubit mask at the time involved tiny bridges (~ 150 nm across) for junctions, and these bridges were connected to freestanding islands 100 μm long and 20 μm wide (the qubit loops), even very small fractional shifts due to expansion or contraction of the underlayer during development could cause bridge breaking (see Figure 4.2). This “resist pulling” problem might be solved by trying a different PMGI developer (we used CD-30, an aqueous alkaline salt developer used for photoresists; others have used IPA/water developer [116]). However, even without swelling effects, the surface tension of the water in the developer was large enough to cause pulling on freestanding features during the post-development blow-dry stage. PMGI is also more complicated in that it doesn’t lift off in acetone like PMMA and MMA do, but requires its own liftoff step in heated n-methylpyrrolidone (NMP).

We then tried a second combination of orthogonal resists, returning to MMA as the underlayer and using ZEP-520A resist as the orthogonal top layer. ZEP is a relatively new high-resolution positive resist which uses anisole as a solvent (allowing it to be orthogonal to MMA for spinning) but uses n-amyl acetate (xylene can also be used) as developer and is not developed by 1:3 MIBK:IPA solution for low exposure doses (at higher doses this no longer holds [117]). ZEP thus provides the desired orthogonality to an MMA underlayer; ZEP also lifts off in acetone, making liftoff again a single-step process. The primary difficulty with ZEP is that features with nearby undercut boxes are often distorted, because ZEP is exposed with a lower dose than PMMA and so undercut boxes would overexpose the ZEP layer. To solve this issue, we turned to cold development.

4.1.2 Cold development

Most groups working on superconducting qubits have used room-temperature development for their lithography, and it has worked fine. However, it has been shown that developing resists at lower temperatures can lead to improved resolution, allowing sub-10 nm features [118, 119, 120]. The theory for cold development is as follows. Electron-beam resist is composed of long-chain linear polymers, which are cut into shorter pieces (“chain scission”) by the electrons in the beam used to expose the resist. The shorter pieces dissolve much more easily in the developer than the longer ones, and so areas where the resist has been exposed to the beam are washed away [121]. However, longer uncut polymers can also be dissolved into the developer, albeit at a considerably slower rate than the cut pieces, allowing development to occur in areas where the resist has not been exposed. Since the dissolving process is thermally activated, developing at lower temperature exponentially reduces the rate at which unexposed resist is developed [122]. The contrast of the resist is increased as a result, giving sharper edges and finer features, even for resists with initially shorter polymer chains [123]. However, the dose must be correspondingly higher because the developer can only remove shorter and shorter pieces of resist as the temperature goes down. The higher doses also improve resolution by reducing dose variation due to shot noise in the beam current [124]. As the dose continues to increase, however, the cut polymer pieces start to reattach in a spiderweb or dendritic pattern, an effect known as “cross-linking”. The cross-linked polymer pieces, owing to their complex shape, are very insoluble in the developer and will not develop.

Since cross-linking ruins lithographic patterns, one cannot simply keep developing

at lower and lower temperatures, and there is an optimal temperature where the benefits of higher dose and reduced development thermal activation are balanced by the onset of cross-linking. For PMMA developed in 1:3 MIBK:IPA, this optimal temperature is about -15° C [120]. We designed a rig to develop resist at this temperature, which enabled us to fabricate superconducting nanobridge junctions for use as magnetometers and amplifiers [125, 59]. The system uses an immersion chiller to cool a bath of isopropanol in a Dewar flask to the appropriate temperature; a magnetic stirring rod circulates the isopropanol and keeps the temperature uniform. The beaker of developer is immersed in the isopropanol bath and held in place by a special bracket.

For ZEP, cold development at the temperature of an ice water bath (1-2° C) is sufficient for a substantial improvement in resolution and edge quality [126]. Importantly for our processes, the cold development also pushes the ZEP dose sufficiently high that undercut boxes written for the MMA underlayer (which is still developed at room temperature, using an orthogonal developer) do not overexpose the ZEP top layer. The beaker of developer simply sits in a holder in an ice water bath, which does not require temperature control. With this arrangement, we perform a rapid cold develop of the ZEP layer, followed by a room temperature MMA develop which can be made as long as necessary to achieve the desired undercut.

4.2 Thin-film deposition

After the resist mask has been developed, it is metallized to form the superconducting circuit structures. There are several standard methods for depositing metal films at our disposal, thermal evaporation, e-beam evaporation, and sputtering. All three methods are performed in vacuum. Thermal and e-beam evaporation both melt the material to be deposited, which then evaporates and coats the substrate. In thermal evaporation, the material is held in a tungsten “boat” or coil which is heated by current flow until the material inside melts. E-beam evaporation works by steering a beam of high-voltage electrons generated by thermionic emission into the target material, which is locally heated and melts. In contrast with the melting-based evaporation techniques, sputtering involves hitting a target with high speed ions (often from an Ar plasma), physically ejecting atoms of the target material, which settle on all surfaces in the chamber including the substrate. Sputtering does not involve melting the target, which makes it suitable for depositing a broader variety of materials than evaporation. Evaporation methods are directional, sending material in a straight line from target to substrate, while sputtering is non-directional, creating a “cloud” of target particles which settle on all surfaces, both horizontal and vertical.

The devices in this thesis are made of aluminum and niobium on intrinsic silicon substrates. Aluminum can be deposited by any of the three techniques above, while niobium must be sputtered because its melting point is too high for efficient evaporation deposition. The junction fabrication techniques used in our samples require directional deposition, so the junctions are made using evaporated aluminum. It is possible to make junctions using only sputtering (allowing for the use of niobium junctions), but the process involves several lithographic steps and will not be discussed in detail. We do use sputtered niobium for ground planes for the paramp samples, as will be described in further detail in section

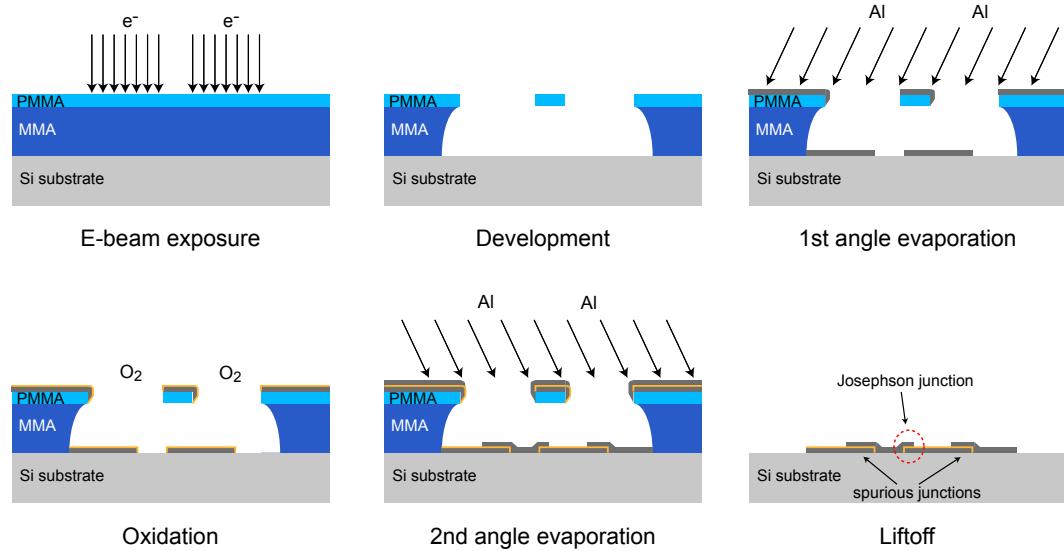


Figure 4.3: Double-angle evaporation.

Using a bilayer resist, one can fabricate Al-AlOx-Al Josephson junctions by performing double-angle evaporation, sometimes called the Dolan bridge technique. The layers are labeled as PMMA, MMA, and Si substrate, but in general they could be any suitable bilayer stack on a substrate. First, a bilayer resist stack is exposed using electron beam lithography (top left) so that a thin bridge is left unexposed. The resist is then developed chemically (top center), washing away the exposed top layer and underlayer and leaving a freestanding bridge, seen here end-on. Note the “scooped out” undercut. Next, the sample is placed in an evaporator and aluminum is deposited at an angle (top right), forming two regions on the substrate disconnected from each other by the shadow of the bridge. The sample is then exposed to an Ar/O₂ gas mixture (bottom left), causing an oxide layer (yellow) to grow on the deposited aluminum. A second evaporation is performed at a different angle (bottom center), such that the aluminum overlaps features on both sides of the original bridge shadow. Finally, the resist is lifted off (bottom right), leaving behind the circuit. The desired Josephson junction is formed by the Al-AlOx-Al tunnel barrier between the two bridge shadows, while several spurious Josephson junctions are formed by Al-AlOx-Al barriers elsewhere. As long as the areas (and thus the critical currents) of the spurious junctions are much larger than that of the intended junction, they will not affect the device behavior.

4.2.2.

Aluminum Josephson junctions are most simply made using the double-angle deposition technique [127], detailed in Figure 4.3. This method requires a freestanding bridge of top-layer resist separating two exposed areas with no resist. By evaporating aluminum at an angle from the substrate normal and perpendicular to the length of the bridge, a

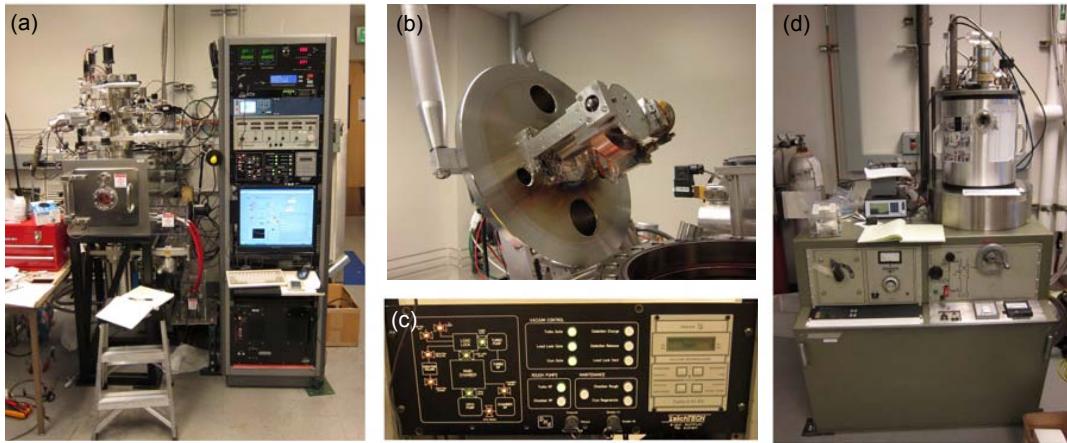


Figure 4.4: SlichTECH e-beam evaporator and NRC thermal evaporator.

The SlichTECH evaporator is shown in part (a), with its vacuum chamber and electronic control rack. Part (b) shows a detail view of the two-axis controllable sample stage with the load lock open. The vacuum control panel and turbo pump controller are shown in (c); the system can be controlled either manually or by computer. The old NRC thermal evaporator is seen in (d), with all manual valves. The NRC evaporator is durable but has inaccurate and often sticky angle control, angle-dependent thickness variations, and poorer base pressure and cycle time due to its single-chamber design.

shifted “shadow” of aluminum is deposited on the substrate. One can then let a controlled amount of O₂ gas (we use a mixture of Ar and O₂) into the vacuum chamber, which causes a layer of aluminum oxide to form on the deposited aluminum. By controlling the exposure time and O₂ pressure, one can control the thickness of the oxide and thus the transparency of the tunnel junction to be formed. Finally, a second layer of aluminum is evaporated at a different angle (often the negative of the first angle), in such a way that the deposited film bridges the gap in the first aluminum layer formed by the bridge shadow. The bridge shadow thus defines one edge of the junction in the bottom layer and the other edge of the junction in the top layer, and the angles determine the amount of overlap between layers (the junction area). Double angle evaporation can be used to make junctions of various sizes ranging from 100 nm or less on a side up to about 4 μm by 1 μm. Junctions larger than this are difficult to make because they require a very long freestanding bridge, which is difficult to achieve reliably, and so are typically made using different methods.

As part of the fabrication work, we designed, built, and commissioned a new evaporator system with angle control and oxidation capabilities to use for device fabrication. The evaporator was christened the “SlichTECH” evaporator, in homage to Joe Aumentado’s “JoeTEK” system which was the inspiration for the design. The SlichTECH is an e-beam evaporation system with a load lock for fast sample cycling and improved base pressure, precisely controllable oxidation pressure, two-axis angle control for the sample, and an Ar ion mill for cleaning surfaces and removing native oxides. It was built to replace an old NRC

thermal evaporation system, which had relatively poor base pressure, unreliable thickness monitoring, and somewhat inaccurate angle and oxidation capabilities. Photos of the two systems are shown in Figure 4.4. The SlichTECH allows improved sample throughput and parameter control, and is automated for use with computer control.

4.2.1 Oxygen plasma cleaning

One of the great frustrations of junction fabrication is the tendency of junction critical currents to decrease over time after fabrication (a phenomenon known as “aging”). Some amount of junction aging is unavoidable, because the oxide barrier in the tunnel junction has some ability to diffuse, and because water and oxygen from room air can migrate into the barrier and change its transparency. Good junctions will age by 10% or less before reaching a steady state; oftentimes, though, junctions can age by 50-100%, making qubit fabrication very difficult¹. In addition, we often observe substantial run-to-run variation in junction critical currents where the fabrication, evaporation, and oxidation parameters were nominally the same. Both this variation and excessive aging present major headaches for getting good samples.

We tried several techniques to combat aging, such as performing controlled sample post-oxidations in vacuum before venting the chamber and designing junctions where the tunnel barrier is overlapped on three sides by aluminum to reduce exposure to air. These methods all met with limited success. Aging can be improved somewhat by using junctions with a lower critical current density (thicker tunnel barrier), but for many applications there is not a great deal of freedom on this axis. However, we eventually determined that the best method for controlling junction aging was to perform a brief (10-20 second) O₂ plasma clean at low power (≈ 10 W at 100 kHz) after development and before evaporation. The theory is that any remaining resist residue on the substrate which was not removed by the developer may contribute to aging, and that the plasma cleaning removes this without damaging the mask. The oxygen plasma is generated by a capacitively coupled parallel plate source, and this means that the plasma etching is directional, preferentially etching vertically. Junctions made with this “pre-ash” technique tend to show consistent aging of about 10% or less. In addition, the run-to-run variation in critical currents is greatly reduced, suggesting that resist residue on the substrate (and perhaps outgassing from that resist) was a cause of absolute variability as well as excessive aging. Recent work in other groups agrees with the hypothesis of leftover resist residue and our observation that pre-ashing improves junction aging and variability [129, 130].

We can also perform oxygen plasma cleaning after the liftoff step, dubbed “post-ashing”. This cleans the substrate surface of any remaining resist residue not removed chemically, which may be a source of loss in qubits and resonators. Post-ashing also has been successful at removing the “black veil of death” [64], theorized to be an aluminum or aluminum oxide film that collapses from the resist sidewalls onto the device and may be a source of dielectric loss. Before and after pictures of post-ashed samples are shown in Figure 4.5. There has been anecdotal evidence from our lab that both pre- and post-ashing

¹Flux qubits are particularly sensitive to aging, as the frequency of a three-junction flux qubit depends exponentially on the critical current of the small qubit junction [128].

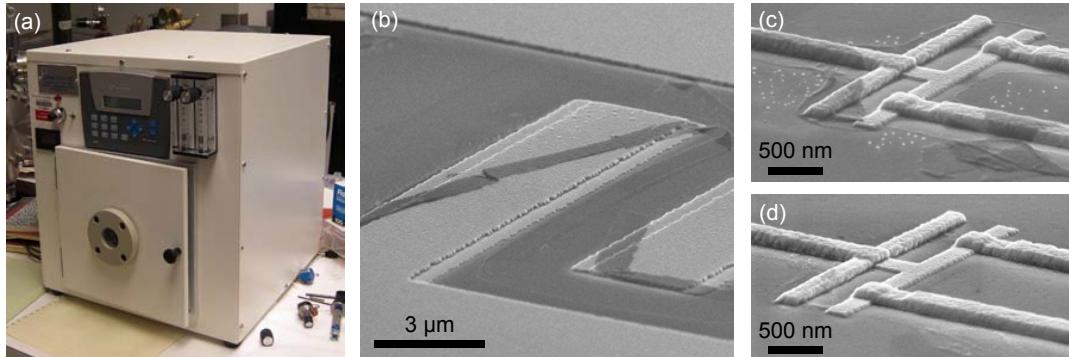


Figure 4.5: O_2 plasma cleaning to remove the “black veil of death”.

Our 100 kHz parallel plate O_2 plasma ashing system (a) is used for pre- and post-ashing of samples. Many samples show pieces of black film after liftoff, nicknamed the “black veil of death” by the Yale group. Part (b) shows “black veil of death” lying on the fingers of a transmon capacitor. The film may be a source of dielectric loss and decoherence in qubits and resonators. A brief oxygen plasma cleaning appears to remove the “black veil of death” without affecting the critical currents of junctions on the sample, as seen in the before (c) and after (d) pictures of transmon junctions.

improve the internal Q of resonators, which should correlate with improved qubit coherence times. We have not yet done systematic tests with qubits to study this effect, but anecdotal evidence from our group suggests that post-ashing can improve T_1 and T_2 noticeably in transmon qubits.

4.2.2 Ground plane fabrication

The paramps require large capacitance (several pF), which we realize using two parallel plate capacitors in series. Each capacitor consists of an evaporated aluminum top plate, a sputtered niobium bottom plate, and a thin-film silicon nitride (SiN_x) dielectric in between. The two capacitors share a common niobium bottom plate, which we refer to as a ground plane², thus forming two capacitors in series. This design eliminates the need for vias through the dielectric layer and simplifies fabrication.

The niobium ground planes are fabricated by electron beam lithography and sputtering in a liftoff process. The e-beam lithography is done in a standard PMMA/MMA bilayer, which gives a substantial natural undercut. When sputtered, this resist profile gives rise to a slow, smooth slope at the edge of the ground plane, rising from the surface of the substrate to the ~ 250 nm thickness of the ground plane over about 3–5 μm . This gentle “shoulder” is important to ensure continuity of the silicon nitride and aluminum

²This capacitor style was initially used in the original JBA designs at Yale [108], and the terminology comes from the fact that under differential excitation of the resonator, the niobium bottom plate is a virtual ground.

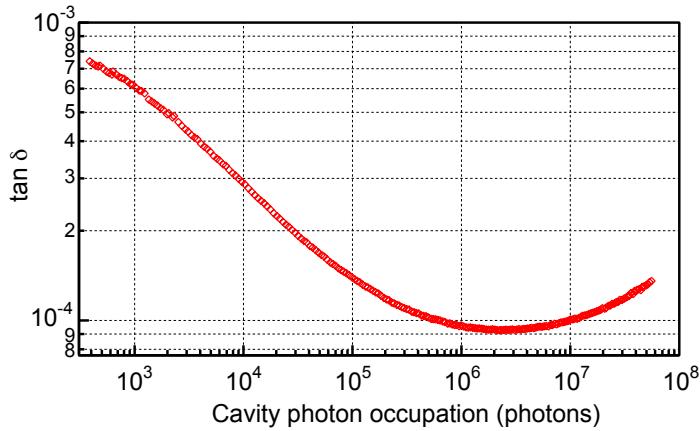


Figure 4.6: Measured loss tangent of microlab SiN_x .

Loss tangent of microlab-grown silicon nitride capacitor dielectric extracted from resonator data. The loss tangent decreases with increasing power, consistent with loss from two-level systems [131]. At very high powers, the loss tangent starts to increase again, perhaps from higher order effects such as local heating or quasiparticle generation.

layers deposited on top, preventing shorts to ground through the nitride and open circuit breaks in the aluminum. Too thick a layer of niobium or too high a dose on the resist mask can cause niobium to climb the resist sidewalls or connect over the top of the resist, ruining the needed shoulder and making liftoff difficult.

After the niobium ground planes have been fabricated, they are coated with SiN_x by plasma enhanced chemical vapor deposition (PECVD). We use the oxford2 machine in the UC Berkeley Microlab (now the Nanolab) to deposit a SiN_x layer between 120 and 180 nm thick with a relative dielectric constant of about 6.75. Thinner layers may not provide full coverage to prevent shorts, while thicker layers may give rise to too much stray inductance in the capacitor. We use SiN_x because its loss properties are superior to SiO_2 [132]. Nitride recipe details are provided in Appendix B.

The large capacitors made using these ground planes have fairly mediocre loss properties due to defects in the SiN_x dielectric. We made some linear resonators to test the loss tangent $\tan \delta$ of the nitride as a function of power. The results are shown in Figure 4.6. At low power $\tan \delta \approx 10^{-3}$, improving to $\tan \delta \approx 10^{-4}$ as the excitation power increases. This power dependence is expected if the loss is due to two-level defects in the dielectric [131]. The substantial loss tangent means that we can't use the parallel plate capacitors as part of a qubit, but they are suitable for use in paramps, which have a very low external Q and operate at higher power. The Maryland group have reported low power loss tangents of 2.5×10^{-5} in high stress nitride [133], more than an order of magnitude better than ours.

Recently, our group has developed a method for making ground planes with a single-crystal silicon dielectric using silicon-on-insulator (SOI) wafers. The loss tangent of these “SOI caps” is about 5×10^{-6} at single photon powers [134], making them suitable not

just for paramps but also perhaps for qubits. We are currently working to develop transmon and phase qubits using SOI cap shunting capacitances in the hopes of improving coherence times.

4.3 Sample design and parameters

The design process for qubit/cavity and paramp samples involves extensive computer simulation. The primary tool in our arsenal is AWR’s Microwave Office (MWO), which provides a suite of capabilities for simulating the behavior of microwave circuits. Every design starts with layout of capacitor and inductor designs for cavity and qubit, which we simulate using Microwave Office’s AXIEM electromagnetic (E-M) simulator. AXIEM does not require a conducting bounding box, giving more freedom in the circuit design for simulation, and runs considerably faster than the other MWO E-M solvers. AXIEM does a good job of capturing unexpected features such as spurious resonances, so it’s good to simulate as much of the circuit as possible to catch potential issues. For most designs, one can efficiently simulate the entire qubit/cavity circuit with input/output coupling and surrounding launch structures and get accurate values for the external quality factor Q_{ext} and resonant frequency ω_{cav} . To estimate the qubit-cavity coupling g , a SPICE extraction can be performed to extract the cavity capacitance (usually accurate to about 10%). Once a MWO design has the desired parameters, it can be exported directly into NPGS for lithography. Fast flux line couplings are usually estimated using FastHenry, an older, standalone program which gives estimates of mutual inductances.

4.3.1 Transmons

The energy structure of a transmon qubit is determined by the critical current of its Josephson junctions and the capacitance shunting them, as described in section 2.1.3. Using AXIEM, we are able to get a reasonable estimate for the shunting capacitance for a given design. We can then determine the junction critical current for a desired qubit frequency using Mathematica. Room temperature resistance measurements on co-fabricated “witness junctions” give us an estimate of the critical current of fabricated qubit junctions using the Ambegoakar-Baratoff relations [72]

$$I_0 = \frac{\pi\Delta(T)}{2eR_n} \tanh \left[\frac{\Delta(T)}{2k_B T} \right], \quad (4.1)$$

where I_0 is the critical current, $\Delta(T)$ is the temperature-dependent superconducting gap, R_n is the normal-state junction resistance, T is the temperature, and e is the elementary charge. The room-temperature junction resistance is a close approximation to R_n . Since we operate at 20-50 mK, $T \ll T_c$ (the superconducting transition temperature) and therefore $\Delta(T) \approx \Delta(0)$ and the tanh term ≈ 1 . Substituting these into (4.1) gives $I_0 \approx \pi\Delta(0)/2eR_n$. For aluminum, $\Delta(0) = 170$ meV.

Table 4.1 gives a listing of the transmon samples used in the quantum jump experiments. The values of E_J , E_C , and coherence times are all quoted at a single operating point which optimizes coherence while still allowing single-shot readout. All of the samples have $E_J/E_C \gg 1$ and so are well in the transmon regime.

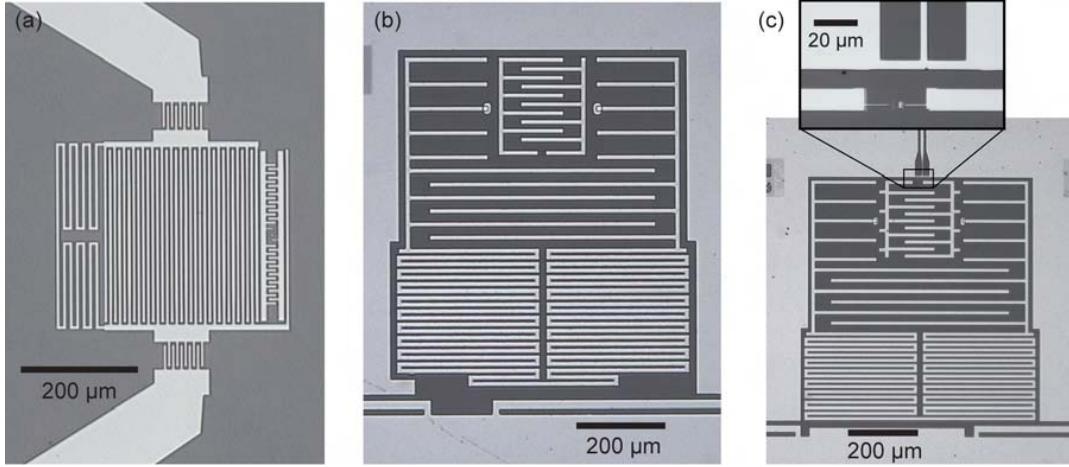


Figure 4.7: Transmon samples.

Optical images of three transmon samples and readout cavities. Panel (a) shows a reflection geometry cavity with capacitive coupling to the microwave feedlines, while panels (b) and (c) are transmission geometry cavities with inductive feedline coupling. The cavities and transmons in (b) and (c) have wider traces and greater spacing between traces compared to (a), which is intended to help reduce capacitive losses. The sample in (c) also has a weakly coupled fast flux line, used for the experiments described in Chapter 8. The inset shows a detail view of the fast flux line and the qubit loop.

Figure 4.7 shows optical images of several different designs for the transmon and its readout cavity, corresponding to the three samples listed in Table 4.1. The first sample, TF051310b, is the initial sample in which quantum jumps were observed and is pictured in part (a). It uses a capacitively coupled cavity in reflection geometry. To improve coherence times, we modified the design to have wider traces and larger gaps between traces, a technique we adopted from the Devoret group at Yale. The newer cavities, pictured in (b) and (c), operate in transmission and are inductively coupled to the microwave feedlines (rather than capacitively coupled). Spurious cross-coupling between the input and output feedlines gives the cavity resonance a Fano lineshape in this design, but this does not degrade per-

Sample name	E_J (GHz)	E_C (MHz)	T_1 (μs)	T_2 (μs)	Sample design
TF051310b	11.4	280	0.29	0.4	Fig. 4.7(a)
TF021111d	14.0	231	1.27	0.4	Fig. 4.7(b)
TF042811b	15.0	235	0.91	1.35	Fig. 4.7(c)

Table 4.1: List of transmon samples.

This table gives parameters for all the samples used in quantum jump experiments. All values are quoted at the best measured qubit operating point.

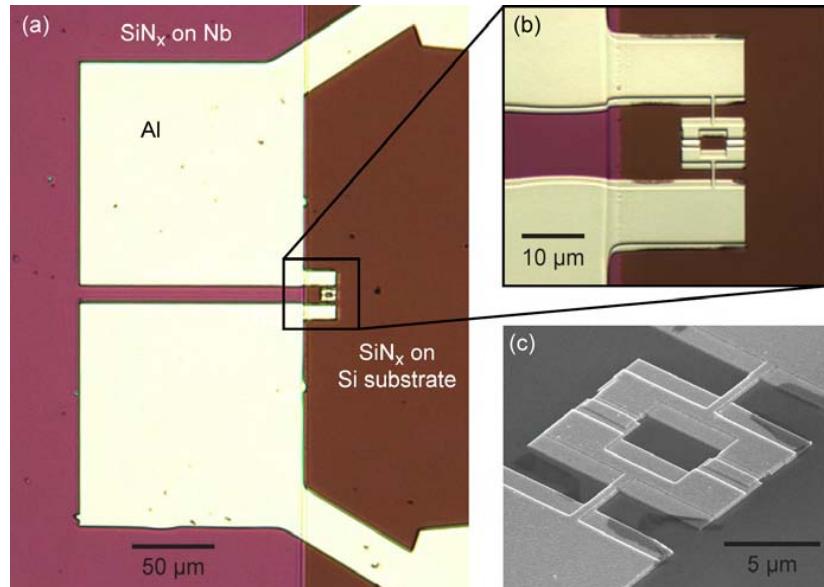


Figure 4.8: Paramp sample.

Panels (a) and (b) are optical micrographs of a typical paramp sample, showing the two large parallel plate capacitors formed by the Al top plate and Nb ground plane with intervening SiN_x dielectric, as well as the small SQUID loop. The slow sloping “apron” at the edge of the Nb ground plane can be seen in (b) as the few- μm -wide purple region to the right of the dark line. This slow slope allows the SiN_x and Al to transition onto the ground plane without breaking. The SQUID loop and junctions are shown in greater detail in the scanning electron micrograph in (c).

formance. For the samples used to study measurement backaction in Chapter 8, we added a weakly-coupled fast flux line as seen in (c).

4.3.2 Paramps

Paramp design is also relatively straightforward. The paramp has $Q \approx 20$ to give large bandwidth, which in turn sets the capacitance through the relationship $Q = \omega Z_0 C$, where ω is the resonant frequency and Z_0 is the microwave impedance of the launch (typically 100Ω for the differential launch used in our experiments and described in section 5.1). The capacitance can be estimated using the standard parallel plate formula $C = A\epsilon_0\epsilon_r/d$, where A is the area of the capacitor plate, d is the dielectric thickness (measured in the Nanolab using interferometry), ϵ_0 is the permittivity of free space, and ϵ_r is the relative permittivity of the dielectric (experimentally, 6.75 for our SiN_x). Once the capacitance has been determined, the resonant frequency is then set by the Josephson inductance of the SQUID junctions $L_J = \Phi_0/2\pi I_c$, where I_c is the flux-dependent SQUID critical current. For a typical paramp sample, $C \approx 6 \text{ pF}$ and $I_c(\Phi = 0) \approx 4 \mu\text{A}$.

Chapter 5

Experimental apparatus

All the experiments in this thesis were carried out in a VeriCold dry dilution refrigerator with a base temperature of 30 - 50 mK. These low temperatures are required so that the thermal energy $k_B T$ is well below the qubit energy splitting $\hbar\omega_q$ and the energy of the readout photons $\hbar\omega_{ro}$. We use a variety of microwave components, both inside and outside the fridge, to manipulate our qubit and readout signals. The VeriCold fridge has five temperature stages, each with a large gold-plated copper plate with tapped holes for mounting components and experimental samples. The two warmest temperature stages are nominally at 70 K and 4 K¹ and are cooled by a two-stage pulse tube refrigerator. The next three stages are the still, a plate between the continuous and discrete heat exchangers, and the mixing chamber. The base temperature of the fridge on delivery was around 11 mK, but due to changes in the mix balance and the heat load of the fridge wiring, the base temperature for these experiments is somewhat higher.

5.1 Sample boxes and launches

We use our own customized microwave launches to connect the paramp and qubit samples to the microwave cabling in the fridge. For the paramps (and initially the qubit/cavity samples as well), we use a modified 6 GHz microstrip 180° rat-race hybrid [135] to convert single-ended microwave signals on the lines into differential excitations at the sample. The circuit design, christened the “Bollywood” launch because of a resemblance to a Bollywood dancer, is detailed in Michael Hatridge’s Ph.D. thesis [136]. Because of its differential output, the Bollywood presents a 100 Ω impedance (rather than 50 Ω) to the sample, which helps with the simultaneous optimization of paramp dynamic range and stability. The initial quantum jump experiments were performed with a Bollywood launch for the qubit/cavity sample as well. However, the Bollywood is unshielded, and we discovered that the qubits on the Bollywood launch suffered from spurious radiative coupling as a result.

For subsequent qubit experiments, we used two different styles of enclosed sample boxes. The first style (“A”) uses edge-launch connectors (Southwest Microwave 292-04A-5)

¹The original pulse tube began to fail around the time of the initial quantum jump experiments, and these stages had temperatures closer to 85 K and 6 K respectively. A retrofit with a new pulse tube has since brought the temperatures back down to 55 K and 4 K.

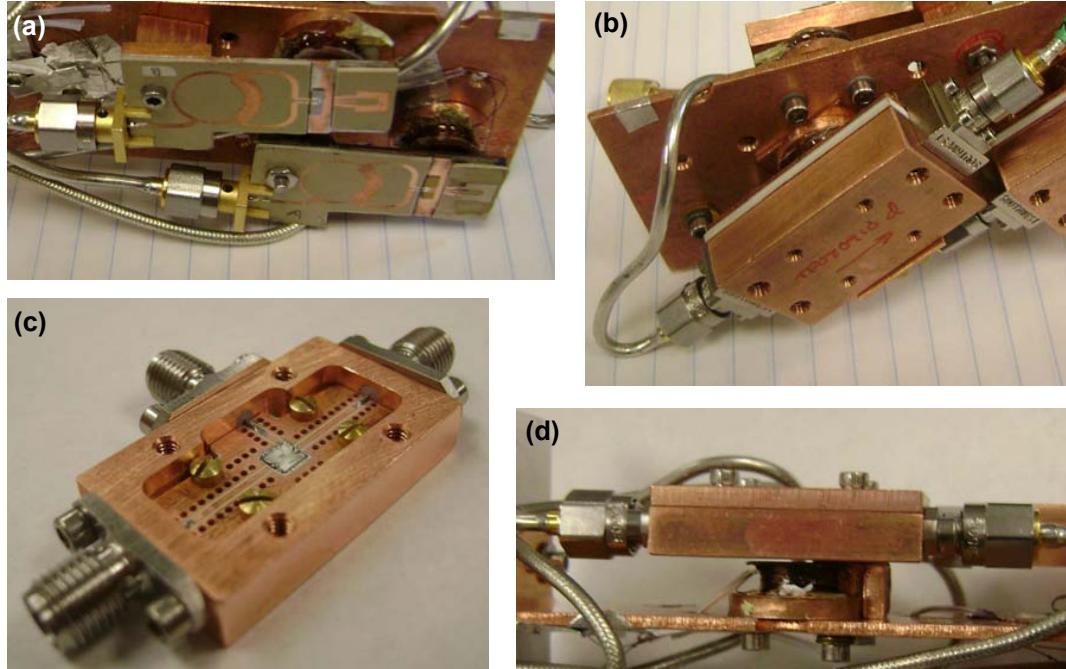


Figure 5.1: Sample boxes and launches.

Panel (a) shows Bollywood launches with samples and flux bias coils. Panel (b) shows a type “A” box with Southwest Microwave connectors and a bias coil. The newer type “B” box with bulkhead connectors and a third port for a fast flux line is seen in panels (c) and (d), with the latter showing the box closed and with a flux bias coil underneath.

on a microwave circuit board and has a clamshell-like lid which can be screwed down on top to enclose the sample. This provides considerably more shielding than the open Bollywood, but the lid does not make tight contact on all surfaces of the edge-launch connectors and so some stray radiation could still enter. The second style of enclosed sample box (“B”) is a milled-out copper block with a flat lid which makes a very good flush seal (this seal can be augmented with indium wire or foil if desired). The microwave circuit board inside is held down by screws, and signals are launched by SMA bulkhead connectors designed for this application (AEP/Radiall 9308-9113-001). Both styles A and B use circuit boards with grounded CPW microwave traces and vias connecting the top and bottom grounds; a recessed square was milled partway through the board to accommodate the sample chips flush with the board surface. Versions of the style B sample box and board were also made with an additional SMA connector for launching fast flux signals, as used in the experiments in Chapter 8. Photographs of Bollywood, “A”-style, and “B”-style launches are shown in Figure 5.1.

The qubit and paramp samples boxes and launches are placed inside fully enclosed superconducting aluminum shields, with dc wires, coaxial cables, and thermalization straps

penetrating the shield through long, narrow holes. This design helps shield environmental noise and reduces trapped flux in the devices by shielding static magnetic fields. Each superconducting shield is surrounded by a Cryoperm can to keep the shield itself from trapping magnetic flux as it cools through T_c . We also place a three-layer mu-metal shield around the outside of the dilution fridge during cooldown and operation to reduce external static and low-frequency magnetic fields at the samples. With this shielding, the primary source of static magnetic fields at the samples is from the permanent magnets in the microwave circulators at base temperature; we try to keep the circulators far from the samples to reduce flux trapping. We also discovered that the superconducting shield boxes mediate spurious radiative coupling between samples inside them, especially those on Bollywood launches. To combat this effect, we put the paramps (which are biased with strong microwave tones) in a separate superconducting shield from the qubit samples, and use enclosed qubit sample boxes as described above.

5.2 Fridge wiring

The experiments in this thesis require extensive electrical wiring and components inside the dilution refrigerator. Microwave-frequency signals are sent over coaxial lines, while low-frequency signals (e.g. for flux biasing) are sent via filtered twisted pairs. We use a number of microwave switches on the base stage, allowing us to try various different experiments in a single cooldown without the increased heat load and decreased space that come with adding more microwave lines. We have also implemented a setup for calibrating amplifier noise performance. Complete schematics of the fridge wiring can be found in Appendix C.

5.2.1 Low-frequency wiring

The fridge has two sets of 16 filtered low-frequency lines, one set for low-current applications with manganin wires and another set for high-current applications with Cu wires from room temperature to 4 K and superconducting CuNi-clad NbTi wires from 4 K to base temperature. In this thesis, these low-frequency lines are used to establish a dc magnetic flux bias in the paramp SQUID loop or the qubit loop, and to operate the hot-cold load switch described in section 5.2.4.

Both the manganin and the Cu/superconducting lines are well shielded and are filtered in several stages. All signals pass through a series of LC π filters (with ~ 1 MHz cutoff frequency) at room temperature, enter the fridge through a hermetic LEMO connector, and continue to further LC filters at 4 K; the signals are carried on twisted-pair looms (Oxford Instruments A8-311, A8-312, A8-313) inside a flexible stainless steel bellows for shielding. Below the 4 K filter, the wires are broken out into eight channels, each with two twisted pairs. Each channel's wires are carried inside a separate braided shield with four-pin Reichenbach connectors. Each twisted pair passes through three further stages of Cu powder filters [7] at 700 mK, 100 mK, and 30 mK. The Cu powder filters remove high-frequency noise from \sim MHz to \sim THz.

We use off-chip superconducting coils made of CuNi- or Cu-clad NbTi wire wound

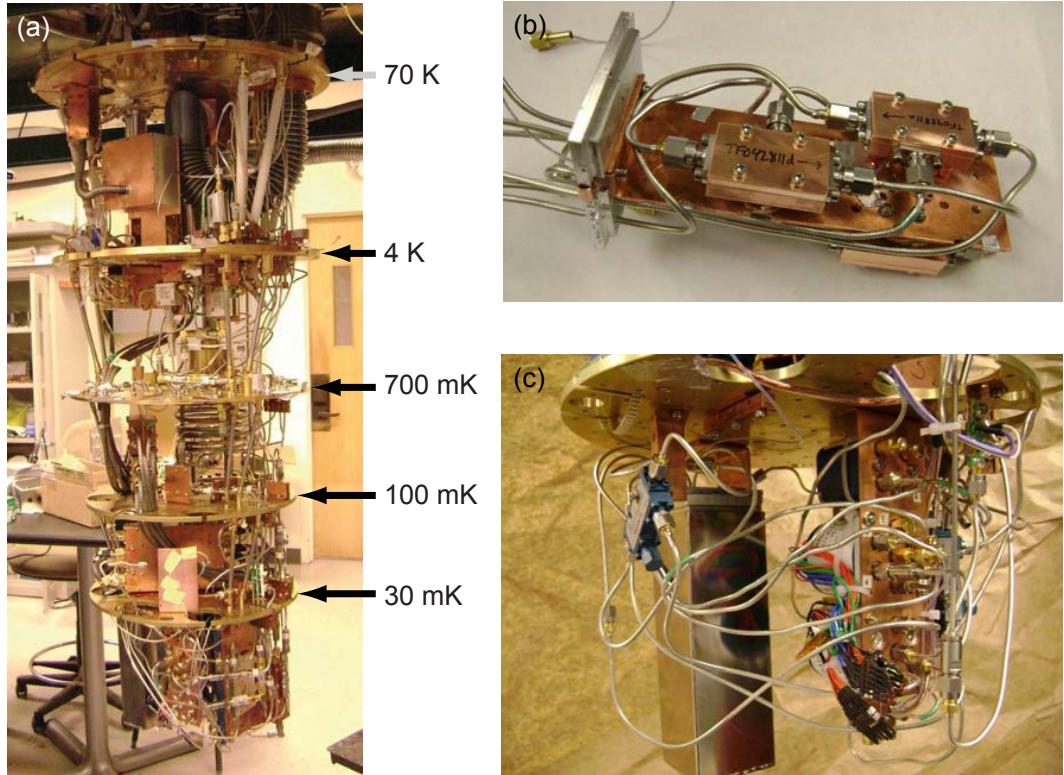


Figure 5.2: VeriCold dilution refrigerator and sample boxes.

Panel (a) is a photograph of the VeriCold fridge, fully wired, with each of the stages labeled by its temperature during operation. The bottom plate is about 25 cm in diameter, ample room for multiple experiments. Panel (b) shows several “B”-style sample boxes mounted to a copper cold finger and connected to captive microwave cables through an aluminum lid. This assembly slides into an aluminum box which screws to the lid to create a continuous superconducting shield. A Cryoperm shield is then placed around the entire assembly to minimize flux trapping. The assembled shielded box can be seen at left in part (c), with the captive microwave cables connected to the microwave switches visible at right.

on Cu bobbins (see figure 5.1) to provide dc flux bias for the paramp and qubit. The coils have 240 turns, and can operate at currents up to ± 15 mA before the mixing chamber plate starts to heat up due to dissipation in the contact resistance of the Reichenbach connectors. The current per flux quantum varies depending on the size of the loop and the sample-coil distance, but generally runs between 5 and 50 mA/ Φ_0 . This is a relatively weak coupling, the lines are well-filtered, and the superconducting coil acts as a giant LC filter, so generally one does not have to be too concerned about flux noise arising from current noise in the coil. However, the paramp is also a very sensitive magnetometer, so it is still important to be careful, as we discovered on one occasion. When biasing a paramp into the gain

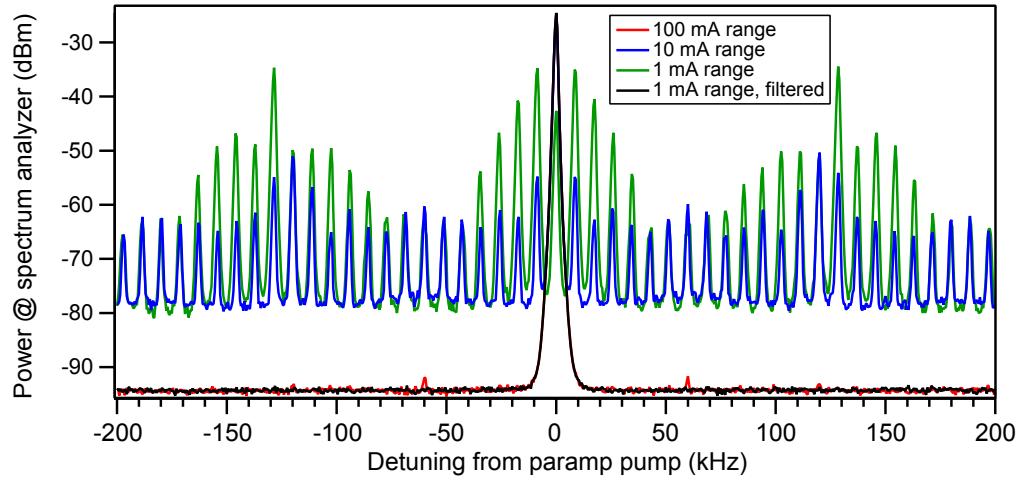


Figure 5.3: Flux bias noise from a Keithley sourcemeter.

These traces show the output frequency spectrum of a paramp biased in the high-gain regime using a Keithley 2400 sourcemeter to generate the current in the flux bias coil. Note many spectral noise components and how they decrease as the Keithley's output range is increased. With a low pass filter ($f_{3\text{dB}} = 0.7$ Hz) on the Keithley output (black trace), the noise is totally removed (green trace shows same range without filter).

regime, we saw unexpected jumping on the output trace. After extensive debugging, we determined that the jumping was due to the paramp amplifying noise from the Keithley 2400 sourcemeter used to establish the dc flux bias. The paramp was acting as a noise sensor, detecting many sharp harmonics in the spectrum of the Keithley output, as shown in Figure 5.3. We found that changing ranges on the Keithley (while outputting the same current) changed the noise output, and we finally solved the problem by putting a big RC lowpass filter with $f_{3\text{dB}} = 0.7$ Hz on the Keithley's output. The lesson is to filter all dc signals carefully, even for microwave circuits.

5.2.2 Microwave wiring and switches

The experiments in this thesis are carried out at millikelvin temperatures and microwave frequencies. This combination presents a challenge, because the high bandwidth coaxial lines needed for the microwave signals also admit thermal noise over a very broad band. Black body radiation from 300 K and 4 K have substantial spectral weight in the microwave regime which must be eliminated before they reach the qubit. Stray black body radiation can cause loss and decoherence in superconducting qubits and resonators [137, 138]. To remove black body radiation from input signals and thermalize them to the base temperature, the coaxial input lines are attenuated in several stages with broadband NiCr attenuators (XMA 2782-6051-10, -20). The attenuation values for each temperature stage are chosen to reduce input black body radiation from the next higher stage to a level

below that of the current stage. Attenuation, like gain, is typically expressed in logarithmic dB units in microwave circuits, with the conversion between linear and dB units given by $A_{\text{dB}} = 10 \log_{10} A_{\text{linear}}$. In addition to the attenuators, we include low-pass and high-pass reactive filters on microwave injection lines to prevent undesired out-of-band noise from reaching our devices. Above the 10-20 GHz range, though, these filters no longer function effectively and the attenuators and intrinsic loss of the coaxial cables must do the job alone. Fortunately, it appears that these are sufficient to thermalize microwave input lines [137].

It's also important to consider the heat load from microwave cables. Unlike low-frequency wires, which can be made very thin to reduce their thermal conductivity, coaxial cables must be of a certain size and geometry. As a result, the thermal conductivity needs to be reduced by an appropriate choice of materials [139]. The coaxial cables between stages (Micro-Coax UT-085B-SS, Coax Co. SC-219/50-SS-SS, SC-219/50-CN-CN, and SC-219/50-SCN-CN) are made of stainless steel, beryllium-copper alloy, or copper-nickel alloy, which have low thermal conductivity. The tradeoff for this reduced thermal conductivity is reduced electrical conductivity and therefore increased cable loss for signals. This is not problematic on injection lines, which are already heavily attenuated, but can be frustrating on output lines carrying weak signals from an experiment to the amplifiers. For cables between points on the same temperature stage thermal conductivity is not an issue, and we use tin-plated copper cables (Micro-Coax UT-085-TP) to minimize loss.

We use a cryogenic high electron mobility transistor (HEMT) amplifier (Low Noise Factory LNF-LNC4.8A) as our initial amplification stage (not counting the paramp). The HEMT amplifier operates over the 4-8 GHz band with a noise temperature of 2.6 K, equivalent to 9 photons of added noise at 6 GHz. This is currently the lowest noise commercially available amplifier in this frequency range. The HEMT is anchored to the 4 K plate of the fridge; unfortunately, it cannot be run at a colder stage due to its power dissipation of 4 mW. Because of this constraint on the HEMT's location, cable losses between the experiment and the HEMT (about 7 dB) significantly degrade the system noise temperature. We eliminated this problem by using superconducting Nb coaxial cables (Coax Co. SC-219/50-Nb-Nb), which have very low loss and very low thermal conductivity², to bring signals from the mixing chamber stage to the HEMT on the 4 K plate.

Our experiments rely on the ability to control the direction of propagation of microwave signals. For this, we use cryogenic 3-port circulators (Pamtech CTH1368K18). Circulators are passive devices which direct incident microwaves to a specific output port while preventing them from reaching any other ports. In a 3-port circulator, incoming signals are directed to output ports in the pattern $1 \rightarrow 2, 2 \rightarrow 3, 3 \rightarrow 1$, while propagation does not occur in the reverse directions $1 \not\rightarrow 3, 3 \not\rightarrow 2, 2 \not\rightarrow 1$. We can use this to separate the input and output signals from a reflection measurement, for example: signals enter at port 1, are directed out port 2, interact with the device under test (a paramp or resonator), reflect

²The Weidemann-Franz law, which states that the electrical conductivity and thermal conductivity are directly proportional to each other, holds for most materials at temperatures below $\sim 10\text{K}$. This is because electronic thermal conductivity is dominant over lattice thermal conductivity at these temperatures. In superconductors, the Cooper pairs have zero entropy and cannot carry heat, so only unpaired electrons can be involved in heat transport. The number of unpaired electrons decreases exponentially with temperature below T_c , so at temperatures well below T_c superconductors are simultaneously good thermal insulators and excellent electrical conductors.

back into port 2, and are directed out port 3. One can also make an isolator, essentially a microwave one-way valve, by terminating one port of the circulator with a matched 50 Ω load. Isolators allow signals out of a sensitive qubit experiment while protecting it from noise or signals which might propagate back down the output lines and cause undesired effects.

As is often the case, the experimental realization of a circulator cannot quite match its Platonic ideal. Real circulators have finite reverse isolation (the ratio of signal which propagates in the allowed versus the forbidden direction), in our case about 20 dB. To get larger isolations, we must use several circulators in series. Each circulator also has an insertion loss of about 0.5 dB, so we pay for the isolation they provide with decreased signal. In addition, real circulators have a finite bandwidth, usually an octave for commercial devices. For the experiments in this thesis, the circulators operate in the 4-8 GHz band. Outside of this band, the circulators are reflective. The Pamtech circulators are unique in that they operate well at low temperatures; in fact, their bandwidth, isolation, and return loss are worse at room temperature than at cryogenic temperatures. Since Pamtech is currently the only source of cryo-compatible circulators, non-Pamtech circulators should not be used at low temperatures because they may not function properly. Because circulators employ permanent magnets to break time reversal symmetry (thus allowing non-reciprocal behavior), one should take care in placing them to minimize the effect of their magnetic fields on qubit and paramp samples.

In addition to circulators, we can also achieve directionality—at the expense of attenuation—with directional couplers, which weakly cross-couple signals between two lines with a preferred direction of propagation for the coupled signals. As with circulators, the directionality is not perfect, and the directivity (ratio of preferred versus opposite direction propagation) is finite, usually about 30 dB. In addition, only a small fraction of the original signal gets coupled to the new line, giving an effective attenuation of typically 10 or 20 dB. However, because of their small size, broad bandwidth, and because they don't generate magnetic fields, directional couplers (Krytar 104020020 and similar) find many uses in our experiments, primarily for coupling in paramp pump tones.

We can multiplex several samples on a single input and output line using microwave switches and/or splitters at base temperature. The exact configuration of switches and splitters depends on the number of samples in the cooldown and whether they are measured in transmission or reflection. For a setup with all reflection samples, a single switch can be used for multiplexing. We use latching 3-pole and 6-pole single-throw switches (Radiall R573423300 and R573423600), which are slightly modified from their stock configuration to allow reliable operation at cryogenic temperatures. For samples in transmission, we need to multiplex both the input signal and the output signal. We accomplish this either by using one switch on the input and one on the output, or by using a four-way splitter (Anaren 44020) on the input. The isolation between switch channels is about 80 dB and the on-off ratio is better than 90 dB, so there is no issue with crosstalk from other samples coming through on the output. In addition, we use a latching DPDT transfer switch (Radiall R577433000) on the output line to switch the paramp and some of its attendant circulators into or out of the amplification chain. This method gives us improved noise temperature when we can't or don't want to use the paramp, and lets us control for the effects of the

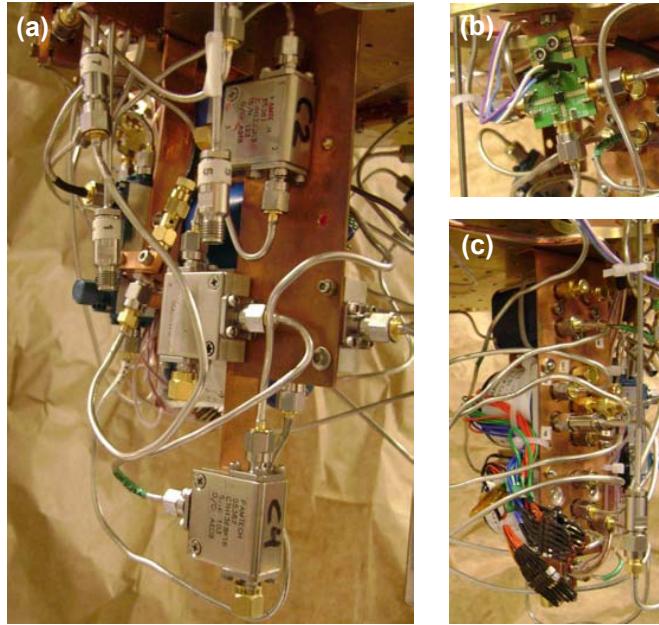


Figure 5.4: Circulators and microwave switches.

Several cryogenic 3-port circulators can be seen in (a), attached to a copper cold finger to ensure thermalization. Part (b) shows a cryo-compatible fast GaAs microwave switch (Hittite) on a circuit board. Three latching microwave switches (Radiall) attached to a copper cold finger are shown in panel (c). The large bundle of red, green, and blue wires carries control pulses from a superconducting loom to the switches.

paramp on the readout trace without having to temperature-cycle the fridge. The Radiall switches are shown mounted to the fridge in Figure 5.4(c).

Each switching operation of these Radiall switches dissipates a small amount of heat, warming the mixing chamber stage by a few tens of mK, but the fridge recovers fully to base temperature within around 10-15 minutes. Care must be taken to make a high-thermal-conductivity connection between the copper finger inside superconducting sample shields and the mixing chamber plate, or else the copper finger may take considerably longer to re-equilibrate after a switch is thrown. Since the switches are latching, and we only seldom operate them (usually they remain in a fixed configuration for days or weeks at a time), this brief heating does not affect the behavior of the samples.

For some purposes, such as the hot-cold load measurement mentioned in Section 5.2.4 and the filter thermalization measurements in Section 5.2.3 and Appendix A, we need a microwave switch that can be switched back and forth at kHz or higher frequencies, much faster than the mechanical Radiall switches can operate. For this we use solid state GaAs switches from Hittite Microwave (HMC 547 and HMC 194). These Hittite switches dissipate as little as $10 \mu\text{W}$ when in use and have nominal switching times of 5-25 ns. The switching

rate is limited in our experiment by the bandwidth of the filtered manganin low-frequency lines used for the control signals, but this could be easily addressed if faster switching was required. The main drawback of the Hittite switches is their insertion loss, which is around 2-3 dB at low temperature, so we only use them when the Radiall switches won't do. Figure 5.4(b) shows a Hittite switch on a microwave board mounted to the mixing chamber plate.

5.2.3 Microwave “roach motel” filters

The standard combination of attenuators and reactive filters used to thermalize injection lines and reject out-of-band signals (as described in the previous section) works well for most applications. However, there are some drawbacks which create a need for a new type of filter. First, it is impossible to pass substantial dc currents (more than a few tens of μA) through a microwave injection line at base temperature given the large attenuation values needed for thermalization. This is an issue for fast flux bias lines, which require the high bandwidth of a coaxial line but must be able to pass $\sim\text{mA}$ of current. One can replace the attenuators with low-pass filters, allowing dc current to pass through while removing the higher-frequency black body radiation propagating down the line. The problem here is that all reactive low-pass filters are re-entrant—with stopband attenuation decreasing to near zero at sufficiently high frequency (usually 20-40 GHz)—and will allow most black body radiation through [140]. Furthermore, reactive filters are reflective in their stopband, so out-of-band photons are not dissipated (as in an attenuator) but rather bounce around in the lines, perhaps finding another route into the experiment. This reflective behavior arises because the impedance of reactive filters is very far from 50Ω in their stopband; this presents yet another problem, because the qubit doesn't see a well-controlled impedance looking out into the fast flux line.

What we need for fast flux lines is a low-pass filter with low loss at dc and a stopband extending to hundreds of GHz/THz frequencies. The filter should present a matched 50Ω impedance in both passband and stopband, and should thermalize down to mK temperatures. We developed a suitable filter, which we christened the microwave “roach motel” or roach filter³. The roach filter is simply a stretch of lossy transmission line, an idea which has been around for quite some time [141, 142]. Lossy, impedance-matched transmission line filters for qubit experiments have been developed before [143, 144, 145], but these filters are fragile, difficult to fabricate, and/or have too low a corner frequency. The roach filter is a simple stripline [135, 146] with a magnetically loaded silicone dielectric (Emerson & Cuming Eccosorb MFS-117) inside a connectorized copper box, shown in Figure 5.5(a). The design is based on one originally used at Yale [108, 147], and was independently developed and published by Dan Santavicca and Dan Prober at Yale [148]. Roach filters are simple and quick to build and their electrical properties are robust to minor fabrication variations. By varying the length of the lossy stripline, one can change the filter cutoff frequency $f_{3\text{dB}}$ between a few hundred MHz and about 1.5 GHz. It should be possible to make higher cutoff frequencies through careful engineering of the copper box and perhaps by using different types of Eccosorb.

³The name derives from the impedance matching in the filter's stopband, which means stopband signals are admitted to the filter and dissipated inside rather than being reflected. This is analogous to the “roach motel” style of cockroach trap, where “the roaches check in, but they don't check out”.

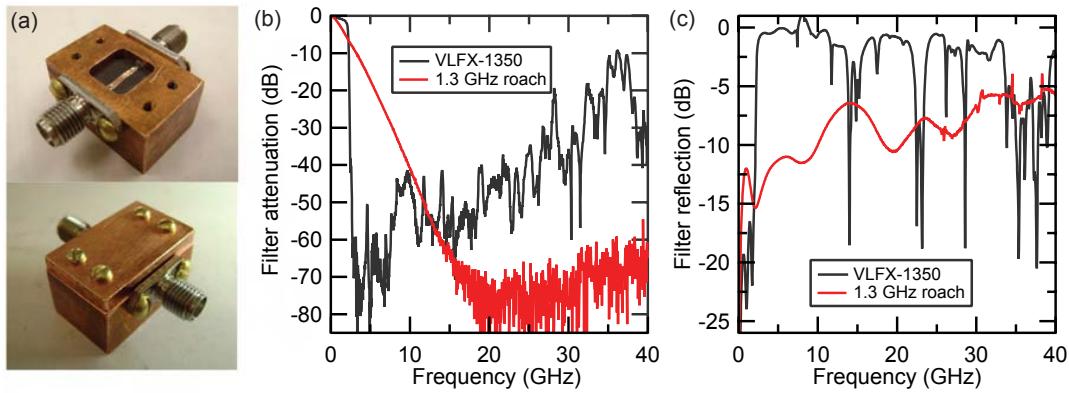


Figure 5.5: Microwave “roach motel” filters.

Part (a) shows a roach filter with $f_{3\text{dB}} = 1.3 \text{ GHz}$. The copper strip center conductor is sandwiched between two pieces of Eccosorb (the top piece has been removed here) and then a lid is screwed on. The two graphs show a comparison of filter attenuation (b) and reflection (c) between a 1.3 GHz roach and a Mini-circuits VLFX-1350 reactive filter. The VLFX has a steeper stopband transition, but its attenuation wanes with increasing frequency to be only 10 dB at 40 GHz. By contrast, the roach filters attenuation is at the network analyzer’s noise floor at all frequencies above 20 GHz. In the stopband, the VLFX is very reflective, but the roach filter maintains its impedance match, with less than 10 dB reflection below 10 GHz and less than 6 dB reflection at 40 GHz.

The insertion loss of the roach filters reaches the noise floor of network analyzer measurements at around 15-20 GHz and remains there until at least 40 GHz, while the return loss is better than 10 dB below 10 GHz and better than 6 dB out to 40 GHz, as shown in Figure 5.5(b) and (c). This compares favorably with the parameters of commercial broadband high-attenuation reactive filters, such as the Mini-circuits VLFX-1350, which has a similar 3 dB corner frequency but becomes re-entrant in its stopband to the point where it only attenuates by 10 dB at 35 GHz. Based on absorption coefficient measurements reported in the literature [149, 150, 151], we expect the Eccosorb dielectric to remain lossy to THz frequencies.

The big remaining question is whether the Eccosorb dielectric works at dilution fridge temperatures. Eccosorb has been used extensively in the astrophysics community as a cryogenic black-body emitter for calibration loads in cosmic microwave background studies [149, 151, 152, 153, 154]. However, these studies were all carried out at temperatures above 1 K, so it was not known whether Eccosorb would thermalize properly and retain its loss characteristics at millikelvin temperatures. We performed experiments at 50 mK showing that the transmission and reflection properties of the roach filters were essentially unchanged from room temperature, and found that the black body noise power emitted by a roach filter in its stopband is equivalent to that emitted by a thermally anchored NiCr attenuator at the same temperature [155]. Details of the experiments and results are presented in Appendix A. With these data, we can confidently use roach filters on fast flux lines, knowing that they

will present a well-thermalized, matched impedance to the experiment and will protect the qubit from black body radiation from higher temperatures.

5.2.4 Hot/cold load setup

To calibrate the system noise temperature of our amplification chain, we use the noise emitted by two 50Ω microwave loads at different known temperatures; we refer to this as a “hot/cold load” setup. By measuring the total output noise of the amplification chain with each of the two loads, we can extract the noise added by the amplifiers and calculate an effective system noise temperature, a technique known as a Y-factor measurement. The method can be seen mathematically as follows: assign the hot and cold loads each known noise temperatures T_H and T_L . If the amplification chain has gain G and system noise temperature T_{sys} referred to the plane of the hot and cold loads, then the output noise power $P_{\text{out},H}$ and $P_{\text{out},L}$ are given by:

$$P_{\text{out},H} = k_B(T_H + T_{\text{sys}})BG \quad (5.1)$$

$$P_{\text{out},L} = k_B(T_L + T_{\text{sys}})BG \quad (5.2)$$

Using these two quantities, we can define the Y-factor $Y \equiv P_{\text{out},H}/P_{\text{out},L}$. Substituting in the previous results then yields:

$$Y = \frac{T_H + T_{\text{sys}}}{T_L + T_{\text{sys}}} \quad (5.3)$$

This quantity is a nice one to measure as it is independent of the gain G , which may be difficult to measure with high precision. Rearranging terms to solve for T_{sys} , we find that

$$T_{\text{sys}} = \frac{T_H - YT_L}{Y - 1} \quad (5.4)$$

In the above analysis, we have assumed that G and B are identical for the measurements of $P_{\text{out},H}$ and $P_{\text{out},L}$. This is certainly true for B , which can be defined with precision by a spectrum analyzer. However, G is subject to slow $1/f$ drifts in typical microwave amplification chains, so that if we wait long times between measuring $P_{\text{out},H}$ and $P_{\text{out},L}$ we may not be able to assume G is the same. To address the problem of gain drifts, we switch rapidly back and forth between the hot and cold loads and detect the output noise power synchronously. The detected Y-factor will then only depend on variations in G on timescales faster than the switching period.

Our experimental setup can be seen in Figure 5.6. We use 50Ω terminations on the 4 K plate and the mixing chamber plate as our hot and cold loads. The noise from the hot load is sent to the mixing chamber plate through a Nb coax to minimize attenuation, which lets us better estimate the noise at base temperature from the hot load. We chop back and forth between the hot and cold loads at 100 Hz using a Hittite GaAs switch (as described in Section 5.2.2) and detect the output noise power synchronously using a spectrum analyzer. We put the spectrum analyzer in zero span time sweep mode, so that we see a square wave

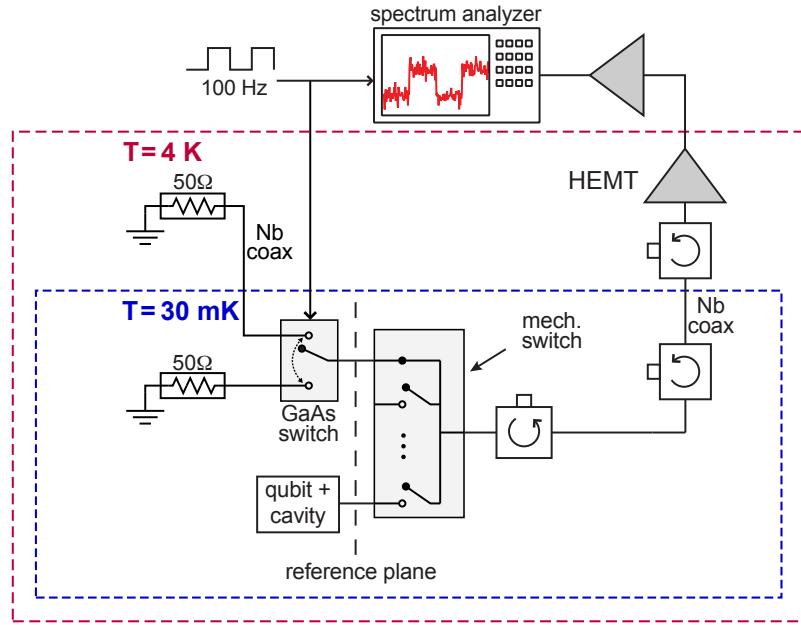


Figure 5.6: Hot/cold load setup for calibrating noise temperature.

Matched $50\ \Omega$ loads on the 4 K plate and mixing chamber plate serve as hot and cold loads for determining system noise temperature. A GaAs switch chops back and forth between hot and cold loads at 100 Hz, and the output noise power is synchronously detected on a spectrum analyzer operating in zero span time sweep mode. The hot load is connected to the mixing chamber plate by a Nb coax to reduce cable loss and give a higher and more accurately calibrated value of $T_{H,\text{eff}}$.

in time from switching between the output noise of the hot and cold loads. The Y-factor is simply the ratio of the powers at the high and low parts of this square wave; we can ignore the absolute power level (as long as it is high compared to the spectrum analyzer noise floor).

As I alluded to in the previous paragraph, we need to be careful in calibrating T_H and T_L in order to determine T_{sys} accurately. We know the physical temperatures of the hot and cold loads very accurately, because they are well thermally anchored to the 4 K and mixing chamber plates, whose temperatures we know from the calibrated fridge thermometry. However, there is attenuation between the loads and the reference plane of the measurement. An attenuation A (in linear power units) between the hot load and the reference plane will reduce the noise at the reference plane to an effective noise $T_{H,\text{eff}}$ according to

$$T_{H,\text{eff}} = T_H(1 - A) + T_{\text{att}}A, \quad (5.5)$$

where T_{att} is the noise temperature of the attenuation. This effect has a substantial con-

tribution in our setup because the Hittite switch has 2-3 dB of attenuation; in order to determine $T_{H,\text{eff}}$, we have to take separate calibration measurements of the Hittite switch insertion loss at base temperature (done on a different fridge cooldown). We have found that the attenuation can vary rapidly across the 4-8 GHz band by as much as 1 dB. As long as $T_{\text{att}} = T_L$, we need not apply the correction for the cold load.

I have so far been talking of the noise temperatures T_H and T_L of the hot and cold loads as being related to their physical temperatures. This is true in most regimes, but when we get to temperatures where $\hbar\omega \approx k_B T$, as is the case in our experiment, we need to look at the full quantum treatment. The symmetrized quantum power spectral density of noise from a matched microwave load can be determined using the fluctuation-dissipation theorem, and is given by [50]:

$$S(\omega) = \frac{\hbar\omega}{2} \coth \left[\frac{\hbar\omega}{2k_B T} \right]. \quad (5.6)$$

We can define an equivalent noise temperature T_{equiv} by dividing $S(\omega)$ by Boltzmann's constant. At high temperatures ($k_B T \gg \hbar\omega$), equation (5.6) reduces to the classic expression for Johnson noise, $S(\omega) = k_B T$ and we get $T_{\text{equiv}} = T$. For our hot load, we are in this regime and so T_H is accurately given by the physical temperature of the 4 K plate.

However, for $k_B T \ll \hbar\omega$ we find that the noise power spectral density saturates to the value of the quantum zero point fluctuations $S(\omega) = \frac{1}{2}\hbar\omega$, and we have $T_{\text{equiv}} = \frac{\hbar\omega}{2k_B}$, no longer related to the physical temperature. For $\omega/2\pi = 6$ GHz, we have $T_{\text{equiv}} \approx 145$ mK. It is this T_{equiv} , and not the physical temperature, that we must use for T_L in our Y-factor expression. The value of T_{att} is also given by T_{equiv} and not the physical temperature of the attenuation. After these corrections, we can modify equation (5.4) for the system noise temperature to read

$$T_{\text{sys}} = \frac{[T_H(1 - A) + \frac{\hbar\omega}{2k_B} A] - Y \frac{\hbar\omega}{2k_B}}{Y - 1}, \quad (5.7)$$

where T_H is the physical temperature of the hot load on the 4 K plate, A is the attenuation of the Hittite switch, Y is the Y-factor, and $\omega/2\pi$ is the frequency at which we are measuring the system noise temperature.

5.3 Room temperature electronics

Our experiments also rely on precise generation and manipulation of microwave and RF signals using room temperature electronics. Schematics of typical room-temperature electronics setups can be found in Chapter 7 and Appendix C.

5.3.1 Pulse and tone generation

We generate our readout and qubit manipulation pulses, as well as our paramp pump tone, using commercial synthesized microwave generators (Agilent 8257C, 8257D). These generators have excellent phase stability, which is important to keep the relative phase between the paramp pump, the readout signal, and the local oscillator (LO) of the

demodulation circuit from drifting while operating the paramp in phase-sensitive mode. For less stable generators, the effect can be mitigated by splitting the output of one generator and using it for readout, pump, and demod LO; however, this is more technically demanding, as one then needs good methods for controlling the amplitudes of these signals independently over disparate ranges without impacting the others.

We shape our readout and qubit pulses by applying shaped pulse envelopes to the IF port(s) of a microwave mixer. Qubit pulses are shaped using the internal IQ mixer of an Agilent 8267C generator, using a 1 GS/s arbitrary waveform generator or “arb” (Tektronix AWG520) to provide the pulse envelope shape. Readout pulses are shaped using a Marki M10310MB mixer, using either a 50 MS/s arb (Agilent AG33220A) or the AWG520 to create the envelope pulse. For both readout and qubit pulses, carrier feedthrough is an issue; we don’t want microwaves impinging on the arbitrary waveform generator outputs (from which the carrier might reflect and be passed to the mixer output), nor do we want leakage to find its way down to the qubit and affect its state. We can solve the first problem by putting roach filters with 800 MHz corner frequencies (see section 5.2.3) on the lines between the arbs and the mixers. The second problem is solved by experimentally determining the dc offset voltage (usually a few tens of mV) at which the carrier feedthrough from LO to RF ports is lowest, and using that as the “off” voltage for the arb output. For the Marki M10310MB mixer, careful optimization of this offset voltage gives an on/off ratio of 39 dB, which for typical experimental parameters means the readout cavity photon population from leakage is below that due to thermal excitations at the base temperature. The offsets of the internal IQ mixer in the Agilent 8267C generator can be nulled using front panel controls to give an on/off ratio of 60 dB. If a better on/off ratio is needed for either pulse type, a Marki IQ-0307LXP IQ mixer can be used for pulse shaping; it can be nulled (using two channels of offset optimization) to better than 70 dB on/off ratio.

Because of the low anharmonicity of the transmon, it is important that the spectral content of qubit pulses not be too broad [76]. For applications requiring very fast qubit pulses, the pulse-shaping techniques described in [76] can compensate for effects of low anharmonicity, but the experiments in this thesis do not require such advanced methods. We use qubit pulses of a minimum 10 ns duration, with the pulse envelopes from the AWG520 shaped by convolution with a Gaussian filter in software. The readout pulse shapes are naturally filtered by the 20 ns rise time of the slower AG33220A arb, and don’t typically need additional software waveform shaping.

5.3.2 Variable attenuators

In addition to time-domain shaping of readout pulses, we need to be able to control their amplitude accurately over at least two orders in magnitude, ranging from very weak to very strong measurement. For pulses shaped with a mixer, this can be accomplished by changing the amplitude of the pulse envelope sent to the IF port (because mixers require a certain fixed LO power to operate, we cannot vary the output amplitude at the RF port simply by varying the LO amplitude). However, reducing the amplitude of the pulse envelope on the IF port has the undesired side effect of reducing the mixer’s on/off ratio by the same amount. One way around this problem is to have the mixer shape pulses with full amplitude IF pulse envelopes and then to attenuate the pulse after the mixer, thereby

preserving the on/off ratio while also giving control over the amplitude of the output pulse. Because we would like fine-grained, repeatable values of readout pulse amplitude, we opted to use digital variable attenuators.

We used two types of digital attenuators in our experiments, a Hittite HMC425 and a Vaunix Lab Brick LDA-602. The Hittite attenuator can be set to give any attenuation between 0 and 31.5 dB in 0.5 dB steps. The attenuator came on a connectorized microwave-optimized demonstration board, and we controlled and powered it using a USB digital I/O board (National Instruments USB-6501). For some experiments, though, the Hittite attenuator did not provide sufficient dynamic range for controlling readout powers, and so we switched to using the Lab Brick attenuator, which can be set to any attenuation between 0 and 63 dB in 0.5 dB steps. The attenuation at a given setting appears to be very stable, with less than 0.01 dB of variation over the course of an hour. The Lab Brick attenuator is only spec'ed to work to 6 GHz, but we tested it on a network analyzer and found that it works admirably to 10 GHz and beyond, with the caveat that as the frequency rises above 7 GHz the maximum attenuation saturates to a value less than the rated 63 dB. This maximum attenuation value changes steadily with frequency from 63 dB at 7 GHz to 43 dB for frequencies above 12 GHz. One additional wrinkle with the Lab Brick attenuators is that they appear to emit digital switching noise at low frequencies, potentially affecting the experiment. To combat this, we use inner-outer dc blocks on both sides of the attenuator, as well as a high pass filter (Mini-Circuits VHF-3800+) on the output.

5.3.3 Demodulation and digitization

The readout signals returning from the fridge carry information about the qubit state as a phase modulation on a roughly 6 GHz carrier. To strip the carrier and just keep track of the phase modulation (whose bandwidth is only $\sim 5\text{-}10$ MHz, set by the readout cavity linewidth $\kappa/2\pi$), we perform homodyne detection by mixing the readout signal with a local oscillator at the same frequency. We demodulate using an IQ mixer (Marki IQ-0307LXP), giving us information about both the amplitude and the phase of the readout signal. For qubit readout, we expect there to be two steady amplitude/phase combinations in the readout signal (corresponding to the two qubit states), with their vector difference defining a line in the I-Q plane (see Figure 2.10, e.g.). By shifting the relative phase of the LO and the readout signal, we can align this line with either the I or Q output quadratures, so that all the qubit state information in the demodulated output signal lies only in one of the two output quadratures. This makes post-processing simple, as we don't have to turn every I-Q pair into an amplitude-phase pair in software to see changes in the qubit state. To accomplish the desired phase shift, we use a phase shifter (Sage 6705-15) on the readout signal right before the mixer.

After demodulation, the signal at the I and Q ports of the mixer is still quite small ($\sim \mu\text{V-mV}$), too small for the resolution of our digitizer card. We boost the amplitude of each output using an ultra-low-noise OPA847 op-amp configured for noninverting voltage gain of 40 with 100 MHz 3 dB bandwidth. The OPA847s are each on their own circuit board (Texas Instruments DEM-OPA-SO-1B) configured for optimal high-frequency performance, including $50\ \Omega$ input and output impedance matching. Both the IQ mixer and the OPA847 following amplifier boards are contained in a metal shielding box with SMA

bulkhead connectors for all signals. The OPA847s are run off dual power supplies with ± 5 V regulators inside the shield box.

After post-amplification by the OPA847s, the demodulated I and Q signals are filtered to reduce noise bandwidth. We use either commercial filters (Mini-Circuits SLP-10.7+) or simple homemade single-pole RC filters (to give a slower rolloff), both with cutoff frequencies in the 5-10 MHz range. After filtering to remove excess noise, the signals are digitized by an 8-bit 2-channel 1 GS/s digitizer (Alazar ATS9870). The digitizer is typically run at 100 MS/s (10 ns per point), which gives a suitably large Nyquist bandwidth of 50 MHz to prevent aliasing. The data from each measurement trace are sent to LabVIEW using asynchronous direct memory access, which allows very high data throughput rates. LabVIEW collects the raw data from each of the two channels and saves it to disk as an array of unsigned 8-bit binary numbers. The digitizer is a PCI Express card, capable of nominal data throughput of 1.4 GB/s, but the process of sending the data through LabVIEW slows down the transfer rate considerably. However, we are able to acquire and save data to the hard drive at a maximum rate of roughly 100 MB/s, corresponding to a roughly 50% duty cycle for data acquisition. We use eSATA hard drives for data storage in order to keep up with this throughput.

Chapter 6

Calibration experiments

This chapter details the experimental realization of the lumped Josephson parametric amplifier (LJPA), or paramp, described in Chapter 3. It also details the calibration procedures for the transmon qubit and its readout cavity.

The paramp is essential to the work presented in subsequent chapters, so it is important that we understand its behavior experimentally. We discuss the techniques for testing and biasing the paramp to its operating points. We also present measurements of amplifier gain, bandwidth, and noise. The data shown in this chapter are taken from several different paramp samples at a variety of operating frequencies, and are selected to provide representative examples of paramp behavior. We tested and used five different paramp samples over the course of the work presented in this thesis, and all five samples displayed very similar gain, noise performance, and tunability.

The techniques for calibrating the qubit-cavity system are general, and we will only sketch them. These calibration methods are used for all the qubit samples in this thesis at every bias point to which they are tuned.

6.1 Paramp sample packaging

The paramp is a low- Q nonlinear microwave resonator measured in reflection. We use Bollywood 180° hybrid microwave board (see Chapter 5) to launch signals to our paramp. The Bollywood turns incoming single-ended microwave signals into a differential excitation and presents a $100\ \Omega$ impedance to the paramp. The paramp is directly coupled to the launch, with no on-chip coupling capacitors, so this impedance sets the Q of the paramp resonance. The differential excitation means that the ground plane used in the split-geometry paramp capacitors is a true virtual ground and so will not radiate. The launch is designed so that the sample is glued directly to a copper surface to aid in thermalization.

The paramp requires a flux bias for tuning, which we generate using an off-chip superconducting coil wound on a copper bobbin. The coil generates fields of up to a few mT at the paramp. The paramp amplifies microwave signals but also acts as a very sensitive low-frequency magnetometer—exhibiting some of the highest flux sensitivity seen to date experimentally [59]—so it is very important to shield it from external magnetic fields. We place the paramp and its flux bias coil inside a superconducting aluminum box to reduce

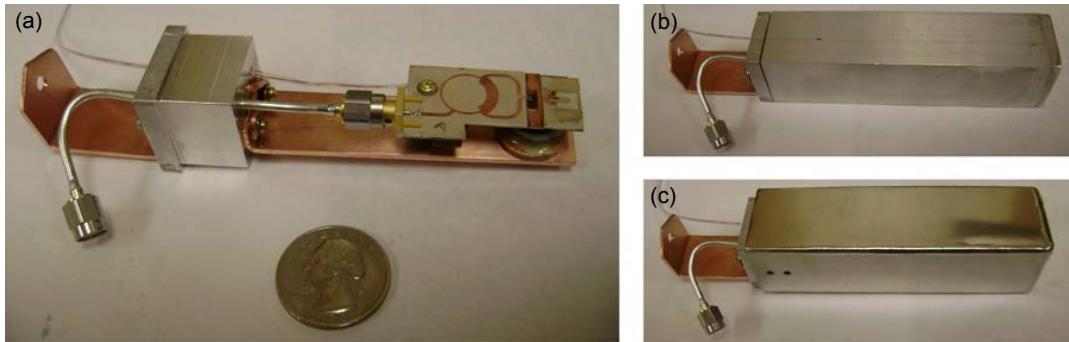


Figure 6.1: Paramp sample box.

The paramp is mounted on a Bollywood launch attached to a copper cold finger as shown in (a). A superconducting coil beneath the sample is used for flux bias. The system is fully enclosed in a superconducting aluminum shield, pictured in (b). Microwave signals, dc current for flux bias, and copper thermalization straps penetrate the shield through high-aspect ratio holes in the lid. A cryoperm shield, seen in (c), is used to reduce flux trapping in the aluminum shield as it cools through T_c .

coupling to external fields. The aluminum box is placed inside a Cryoperm can to reduce flux trapping in the box and sample as they cool through T_c . The microwave line and coil wires enter the aluminum box through high-aspect-ratio holes to maintain shielding effectiveness. One must be careful to filter the leads to the flux bias coil at very low frequency to prevent noise pickup. Further details of the sample boxes and coil filtering are in sections 5.1 and 5.2.1. Figure 6.1 shows an image of the sample box with Bollywood launch, flux bias coil, thermalization straps, and shields.

6.2 Paramp biasing procedure

Upon reaching base temperature, we do some initial tests and then adjust the paramp bias parameters to give the desired gain. One of the advantages of the LJPA paramp is its wide frequency tunability; typical samples can be tuned to provide gain anywhere in the 4-8 GHz band by tuning with a magnetic field. The first test we usually complete is to check that the paramp resonance can be tuned into the desired frequency range. Figure 6.2 shows the phase response of a typical paramp as a function of flux bias, measured in terms of current through the flux bias coil. The periodic response comes from tuning through several flux quanta; a flux quantum typically corresponds to a few mA of coil current in our setup. Since the flux sensitivity of the paramp can be as low as $10^{-8} \Phi_0/\sqrt{\text{Hz}}$, the current noise on the flux bias coil must be no more than a few pA/ $\sqrt{\text{Hz}}$. This typically requires substantial filtering when using a digital current source (see section 5.2.1 for further details).

The astute reader will notice a number of horizontal bands in this figure, indicating variations in reflected phase with frequency that do not depend on flux bias. The bands

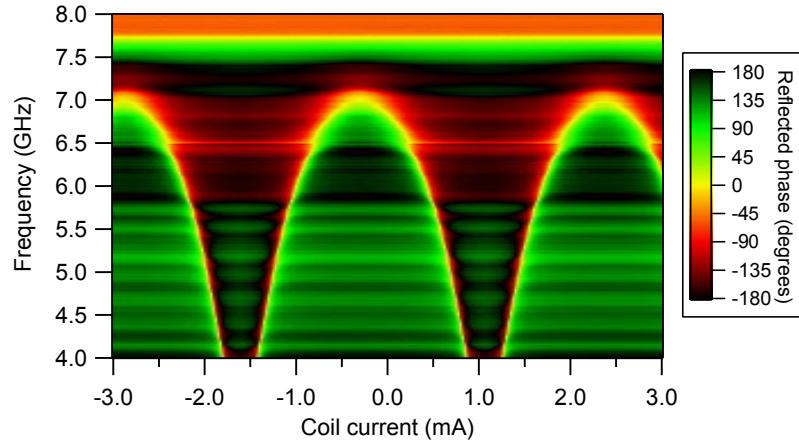


Figure 6.2: Typical paramp tuning with flux.

This plot shows the reflected phase of the paramp as a function of applied magnetic flux, parameterized in terms of the current in the flux bias coil. The yellow band which signifies the paramp's resonant frequency tunes between 4 and 7 GHz in this scan and repeats over several flux quanta. The ripple-like horizontal banding is due to interference effects from the biasing circuit, and the large phase excursions above 7.5 GHz are due to the circulators in the output chain. A small spurious resonance in the launch board is visible at 6.5 GHz, but does not affect the performance of the paramp.

appear to be periodic in frequency. These bands are small-amplitude ripples in the reflected phase caused by interference effects arising from the finite directivity of the directional coupler used to bias the paramp. The directional coupler sends the pump tone from the side port almost entirely towards the paramp, but a small fraction, usually about 20 dB lower in power, is coupled in the direction away from the paramp. This leakage interferes with the reflected signal from the paramp and causes small phase and/or magnitude shifts, depending on the relative phase between the two signals. As the frequency changes, the relative phase will change as well. The frequency periodicity of these phase modulations given by $v/2\ell$, where v is the group velocity of microwaves on the coaxial line and ℓ is the length of line between the directional coupler and the paramp. We typically would like to increase the spacing between these “phase wiggles” so that the phase response of the paramp is locally flat in frequency; this entails putting the directional coupler as close to the paramp as is practical. In addition to the small phase modulations, there are two other horizontal features in Figure 6.2. The first is a small spurious environmental resonance at 6.5 GHz. This is atypical, but did not affect the performance of the paramp. The second is the slow phase rolloff above about 7.25 GHz; this is caused by the circulators on the output amplification chain, which display this sort of phase response near the edge of their band.

Once we have determined that the paramp tunes properly with flux, we can choose a rough flux bias for our frequency of interest. This is done using a vector network analyzer in frequency sweep mode, looking at the reflected phase of the paramp and tuning the flux

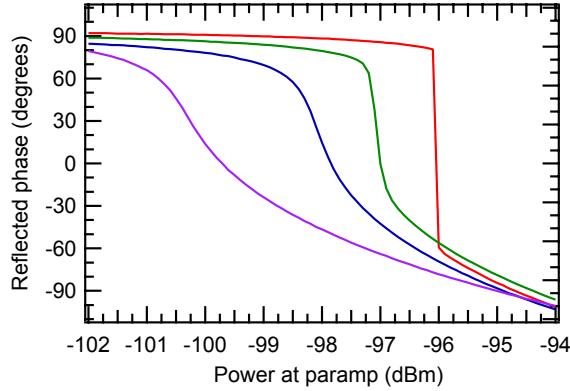


Figure 6.3: Paramp transfer function.

We show four linecuts of the paramp reflected phase versus pump power. These curves represent the transfer function of the amplifier, which maps an amplitude modulation on an incident signal to a phase modulation of the reflected pump. The red curve is in the bistable regime and shows an abrupt phase jump. The green, blue, and purple curves are cuts in the paramp regime with decreasing paramp gain. The steepness of the transfer function gives the gain, while the width in power gives the dynamic range. The phase shift as one goes through the resonance in the power direction is about 180° . The curves have been shifted vertically to show the similarity in their overall phase shifts. The shape of these curves compares well with the theoretically calculated ones in Figure 3.6.

until the steepest part of the phase response is at the desired frequency. Since the frequency of this steepest phase response is a function of pump power as well as flux bias, we perform this search at powers roughly equivalent to a typical pump power during operation, about -95 dBm at the paramp.

With the flux roughly tuned up, we switch from a frequency sweep to an upward power sweep at our desired pump frequency, centered around our estimated pump power. This shows us the response of the pump phase with pump power, which is the transfer function of our paramp¹. Figure 6.3 shows some sample traces of this transfer function. If the phase response has a discontinuity, we are in the bistable regime and we tune the flux to decrease the resonant frequency of the paramp until the phase response becomes smooth and continuous. If there is little or no slope in the phase response, we tune the flux to increase the resonant frequency. The gain is directly related to the steepness of the phase response curve, as described in Chapter 3. The dynamic range is proportional to the width of the phase response in power.

We note that the reflected phase changes by $\sim 180^\circ$ as we go through the paramp resonance. One typically thinks of resonators measured in reflection having 360° of phase

¹The fact that there is any phase response with power is evidence of the nonlinearity of the paramp. A linear resonator exhibits no change in phase response with drive amplitude. Thus, as would be expected, the transfer function of our paramp requires nonlinearity to exist.

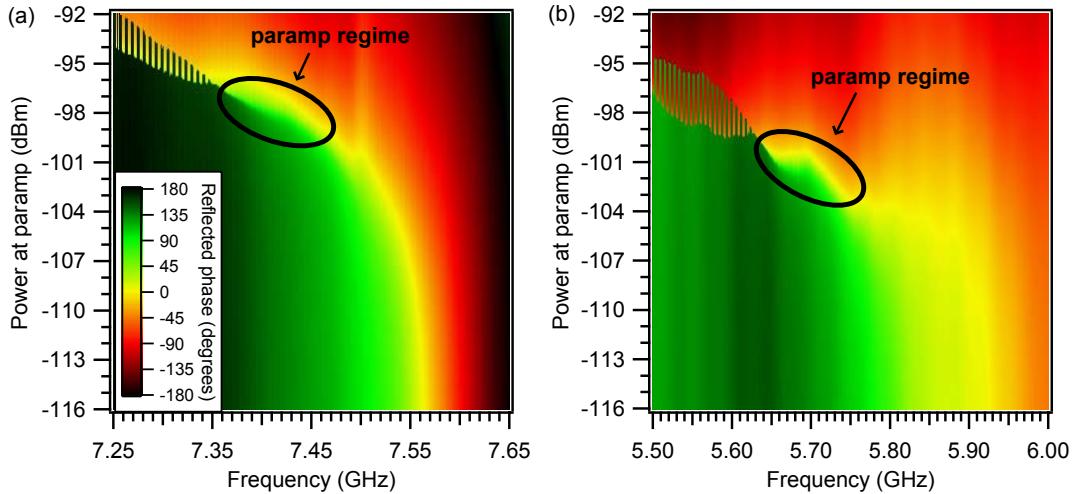


Figure 6.4: Paramp phase response.

These two plots show the reflected phase of the paramp as a function of pump power and frequency. The plots are made by interleaving increasing and decreasing power sweep data so that regions of bistability appear “toothed”. The region of bias parameter space useful for parametric amplification, which occurs near but not inside the bistable region, is labeled. Part (a) shows a fairly well-behaved paramp response, similar to theory. Part (b) shows the paramp response in the presence of embedding circuit impedance fluctuations, which cause non-monotonic behavior of the resonance in the paramp and bistable regimes. One must be more careful when choosing a paramp bias point in this situation.

shift as one crosses the resonance. However, this only holds true if we sweep through the resonance in frequency. In our case, we are sweeping in power and not frequency, and thus the phase shift is not necessarily 360° ; a purely linear resonator would show no phase shift in power, for example.

It is useful to make a map of the phase response of the paramp with respect to pump power and frequency. This shows the regions of bistability and the regions of paramp gain, and helps diagnose any issues with the paramp bias. Two sample phase response plots are shown in Figure 6.4. The data for these plots is taken by sweeping the excitation power both upwards and downwards and then interleaving the up and down sweep directions as we step in frequency. At low powers, the resonance is essentially linear and independent of power. As the power increases, the resonance bends to lower frequency and sharpens, eventually becoming bistable beyond a the critical point at (ω_c, P_c) . The bistable region is readily evident from the “toothed” effect caused by interleaving the up and down sweep traces. One should always check that the chosen paramp bias point is not in or near a region of bistability or the paramp will not function as desired.

At the bias point in Figure 6.4 (b), the width in power of the bistable region does not grow smoothly as we go to lower frequencies (as is expected theoretically), but instead

grows and shrinks non-monotonically. This effect is due to variations in the impedance of the embedding circuit, in our case the Bollywood launch, likely due to spurious resonances. The embedding circuit impedance shunts the paramp resonator and directly sets its Q , which in turn sets the frequency and power of the critical point (ω_c, P_c). If the embedding circuit impedance changes with frequency, then the detuning from the critical point $\omega - \omega_c$ and thus the width in power of the bistable region will vary non-monotonically with frequency as well [108, 156].

The same impedance variations which affect the region of bistability also affect the paramp gain; in their presence, the gain when biased in the paramp regime is no longer necessarily a monotonic function of the flux bias or pump frequency. This is one reason it is helpful to take 2D phase response plots, because they help identify the best paramp bias points in a reliable way. For example, flux tuning may increase, then decrease, then increase the gain again, so if one does not search thoroughly it is possible to miss finding the optimal bias point. If the impedance variation is very pronounced, the gain profile can have strange behavior, as detailed in section 6.3.1. In truly extreme cases we have observed two subregions of bistability separated by a monostable region. Using the 2D plot can help to avoid accidentally biasing in such a regime, which might be unsuitable for paramp use because of the proximity of bistable regions on both sides.

6.3 Paramp performance

Having found a suitable flux bias for operation, we can now optimize the paramp's properties for our desired application. To do this, we use a directional coupler to combine test signals from the vector network analyzer with a pump tone from a separate microwave generator. The network analyzer is used to calibrate the gain and bandwidth of the amplifier as a function of pump power and frequency. We determine the paramp noise performance with a spectrum analyzer, using the network analyzer as a CW signal source.

6.3.1 Paramp gain

We first take a network analyzer frequency sweep over the frequency range of interest with the paramp pump off to calibrate the loop gain of the system in the absence of the paramp pump. All further traces are then taken dividing out this loop gain so that with the paramp pump on, the normalized network analyzer traces now give a direct measurement of the paramp gain. We then explore the phase space of pump power, and to a lesser extent fine-tuning of the flux bias, to optimize the gain and bandwidth as desired. Typically, we get around 20 dB of gain with 10-15 MHz full 3dB bandwidth.

Initially we use a very small signal power, usually around -140 dBm at the sample, to ensure we are not exceeding the dynamic range of the paramp. However, it is important to know when the paramp starts to saturate, so we can find its 1 dB compression point (the point at which gain is reduced by 1 dB from nominal, a standard measure of saturation onset in microwave amplifiers) by increasing the signal power and measuring the gain. Typical values for the input signal power at the 1 dB compression point are around -135 to -130 dBm when biased in the high-gain regime. Figure 6.5 shows three sample gain traces at

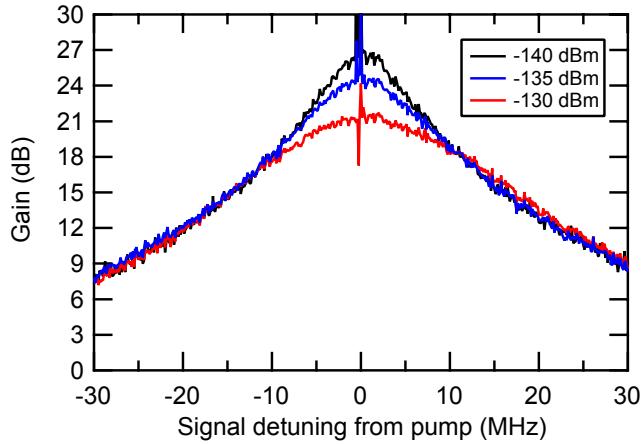


Figure 6.5: Paramp gain and bandwidth.

This figure shows three typical gain traces with identical pump power of -100 dBm and varying signal power as indicated in the legend. As the signal power increases, the paramp saturates and the gain is reduced.

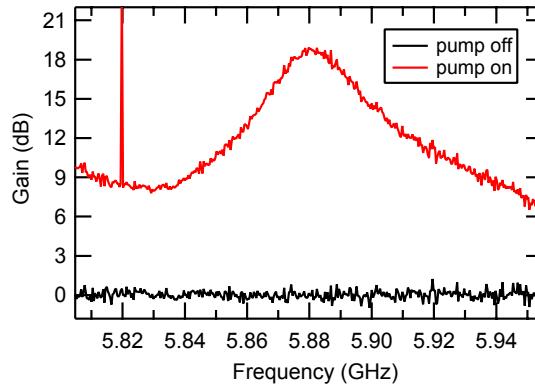


Figure 6.6: Unusual gain profile.

This paramp sample suffered from substantial impedance variations in the embedding circuit, and has a very unusual gain response as a result. The pump is clearly visible at 5.82 GHz, while the main gain peak shows up 60 MHz detuned (a second symmetrical gain peak occurs at - 60 MHz detuning, not shown).

different signal powers. For the lowest power signals, we have around 26 dB of gain with 10 MHz full bandwidth; as the signal power increases, the gain is reduced due to saturation, although the bandwidth remains substantial.

The embedding circuit impedance variations mentioned in the preceding section

can produce some strange effects, including a “double-humped” gain profile at large detunings, with gain comparable to or greater than in the central peak at optimal bias. The impedance variations are on frequency scales smaller than resonator linewidth, so we end up coupling a low- Q mode at the pump frequency with higher- Q modes detuned to either side. Because gain increases with Q , these impedance variations depress the central peak gain while increasing the gain detuned to the sides. Figure 6.6 shows such a bias point, with the maximum gain occurring 60 MHz detuned from the pump (sharp needle). This bias point was used successfully in some experiments, despite its unusual shape.

6.3.2 Noise performance

Once we have chosen a satisfactory gain point, we can check the noise performance of the amplifier. For the HEMT amplifier at 4 K, we used a hot-cold load technique described in section 5.2.4 to determine system noise temperature. Unfortunately, we cannot do this with the paramp. The thermal noise power $k_B T_B$ in a 20 MHz band from a 4 K hot load is about -120 dBm, so exposing the paramp to noise from the hot load will saturate it and lead to an inaccurate noise assessment.

To get around this difficulty, we use the method of signal-to-noise ratio (SNR) improvement, detailed in Figure 6.7. First, we measure the system noise temperature T_{sys} with the paramp off using the standard hot-cold load method. This noise temperature is calibrated to the reference plane of the 6-way switch used to change between the hot-cold load and the qubit/cavity samples of interest. We can then inject a signal, again with the paramp off, and measure the ratio of the signal peak to the surrounding noise floor using a spectrum analyzer. This is then repeated with the paramp pump on. Since the signal has the same amplitude at the reference plane in both cases, the improvement in SNR from turning the paramp on must be due to a reduction in the system noise. We can simply divide the value of T_{sys} measured with the paramp off by the SNR improvement (in linear power units) to give the value of T_{sys} with the paramp on.

Because the attenuation of the cold fridge injection lines is difficult to calibrate accurately, we use a qubit and cavity to calibrate the power of our input signal at base temperature. We know that the power radiated from a cavity in transmission is given by $P_{\text{rad}} = \bar{n} \hbar \omega \kappa$, where \bar{n} is the cavity photon occupation, ω is the signal frequency, and κ is the cavity linewidth. The value of \bar{n} can be extracted very precisely using the ac Stark shift, as discussed in sections 2.3.3 and 6.4.3, while ω is set by the signal generator and κ is easily measured independently. This precise calibration of the power at base temperature can be used to perform an independent check on the extracted noise temperature without requiring the measurement from the hot-cold load. Because we know the precise signal power both at base temperature and at room temperature, we can extract the overall gain G of the amplification chain between the reference plane and the spectrum analyzer. We can then take the noise power spectral density measured at the spectrum analyzer and divide by G to find the noise power spectral density at the reference plane at base temperature; the noise power spectral density can be converted into a noise temperature by dividing by k_B .

These methods require us to resolve the height of the signal peak above the noise floor, which means that the signal and pump must be at different frequencies. As a consequence, we can only extract the noise temperature in phase-preserving amplification mode.

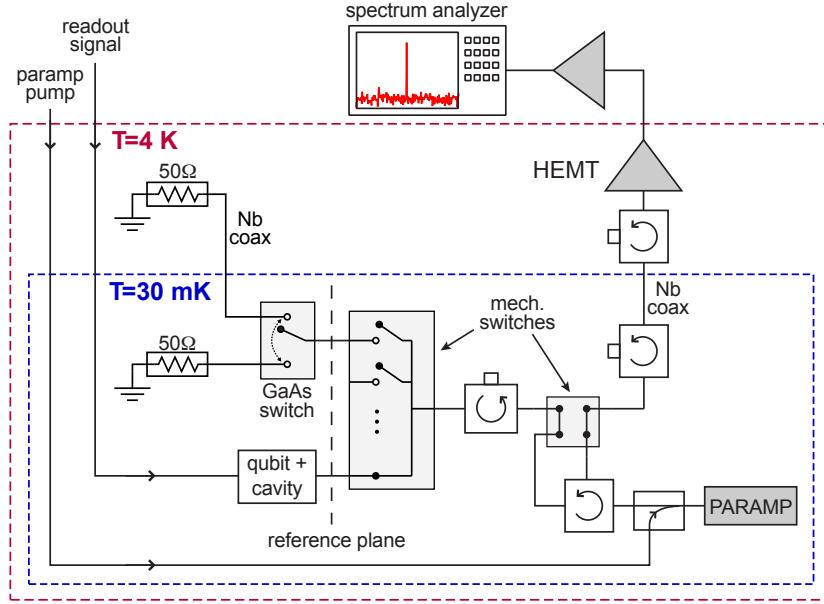


Figure 6.7: Paramp noise measurement scheme.

This experimental setup allows us to calibrate the system noise temperature T_{sys} with the paramp on by measuring the improvement in signal-to-noise ratio over when the paramp is off. We know T_{sys} with the paramp off from using a hot-cold load. We then inject a signal of known amplitude and look at the ratio of signal peak to noise floor at the output using a spectrum analyzer. When we turn the paramp on and off, we see this ratio change and can deduce the reduction in T_{sys} when the paramp is on. The qubit and cavity are used to provide an accurate calibration of signal power using the ac Stark shift.

There is not an analogous procedure for measuring noise temperature in phase-sensitive mode, but we can make a rough estimate by subtracting $\hbar\omega/2k_B$ from the noise temperature estimated in the phase-preserving mode.

The system noise temperature we extract is shown in Figure 6.8 as a function of signal power. For small signals, where the amplifier is not saturated, our T_{sys} approaches the standard quantum limit for phase-preserving amplification of $\hbar\omega/2k_B$. The noise temperature appears to be consistently slightly higher than the quantum limit, which can be attributed to loss between the reference plane and the paramp, primarily due to the circulators. As the signal power increases, the amplifier starts to saturate and the noise performance degrades accordingly. This knee occurs at powers which correspond roughly to cavity photon occupations $\bar{n} \approx 1 - 2$ in a typical circuit QED setup. Even with a signal power 10 – 15 dB above the level where the paramp starts to saturate, the value of T_{sys} is still considerably lower than that without the paramp, which is $T_{\text{sys}} \approx 7\text{ K}$ with minimal circulators, a state-of-the-art HEMT, and superconducting coaxial cables between the base stage and HEMT stage (see section 5.2.2 for details).

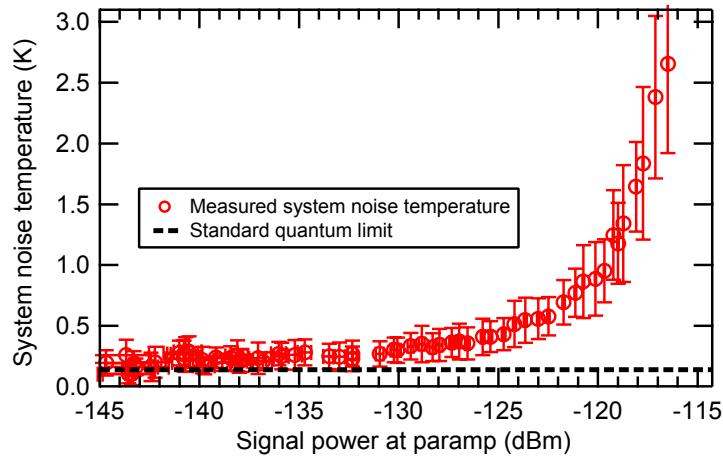


Figure 6.8: System noise temperature with paramp.

The system noise temperature with the paramp on is shown here as a function of the signal power. For low power signals, the noise is almost at the quantum limit. As the signal power increases and the paramp saturates, the noise temperature degrades, but even deep into the saturated regime it is still considerably better than the best value achieved without the paramp of $T_{\text{sys}} \approx 7$ K.

6.3.3 Saturated regime operation

Because most signals of interest (e.g. from a qubit readout cavity) are of high enough power that the paramp is no longer in the linear gain regime, it is helpful to understand the behavior of the paramp in the saturated regime as well. Here we use the paramp not as a linear amplifier in the traditional sense, but rather as a continuous digital detector. One could imagine an op-amp configured to have very large closed-loop gain; its output would always be railed high or low, depending on the input signal, but would change rapidly to follow that input signal. If the signal of interest only has two values, as is the case with the readout signal from a qubit, then such an instrument could be very useful as an amplifier even though its output is not a linear function of its input. The principle of saturated regime operation for the paramp is explained in more detail in section 3.4.4.

Figure 6.9 shows the measured average phase of the reflected paramp pump as a function of pump power, where the input signal comes from a qubit in a cavity (for power calibration purposes). The paramp is operating in phase-sensitive mode. We prepare the qubit either in the ground (blue, open circles) or excited (red, filled circles) states; the resulting pointer states which form the paramp input signal are approximately 180° phase shifted from each other. In part (a), the signal power is about -132 dBm, while in part (b) it is about -118 dBm. In both cases, we can see that there is a range of pump powers for which there is a large difference in the output phase shift signal depending on the phase

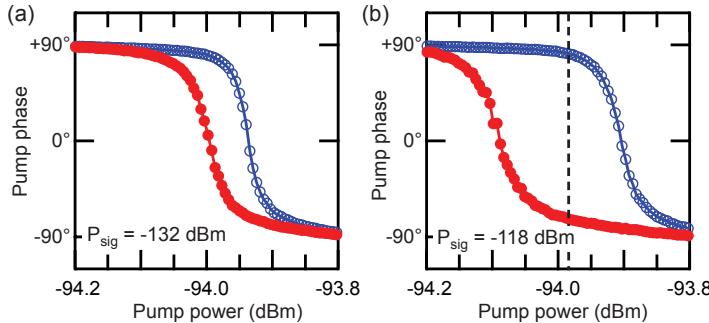


Figure 6.9: Saturated regime paramp response.

These curves show the average reflected phase of the pump in phase-sensitive mode as a function of pump power. We use two input signals of the same amplitude but approximately 180° out of phase with each other (red and blue traces/circles), both for weak signals (a) at the edge of the linear regime and strong signals (b) deep in the saturated regime. In both instances we can find a value of the pump bias which gives clearly distinguished values of the reflected pump phase depending on the input signal.

of the input signal. The amplifier is already starting to saturate in part (a), while in part (b) it is well into the saturated regime; the increase in the input signal amplitude has not brought about a correspondingly large increase in output phase shift. If we set the pump power at the point denoted by the dashed vertical line, the reflected paramp phase will swing almost 180° depending on the two-state input signal. In this regime, we can use the paramp as a continuous digital detector.

6.4 Calibration of the qubit/cavity system

We perform a number of tests and calibrations on our qubit/cavity system to extract the relevant experimental parameters, assess the coherence properties of the qubit, and calibrate the measurement strength. This section gives a brief overview of a standard set of calibrations performed on each sample at each qubit bias point.

6.4.1 Qubit and cavity spectroscopy

Our first task is to extract the parameters E_J and E_C of the qubit, ω_{cav} and κ of the cavity, and the coupling g . Note that for simplicity we will use the notation g and Δ to refer to the quantities denoted g_{01} and Δ_0 , respectively, in section 2.3.3.

The cavity frequency can be seen almost immediately in a network analyzer measurement, and we use this rough frequency to set the measurement tone for qubit spectroscopy. Here we turn on a weak measurement tone, with the power chosen to give $\bar{n} \ll 1$, and examine the cavity response in the presence of a weak qubit spectroscopy tone which is swept in frequency. The output signal will change when the qubit spectroscopy tone is at the frequency of the qubit line. We use this method to extract the qubit energy splitting

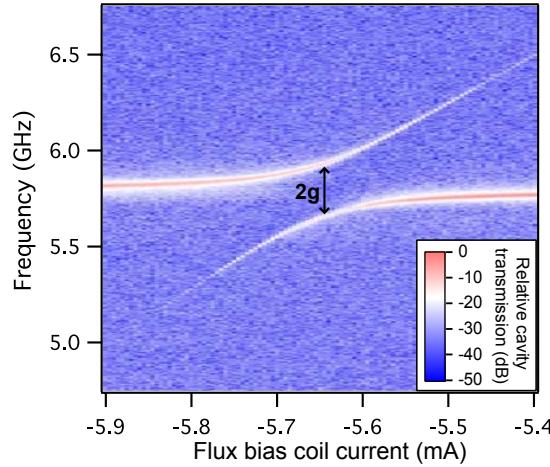


Figure 6.10: Avoided crossing of qubit and resonator.

This plot shows the magnitude response of a cavity measured in transmission as a transmon qubit is tuned through the cavity frequency. We see the expected avoided crossing with a splitting given by $2g$. The transmon line can be seen in this cavity measurement because the eigenstates of the Jaynes-Cummings Hamiltonian are a superposition of cavity and qubit states; as the qubit detunes farther from the cavity, the photonic part of the eigenstate is reduced and so the line weakens.

ω_{01} . If we turn up the power of the qubit spectroscopy tone, we can excite two-photon transitions between transmon states $|0\rangle$ and $|2\rangle$; the cavity will show a response when the spectroscopy tone is at frequency $\omega_{02}/2$. We can repeat this process to perform three-photon spectroscopy of the $0 \rightarrow 3$ transition as well².

There are two methods for calibrating g . One method is to measure the cavity response as we tune the transmon frequency through the cavity frequency; this will give us an avoided crossing as shown in Figure 6.10. We can see the relatively broad, horizontal resonator line hybridizing with the narrow qubit line as the system enters the resonant limit. At the point of narrowest approach between the two lines, they are separated by a frequency of $2g$.

A superior way to calibrate g is what we call the “punchout” method, which is a measurement of the Lamb shift. For small cavity photon occupation \bar{n} , the bare cavity frequency is dressed by the presence of the qubit in its ground state, which shifts the cavity resonant frequency by an amount

$$-\chi_{01} = \frac{-g^2}{\Delta}. \quad (6.1)$$

²Technically, the spectroscopic values measured are the Lamb shifted qubit transition frequencies $\tilde{\omega}_{ij}$ and not the true uncoupled qubit transition frequencies ω_{ij} . For sufficiently large Δ , this does not have much effect on the calibrations described in this section. However, for maximum precision one should correct for the Lamb shifts on the spectroscopic frequencies, which is done self-consistently using progressively finer estimates of E_J , E_C , and the g_{ij} .

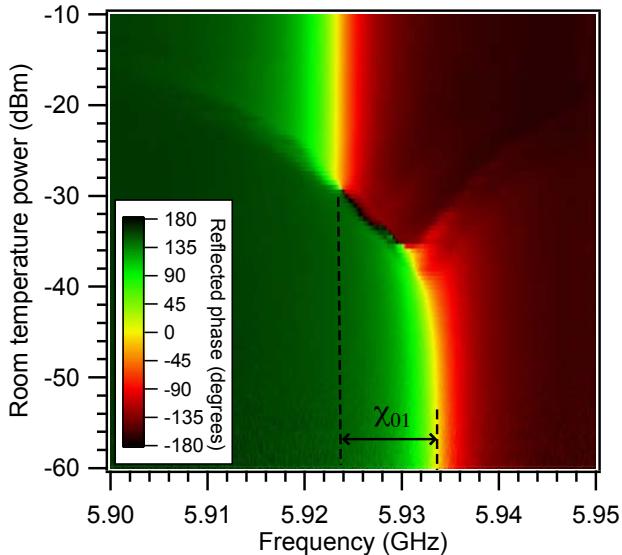


Figure 6.11: Qubit “punchout” calibration.

This plot of cavity phase response versus drive power shows the “punchout” calibration method. At low powers, the cavity frequency is dressed by the qubit. As the power increases, the cavity becomes nonlinear due to the presence of the qubit, until finally the qubit is “punched out” and the cavity shows its bare frequency ω_{cav} . The difference between the bare and dressed cavity frequencies is the Lamb shift χ_{01} , which gives us g by equation (6.1). In our qubit experiments, we always operate with sufficiently low readout power that the cavity remains linear.

As we drive the cavity at higher and higher powers, the cavity inherits the nonlinearity of the qubit and the cavity resonance shifts. Finally, at very high photon numbers there are so many quanta of excitation in cavity and qubit that the cavity behaves classically and no longer shows the effects of the qubit’s presence [157]. In this regime, we simply see a response at the bare cavity frequency ω_{cav} ; the qubit has been “punched out”. Figure 6.11 shows this effect, with the cavity response changing from the dressed frequency at low power to the bare frequency at high power.

We take linecuts of the vector cavity response both at low powers (well below the punchout threshold) and at high powers (above the punchout threshold) and fit the real part of the cavity response to a Lorentzian function to extract the resonant frequency and linewidth for both cases. The resonant frequency with the qubit punched out is just ω_{cav} , and the difference between the resonant frequencies with and without punchout gives us the Lamb shift shown in (6.1). Using our knowledge of ω_{01} from qubit spectroscopy (which we must also correct for the Lamb shift by an amount identical to what is measured from the punchout), we can extract the value of g . We find that the linewidth κ is the same at both high and low powers.

This punchout technique is straightforward and has the advantage that one does not require an avoided crossing between qubit and cavity to calibrate g . If one exists, we can use it as a less-trusted cross-check of the extracted value of g . We also note that if we prepare the qubit in state $|1\rangle$ and perform a punchout, the power threshold at which the punchout occurs is different than if the qubit is in state $|0\rangle$. This effect has been used experimentally to provide a high-fidelity single shot readout [158] (albeit one which is not QND) and can be well explained theoretically by numerical diagonalization of the generalized Jaynes-Cummings Hamiltonian for a multilevel system [93].

Armed with measurements of the first three transmon frequencies, ω_{cav} , and g , we can use Mathematica to solve for the qubit energy scales E_J and E_C . If we know E_J and E_C , we can determine all of the g_{ij} for higher levels numerically and solve for the dispersive shift $\chi(\bar{n})$ using the expressions given at the end of section 2.3.3. Since the estimates of E_J and E_C depend on the accuracy of our knowledge of the Lamb shifts and thus on the g_{ij} , we can perform this procedure iteratively for increased accuracy if desired. Generally, such a procedure is not needed if we are well into the dispersive regime $|\Delta| \gg g$.

To calibrate the qubit's sensitivity to flux, we can perform qubit spectroscopy as a function of flux bias in the external coil and then fit the resulting spectrum to the theoretical energy spectrum of the qubit. We fix the value of E_C to be the one extracted by the method above at a given bias point, allowing the maximum E_J and the flux periodicity to vary until the fit is optimized. One does not need to be able to sweep through a full flux quantum to yield an accurate fit.

6.4.2 Rabi, Ramsey, T_1 , and tomography

Once we have established the qubit and cavity frequencies, we can proceed to examine the qubit coherence, characterized in terms of the relaxation time T_1 and the dephasing time T_2 . We measure these coherence times using several standard methods known from nuclear magnetic resonance and atomic physics. All of these measurements are ensemble measurements, requiring many repetitions of a given experiment. The delay time between repetitions is chosen to ensure that the qubit has relaxed back to the thermal ground state (which is $|0\rangle$, since we operate with $\hbar\omega_q \gg k_B T$) before the next repetition.

Typically we start by measuring Rabi oscillations of the qubit. We send a pulse at the qubit frequency of varying duration and then measure the state of the qubit after the pulse has stopped. We repeat this experiment a number of times for each value of the Rabi pulse duration and average the results from each pulse duration together, giving us decaying sinusoidal Rabi oscillations as shown in Figure 6.12(d). The frequency of these Rabi oscillations depends on the amplitude of the Rabi drive and on the detuning of the Rabi drive from the qubit frequency $\Delta\omega$ as follows:

$$\Omega_{\text{Rabi}} = \sqrt{\Omega^2 + (\Delta\omega)^2}, \quad (6.2)$$

where Ω is the Rabi frequency at zero detuning.

We can use Rabi oscillations for several purposes. First, they provide a probe of the qubit frequency, which can be found by fixing the amplitude of the Rabi drive tone and varying its frequency until we minimize the frequency of the Rabi oscillations. Secondly,

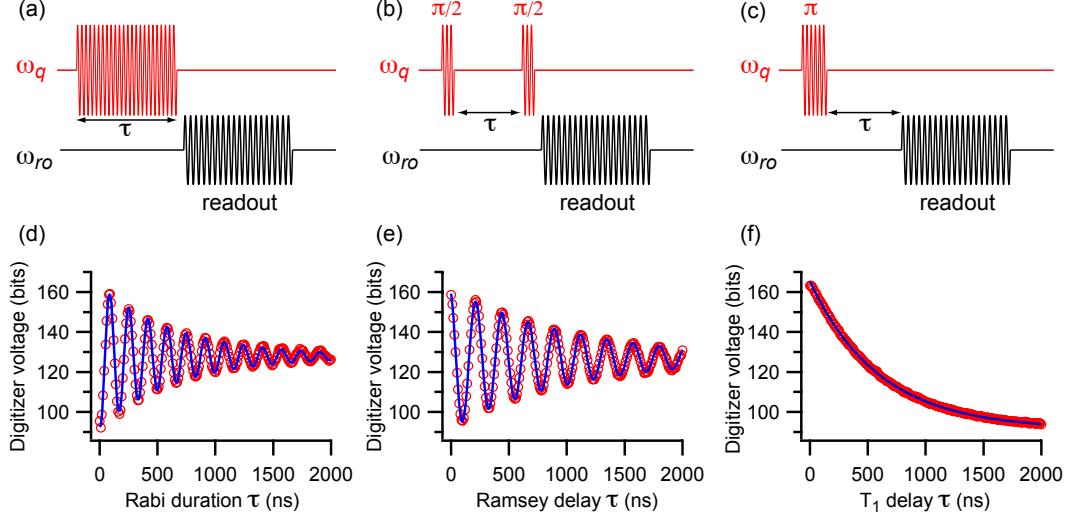


Figure 6.12: Qubit coherence measurements: Rabi, Ramsey and T_1 .

Parts (a), (b), and (c) show timing diagrams for the Rabi, Ramsey, and T_1 experiments respectively. The red trace indicates pulses at the qubit frequency ω_q , while the black indicates pulses at ω_{ro} , the readout frequency. Panels (d), (e), and (f) show typical experimental Rabi, Ramsey, and T_1 traces. Each data point (red circles) represents the average measured voltage of 3×10^4 identical experiments for a given value of τ . The excited state corresponds to higher digitizer voltage and the ground state corresponds to lower digitizer voltage. These data do not have an absolute calibration to qubit state probability, and such a calibration is not strictly required to extract the information we need about the qubit frequency, coherence, and pulse coupling. If desired, though, one can perform such a calibration using the methods described in [44]. The blue lines are fits to a sine wave with an exponentially decaying envelope in parts (d) and (e), and a fit to a decay exponential in (f). These data were taken with the paramp on to speed data acquisition.

Rabi oscillations give us a calibration for the pulse duration and amplitude required to make a pulse that takes $|0\rangle$ to $|1\rangle$, called a π pulse because it corresponds to a rotation of the Bloch vector by π .

The decay of Rabi oscillations can also provide a useful probe of qubit dephasing. Rabi oscillations are an ensemble measurement, and their decay results from noise-induced variations in the Rabi frequency, which causes individual measurements to have different phases at longer times. We can quantify this effect and relate the Rabi decay envelope to the qubit relaxation and dephasing times. The expression for the Rabi decay rate Γ_{Rabi} is given by [159]:

$$\Gamma_{\text{Rabi}} = \frac{3 - \cos^2 \eta}{4} \Gamma_1 + \Gamma_\phi \cos^2 \eta + \frac{1}{2} \Gamma_\nu \sin^2 \eta, \quad (6.3)$$

where $\Gamma_1 = 1/T_1$, $\Gamma_\phi = 1/T_2 - 1/2T_1$ is the low-frequency dephasing rate, and $\Gamma_\nu =$

$\pi S_{\delta\omega_q}(\Omega_{\text{Rabi}})$, where $S_{\delta\omega_q}(\omega)$ is the power spectral density of qubit frequency fluctuations at frequency ω . We can alternately define $\Gamma_\phi = \pi S_{\delta\omega_q}(\omega \rightarrow 0)$. The angle η is defined in terms of the Rabi frequency and detuning by $\cos \eta = \Delta\omega/\Omega_{\text{Rabi}}$ and $\sin \eta = \Omega/\Omega_{\text{Rabi}}$.

Depending on the Rabi detuning, one can use the Rabi decay as a probe of the qubit noise. The envelope of the Rabi oscillations can also change from exponential to Gaussian depending on the spectrum of the noise; the expression in (6.3) is valid for the case of exponential decay, where the noise spectrum is not singular at low frequencies [159, 160].

We can also probe the qubit by measuring Ramsey fringes, giving us information about the qubit frequency and dephasing time T_2 . We send two $\pi/2$ pulses to the qubit with a varying time delay between the pulses, then measure the state of the qubit. We repeat this experiment many times for each value of the time delay and plot the average measured qubit state versus time delay. If the frequency of the qubit drive is detuned from the true qubit frequency, the plot will show oscillations called Ramsey fringes. A typical Ramsey fringe is shown in Figure 6.12(e). The decay envelope of the Ramsey fringes with delay time is exponential, with a time constant given by the dephasing time T_2 . The frequency of oscillations is equal to the detuning between the qubit drive frequency and the true qubit frequency. Because of this feature, Ramsey fringes are a very precise way to measure the qubit frequency. We can measure low-frequency fluctuations in the qubit frequency, caused for example by flux noise, by taking a series of Ramsey fringes to extract the qubit frequency as a function of time. Ramsey fringes are a more sensitive way to extract the qubit frequency than Rabi oscillations, since for small detunings the Ramsey frequency depends linearly on the detuning, while the Rabi frequency has a only a quadratic detuning dependence.

Having calibrated the qubit frequency and the π pulse amplitude using Rabi and Ramsey experiments (sometimes applied iteratively to refine the extracted values), we measure the relaxation time T_1 of the qubit. To do this, we prepare the qubit in the excited state with a π pulse, wait for a variable amount of time, and then measure the qubit state. This is repeated many times for each value of the wait time, allowing us to see the ensemble decay of the qubit state. The excited state population decays exponentially in time with time constant T_1 , as seen in Figure 6.12(f).

For Rabi, Ramsey, and T_1 experiments, we use curve fitting to extract the coherence parameters of the qubit, the qubit frequency, and the Rabi drive coupling relating the amplitude of the Rabi drive to the zero-detuning Rabi frequency. These data can in turn give us information about the coupling of the qubit to its environment and the amount of environmental noise the qubit sees.

Because the qubit eigenstates of the Jaynes-Cummings Hamiltonian in the dispersive regime have a small photonic part, there is a possibility for the qubit to decay by the emission of its photonic component from the cavity. This phenomenon is known as the Purcell effect, and sets a limit on the T_1 of a qubit coupled to a cavity. The rate at which the qubit decays due to the Purcell effect is

$$\gamma_{\text{purc}} = \kappa \frac{\Delta^2}{g^2}. \quad (6.4)$$

We can observe this effect experimentally by measuring the T_1 of the qubit as a function of its detuning from the cavity, thus changing γ_{purc} . The results of such a measurement are

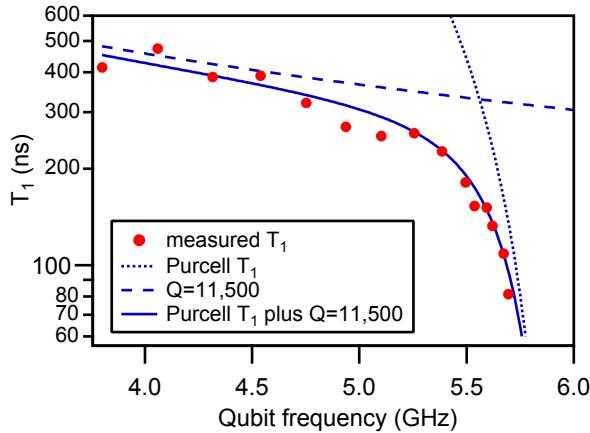


Figure 6.13: Purcell effect.

The red dots show the fitted value of T_1 (see Figure 6.12(f) for a sample fit) as a function of qubit frequency with $\omega_{\text{cav}} = 5.923$ GHz. The dotted blue line gives the theoretical prediction for the Purcell T_1 from (6.4), the dashed blue line shows a line representing a Q for qubit oscillations of 11,500 due to other loss mechanisms, and the solid blue line shows the combined effects of these two types of decay.

shown in Figure 6.13. The red dots are measured values of T_1 , while the blue line shows the expected T_1 based on the Purcell effect in combination with qubit decay from an unknown additional channel giving a Q of 11,500 for the qubit. This phenomenological limiting Q has been reported in other transmons and is likely due to dielectric loss in the transmon capacitors [43].

When the readout cavity has multiple modes, the Purcell decay rate becomes more complicated than the simple expression given above. In general, the presence of higher cavity modes makes γ_{purc} considerably higher for $\omega_q > \omega_{\text{cav}}$, while reducing it slightly for $\omega_q < \omega_{\text{cav}}$. This multi-mode Purcell effect has been theoretically modeled and agrees well with experimental measurements including the effects of an intrinsic qubit Q [43]. To avoid the multi-mode Purcell effect, we make our resonators in a lumped geometry that has no higher harmonics below 20 GHz.

We can also perform quantum state tomography on the qubit to determine the location of the qubit state on the Bloch sphere [161]. We prepare a desired state, apply a rotation of a given angle about a given axis in the x-y plane, and measure the qubit state. We repeat this process many times for each of a number of different choices of rotation axis and angle. Each axis/angle pair is equivalent to measuring a different projection of the qubit state (thus the use of the term “tomography”), so by choosing a variety of axis/angle combinations we can reconstruct the full density matrix of the measured ensemble³. This technique can be extended to characterize the density matrices of multiple qubits [20, 162]

³Technically, one need only perform four rotations, such as a $\pi/2$ x rotation, a $\pi/2$ y rotation, a π rotation about either x or y, and an identity (null) rotation, to be able to extract the full density matrix.

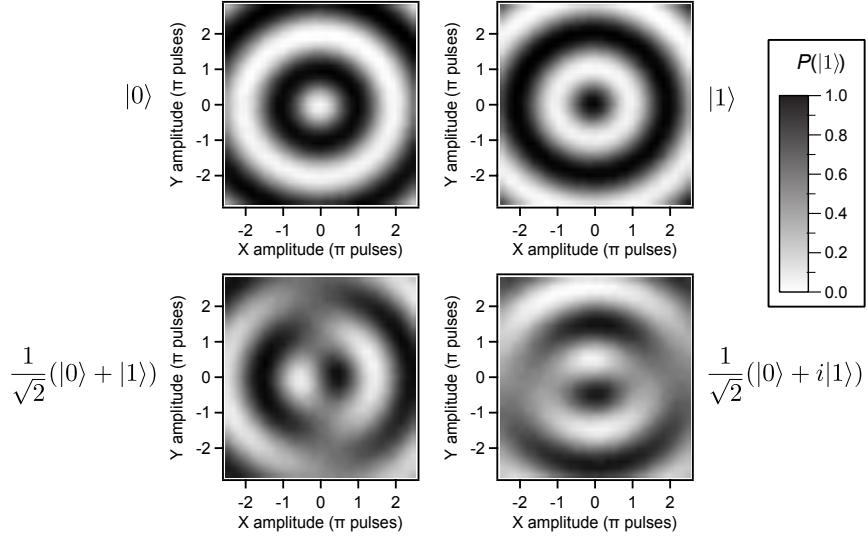


Figure 6.14: Quantum state tomography.

These four panels show quantum state tomography of a transmon qubit for each of four quantum states, $|0\rangle$, $|1\rangle$, $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, and $\frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$. The tomography for the two eigenstates ($|0\rangle$ and $|1\rangle$) show no angular dependence, while that for the two superposition states show strong angular dependence. This angular dependence is a signature of phase coherence of the prepared superposition states. The extreme values of $P(|1\rangle)$, the probability to measure $|1\rangle$ after the tomographic rotation, are close to 0 and 1 in all four instances, showing we can prepare a pure quantum state.

and/or of states of the microwave field [82].

Figure 6.14 shows quantum state tomography of a qubit prepared in four different states, the ground state, excited state, and two equal superposition states which differ by a phase factor of $e^{i\pi/2}$. The horizontal and vertical axes represent the magnitude of rotation in the x and y axes, respectively, while the color scale indicates the probability of measuring the qubit in the excited state after the rotation. For the ground and excited states, the tomography has no angular dependence. Any radial linecut from the center just shows a Rabi oscillation. For the superposition states, however, there is a strong angular dependence. We see no radial dependence when rotating around an axis parallel to the state vector, but we do if we rotate around an axis perpendicular to the state vector. The two superposition state images show that we can prepare superposition states of arbitrary and well-defined phase.

6.4.3 Photon number and ac Stark shift

The strength of our measurement depends on the average cavity photon occupation \bar{n} . As described in section 2.3.3, we can calibrate this in an elegant way using the ac Stark

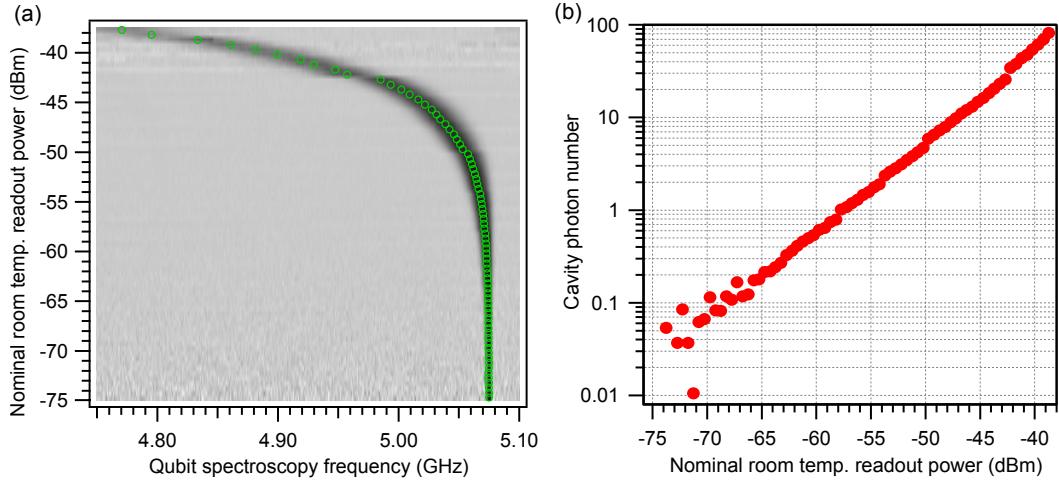


Figure 6.15: ac Stark shift and photon number calibration.

Part (a) shows qubit spectroscopy in the presence of readout photons as a function of readout power. The qubit line shifts in frequency and broadens in response to the ac Stark shift and associated dephasing. The green circles show the qubit frequencies at each readout power extracted from curve fits. Part (b) shows the cavity photon occupation \bar{n} calculated from the ac-Stark-shifted frequencies found in part (a). The occasional “steps” seen in the photon number are due imperfections in the digital attenuator used to set the readout power.

shift. The ac Stark shift dictates that the qubit frequency will depend on \bar{n} , with the shift between the qubit frequency in the absence of photons and in the presence of photons (known as the cavity pull), given by

$$\zeta = (S_1 - S_0)\bar{n} + (K_1 - K_0)\bar{n}^2, \quad (6.5)$$

where the S_i and K_i are described in section 2.3.3. The values of the S_i and K_i depend on the transmon E_J and E_C , as well as the bare cavity frequency ω_{cav} and the qubit-cavity coupling g . All four of these parameters are extracted from spectroscopic data as described in section 6.4.1.

We can make an experimental measurement of ζ using qubit spectroscopy. We turn on the readout cavity drive, wait several microseconds, and then turn on a weak qubit spectroscopy pulse lasting for several more microseconds while the cavity drive is still on. We step the frequency of the qubit spectroscopy and look for differences between the cavity responses before and during the spectroscopy pulse. We repeat this for a range of cavity drive powers set by changing the attenuation of a digital variable attenuator of the type discussed in section 5.3.2. The digital attenuator enables us to return to a given drive power in a repeatable fashion.

The resulting spectrum is shown in Figure 6.15(a). The light regions show regions where the qubit did not respond to the spectroscopy tone, while the dark regions show the

regions where it did. As the cavity drive power increases, the qubit line shifts to lower and lower frequencies and becomes broader. The frequency shift is downward (as oppose to upwards) because the qubit frequency is below that of the cavity for this sample. The increase in qubit linewidth with increasing readout power is a manifestation of the measurement-induced dephasing of the qubit [42]. We can perform curve fits on the spectroscopy response at each value of cavity drive power to determine the qubit center frequency and linewidth. The extracted values of the qubit frequency are overlaid on part (a) as green circles.

We can use equation (6.5) to calculate the value of \bar{n} at a given cavity power using these extracted qubit frequencies. Figure 6.15(b) plots the value of \bar{n} found by this method against the nominal room temperature cavity readout power. As we expect, the signal is quite linear for small \bar{n} , deviating slightly as \bar{n} becomes large. This is in keeping with the expected theoretical relationship between the drive power and the cavity occupation, which begins to increase nonlinearly for very strong cavity drive [93].

It is important to note that this method for calibrating the photon number is only useful for calibrating the photon number \bar{n}_g when the qubit is in the ground state. In general, the value of \bar{n} will depend on the cavity drive frequency and the qubit state as well as the cavity drive power. If we bias the cavity at a point midway between $\tilde{\omega}_{\text{cav}}(|0\rangle)$ and $\tilde{\omega}_{\text{cav}}(|1\rangle)$, however, the photon occupation will be the same in both states for low \bar{n} .

Chapter 7

Quantum jumps

This chapter details the experimental observation of quantum jumps in a superconducting qubit. We begin by laying out some historical background on the theory and observation of quantum jumps, then proceed to describe the experimental method. We then describe the experimental results and provide evidence that we are indeed seeing quantum jumps of the qubit. The chapter concludes by examining the fidelity of qubit readout using this technique. This chapter will be concerned primarily with the initial quantum jump experiments, while details of subsequent quantum jump experiments used to examine measurement non-idealities and the quantum Zeno effect will be presented along with those data in Chapter 8.

7.1 What are quantum jumps?

According to the postulates of quantum mechanics, the act of measurement collapses the state of a quantum system into an eigenstate $|k\rangle$ of the measured observable \hat{A} :

$$\hat{A}|k\rangle = a_k|k\rangle \quad (7.1)$$

If the measurement is quantum non-demolition, then the state $|\Psi\rangle$ of the system immediately after the measurement is given by $|\Psi\rangle = |k\rangle$. In the absence of further measurement, the system will evolve according to the time-dependent Schrödinger equation $i\hbar\frac{\partial}{\partial t}|\Psi\rangle = \hat{H}|\Psi\rangle$. In many cases (such as in our qubit experiments) the measured observable \hat{A} is the energy, so $|k\rangle$ is an eigenstate of \hat{H} . In this case, $|\Psi\rangle$ will remain unchanged by time evolution (except for a phase factor).

Now consider the case where environmental noise or an applied drive enters as a perturbation to the Hamiltonian. This perturbation will in general give rise to non-zero matrix elements between the eigenstates of the unperturbed Hamiltonian, so that the time evolution of $|\Psi\rangle$ is no longer simply a phase factor. In general, we can write that at a time τ after the measurement,

$$|\Psi(\tau)\rangle = c_k(\tau)|k\rangle + \sum_{k' \neq k} c_{k'}(\tau)|k'\rangle, \quad (7.2)$$

where $c_{k'}(0) = \delta_{kk'}$. If we perform a second QND measurement at time τ , we will again collapse the quantum system to one of its eigenstates. If τ is small, in the sense that $|c_k(\tau)| \gg |c_{k'}(\tau)|$ for $k' \neq k$, then the measurement will collapse the system to $|k\rangle$ again with high probability. We can repeat this process of measuring at time intervals of τ and will generally get a sequence of results showing collapse to state $|k\rangle$. However, every so often (with probability $1 - |c_k(\tau)|^2$ per measurement) the system will instead collapse to a different state $|k'\rangle \neq |k\rangle$. Repeated measurements will then show state $|k'\rangle$ for a period of time, until the system is projected to a different state again, and so on. When we see a change between repeated measurements showing $|k\rangle$ and repeated measurements showing $|k'\rangle$, we say that a quantum jump has occurred. This is the classic interpretation of quantum jumps, due to Cook [163].

As seen from this definition, the observation of quantum jumps requires a method for performing quantum non-demolition measurements on the system in question at a rate considerably faster than the system's free evolution. The definition given uses discrete measurements at fixed points in time, but in our experiments the measurement is continuous (in the sense the readout cavity is being continuously irradiated with microwaves). In this instance, instead of a time between discrete projective measurements, we have a measurement rate Γ_m at which information about the qubit state can be determined [50]. The measurement time can be taken to be $1/\Gamma_m$ and then the quantum jump formulation for discrete measurements applies [91].

7.2 Historical background

This notion that quantum systems evolve instantaneously by “jumping” between eigenstates was first proposed by Bohr in 1913 [26]. For three quarters of a century, the concept of “quantum jumps” remained a purely theoretical curiosity, and a subject of substantial debate. Schrödinger wrote a lengthy piece denying their existence in 1952, commenting that the notion of monitoring the state of a single quantum particle was ludicrous and that quantum jumps are the modern equivalent of epicycles, “fudge factors” used by the Greeks to reconcile astronomical observations of retrograde planetary motion with the geocentric theory of the universe [164, 165]. By the early 1980s, though, advances in atomic physics allowed for the trapping and cooling of single ions [27, 28], raising the possibility that the question of quantum jumps might be settled experimentally. A few years later, in 1986, three groups simultaneously reported the observation of quantum jumps between the electronic states of individual trapped ions [29, 30, 31]. Since that time, quantum jumps have been observed in a variety of other systems, beginning with the electronic states of single terrylene molecules embedded in a crystal [32]. Single electrons in cyclotron orbits are seen to undergo quantum jumps between Landau levels [33], and single microwave photons are suddenly created and annihilated by thermal processes inside a Fabry-Perot cavity [34]. In solid state systems, quantum jumps have been observed in a microscopic defect in a Josephson junction [35], while more recently, work showed that the state of a single nuclear spin in a diamond NV center undergoes quantum jumps [36], as does the electronic state of an electron in an indium gallium arsenide quantum dot [37]. Quantum jumps from spin flips of a single trapped proton in a magnetic field were reported recently as well, paving

the way for precision test of matter-antimatter asymmetry [38].

All of these experiments observed quantum jumps in microscopic quantum degrees of freedom (single atoms, molecules, electrons, protons, or photons) with long (\sim ms to s) relaxation times, both in isolation and as part of a larger solid state system. The paper of Yu *et al.* claims that their quantum jumps are of the qubit-TLS system, and thus are macroscopic quantum jumps [35]. However, we maintain that their observed jumps must be attributed purely to the microscopic TLS defect (or to some other unknown mechanism), with the qubit functioning solely as a measurement apparatus, for several reasons. First, the observation of quantum jumps requires multiple QND measurements of the quantum system under study before it changes state. In the work by Yu *et al.*, measurements were made every 10 ms, while they report a qubit T_1 of 600 ns. Secondly, their measurement technique is not QND because it causes the phase qubit to transition to the voltage state¹ with each measurement, at which point there is no qubit anymore. Since the measurement completely destroys the phase qubit, it cannot hold quantum information between measurements, so any quantum state information must be held solely in the TLS defect. Finally, the reported quantum jump lifetimes are of the order of seconds, much longer than those any reported superconducting qubit. The jumping dynamics observed by Yu *et al.* must be purely those of the TLS defect (or of some other part of the experiment), with the phase qubit serving merely as a readout element.

The work presented in this chapter represents the first observation of quantum jumps in a superconducting qubit, which a macroscopic quantum system [62]. It provides a further confirmation of the quantum nature of superconducting qubits, which had been shown previously through entanglement and violation of Bell's inequality [18, 19, 20, 21, 22, 23, 166]. Our readout method gives the ability to monitor continuously with high fidelity to observe single-shot qubit dynamics. Because the relaxation time of superconducting qubits is much shorter ($\sim \mu\text{s}$) than that of the other systems in which jumps have been observed, the experiment relies crucially on the high bandwidth and quantum-limited noise performance of the paramp.

7.3 Experimental design

Quantum jumps occur whenever strong projective measurements are performed at a rate much faster than the natural evolution dynamics of the qubit state. In circuit QED, quantum jumps occur from the action of the readout photons in the cavity (given a sufficient number of photons \bar{n} and a nonzero dispersive shift 2χ). The readout photons become strongly entangled with the qubit state; if we then measure the phase of the photons, the qubit state also collapses to the corresponding eigenstate. The difficulty in observing these jumps comes because the number of readout photons is necessarily small, so the signal is drowned by the noise added in the amplification chain. The solution to this problem is the introduction of the superconducting parametric amplifier (paramp) described in Chapters 3 and 6, which lowers the noise temperature of the amplification chain enough to enable continuous high-fidelity readout. No matter the qubit or readout cavity design, all of the quantum jump experiments in this thesis make use of a paramp for this purpose.

¹For a detailed explanation of this phase qubit readout method, see refs. [48] and [71].

The basic experimental setup (on which a few variations will be discussed in this and later chapters) is shown in Figure 7.1. The key ingredients are a qubit coupled to a readout cavity in a circuit QED architecture and a paramp to amplify the readout signal. This particular setup was used for the initial quantum jump experiments. We implement the readout cavity as a quasi-lumped-element linear resonator consisting of a meander inductor of $L = 1.25$ nH (orange) in parallel with finger capacitors of $C = 575$ fF (blue) [167]. We chose this design over the traditional transmission line resonators used in most circuit QED experiments to this point for several reasons. First, the lumped element cavity doesn't have higher harmonic modes and so should not suffer from the multi-mode Purcell effect [43], where higher cavity modes cause additional qubit relaxation. Second, the smaller footprint of the lumped-element cavity allows more cavities to be placed on a single chip, which may help with scaling up to multi-cavity designs. The qubit is a transmon (described in detail in section 2.1.3 and in [13]) which is capacitively coupled to the cavity; the qubit parameters at the operating point are $E_J = 11.4$ GHz and $E_C = 280$ MHz, with a qubit frequency of 4.753 GHz. The coupling strength between the qubit and the cavity is $g/2\pi = 109$ MHz. The cavity has a bare resonant frequency $\omega_{\text{cav}} = 5.923$ GHz with a linewidth $\kappa/2\pi = 4.9$ MHz. For this experiment, the dispersive shift $2\chi/2\pi = 4.4$ MHz. Since $2\chi \approx \kappa$, there will be about a 180° phase shift between the reflected readout cavity signals corresponding to the two qubit states. We use the ac Stark shift for a multilevel qubit to calibrate the cavity photon occupation, as described in sections 2.3.3 and 6.4.3.

The experiment proceeds as shown by the arrows in the figure. Readout photons at 5.932 GHz from the input port are directed through a circulator into the readout cavity, where they interact with the qubit and acquire a phase shift depending on the qubit state. The readout photons then leave the cavity and travel to the paramp through a series of three circulators, which isolate the readout cavity and qubit from the strong pump tone used to bias the paramp to its gain point². The readout photons then combine with the paramp pump, also at 5.932 GHz, which enters through the -20 dB weakly coupled port of a directional coupler. We adjust the phase of the pump beforehand so that the readout photons and pump interfere either constructively or destructively, depending on the phase shift of the readout photons (and thus on the qubit state). Since the readout photons corresponding to the two qubit states are about 180° phase shifted from each other, this interference between readout signal and pump results in a modulation of the pump amplitude and thus a large phase modulation of the reflected pump, as described in section 3.4.3. By mapping the readout phase shift signal (carried by a few photons) onto the phase shift of the reflected pump (carried by hundreds of photons), we have effectively amplified the readout signal (specifically, this realizes a phase-sensitive amplification as described in Chapter 3). The amplified signal is reflected out of the paramp and directed through a circulator to further cryogenic and room temperature amplification stages. The amplified signal is then mixed down to zero frequency (converting the phase shift signal into a quadrature voltage signal) and digitized. By shifting the relative phase of the local oscillator, we can put the entire readout signal into one quadrature. Experimental details

²The pump power is about -90 to -95 dBm at the paramp, so three circulators (each with about 20 dB of isolation) will reduce the pump leakage to about -150 dBm at the readout cavity, a power which corresponds to a cavity occupation of roughly 0.01 photons. This power is too small to affect the qubit in any significant way.

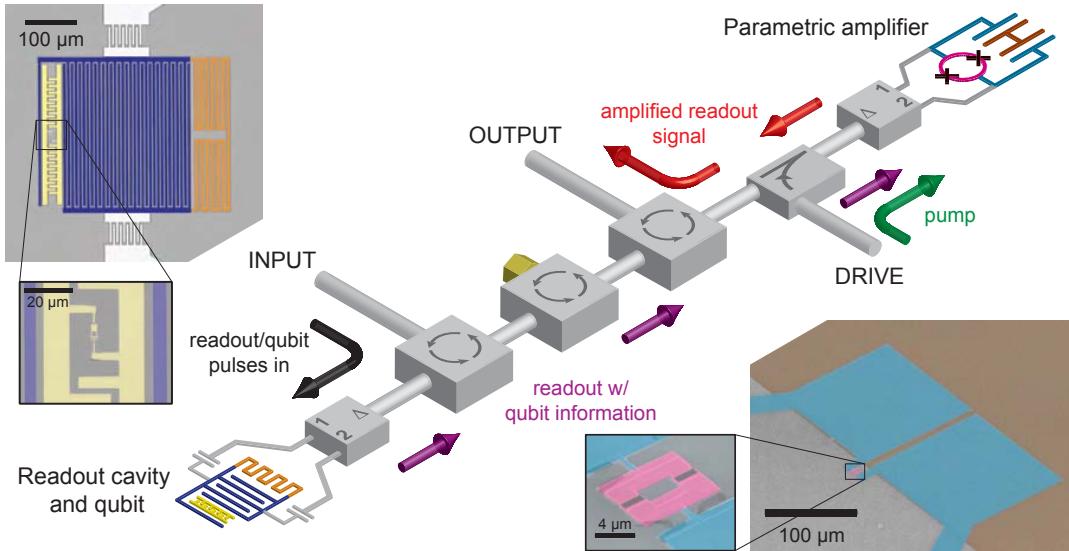


Figure 7.1: Base temperature apparatus.

Readout photons enter the from the input port (black arrow) and are directed through a circulator into the readout cavity, where they interact with the qubit and acquire a phase shift that depends on the qubit state. Upon leaving the cavity, the readout photons (purple arrows), now carrying information about the qubit state in their phase, travel through three microwave circulators, which isolate the qubit from the strong pump tone used to bias the paramp. The readout photons combine with the paramp pump (green arrow) which enters from the weakly coupled port of a directional coupler and interact in the parametric amplifier so that the phase of the readout photons is encoded in the phase of the reflected pump tone which forms the output signal (red arrow). Since the pump has many more photons than the initial readout signal, this effects an amplification of the readout signal. The output is directed on to further cryogenic and room-temperature amplification stages (not shown) and then is mixed down to zero frequency (converting the phase shift signal into a voltage signal) and digitized. Both the readout cavity and paramp are excited using Bollywood 180° hybrids, which turn the single-ended input signals into differential ones. The false color images show the qubit/cavity sample in the upper left and the paramp sample in the lower right. The cavity consists of a meander inductor (orange) shunted by an interdigital capacitor (blue); input signals are coupled in through the interdigital capacitors at the top and bottom. The transmon qubit (yellow) is capacitively coupled to the cavity. The inset view shows the qubit loop. The paramp sample consists of a SQUID loop (pink) with Josephson junctions (black) shunted by two capacitors in series formed by Al top layers (cyan) on a Nb ground plane (brown) coated with SiN_x dielectric.

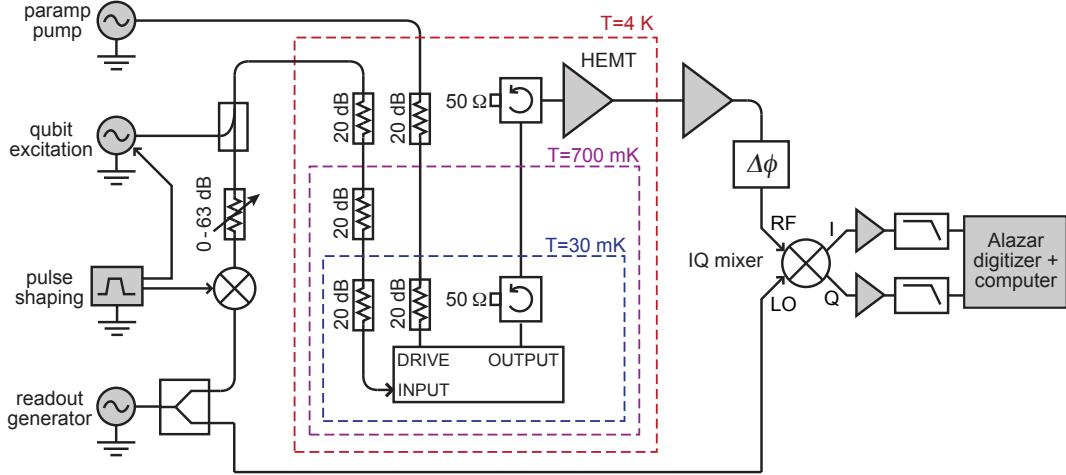


Figure 7.2: Overview of experimental setup.

Qubit and readout pulses are formed by mixing the output of microwave generators with pulse envelope shapes made by an arbitrary waveform generator. The readout power is set using a variable attenuator, and both qubit and readout pulses are combined and sent to the INPUT port of the experiment by a heavily attenuated input line. The paramp pump is supplied to the DRIVE port from a separate generator. After interacting with the qubit (see Figure 7.1), the amplified signal from the paramp leaves the OUTPUT port, is further amplified by a cryogenic HEMT and low-noise room-temperature microwave amplifiers, and is then mixed down to zero frequency with an IQ mixer. A phase shifter is used to adjust relative phase and put the information all in one quadrature of the mixer output. The mixer output is amplified, low-pass filtered at 11 MHz, and digitized at 100 MS/s.

are in Chapter 5, while details of the amplification mechanism are presented in Chapter 3. Fully detailed diagrams of the experimental setup can be found in Appendix C.

7.4 Observation of quantum jumps

Using this experimental apparatus, we can look at individual time traces of the mixed-down voltage signal and see quantum jumps of the qubit state. In these traces, different dc voltage levels correspond to different phases of the readout signal, which in turn correspond to different qubit states. We prepare the qubit in either the excited state or ground state before the readout turns on by applying an appropriate microwave pulse, as shown in Figure 7.3(a)³. In this instance, we prepared two traces in the excited state with

³Before these pulses are applied, the qubit is initialized in the thermal ground state (which is $|0\rangle$ to a good approximation since $\hbar\omega_q \gg k_B T$) by letting it equilibrate for a period of 50 or 100 μs , much longer than the qubit T_1 .

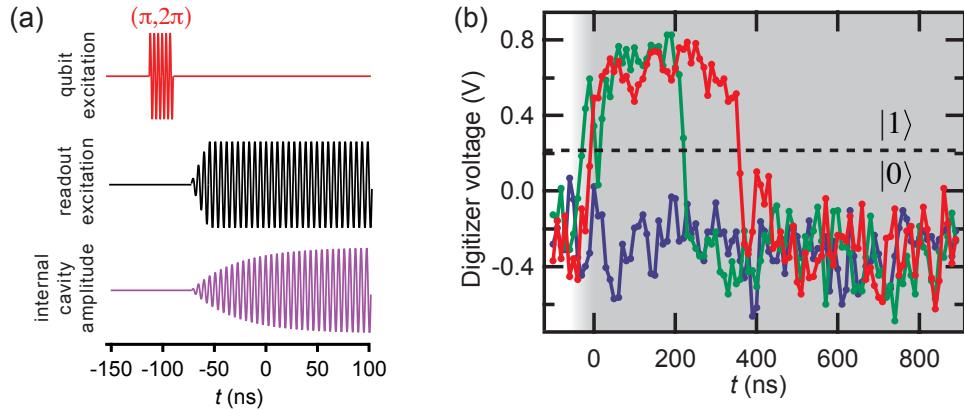


Figure 7.3: Individual measurement traces.

Part (a) shows the timing diagram for the data traces in (b). The qubit is prepared in either the excited or ground state with a π pulse or a 2π pulse, respectively (red). The readout is then energized (black) and the internal cavity amplitude (purple) rises, effecting a measurement. We select time $t = 0$ to be two cavity time constants after the readout is first energized. Part (b) shows three individual time traces, two prepared with a π pulse (red and green) and one prepared with a 2π pulse (blue). Data points are recorded every 10 ns. When the readout turns on (grey background), the excited state and ground state traces separate and settle to different voltages. As the measurement continues, the excited state traces decay stochastically to the ground state with an abrupt quantum jump. The qubit states are easily resolved from each other at all times during readout.

a π pulse (red and green) and one in the ground state with a 2π pulse (blue). Figure 7.3(b) shows what happens during the subsequent readout. When the readout is off, the voltage is the same level for all state preparations (this level is not quite 0 V, due to offsets in the mixer and its following amplifiers). However, as the readout turns on (grey background), the ground and excited state traces separate to the point where they can be clearly distinguished at any given time point (data points are taken every 10 ns). The dotted line in the figure shows a discrimination threshold for determining the state of the qubit based on statistics from many traces. As we know from ensemble measurements, the qubit will relax from the excited state to the ground state with a characteristic time constant T_1 . However, the continuous measurement forces the qubit to remain in an energy eigenstate. As a result, the qubit transitions stochastically to the ground state with an abrupt quantum jump, as can be seen in the red and green traces in Figure 7.3(b). No jump is evident in the blue trace because it began in the ground state.

There are many possible things which may cause a measurement trace to jump between two stable states in an experiment. For example, the paramp is bistable for some bias values, a fact utilized for bifurcation readout in the past [46]. To ensure that the jumps we are seeing are truly quantum jumps of the qubit, we must perform some checks. The

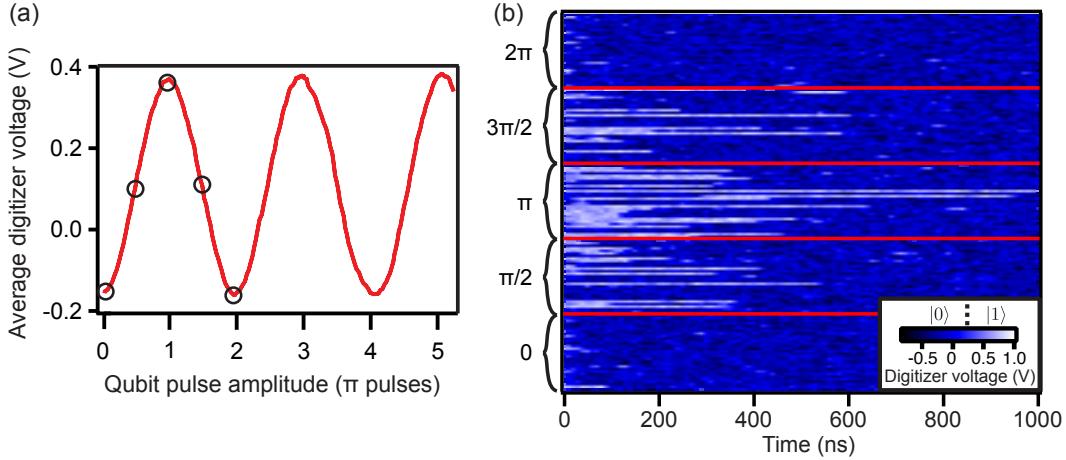


Figure 7.4: Jumps with varied state preparation.

Part (a) shows Rabi oscillations of the qubit as a function of the qubit excitation pulse amplitude. The data plotted are the average voltages of 10^4 traces in the 10 ns interval starting at $t = 40$ ns (as defined in Fig. 7.3(a)), where fidelity was optimal. Part (b) shows a color plot of 30 individual time traces for each of 5 initial qubit state preparations, corresponding to the circled points in (a). The white color represents the excited state, while the blue indicates ground state. After a π pulse, the qubit is primarily in the excited state at the start of the readout and jumps down stochastically in time. After no pulse or a 2π pulse, the qubit starts out in the ground state and stays there. Preparations of equal excited and ground state populations with a $\pi/2$ or $3\pi/2$ pulse show some traces in the excited state and some in the ground state at the start of the readout, in a fraction about midway between the 0 pulse and π pulse populations. The observed dwell times in the excited state are independent of state preparation.

simplest check is to prepare the qubit in different states to see if the measurement trace is correlated appropriately to the state preparation. Figure 7.4 shows data from such a test, where we prepared the qubit with pulses varying between 0 and $\sim 5\pi$. In part (a), we sweep the amplitude of the qubit excitation pulse and measure the average voltage across 10^4 traces for a 10 ns interval where the readout fidelity is maximized. This average voltage signal undergoes Rabi oscillations between ground and excited states as the excitation power is increased; this tends to rule out the idea that the qubit excitation pulse itself is causing the jumps. Part (b) shows a color plot of 30 individual time traces (running horizontally) for each of the five bias points circled in part (a); blue represents low voltage and white high voltage. We find that the trace voltages correlate with the prepared qubit state; when preparing $|0\rangle$ (either with no pulse or with a 2π pulse), we see blue at the start of measurement, while preparing $|1\rangle$ gives us mostly white at the start. When we prepare an equal superposition of $|0\rangle$ and $|1\rangle$, we see both blue and white in similar proportions.

There are a few details to note about this figure. First, even when we prepare

$|1\rangle$ with a π pulse, we measure $|0\rangle$ (blue) at the start of the traces a fair amount of the time. This loss of fidelity is due to the qubit relaxing before the readout can ring up, and will be discussed in more detail in section 7.5. We see the same effect after a $\pi/2$ pulse; here, the fraction of traces measured as $|1\rangle$ (white) at the start of the readout is somewhat below half, and is about halfway between what we get when preparing $|0\rangle$ and $|1\rangle$. A second important feature is that the typical high and low voltages measured after a $\pi/2$ pulse are the same as for other state preparations; only the fraction of traces showing high voltage versus low voltage changes. This means that the oscillations in average voltage seen in part (a) are due to changes in the fraction of traces showing high versus low voltage ($|0\rangle$ versus $|1\rangle$), not due to changes in the typical readout voltage levels. Finally, the typical lifetime of the excited state traces before decay appears to be independent of the qubit preparation pulse, as expected if the jumps represent T_1 decay of the qubit. We will examine the decay times in more detail shortly.

To see the robustly bimodal nature of the readout voltages more clearly, we can histogram many individual time traces. Figure 7.5 shows histograms of voltage versus time; each panel represents 2×10^4 individual traces for a given qubit state preparation. The voltage signal is well separated once the readout has rung up, and we can identify a threshold voltage for distinguishing between ground and excited states that is independent of qubit state preparation. Since the histogram represents many identical trials of the same experiment, we can talk about an ensemble qubit population in this instance. The excited state population decreases in time but remains at the same mean voltage, indicating that the individual traces decay with discrete jumps from $|1\rangle$ to $|0\rangle$ rather than slowly evolving to $|0\rangle$. We can also see the initial ground and excited state populations are correlated appropriately to the prepared qubit state. Finally, we see a small but noticeable excited state population at large times and when preparing the ground state. This is simply a steady-state thermal population of the qubit.

We have successfully checked that the jumps we observe are between two different voltages which are independent of qubit state preparation, but that the fraction of traces observed at a given voltage is well correlated to the qubit state preparation. This suggests that the jumps are indeed due to the qubit. However, we should also check that the jumps occur at appropriate times. Individual quantum jumps happen stochastically, but the distribution of times at which they occur is well understood [168]. For jumps from decay events, the time t_{jump} when each jump occurs follows an exponential distribution with the time constant given by the qubit T_1 ⁴.

We can find t_{jump} for each data trace by detecting when the voltage goes below the $|0\rangle/|1\rangle$ discrimination threshold. In Figure 7.6(a), we histogram the extracted values of t_{jump} for 2×10^4 traces. The times follow an exponential distribution with time constant $\tau_{\text{jump}} = 310$ ns, which is within the fitting error of the independently measured qubit T_1 of 320 ns. This correspondence again strongly suggests that the jumps we are seeing are those of the qubit. In part (b) of the figure, we show a 3D version of Figure 7.5(c), highlighting the exponential decay of the excited state population (nearer peak) with time, and the corresponding growth of the ground state population (far peak). The excited

⁴For very strong measurement, the time constant may be different due to the quantum Zeno effect, which we will discuss in the next chapter

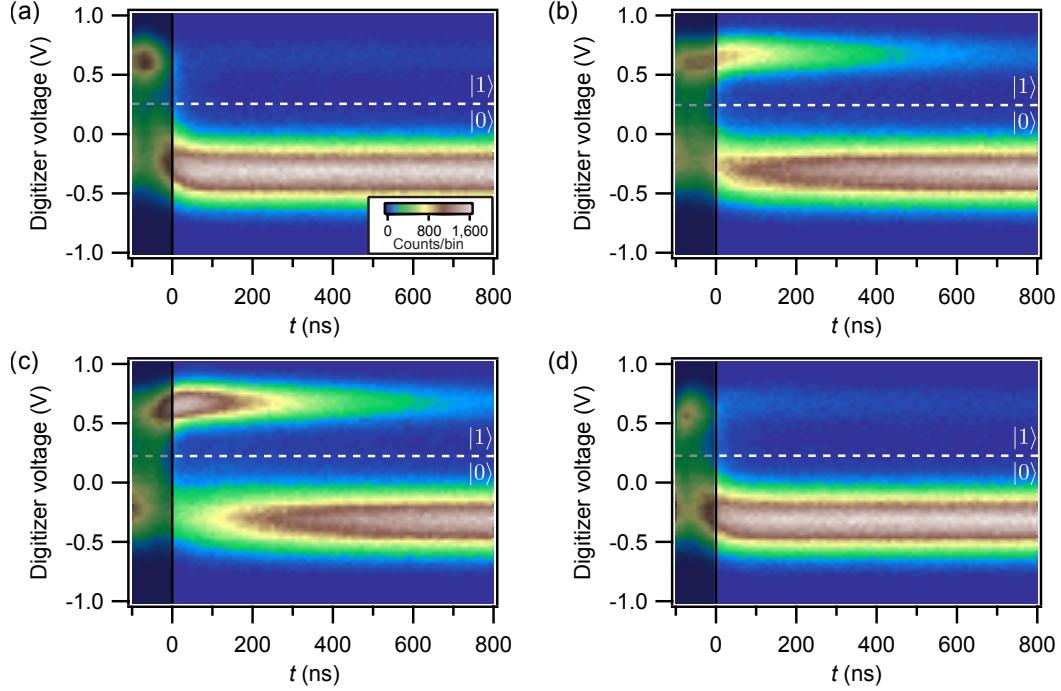


Figure 7.5: Histograms of many individual traces.

These color plots show histograms of the readout signal voltage as a function of time for 2×10^4 individual readout traces. The four panels correspond to four different qubit state preparations; (a) no pulse, (b) $\pi/2$ pulse, (c) π pulse, and (d) 2π pulse. We can draw a clean discrimination threshold (dotted white line) between ground and excited state voltages. The relative fraction of traces in ground versus excited states depends on the state preparation. As we go farther into the readout, the excited state population decays back to the ground state with quantum jumps, reflected in the unchanging mean voltages but changing weights of the ground and excited state parts of the histogram. The grayed-out region shows the voltage transients during qubit excitation and readout ring-up, before we consider the readout to be “on” at $t = 0$ ns.

state population, as measured by the histogram weight at voltages above the discrimination threshold, decays exponentially with a time constant of $\tau = 290$ ns, which again correlates well with both τ_{jump} and T_1 .

We make a brief comment here on determination of decay time constants. For the data in Figure 7.6(a), we did a least-squares fit of a curve to a histogram. This is an acceptable way to estimate the decay time constant, but note that for such a fit to be correct one must weight each bin by its standard deviation (given by the square root of the bin occupation). However, this method is not optimal because it uses coarse histogram bins, rather than the more fine-grained information about individual jump times which we have access to. A more robust and accurate way of estimating a parameter such as a

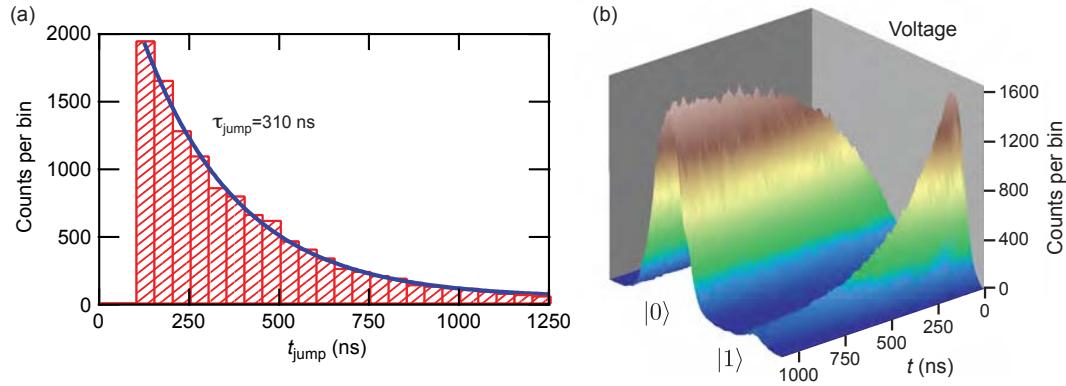


Figure 7.6: Jump times and population decay.

In part (a), we show a histogram of values of t_{jump} extracted from 2×10^4 traces prepared with a π pulse. Fitting the decay times to an exponential distribution (blue line) yields a time constant $\tau_{\text{jump}} = 310 \text{ ns}$. This is within fitting error of the independently measured qubit T_1 of 320 ns. Part (b) shows the data from Figure 7.5(c) as a 3D plot, allowing us to see more clearly the exponential decay of the excited state population and corresponding rise in ground state population with time.

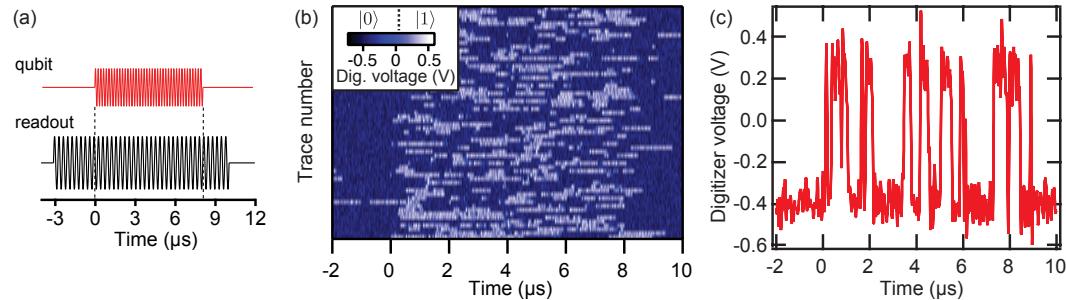


Figure 7.7: Jumps with simultaneous readout and excitation.

Part (a) shows a timing diagram with simultaneous readout and qubit excitation. Part (b) shows 60 individual time traces. The qubit is usually in the ground state when the qubit excitation tone is off, and then jumps stochastically back and forth between ground and excited states when the excitation tone is on. Once the excitation turns off again at 8 μs , we see essentially no more jumps up to the excited state. Part (c) shows one individual time trace with repeated jumps between ground and excited states.

decay constant is to use maximum likelihood estimation, which will be discussed in the next chapter. The maximum likelihood estimate for τ_{jump} is 296 ns, again consistent with the measured qubit T_1 . Maximum likelihood estimation will be described in detail in the next chapter.

Finally, we can look at quantum jumps in the presence of a simultaneous qubit excitation tone (we set the frequency of this tone to the appropriately ac-Stark-shifted qubit frequency). The qubit excitation should cause the qubit state to evolve in time, but the measurement pins the qubit in an eigenstate. Under these two competing influences, we expect the qubit to jump back and forth between the ground and excited states. This is just what we see experimentally, as shown in Figure 7.7. With just the readout tone on initially, the qubit is primarily in the ground state (with the occasional jump visible). However, when the qubit excitation tone is on as well, we see frequent jumps back and forth between the ground and excited states. With very weak measurement, we would see Rabi oscillations of the qubit if we averaged many of these traces together. However, because of the strong measurement the qubit is rapidly dephased and the up/down jumps from multiple traces do not average together to give Rabi oscillations. This transition from smooth qubit state evolution to jump-like evolution has been seen in ensemble measurements in the frequency domain by the Saclay group [23], and we have reproduced these results. As the measurement grows stronger, causing a transition to the jumping regime, it actually inhibits qubit evolution, a phenomenon known as the quantum Zeno effect [169]. We show measurements of the quantum Zeno effect in the next chapter.

7.4.1 Jumps without a paramp

The jumps we observe correlate well with the qubit state preparation and excited state lifetime, suggesting that they are indeed quantum jumps of the qubit. However, since we know that the paramp can also exhibit bistability, it would be nice to see quantum jumps without the paramp present. This is quite difficult to do, as the noise performance of the HEMT amplifier is much worse than that of the paramp. However, we were able to reduce the system noise temperature to 7 K by using a Nb coaxial cable to transmit the output signal between the mixing chamber plate and the 4 K plate (details are in Chapter 5). With this reduced noise temperature, we could look for jumps both with and without the paramp in the circuit.

Data from this experiment⁵ are shown in Figure 7.8. We turn on the readout and apply a π pulse to the qubit (at the ac-Stark-shifted qubit frequency) after a 3 μs delay, as shown in part (a). We did this for a strong measurement power, $\bar{n} = 60$, both (b) with the paramp on and (c) with the paramp off and switched out of the circuit. While the data without the paramp are much noisier, one can see that both data sets look similar. With the readout on, there is some spurious excited state population at all times (more details in Chapter 8). When the π pulse occurs, a majority of the traces show a transition to the excited state, with subsequent relaxation to the ground state (the π pulse efficiency is low due to the strong measurement). The ground and excited state populations as a function

⁵These data we obtained with a different qubit sample (TF02111d) than the data in the previous section. This sample has a transmission-mode cavity (so qubit pulses are filtered by the cavity and do not impinge on the paramp) and a longer T_1 of 1.27 μs .

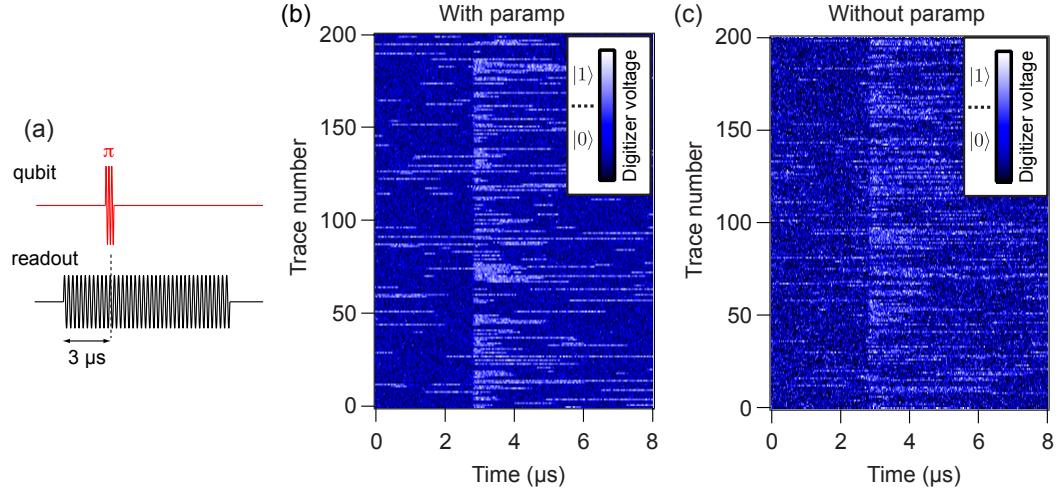


Figure 7.8: Quantum jumps with and without paramp.

Part (a) shows the timing diagram for parts (b) and (c); the readout is turned on, and the qubit is excited with a π pulse at the ac-Stark-shifted qubit frequency $3 \mu\text{s}$ into the readout. Parts (b) and (c) show 200 individual time traces with readout power $\bar{n} = 60$; part (b) shows data taken using the paramp, while the data in (c) were taken with the paramp off and removed from the circuit. The jump timescales and behavior in (c) are commensurate with those in (b), albeit with considerably more noise.

of time, as well as the excited state lifetimes, appear to be commensurate between the two data sets. We conclude that the jumps are not due to the presence of the paramp, but that the paramp merely serves to improve our noise performance.

7.4.2 Three-level jumps

Our quantum jump readout allows us to distinguish between ground and excited states, but we must take into account that the transmon qubit is really a multilevel system. We can attempt to make measurements to distinguish between $|0\rangle$, $|1\rangle$, and $|2\rangle$, which can work because the low anharmonicity of the transmon means that the qubit-state-dependent cavity frequencies $\tilde{\omega}_{\text{cav}}(|0\rangle)$, $\tilde{\omega}_{\text{cav}}(|1\rangle)$, and $\tilde{\omega}_{\text{cav}}(|2\rangle)$ are roughly evenly spaced. If we choose our readout frequency to be $\omega_{\text{ro}} \approx \tilde{\omega}_{\text{cav}}(|1\rangle)$, we should be able to see three different phase shifts, corresponding to the three different qubit states⁶. Such a measurement would let us observe the qubit dynamics outside the $\{|0\rangle, |1\rangle\}$ subspace. It would also be another strong piece of evidence that these really are quantum jumps of the qubit and not a spurious experimental bistability, since we would see tristable as well as bistable readout traces.

Figure 7.9(a) shows three individual time traces⁷ from a measurement with simu-

⁶Note that for the measurements discussed earlier in this chapter, we have chosen a readout frequency fairly close to $\tilde{\omega}_{\text{cav}}(|0\rangle)$.

⁷These data were obtained with qubit sample TF042811b.

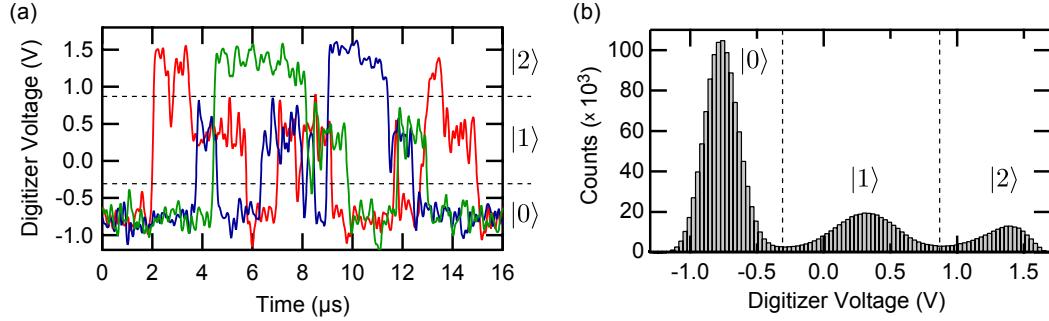


Figure 7.9: Three-level jumps.

Part (a) shows three individual time traces from a measurement with simultaneous strong driving of the $0 \rightarrow 1$ transition. The signal exhibits three stable voltages, corresponding to the qubit states $|0\rangle$, $|1\rangle$, and $|2\rangle$, with clear discrimination thresholds. Most relaxation events occur one level at a time, with the qubit “stepping down” the energy ladder. Part (b) shows a histogram of the voltages of 10^4 individual time traces at all time points, again showing the tristability of the readout signal and the discrimination thresholds.

taneous strong driving of the $0 \rightarrow 1$ transition. The traces jump back and forth between three stable levels, corresponding to the qubit states $|0\rangle$, $|1\rangle$, and $|2\rangle$. The $|2\rangle$ state is populated due to spurious transitions arising from measurement non-idealities (more details in Chapter 8). We can verify which voltage corresponds to each state by preparing the qubit state in one of the three states before the readout with either no pulse, a pulse at ω_{01} , or a pulse at $\omega_{02}/2$, and noting the voltage when the readout turns on. We observe that the decay transitions tend to occur “stepwise”, going one level at a time. In part (b), we histogram the voltages for all time points across 10^4 of these traces. The histogram clearly shows the three stable voltage levels, as well as discrimination thresholds for state determination.

When operating the paramp in phase sensitive mode, we can rotate the phase of the pump generator to change the phase of the amplified signal quadrature. In this experiment, there are three different readout pointer states depending on the state of the qubit. By simply varying the quadrature we amplify, we can change the output signal so that it distinguishes between $|0\rangle$ and $\{|1\rangle, |2\rangle\}$ or individually between $|0\rangle$, $|1\rangle$, and $|2\rangle$. Figure 7.10(a) shows a cartoon of this process, where each qubit state corresponds to one of three complex readout amplitudes $\alpha_{|0\rangle}$, $\alpha_{|1\rangle}$, and $\alpha_{|2\rangle}$. The signal into the paramp is just the projection of these complex amplitudes onto the amplified quadrature of the paramp. Two example quadratures are shown, one which distinguishes all three states and another which just distinguishes two.

Even though the paramp is in the saturated regime, we can choose a bias point so that the middle of the three readout states (corresponding to $|1\rangle$ in the data shown above) is right on the steep part of the paramp transfer function, with the other two readout states on the two saturated ends of the transfer function. This configuration is shown schematically in Figure 7.10(b). In this way, we can make a continuous three-state detector,

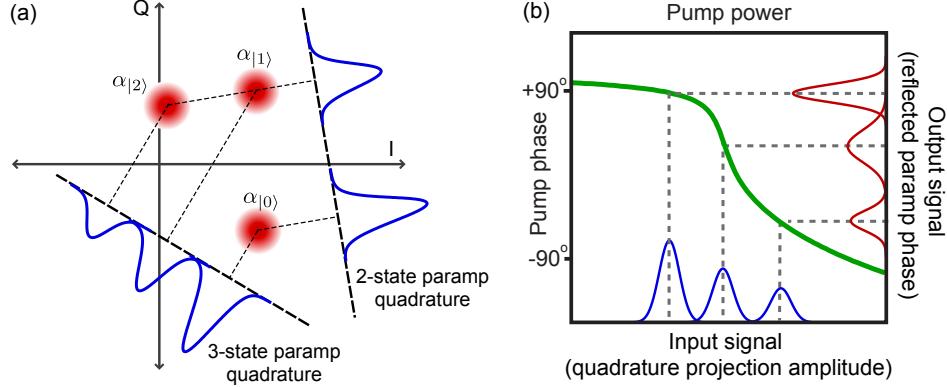


Figure 7.10: Paramp biasing for three-level jumps.

Part (a) shows the complex amplitudes $\alpha_{|0\rangle}$, $\alpha_{|1\rangle}$, and $\alpha_{|2\rangle}$ of the readout pointer states in the IQ plane. The projection of these states onto the amplified quadrature of the paramp is different depending on the quadrature chosen; one can choose this quadrature to distinguish between either $|0\rangle$ and $\{|1\rangle, |2\rangle\}$ or between all three states individually. The paramp quadratures are shown displaced from the origin for clarity. Part (b) shows the paramp transfer function, whose output, shown in red, can distinguish between three input values, shown in blue, even in the saturated regime. To distinguish the three inputs efficiently, the middle value to be amplified should be biased near the middle of the paramp transfer function.

at the expense of the signal-to-noise ratio for discriminating between states $|0\rangle$ and $|1\rangle$. If all we care about is whether the qubit is in the ground state or an excited state, we can usually improve our state discrimination by operation in the two-state rather than three-state regime.

7.5 Signal-to-noise ratio and fidelity

Having established that the jumps we see are indeed quantum jumps of the qubit, we can now try to quantify how well our technique lets us know the qubit state. We can look at two figures of merit for our measurement technique. The first is our ability to discriminate between the readout pointer states at a given time point, which we refer to as the signal-to-noise ratio (SNR). The second is our ability to prepare the qubit in a desired state and then obtain a measurement result corresponding to that state, which we refer to as the single-shot fidelity F . The SNR obviously enters into the fidelity; if we are unable to distinguish between our two readout pointer states, we will be unable to measure the qubit state reliably.

If the SNR is poor enough to be the main limit on fidelity (as is the case in typical circuit QED experiments), averaging in time will improve fidelity, but only until we reach the limits set by T_1 decay or errors in qubit state preparation [94]. In our quantum jump

experiments, the SNR is sufficiently high that the ability to distinguish between readout pointer states should not be the limiting factor on our total fidelity. We will describe this in more detail in section 7.5.3.

7.5.1 Theoretical SNR

We now present calculations of the expected SNR for resolving readout pointer states. The signal depends on the difference between the two coherent states coming from the readout cavity corresponding to the two qubit states. To extract the signal, we need to know the amplitude of these states and the angle between them. The noise comes from standard noise theory as detailed in Chapter 3.

For a readout cavity with an average photon occupation \bar{n}_{res} when driven on resonance, the power radiating out of the cavity is given by $P_{\text{rad}} = \bar{n}_{\text{res}}\hbar\omega\kappa$, where ω is the readout frequency and κ is the cavity linewidth. When the cavity is driven in the reflection geometry, the steady state incident power is the same as the steady state reflected power which forms the output. This output power is $P_{\text{out}} = P_{\text{rad}}/4$, where the factor of four arises due to the interference between the incident and radiated signals [135]. Assuming that the internal loss of the cavity is small (in the sense that $Q_{\text{int}} \gg Q_{\text{ext}}$), P_{out} is actually the same at all bias frequencies, not just the resonance frequency. This power corresponds to a peak voltage amplitude $V_{\text{out},r}$ through the relation $|V_{\text{out},r}| = \sqrt{2}\sqrt{50P_{\text{out}}}$, where we have used the fact that $P = V_{\text{rms}}^2/R$, and that the amplitude of a sinusoidal signal is $\sqrt{2}$ times its RMS value. Substituting for P_{out} gives

$$|V_{\text{out},r}| = \sqrt{25\bar{n}_{\text{res}}\hbar\omega\kappa}. \quad (7.3)$$

Depending on the qubit state, this output voltage will have one of two phases. The readout signal changes between two points in the IQ plane corresponding to these two phases with the given amplitude $|V_{\text{out},r}|$. Thus the peak amplitude of the readout signal is given by half the vector difference between $V_{\text{out},r}$ with one phase and $V_{\text{out},r}$ with the other phase. We can calculate the magnitude of this vector difference, giving

$$|V_{\text{sig},r}| = |V_{\text{out},r}| \sin \frac{\theta}{2} = \frac{1}{2}|V_{\text{out},r}|\sqrt{2 - 2 \cos \theta}, \quad (7.4)$$

where the r subscript indicates reflection geometry. This result is derived from simple trigonometric arguments. We can substitute in the expression in (7.3) to give:

$$|V_{\text{sig},r}| = \sqrt{12.5 \times \bar{n}_{\text{res}}\hbar\omega\kappa(1 - \cos \theta)}. \quad (7.5)$$

The signal is maximized for $\theta = \pi$, which is achieved when the dispersive shift 2χ is equal to κ , with the readout frequency set halfway between $\tilde{\omega}_{\text{cav}}(|0\rangle)$ and $\tilde{\omega}_{\text{cav}}(|1\rangle)$. The signal at this optimal bias point is

$$|V_{\text{sig},r}| = \sqrt{25\bar{n}_{\text{res}}\hbar\omega\kappa}. \quad (7.6)$$

Having solved for the signal in reflection geometry, we turn to an asymmetric transmission geometry with the signal exiting from the strongly coupled port. With the same caveat of low internal cavity loss, we find that the power radiated from the cavity is again $P_{\text{rad}} =$

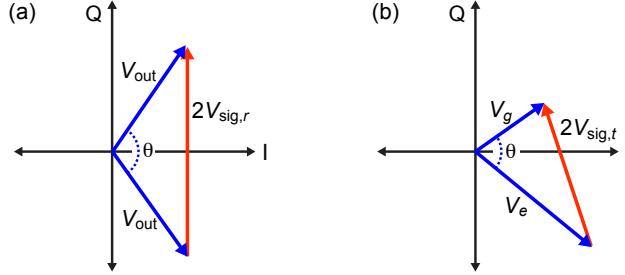


Figure 7.11: Readout signal cartoon.

The signal to the paramp in reflection (a) and transmission (b) geometries is determined by the vector difference between the two pointer states in the IQ plane. The amplitudes of these pointer states are given by equations (7.3), (7.7), and (7.8). The signal amplitude is half the peak-to-peak excursion, and thus half of the vector difference between the pointer states.

$\bar{n}\hbar\omega\kappa$. However, in transmission there is no interference between incoming and outgoing waves, so $P_{\text{rad}} = P_{\text{out}}$. Furthermore, P_{rad} depends on \bar{n} , which decreases as we move away from the cavity resonance frequency (assuming a fixed cavity drive power). This means that both the phase and amplitude of the output signal may change depending on the qubit state. In analogy with equation (7.3), we can write the peak voltage amplitudes corresponding to the ground and excited states as

$$|V_{g,t}| = \sqrt{100\bar{n}_g\hbar\omega\kappa} \quad (7.7)$$

$$|V_{e,t}| = \sqrt{100\bar{n}_e\hbar\omega\kappa}, \quad (7.8)$$

where \bar{n}_g and \bar{n}_e are the ground and excited state cavity photon occupations, respectively. The output signal is again half of the vector difference between $V_{g,t}$ and $V_{e,t}$ in the IQ plane. Using the law of cosines, we can express the signal amplitude as

$$|V_{\text{sig},t}| = \frac{1}{2}\sqrt{|V_{g,t}|^2 + |V_{e,t}|^2 - 2|V_{g,t}||V_{e,t}|\cos\theta}, \quad (7.9)$$

where here the t subscript indicates transmission geometry. If we substitute in the results from equations (7.7) and (7.8), this becomes

$$|V_{\text{sig},t}| = \sqrt{25 \times \hbar\omega\kappa(\bar{n}_g + \bar{n}_e - 2\sqrt{\bar{n}_g\bar{n}_e}\cos\theta)}. \quad (7.10)$$

We find that this signal is maximized when $\theta = \pi/2$, corresponding to $\kappa = 2\chi$ and readout frequency midway between $\tilde{\omega}_{\text{cav}}(|0\rangle)$ and $\tilde{\omega}_{\text{cav}}(|1\rangle)$. These are the same parameters for which the signal was maximized in reflection geometry⁸. At this bias point, $\bar{n}_g = \bar{n}_e = \frac{1}{2}\bar{n}_{\text{res}}$. If

⁸It is not immediately evident that equation (7.10) is maximized for $\theta = \pi/2$, but arises because \bar{n}_g and \bar{n}_e are not independent of θ . However, it should not be surprising that the same relation between κ and χ gives the maximum signal for both reflection and transmission. The proof is left as an exercise to the reader. Hint: draw pictures in the IQ plane.

we substitute in this relationship, equation(7.10) reduces to

$$|V_{\text{sig},t}| = \sqrt{25\bar{n}_{\text{res}}\hbar\omega\kappa}, \quad (7.11)$$

which is exactly the same result for maximum signal in the reflection case given in equation (7.6). We see that in both transmission and reflection, the output signal is maximized when $2\chi = \kappa$; this is why we aim for this regime with our sample bias parameters.

Having derived expressions for the signal, we now examine the noise. The intrinsic noise floor is set by quantum fluctuations and can be expressed as a noise temperature $T_Q = \hbar\omega/2k_B$. The corresponding voltage noise is given by $V_Q = \sqrt{50k_B T_Q B}$, where B is the noise bandwidth. The noise bandwidth is set by our integration time; the longer we average the readout signal, the lower the noise bandwidth and thus the lower the noise. In practice, B is effectively given by the bandwidth of the filters on the demodulated signal (see section 5.3.3). However, the paramp bandwidth also plays a role, because the noise on the output signal consists primarily of amplified paramp output noise. If the paramp bandwidth is smaller than the demod filter bandwidth, then B will depend on both⁹. When setting B to minimize noise, we must also consider the bandwidth of our desired signal, which is given by the cavity amplitude response rate $\kappa/2$. If we make B larger than the signal bandwidth, we introduce unnecessary noise, while if we make B smaller we will begin to lose signal¹⁰.

If our amplification chain has a system noise temperature T_{sys} , then the effective voltage noise referred to the input of the chain is given by

$$V_{\text{sys}} = \sqrt{50k_B(T_Q + T_{\text{sys}})B}, \quad (7.12)$$

combining both the intrinsic quantum noise and the noise added by amplification. We can now derive the theoretically expected signal-to-noise ratio. In the case of noiseless amplification (paramp in phase-sensitive mode), we have $T_{\text{sys}} = 0$. If we maximize the signal (which involves both having $\kappa = 2\chi$ and choosing the amplified quadrature of the paramp appropriately), its amplitude is given by the expressions in equations (7.6) and (7.11). We then calculate the maximum achievable signal to noise ratio given a noise bandwidth B :

$$\text{SNR}_{\text{max}} = V_{\text{sig}}/V_{\text{sys}} = \sqrt{25\bar{n}_{\text{res}}\hbar\omega\kappa/50k_B(\hbar\omega/2k_B)B} = \sqrt{\bar{n}_{\text{res}}\kappa/B}. \quad (7.13)$$

This is a pleasing result; the maximum SNR is just the square root of the photon flux out of the cavity on resonance divided by the noise bandwidth (which depends on the integration time). We can also derive expressions for the SNR in reflection or transmission for arbitrary θ and \bar{n} (assuming noiseless amplification) using (7.5), (7.10), and (7.12):

$$\text{SNR}_r = \sqrt{\bar{n}_{\text{res}}\kappa(1 - \cos\theta)/2B} \quad (7.14)$$

$$\text{SNR}_t = \sqrt{\kappa(\bar{n}_g + \bar{n}_e - 2\sqrt{\bar{n}_g\bar{n}_e}\cos\theta)/B}. \quad (7.15)$$

⁹Ref. [95] gives methods for calculating B given the demod filter and/or paramp frequency response.

¹⁰In this section, we assume that the signal and noise are filtered with a boxcar frequency response. The question of optimal filtering of the readout signal to give maximum information about the qubit state is complex, and will be considered further in section 7.5.3.

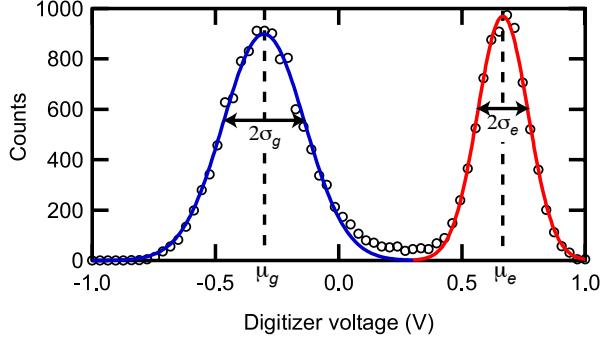


Figure 7.12: Readout SNR.

The black circles are a voltage histogram of 2×10^4 readout traces at $t = 200$ ns, taken from a linecut of Figure 7.5(c). The blue and red curves are Gaussian fits to the ground and excited state peaks, respectively. The means (μ_g, μ_e) and standard deviations (σ_g, σ_e) of each fit are shown. The data are slightly above the fits in the region around $V=0.25$ V; this effect is due to traces making a transition between excited and ground states.

As would be expected, increasing the cavity photon occupation makes the SNR go up, as does decreasing the noise bandwidth (increasing the integration time). One should be careful to note that κ is an angular frequency, while B is not. For the qubit described in section 7.3, measured in reflection with $\kappa/2\pi = 4.9$ MHz, $\theta = 155^\circ$ and $B \approx 15$ MHz¹¹, we get a theoretical maximum SNR $\approx 1.4\sqrt{\bar{n}_{\text{res}}}$.

7.5.2 Experimental measurements of SNR

The SNR can be improved by increasing \bar{n}_{res} , but in practice higher order effects start to reduce the dispersive shift [93] and cause qubit state mixing (see Chapter 8 for details), setting a practical upper bound on \bar{n}_{res} for our measurement. In our initial experiment, we used a maximum photon occupation of $\bar{n}_{\text{res}} = 47$. For this level of excitation, we compute an optimal SNR ≈ 9.6 .

We can compare this number with the SNR we achieve experimentally. Figure 7.12 shows a voltage histogram of 2×10^4 individual traces for $\bar{n}_{\text{res}} = 47$; the data are from a vertical linecut of Figure 7.5 (c) at $t = 200$ ns. There are two clearly resolved peaks, corresponding to the qubit in ground (negative voltage) and excited (positive voltage) states. We can fit each peak to a Gaussian function (blue and red lines) and extract the means (μ_g, μ_e) and standard deviations (σ_g, σ_e) for each fit. Note that the widths of the two peaks are different because the paramp is not linear for this high level of excitation. We define the measured SNR as

$$\text{SNR}_{\text{meas}} = |\mu_g - \mu_e| / (\sigma_g + \sigma_e). \quad (7.16)$$

This definition of SNR is consistent with the analytical expressions given in the previous

¹¹This value of the noise bandwidth was determined from the frequency response of the filter on the demodulated signal, as well as the bandwidth of the paramp.

section¹². For the data shown $\text{SNR}_{\text{meas}} \approx 3.75$, a factor of 2.6 smaller than the ideal case. This implies that $T_{\text{sys}} \approx 5.6 T_Q = 0.8$ K and the added noise is not zero. Nevertheless, this is an improvement of more than an order of magnitude in noise temperature over a typical state-of-the-art microwave amplification chain using cryogenic semiconductor amplifiers, where $T_{\text{sys}} \approx 10 - 30$ K. The experimental parameters give a signal power $P_{\text{out}} = \bar{n}_{\text{res}} \hbar \omega \kappa / 4 = -118.5$ dBm. The extracted value of T_{sys} is commensurate with the values measured for similar paramp input powers as presented in Chapter 6.

For the data in Figure 7.12, the paramp is deeply in the saturated regime. In the linear regime, the output SNR grows linearly with the input SNR from the cavity; the output signal grows, while the output noise stays constant. When the paramp saturates, the amplitude of the output signal no longer grows with increasing input signal amplitude. Input noise continues to be amplified, although the noise gain decreases gradually as the paramp moves farther into the saturated regime. As a result, the output SNR plateaus and grows only sub-linearly with increasing input SNR; the output signal is fixed, while the output noise is slowly decreasing. Thus when the paramp is deeply saturated, the output SNR will be considerably lower than the input SNR. This is another way of understanding the increase in paramp noise temperature in the saturated regime (see discussions in Chapters 3 and 6). It is important to note, however, that the output SNR is still a monotonically increasing function of input SNR, so that operation in the saturated regime is advantageous for maximizing readout fidelity. Improving the dynamic range of the paramp will prevent saturation and should allow further improvements in SNR.

7.5.3 Measurement fidelity

Having determined our experimentally achieved SNR, we turn to the examine our overall qubit measurement fidelity. We prepare the qubit in either the ground or the excited state and then measure its state. By taking a number of measurements, we can find an ensemble probability to measure the qubit in the ground state when it was prepared in the excited state, denoted $P(s_0|q_{|1})$, and a corresponding probability $P(s_1|q_{|0})$ to measure excited state when the qubit was prepared in the ground state. The total fidelity is defined by subtracting these errors:

$$F = 1 - P(s_0|q_{|1}) - P(s_1|q_{|0}) \quad (7.17)$$

These probabilities are determined by the probability of the qubit actually being in the desired state $|i\rangle$ when measured, which we term the state-specific process fidelity $F_{|i\rangle}$, and our ability to determine the qubit state from the readout signal, which we call the readout efficiency η_{ro} . $F_{|i\rangle}$ depends on the quality of our ground state initialization and qubit gates, as well as on qubit excitation and decay rates and the delay time between state preparation and readout; the theoretical expression has a complicated form. The circuit QED readout method has been shown experimentally to map the qubit state to the phase of the readout signal with unit efficiency [44], in which case the readout efficiency is just our ability to distinguish between the two states of the readout signal at a given time point. This is

¹²We checked that these definitions are consistent by creating simulated noisy readout traces (with signal and noise amplitudes as given by the expressions in the previous section) and extracting SNR_{meas} .

determined by the SNR, and for the definition of SNR given in the preceding sections is $\eta_{\text{ro}} = \text{erf}(\text{SNR}/\sqrt{2})$ [94]. This formula implicitly assumes an effective signal averaging time of $1/B$, where B is the noise bandwidth that enters into the SNR. Because η_{ro} is so high in our experiment (almost always above 95% and typically above 99%), and because the T_1 of our qubit is relatively short, averaging further in time does not improve the overall fidelity. The effects on readout fidelity of averaging or filtering a continuous readout signal have been studied carefully by Gambetta and co-workers in the Yale group [94]. They set limits on the achievable fidelity for a given signal-to-noise ratio and qubit decay time. The reader is cautioned that they define their signal-to-noise ratio r_{sn} differently from how we define our SNR. The conversion is given by $r_{sn} = T_1 B(\text{SNR})^2$ in the case of simple average filtering (the relation is more complex or even not analytical in the case of other filters). Our experimental value $\text{SNR}_{\text{meas}} = 3.75$ given in the previous section translates to a value $r_{sn} = 45$ for the signal-to-noise ratio definition in [94].

For the qubit sample presented in this chapter, our single-shot fidelity was around 70%, which is within the error bars of the theoretically predicted value. The 30% fidelity loss is almost entirely due to T_1 decay. Using a different sample¹³ with longer $T_1 = 910$ ns and $\bar{n} = 80$ photons, we achieved our best single-shot fidelity of $81 \pm 2\%$, which agrees well with the theoretically predicted value of $82 \pm 3\%$. The individual error probabilities for equation (7.17) are found experimentally to be $P_{\text{ex}}(s_0|q_{|1}) = 0.17 \pm 0.02$ and $P_{\text{ex}}(s_1|q_{|0}) = 0.015 \pm 0.004$. These numbers are very close to the theoretically predicted values $P_{\text{th}}(s_0|q_{|1}) = 0.16 \pm 0.03$ and $P_{\text{th}}(s_1|q_{|0}) = 0.018 \pm 0.01$. The loss of fidelity (both theoretically and experimentally) is primarily due to T_1 decay and to the thermal excited state population of the qubit. The good agreement between theoretical and experimental fidelity values means that our assumption of perfect mapping between qubit state and readout photon state is reasonable.

Although these fidelity numbers are not quite as high as some achieved in other superconducting systems [47, 48], the fidelity as quoted above does not fully capture the power of our measurement technique. The switching phase qubit readout [48] has better than 90% fidelity and is very rapid (readout takes a few ns). However, the readout is not QND because the qubit junction switches to the voltage state if state $|1\rangle$ is measured, thereby destroying the qubit. The latching bifurcation readout used in [47] is QND and has 86% fidelity (which can be increased to 92% with the use of an additional “shelving” qubit state), but measurements can only be made once every 250 ns. This is too slow to resolve quantum jumps when the qubit lifetime is of order $\sim 1 \mu\text{s}$. In addition, use of the “shelving” state to increase fidelity takes the qubit out of the $\{|0\rangle, |1\rangle\}$ manifold and so continuous qubit dynamics cannot be observed.

In contrast, our ability to determine the qubit state when the readout is on is limited only by η_{ro} and the rate $\kappa/2$ at which the readout signal amplitude responds to changes in the qubit state. We can perform continuous monitoring of the qubit state with error rates in qubit state determination at any given point in time of around 2-4%, depending on averaging and readout strength. This enables studies of qubit dynamics (such as those presented in this chapter and the next) which are not possible using the other readout methods described in the previous paragraph.

¹³Sample TF042811b: see Chapter 4.

Since our fidelity is essentially set by T_1 decay between state preparation and readout, we could improve the fidelity by reducing this delay time and/or increasing T_1 . It is possible to prepare transmon qubit states with better than 99% efficiency [76, 170], and the recent development of transmon qubits with very long lifetimes [17] should reduce losses from decay. Using optimized state preparation pulses with a qubit with $T_1 = 10 \mu\text{s}$, we expect a fidelity above 95% given the same pulse timing and SNR as were used in the experiments above.

Chapter 8

Measurement backaction

With the ability to perform continuous monitoring and see quantum jumps, we have a powerful tool for probing the dynamics of our qubit. As with all things quantum, the act of measurement is bound to affect the system under study. This chapter looks at the backaction of the circuit QED measurement technique. We first present a method for efficient automated analysis of individual qubit measurement traces, and then use this method to study the behavior of the qubit during measurement. We observe and describe two forms of measurement backaction, spurious qubit state mixing and the quantum Zeno effect.

8.1 Automated qubit state extraction

The techniques for continuous monitoring of the qubit state presented in the previous chapter let us observe the dynamics of the qubit with high accuracy in a single time trace. We can now study ensemble behavior (for example decay times, as shown in Figure 7.6) by extracting the behavior in individual time traces and looking at ensemble statistics of the extracted parameters. This allows us to see effects and quantify behavior in ways which are much harder or impossible to do using the traditional method of averaging together many time traces and then extracting parameters from the average. However, the disadvantage of this ability to extract the qubit dynamics from individual time traces is that we need to save and process tens of thousands of individual time traces, rather than just a single averaged time trace, at each experimental bias point.

With data sets of many gigabytes for typical experiments, we need a reliable automated method for analyzing individual time traces and determining ensemble behavior of interest. We devised a suitable algorithm to determine the qubit state as a function of time during a measurement trace, noting when jumps occur and how long the qubit remains in each state before jumping. The algorithm also aggregates this information over many traces with identical experimental parameters and determines ensemble qubit behavior such as excited state population and excitation and decay rates. With only a small amount (~ 10 parameters) of human input at the very beginning of operation, the extraction algorithm can run unattended for days and achieve accurate results processing tens of gigabytes of data over a wide range of experimental parameters. The algorithm currently only works for

analyzing jumps between two qubit states. It could be extended to analyze jumps between three states (for data such as that shown in section 7.4.2), but because of the substantial work involved this capability has not been implemented yet.

8.1.1 Determining the qubit state

The first task, from which all others follow, is the extraction of the qubit state as a function of time in each individual measurement trace. We accomplish this by removing noise in a controlled way to boost the SNR, then performing level detection to determine the initial state and the positions of all transitions between states.

The noise power in the signal is determined by the noise bandwidth B , typically set by the 10-15 MHz cutoff frequencies of the filters between the demod box and the Alazar card (see section 5.3.3). The value of B sets the SNR as described in the previous chapter. However, the qubit usually remains in a given state longer than $1/B$, suggesting that we can filter the recorded signal in software to further reduce B and improve our SNR without losing too much information about the qubit state. For an automated system, we need to have an algorithmic way of determining the optimal level of filtering for a given data set.

For each set of time traces (usually 10^4 traces, recorded over a period of about 0.1-1 sec, and corresponding to one particular experimental bias point), we histogram the voltages of all the data points during readout (typically $\sim 10^7$ total points)¹. We then smooth the histogram slightly and take its derivative, using the zero-crossings to find the locations of the two peaks corresponding to ground and excited state voltages, as well as the location of the minimum between them, which serves as the state discrimination threshold². The height of this minimum is an indicator of the SNR; lower height indicates that the peaks are better resolved. To optimize the SNR, we apply a filter to the individual time traces, repeat this histogram/derivative procedure, and find the height of the minimum again. If the minimum is decreased relative to what it was before the last round of filtering, we have improved the SNR. We can repeat this procedure until the SNR is no longer improved (i.e. the inter-peak minimum is no longer decreased) by additional filtering, at which point we have determined the optimal filtering. Further filtering (“overfiltering”) means that the noise bandwidth B has become low enough that we are losing substantial signal information. We show three sample histograms—unfiltered, optimally filtered, and overfiltered—in Figure 8.1.

There are several ways one could perform this filtering. Because we have a great deal of data to process, fast execution time is important, so we filter using a simple, speedy boxcar smoothing function (a moving average) [171]. The boxcar width can be chosen as desired at the start of extraction. The algorithm uses successive applications of the boxcar smoothing function at each iteration, so that optimizing the SNR means optimizing the number of passes of boxcar smoothing applied to the raw data traces. One could in principle

¹Raw data traces typically have “dead” spots at the beginning and end, before the readout has turned on and after it has turned off. We use only data taken while the readout is on in our automated analysis. We specify the data range of interest by giving the algorithm turn-on and turn-off times for the readout signal, which are chosen to be consistent throughout a given experiment.

²For traces with initially poor SNR, the two peaks may overlap so much that there is no minimum between them, especially if one peak is much higher than the other. In this case, the algorithm performs its iterative filtering until a minimum appears, or until a maximum number of filtering operations is reached. Once the minimum appears, the algorithm proceeds as described.

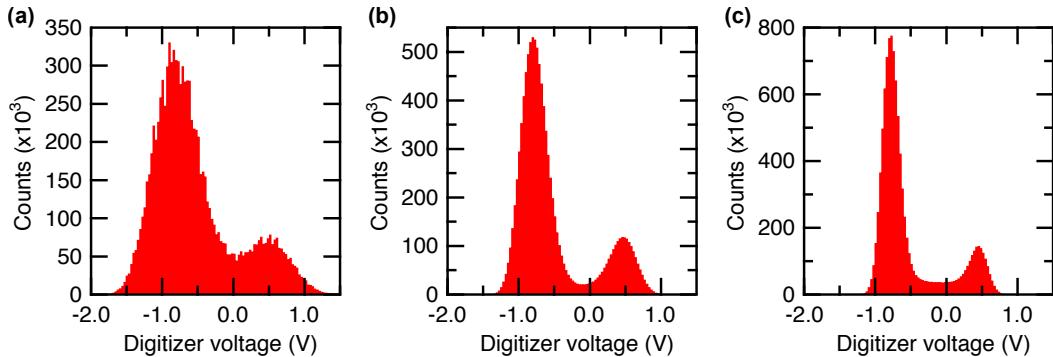


Figure 8.1: Optimal smoothing of time traces.

We show histograms of 10^7 data points from 10^4 traces at a given experimental bias point. Part (a) shows a histogram of the original data, with a measured SNR of 2.1. Part (b) is a histogram of the same data after the individual traces have been optimally smoothed. The measured SNR has now increased to 3.4. Part (c) shows the effects of oversmoothing. There are now a number of data points which do not fall into either of the Gaussian peaks; these points correspond to state transitions which have been smeared out in time by the smoothing process. Because of this smearing, our ability to estimate the transition times (and to see short-duration dwell events in a given state) is degraded.

use any form of filtering, for example a software Chebyshev or Butterworth filter with a variable cutoff frequency. In this case, finding the optimal filtering would amount to finding the filter cutoff frequency and/or order which maximizes SNR. However, performance testing of the automatic extraction algorithm show that it produces quite accurate results using just simple boxcar smoothing (details are given in section 8.1.4).

We note that any filtering process will remove some signal information as well. The boxcar smoothing described above hurts our ability to resolve jumps with very short duration because the filtering has removed the high-frequency components of the signal³. Because such ‘missed’ events are short, their absence has little effect on our determination of ensemble qubit population. We have to be more careful of this phenomenon when estimating transition rates, but the methods described in section 8.1.2 compensate for ‘missed’ short events.

Once we have performed our optimal smoothing, we need to determine a voltage threshold for discriminating between the qubit states. Rather than one threshold, we use two different thresholds with hysteretic ‘Schmitt trigger’ behavior [172]. This method allows us to eliminate spurious event counts caused by noise on the individual traces as they pass through the threshold. To choose our thresholds, we perform Gaussian fits on the histogram to determine the means μ_h and μ_l and standard deviations σ_h and σ_l of the high and low voltage peaks respectively. We also calculate the signal to noise ratio

³The cavity linewidth κ sets a limit on our ability to see very fast qubit dynamics, but added filtering can exacerbate this effect.

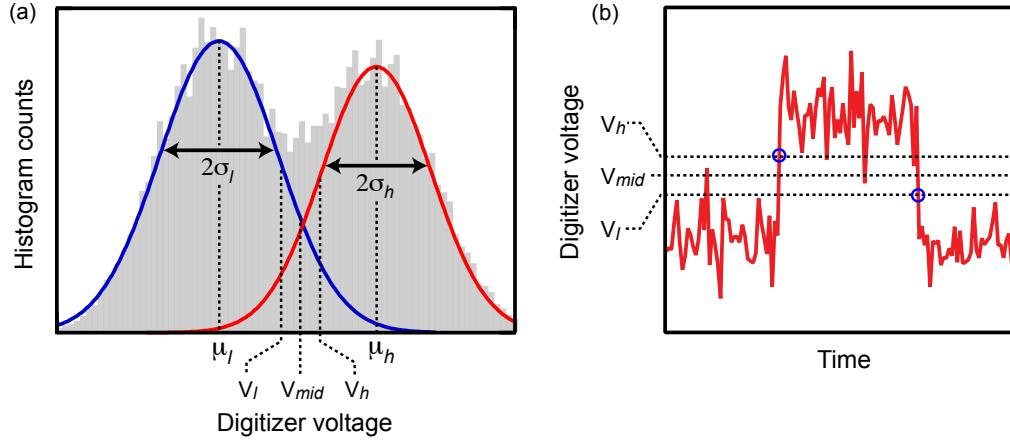


Figure 8.2: Hysteretic data thresholding.

Part (a) shows a cartoon histogram of the voltages in a data set (grey bars) with Gaussian fits to the high and low voltage peaks (red and blue lines). The means μ_h and μ_l and standard deviations σ_h and σ_l of the fits are shown. We choose the hysteretic threshold voltages V_h and V_l as described in the text. Part (b) shows a cartoon of a trace being analyzed using the hysteretic thresholding method. State changes (blue circles) are only registered for upward transitions through V_h from the low-voltage state or downward transitions through V_l from the high-voltage state. This method rejects spurious level-crossings of V_{mid} which do not reach the appropriate hysteretic threshold.

$\text{SNR} = (\mu_h - \mu_l)/(\sigma_h + \sigma_l)$, as in section 7.5.2. The two thresholds V_h and V_l sit on either side of the midpoint voltage $V_{mid} = (\mu_h + \mu_l)/2$ at voltages given by $V_{h,l} = V_{mid} \pm \sigma_{h,l}^2 / (\mu_h - \mu_l)$. These thresholds are shown schematically in Figure 8.2(a). The data are taken in such a way that $\sigma_h \approx \sigma_l$, so we can write these thresholds approximately as $V_{mid} \pm \sigma_{h,l}/(2 \times \text{SNR})$, which gives some intuition as to the choice; for an SNR of 1 we choose thresholds midway between V_{mid} and the $\mu_{h,l}$, with the width of the hysteretic region diminishing as SNR improves. Unlike the method described in [172], we do not perform iterative analysis to refine these thresholds, but simply use the initial starting values. Empirically, this has not degraded the algorithm's performance.

The optimal smoothing and threshold determination steps are performed anew with each set of time traces to be analyzed. This is done because slow phase drifts in the paramp pump can cause the voltage levels to drift up and down with time; they are constant over the ~ 1 second interval to take a given data set, but separate data sets taken several minutes apart can have mean levels sufficiently different to affect the analysis. In addition, changing experimental bias parameters may cause the qubit to jump up and down at very different rates or change the intrinsic SNR, both of which can affect the optimal number of smoothing operations. As a result, redoing the optimal smoothing and threshold determinations for each data set greatly improves the robustness of the state extraction algorithm.

With thresholds established, we can now proceed to determine the state of the qubit as a function of time during each individual time trace. We use the nomenclature s_{hi} and s_{lo} to refer to the voltage state of the signal, with “up” meaning a transition from s_{lo} to s_{hi} and “down” meaning a transition in the opposite direction. The algorithm deals only with s_{hi} and s_{lo} ; the experimenter must note which voltage state (s_{hi} or s_{lo}) corresponds to which qubit state ($|0\rangle$ or $|1\rangle$). This mapping is usually obvious from the data. All the data for a given experiment are taken with the same mapping between the voltage state and qubit state to avoid confusion.

The algorithm records the voltage state as a function of time during each trace, as well as the durations of all dwell periods in s_{lo} and s_{hi} . The first step is to determine the initial state of the qubit in the readout trace. Starting from $t = 0$ at the beginning of the trace, we look for the first level crossing of V_{mid} and find the mean voltage of the data points up until that first crossing. If the mean is above V_{mid} , the initial state is s_{hi} , if below, it’s s_{lo} . We then go back to the start of the trace and look for transitions through the hysteretic thresholds, as shown in Figure 8.2(b). If the initial state is s_{lo} , we look for upward-rising level crossings of V_h , while for s_{hi} we look for downward-falling level crossings of V_l . If a crossing is found, we note the dwell time since the last state change, change the current state in the extracted time record appropriately, and then repeat the process of looking for level crossings. This continues until the end of the measurement trace is reached. In some cases, there are many transitions during a readout trace, while in others there may be none at all. Either way, for every trace one dwell time is cut short by the end of the readout. We record these dwell times but mark them as “censored” [173]. We can use such censored data points in our extraction of qubit transition rates, but we must handle them differently from non-censored data, as will be described in section 8.1.3.

Figure 8.3 shows the performance of the state extraction algorithm for 100 individual time traces. The left panel shows the raw input data, the middle panel shows the same data after optimal smoothing, and the right panel shows the extracted qubit state. The extracted state trace accurately reproduces the features visible in the raw data, although it occasionally misses events with short dwell times. The raw SNR of the data shown is 2.1.

As can be seen (and will be quantified in section 8.1.4), this method for extracting the qubit state works very well. We should mention that there are a number of other methods for extracting state information from noisy telegraph signals such as our jump data. One of our collaborators has successfully used hidden Markov models [174] to perform state extraction with low-SNR tunneling data from quantum dots [175]. This work shows the hidden Markov method to be more accurate in extracting rates than non-hysteretic thresholding and cumulative sum testing [176] (another popular jump-finding algorithm), but the advantage over level-finding disappears for data with reasonable SNR. Another option related to cumulative sum testing is edge detection [177], which finds transitions in a data set and can be used when the mean voltages for s_{hi} and s_{lo} vary significantly within a given data set. These other methods are all more computationally intensive and algorithmically complex than the hysteretic level-finding described above, so we chose not to pursue them in light of the already high accuracy of the simple method.

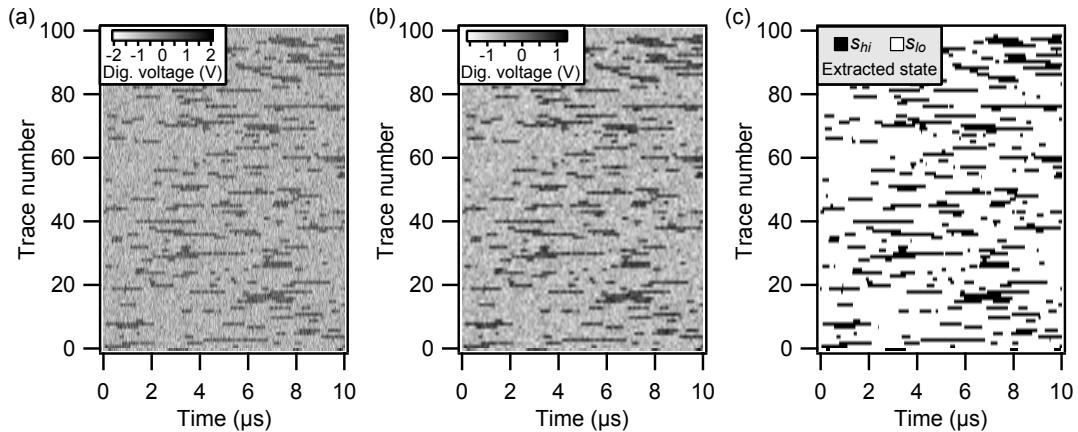


Figure 8.3: Automated qubit state determination.

Panel (a) shows the raw data from 100 individual time traces with an SNR of 2.1. Panel (b) shows the same data after optimal smoothing, which has increased the SNR to 3.4. Panel (c) shows the qubit state as determined by the automated state extraction algorithm. The extracted state traces accurately reproduce the behavior seen in the raw data. These data are from the same data set used to make the histograms in Figure 8.1.

8.1.2 Jump time distribution

The decay events of a qubit are exponentially distributed in time with a time constant T_1 (alternatively, we can parameterize the distribution in terms of a decay rate $\Gamma_1 = 1/T_1$). In an ensemble of exponentially distributed decay times, most decays occur at short times, with fewer and fewer occurring in a given interval as time progresses. However, in our extracted data we see very few decays at short times. This is a result of the limited bandwidth of the experiment B_{det} , set by the readout cavity linewidth κ , the paramp bandwidth, and the filtering and smoothing of the demodulated readout signal. The readout signal shows little or no response to events where the qubit jumps and then jumps back in a time less than $\sim 1/B_{det}$, and we don't detect them with our extraction algorithm. On its face, this does not seem to be a problem for estimating the decay rate Γ_1 ; we might just simply fit the events we do observe to an exponential and get our answer.

The problem here is a subtle one. Imagine we have two dwell times τ_1 and τ_3 in s_{lo} , separated by a dwell τ_2 in s_{hi} so short that the readout signal does not register it. What were actually two dwell times in s_{lo} of length τ_1 and τ_3 are seen in the readout record as a single dwell time of length $\tau_1 + \tau_2 + \tau_3$. This spurious combining of separate events skews the distribution of observed dwell times to longer times. The overall statistics may still look reasonably exponential, but an exponential fit will give us a time constant longer than the true one. The fractional error in the estimated time constant can be as much as 20-30% for our experimental parameters.

A solution to this problem for a Poissonian two level system (such as a qubit) was developed by Naaman and Aumentado [178]. They create a model of the qubit and readout

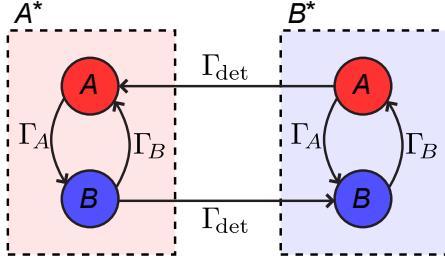


Figure 8.4: State diagram for finite bandwidth detection.

This figure shows a system state diagram for finite bandwidth detection. The qubit is in either state A or B (circles), while the readout is in state A^* or B^* . For each qubit/readout state, the likelihood of transitioning to a different qubit/readout state is determined by the rates Γ_A , Γ_B , and Γ_{det} as indicated. Adapted from ref. [178].

where the states of the the qubit (denoted A and B) are allowed to be independent of the states of the readout (denoted A^* and B^*). A diagram of this model, adapted from their paper, is shown in Figure 8.4. For each readout state, the qubit can be in either state A or B . Transitions occur between A and B with rates Γ_A and Γ_B . However, when the qubit and readout “disagree”, i.e. states (B, A^*) and (A, B^*) , the readout can also make a transition (to (B, B^*) or (A, A^*) , respectively) with rate Γ_{det} . We assume that the readout does not change states unless the qubit has changed states (no false positives).

To take an example, imagine the readout record has just made a transition to A^* , so the system will be in (A, A^*) . After some time, a qubit transition will occur and we go from (A, A^*) to (B, A^*) . As we wait for the readout to register by transitioning to (B, B^*) , there is some probability that the qubit will change states again and return to (A, A^*) . If so, this is a “missed” event. It is possible for several such cycles to occur before a transition to (B, B^*) occurs. At this point, the cycle repeats itself. The important point to note is that the readout record gives us access to transition rates between A^* and B^* , but NOT to rates between A and B . However, we can write down analytic expressions relating the two to allow us to determine Γ_A and Γ_B , the true qubit transition rates.

If we are in readout state A^* , we can write down rate equations for the probabilities of being in qubit states A and B respectively:

$$\dot{P}_A = -\Gamma_A P_A + \Gamma_B P_B, \quad (8.1)$$

$$\dot{P}_B = -(\Gamma_B + \Gamma_{\text{det}}) P_B + \Gamma_A P_A. \quad (8.2)$$

The probability density for transitioning to readout state B^* at time t after arriving in A^* is given by $P_B(t)\Gamma_{\text{det}}$. Solving equations (8.1) and (8.2) for P_B and using the boundary conditions⁴ $P_A(0) = 1$ and $P_B(0) = 0$, we come up with an analytical formula for the

⁴These boundary conditions are simply the statement that the readout never makes a transition to state A^* when the qubit is in state B .

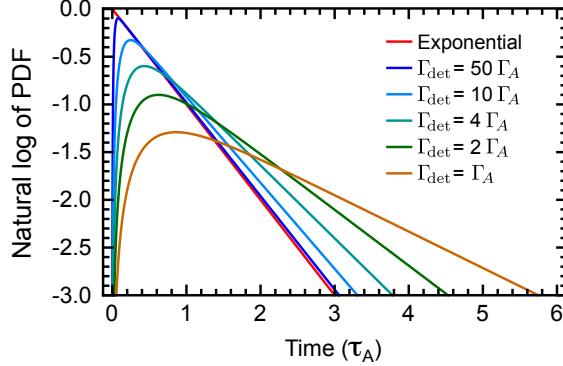


Figure 8.5: Expected dwell time distributions.

We plot an exponential probability distribution function (PDF) with rate Γ_A as well as $h(t)$ for various values of Γ_{det} with $\Gamma_A = \Gamma_B$. On these semi-log axes, exponential decay appears as a straight line with slope proportional to decay rate. The time axis is in units of $\tau_A = 1/\Gamma_A$. For $\Gamma_{\text{det}} \gg \Gamma_A$, the behavior of $h(t)$ closely approximates the exponential. However, for smaller Γ_{det} the long-time exponential tail of $h(t)$ has a shallower slope than the exponential. The slope is 98%, 90%, 76%, 59%, and 38% of the true Γ_A slope, respectively, for the 5 curves with decreasing Γ_{det} . The probability of observing short dwell time events decreases with Γ_{det} as expected.

distribution of dwell times in state A^* :

$$h(t) = \frac{2}{\theta} \Gamma_A \Gamma_{\text{det}} e^{-\lambda t/2} \sinh\left(\frac{\theta t}{2}\right), \quad (8.3)$$

where $\lambda = \Gamma_A + \Gamma_B + \Gamma_{\text{det}}$ and $\theta = \sqrt{\lambda^2 - 4\Gamma_A\Gamma_{\text{det}}}$. Examining the t dependence of this function, we note that the sinh term tends to reduce the probability of seeing short-time events, as we expect. The overall behavior can be slightly difficult to glean by inspection, so several different curves of $h(t)$ for varying values of Γ_{det} are plotted on semi-log axes in Figure 8.5. We set $\Gamma_A = \Gamma_B$ and vary Γ_{det} ; a curve plotting an exponential decay with rate Γ_A is shown for comparison. While the sinh term reduces the probability of seeing short-duration events, at long times $h(t)$ looks exponential again. However, the rate of this long-time exponential behavior is substantially different from Γ_A , with the effect growing worse as Γ_{det} decreases. This is the effect we noted at the start of this section; now that we have the form of $h(t)$, though, we can use it to extract the true Γ_A .

The expression for $h(t)$ is properly normalized to be a probability distribution function (PDF), with $\int_0^\infty h(t)dt = 1$. We can also define a survival function $s(t)$, which is the probability of an event occurring at any point after time t , expressed mathematically as $s(t) = \int_t^\infty h(t')dt'$. The functional form of $s(t)$ is:

$$s(t) = \frac{e^{-\lambda t/2}}{\theta} \left[\lambda \sinh\left(\frac{\theta t}{2}\right) + \theta \cosh\left(\frac{\theta t}{2}\right) \right] \quad (8.4)$$

The reason for our interest in $s(t)$ will become clear in the next section.

8.1.3 Maximum likelihood estimation

One of the classic problems of statistics is to determine the shape of a probability distribution given a random set of observations drawn from that distribution. In the case where the distribution is known to have a specific parameterized functional form, the problem becomes the estimation of the parameters which characterize the distribution. For example, an exponential distribution is characterized by a single parameter, the time constant⁵ τ , while a normal or Gaussian distribution is characterized by two parameters, the mean μ and standard deviation σ . In our case, we are interested in estimating the qubit transition rates Γ_A and Γ_B (for which we must also estimate Γ_{det}) given a set of observed qubit state dwell times.

Of the many statistical methods used for parameter estimation, the technique known as maximum likelihood estimation is generally the most robust method for generating unbiased parameter estimates if we know the functional form of the underlying distribution. Maximum likelihood estimation is relatively simple to understand. Let X be a random variable from a distribution $p(\theta, X)$ characterized by a parameter θ . If we are given a set of observations $\{x_i\}$ drawn from this distribution with $i = \{1, 2, \dots, n\}$, we can write down the probability of receiving this particular set of observations, which is just the product of each of the individual observation probabilities. This total probability for a particular set of observations is called the likelihood function and denoted with \mathcal{L} :

$$\mathcal{L} = \prod_{i=1}^n p(\theta, x_i). \quad (8.5)$$

The value of the likelihood function depends on the parameter θ . The maximum likelihood principle says simply that the best estimate for θ , called the maximum likelihood estimate or MLE and denoted $\hat{\theta}$, is the one for which the likelihood function is largest.

Maximum likelihood estimation has a number of nice properties, which are enumerated and rigorously proven in standard statistics books [179, 180, 181], and which we touch on briefly here. First, maximum likelihood estimators are called “invariant”, meaning that they are independent of the particular parameterization of the distribution. In other words, if $\hat{\theta}$ is the MLE of the parameter θ for a given distribution, then the MLE of a parameter $\theta^* = g(\theta)$ (where g is an arbitrary function) is just $\hat{\theta}^* = g(\hat{\theta})$ ⁶. Second, maximum likelihood estimates are unbiased, meaning that in the limit of a large number of data points the MLE $\hat{\theta}$ tends to the true value of the parameter θ . Third, the maximum likelihood estimate is asymptotically normal, meaning that for large n , the value of $\hat{\theta}$ will be normally distributed about the true parameter value θ with a variance σ_θ^2/n which can

⁵One can equivalently parameterize an exponential distribution with a rate constant $\gamma = 1/\tau$; in either case, though, only one parameter is required.

⁶For example, the exponential distribution can be parameterized in terms of a time constant τ or a rate constant $\gamma = 1/\tau$; the invariance of the maximum likelihood estimator means that it does not matter which parameterization we use.

be explicitly calculated⁷. The property of asymptotic normality allows us to put confidence intervals on maximum likelihood parameter estimates given sufficiently large data sets. Finally, the maximum likelihood estimator is what is known as a uniformly minimum variance unbiased estimator. This means that of all the possible unbiased estimation methods for a parameter θ of a distribution, maximum likelihood estimation always has the smallest uncertainty in the final estimate $\hat{\theta}$.

The actual process of maximum likelihood estimation just involves solving for the value of $\hat{\theta}$ which maximizes equation (8.5). A maximum will occur at a point where

$$\frac{\partial \mathcal{L}}{\partial \theta} = 0 \quad (8.6)$$

Note that equation (8.6) represents a necessary, but not sufficient, condition for maximization. Because the derivative will be zero at both minima and maxima, and because there may be several maxima, in general we may need to check the values of \mathcal{L} at each extreme point to find which one is the global maximum. In addition, maxima may lie on the boundaries of the parameter space in places where $\frac{\partial \mathcal{L}}{\partial \theta} \neq 0$, so it is important to check boundary cases as well.

To solve for $\hat{\theta}$, we can simply solve equation (8.6). As a practical matter, though, the problem of maximizing \mathcal{L} is generally solved by instead maximizing the log-likelihood function $L = \ln \mathcal{L}$:

$$L = \sum_{i=1}^n \ln[p(\theta, x_i)] \quad (8.7)$$

Since $\ln(x)$ is a monotonic function of x , the value of θ which maximizes \mathcal{L} will also maximize L , so we can use either formulation of the problem. However, the math of maximizing L is typically much simpler, since it is a sum of terms instead of a product. There is another practical reason for using L instead of \mathcal{L} , related to the constraints of numerical calculations. For large n , \mathcal{L} rapidly becomes incredibly small. If the probability for each data point were 0.5 (a high estimate in most cases!) and $n = 10^4$, we find $\mathcal{L} = 2^{-10,000}$, a number so small that it cannot be represented as a standard double precision floating-point number. Since L sums numbers rather than multiplying them, the log-likelihood function is much better behaved with large numbers of data points; in the above example, $L = 10^4 \ln(0.5) \approx -6930$, a much more tractable number. In general, if there are enough data points to yield reasonable parameter estimates, one typically has to use L .

Many probability distributions have multiple parameters, which we can express by writing θ as a vector $\vec{\theta}$ with K components, one for each parameter. In this case, maximizing L requires finding extreme points in multidimensional space, with derivatives with respect to each parameter θ_k satisfying

$$\frac{\partial L}{\partial \theta_k} = 0, \quad k = \{1, 2, \dots, K\}. \quad (8.8)$$

⁷The value of σ_θ^2 is the inverse of the Fisher information $\mathcal{I}(\theta)$, given by $\mathcal{I}(\theta) = E_\theta \left[\frac{\partial}{\partial \theta} \ln p(\theta, X) \right]^2 = \int_{-\infty}^{\infty} \left[\frac{\partial}{\partial \theta} \ln p(\theta, x) \right]^2 p(\theta, x) dx$. The notation E_θ denotes the conditional expectation value over X with respect to $p(\theta, X)$ for a given θ [181].

We will use the notation of $\vec{\theta}$ to maintain full generality, and because it applies to our case of interest.

It is instructive to demonstrate how maximum likelihood estimation works on a specific probability distribution. Let's take the example of an exponential distribution with random variable T and decay time parameter τ , given by $p(\tau, T) = \frac{1}{\tau} e^{-T/\tau}$. Given a set of observations $\{t_i\}$, we construct the likelihood function:

$$\mathcal{L} = \prod_{i=1}^n \frac{1}{\tau} e^{-t_i/\tau}. \quad (8.9)$$

Taking the log of \mathcal{L} gives us the log-likelihood function, which takes the nice form:

$$L = \sum_{i=1}^n \left(-\ln \tau - \frac{t_i}{\tau} \right) = -n \ln \tau - \frac{1}{\tau} \sum_{i=1}^n t_i. \quad (8.10)$$

Next we take the derivative of L with respect to τ , which gives:

$$\frac{\partial L}{\partial \tau} = \frac{-n}{\tau} + \frac{1}{\tau^2} \sum_{i=1}^n t_i. \quad (8.11)$$

The extreme points of L will occur when $\frac{\partial L}{\partial \tau} = 0$. Setting the above expression equal to 0 and factoring out $1/\tau^2$ leaves us with:

$$\frac{1}{\tau^2} \left(-n\tau + \sum_{i=1}^n t_i \right) = 0. \quad (8.12)$$

One solution to this equation is $\tau = \infty$, which is a non-physical solution. The other solution is the maximum, and gives us the MLE $\hat{\tau}$ for τ :

$$\hat{\tau} = \frac{1}{n} \sum_{i=1}^n t_i. \quad (8.13)$$

The result we get is that $\hat{\tau}$ is the sum of all observed times divided by the number of observations, in other words the mean observed decay time. This is a simple, clean result which makes intuitive sense⁸.

If we have only a certain window in which we can observe events, for example the finite duration of a qubit readout pulse, we may have some data which is censored. Censoring broadly refers to observations where we do not know the exact observed value, merely a range in which it lies. Censored data is common in the statistics subfields of survival analysis and reliability analysis, and methods for using censored data are very well developed [173].

Our quantum jump data exhibit right censoring, which is when we know that the value of the censored data point is greater than or equal to a known value. Considering the

⁸For a Gaussian distribution, the estimator $\hat{\mu}$ of the mean is simply the mean of the observed data points, and the estimator $\hat{\sigma}^2$ of the variance is the variance of the observed data points. These simple cases are nice checks, but don't fully showcase the power of maximum likelihood estimation for non-intuitive cases.

case of a readout pulse, the qubit is in a given state when the readout turns off, but does not necessarily decay right then. All we can know about the dwell time of the qubit in that state is that it was at least as long as what we observed during readout. However, despite the fact that we have incomplete information about this data point, we can still make use of it.

We can modify equation (8.5) to account for right censoring by introducing the variables δ_i and the function $P(\vec{\theta}, X > x)$. The δ_i are variables which take either the value 0 (if the i^{th} observation is uncensored) or 1 (if the i^{th} observation is censored). The function $P(\vec{\theta}, X > x)$ is the survival function, given by integrating the PDF $p(\vec{\theta}, X)$ over all values of X greater than a specific one x :

$$P(\vec{\theta}, X > x) = \int_x^\infty p(\vec{\theta}, x') dx'. \quad (8.14)$$

Armed with these definitions, we construct the likelihood function as before by multiplying together the probabilities of each individual observation. For uncensored observations, we use the same probability as before, $p(\vec{\theta}, x_i)$. For censored observations, however, the probability is now the survival function $P(\vec{\theta}, X > x_i)$. We need to construct a likelihood function that is able to represent either one of these cases based on whether a given data point is censored or not. A suitable function takes the form:

$$\mathcal{L} = \prod_{i=1}^n p(\vec{\theta}, x_i)^{1-\delta_i} P(\vec{\theta}, X > x_i)^{\delta_i}. \quad (8.15)$$

If $\delta_i = 0$ for a given i (an uncensored data point), the second factor is just 1 and so we just recover the original probability $p(\vec{\theta}, x_i)$. If $\delta_i = 1$, the first factor becomes one and the probability for that point is then given by $P(\vec{\theta}, X > x_i)$, as desired. The corresponding log-likelihood function is then:

$$L = \ln \mathcal{L} = \sum_{i=1}^n \left[(1 - \delta_i) \ln[p(\vec{\theta}, x_i)] + \delta_i \ln[P(\vec{\theta}, X > x_i)] \right]. \quad (8.16)$$

This log-likelihood function allows us to use all data points, both censored and uncensored, to estimate the parameters. Let's return to the example of the exponential distribution, but this time allow data points to be censored. First, we must calculate the survival function for the exponential distribution, which is given by $P(\tau, T > t) = \int_t^\infty \frac{1}{\tau} e^{-t'/\tau} dt' = e^{-t/\tau}$. We can then write down the likelihood function:

$$\mathcal{L} = \prod_{i=1}^n \left(\frac{1}{\tau} e^{-t_i/\tau} \right)^{1-\delta_i} \left(e^{-t_i/\tau} \right)^{\delta_i}. \quad (8.17)$$

Taking the logarithm gives us the log-likelihood function:

$$L = \sum_{i=1}^n \left[(1 - \delta_i) \left(-\ln \tau - \frac{t_i}{\tau} \right) + \delta_i \left(\frac{-t_i}{\tau} \right) \right] = -\ln \tau (n - \sum_{i=1}^n \delta_i) - \frac{1}{\tau} \sum_{i=1}^n t_i. \quad (8.18)$$

To find the maximum, we take the derivative of L with respect to τ :

$$\frac{\partial L}{\partial \tau} = -\frac{1}{\tau}(n - \sum_{i=1}^n \delta_i) + \frac{1}{\tau^2} \sum_{i=1}^n t_i. \quad (8.19)$$

A maximum will occur when $\frac{\partial L}{\partial \tau} = 0$. We set the above derivate equal to zero and factor out $1/\tau^2$ as before, giving us:

$$\frac{1}{\tau^2} \left(-\tau(n - \sum_{i=1}^n \delta_i) + \sum_{i=1}^n t_i \right) = 0. \quad (8.20)$$

Again we ignore the pathological $\tau = \infty$ solution. This equation contains the term $\sum \delta_i$, which we recognize to be simply the total number of censored data points n_{cens} . Using the definition

$$\sum_{i=1}^n \delta_i = n_{\text{cens}}, \quad (8.21)$$

we substitute in to equation (8.20) to get our MLE $\hat{\tau}$, which is given by

$$\hat{\tau} = \frac{1}{n - n_{\text{cens}}} \sum_{i=1}^n t_i. \quad (8.22)$$

Our result is almost as simple as the result for $\hat{\tau}$ in the case without censoring. The MLE for τ is the sum of all of the individual observed times divided by the total number of uncensored observations. In this way, the estimate uses both censored and uncensored data (because all the t_i are included), giving us more information than if we just used uncensored observations. While simple, equation (8.22) is not intuitive, and begins to demonstrate the power of maximum likelihood estimation.

We are interested in determining the rate of transitions between the two qubit states, which we can do using the model described in section 8.1.2. There we derived a probability distribution function $h(t)$ and survival function $s(t)$, defined in equations (8.3) and (8.4), for the observed jump times, given some underlying rate parameters. To keep our notation consistent with that used for the qubit state extraction, we make the identification between states $A \leftrightarrow s_{lo}$ and $B \leftrightarrow s_{hi}$, with $\Gamma_A \leftrightarrow \Gamma_{\text{up}}$ and $\Gamma_B \leftrightarrow \Gamma_{\text{dn}}$.

Our task is now to find the true values of Γ_{up} and Γ_{dn} , for which we will use maximum likelihood estimation. We are given two sets of observations, the dwell times $\{t_i\}$ in s_{lo} and the dwell times $\{t_j\}$ in s_{hi} , as well as lists $\{\delta_i\}$ and $\{\delta_j\}$ of censoring variables for each. The parameters Γ_{up} , Γ_{dn} , and Γ_{det} are the same for both sets of observations, so we should construct our likelihood function to enable us to use both data sets simultaneously to optimize all three parameter values. If there are n_{lo} observed dwell times in s_{lo} and n_{hi} in s_{hi} , we can write:

$$\begin{aligned} \mathcal{L} &= \prod_{i=1}^{n_{lo}} h(\Gamma_{\text{up}}, \Gamma_{\text{dn}}, \Gamma_{\text{det}}; t_i)^{1-\delta_i} s(\Gamma_{\text{up}}, \Gamma_{\text{dn}}, \Gamma_{\text{det}}; t_i)^{\delta_i} \\ &\quad \times \prod_{j=1}^{n_{hi}} h(\Gamma_{\text{dn}}, \Gamma_{\text{up}}, \Gamma_{\text{det}}; t_j)^{1-\delta_j} s(\Gamma_{\text{dn}}, \Gamma_{\text{up}}, \Gamma_{\text{det}}; t_j)^{\delta_j} \end{aligned} \quad (8.23)$$

where we explicitly write out the dependence of h and s on the three Γ parameters. This likelihood function is constructed by multiplying together the individual likelihood functions for each set of observations to give a joint likelihood for maximization. In this way, we use all of the data in a single estimation. Importantly, note that the order of Γ_{up} and Γ_{dn} is interchanged for the t_j relative to the t_i . This is because the rate of leaving s_{lo} is the same as the rate of leaving s_{hi} with the Γ_{up} and Γ_{dn} interchanged, as is clear from the symmetry of the model.

Given this rather large likelihood function, we make the log-likelihood function L , which takes the form:

$$\begin{aligned} L = & \sum_{i=1}^{n_{lo}} \left[(1 - \delta_i) \ln[h(\Gamma_{\text{up}}, \Gamma_{\text{dn}}, \Gamma_{\text{det}}; t_i)] + \delta_i \ln[s(\Gamma_{\text{up}}, \Gamma_{\text{dn}}, \Gamma_{\text{det}}; t_i)] \right] \\ & + \sum_{j=1}^{n_{hi}} \left[(1 - \delta_j) \ln[h(\Gamma_{\text{dn}}, \Gamma_{\text{up}}, \Gamma_{\text{det}}; t_j)] + \delta_j \ln[s(\Gamma_{\text{dn}}, \Gamma_{\text{up}}, \Gamma_{\text{det}}; t_j)] \right] \end{aligned} \quad (8.24)$$

Already this is beginning to look considerably more complicated than the case of the simple exponential distribution! We now plug in the functional forms for $h(t)$ and $s(t)$ which are shown in equations (8.3) and (8.4). We must be careful, because they are parameterized in terms of the variables θ and λ , which are combinations of the three gammas. While $\lambda = \Gamma_{\text{up}} + \Gamma_{\text{dn}} + \Gamma_{\text{det}}$ is invariant under a swap of Γ_{up} and Γ_{dn} , θ is not. We therefore must define two different values, $\theta_{\text{up}} = \sqrt{\lambda^2 - 4\Gamma_{\text{up}}\Gamma_{\text{det}}}$ and $\theta_{\text{dn}} = \sqrt{\lambda^2 - 4\Gamma_{\text{dn}}\Gamma_{\text{det}}}$ to be used in the appropriate cases. Note that the use of λ and θ is simply a change of variables used to make the algebra easier; they do not represent additional parameters to be maximized, because each is fully specified by the three gammas⁹.

We can write out the first line of equation (8.24) using the full definitions of $h(t)$ and $s(t)$. This substitution gives us:

$$\begin{aligned} & \sum_{i=1}^{n_{lo}} (1 - \delta_i) \left[\ln \left(\frac{2\Gamma_{\text{up}}\Gamma_{\text{dn}}}{\theta_{\text{up}}} \right) - \frac{\lambda t_i}{2} + \ln \left(\sinh \left(\frac{\theta_{\text{up}} t_i}{2} \right) \right) \right] \\ & + \sum_{i=1}^{n_{lo}} \delta_i \left[-\ln(\theta_{\text{up}}) - \frac{\lambda t_i}{2} + \ln \left[\lambda \sinh \left(\frac{\theta_{\text{up}} t_i}{2} \right) + \theta_{\text{up}} \cosh \left(\frac{\theta_{\text{up}} t_i}{2} \right) \right] \right] \end{aligned} \quad (8.25)$$

The second line of equation (8.24) is the same, except we interchange every θ_{up} for θ_{dn} , n_{lo} for n_{hi} , and i for j . This gives us:

$$\begin{aligned} & \sum_{j=1}^{n_{hi}} (1 - \delta_j) \left[\ln \left(\frac{2\Gamma_{\text{dn}}\Gamma_{\text{up}}}{\theta_{\text{dn}}} \right) - \frac{\lambda t_j}{2} + \ln \left(\sinh \left(\frac{\theta_{\text{dn}} t_j}{2} \right) \right) \right] \\ & + \sum_{j=1}^{n_{hi}} \delta_j \left[-\ln(\theta_{\text{dn}}) - \frac{\lambda t_j}{2} + \ln \left[\lambda \sinh \left(\frac{\theta_{\text{dn}} t_j}{2} \right) + \theta_{\text{dn}} \cosh \left(\frac{\theta_{\text{dn}} t_j}{2} \right) \right] \right] \end{aligned} \quad (8.26)$$

⁹Recall as well that the invariance property of maximum likelihood estimation means that we are free to choose among these different parameterizations when describing our probability density without affecting the outcome of the maximum likelihood estimation.

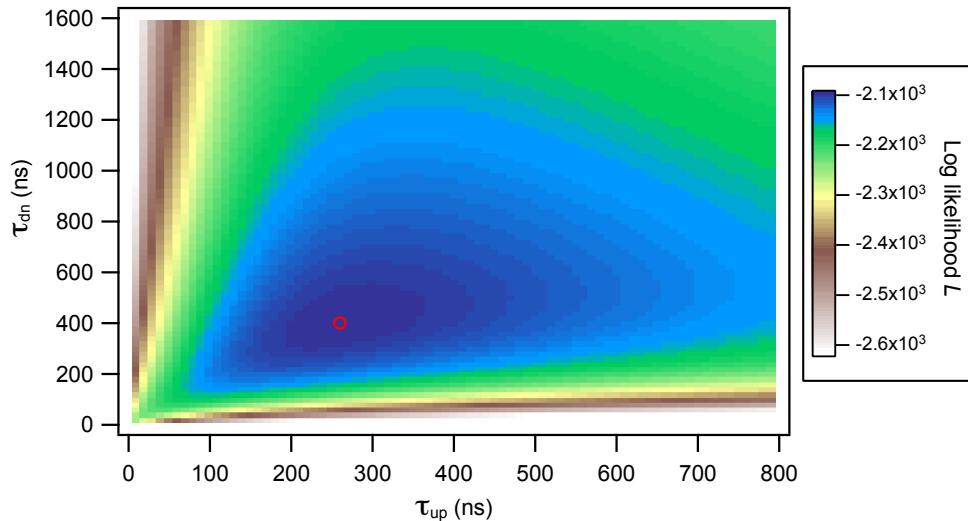


Figure 8.6: Log-likelihood in parameter space.

This is a color plot of the log-likelihood function L in equation (8.24) as a function of $\tau_{\text{up}} = 1/\Gamma_{\text{up}}$ and $\tau_{\text{dn}} = 1/\Gamma_{\text{dn}}$, with $\Gamma_{\text{det}} = 0.01 \text{ ns}^{-1}$, using dwell times extracted from 10^3 measurement traces. The likelihood is maximized for this value of Γ_{det} at the point marked by the red circle. The function L actually exists in three dimensions, and so we must search in Γ_{det} as well to determine the true maximum likelihood. The function is smooth and well-behaved except at the boundaries, where it diverges.

Looking at the expressions in (8.25) and (8.26), it is readily apparent that we won't be able to solve analytically to determine the maximum likelihood parameter estimates $\hat{\Gamma}_{\text{up}}$, $\hat{\Gamma}_{\text{dn}}$, or $\hat{\Gamma}_{\text{det}}$. We instead resort to numerical optimization. By providing a reasonable initial guess for our parameters, we can use numerical methods to "hill climb" to a maximum of L in parameter space and determine $\hat{\Gamma}_{\text{up}}$, $\hat{\Gamma}_{\text{dn}}$, and $\hat{\Gamma}_{\text{det}}$ in this way. Figure 8.6 shows the numerical value of L for a real data set with 10^3 traces. The function is smooth and well-behaved, except near the boundaries $\Gamma = 0$ where it diverges. The maximum likelihood point is marked in red.

There are a number of methods in the literature for performing multidimensional numerical optimizations of this kind [182, 183]. However, some of the best methods are notoriously finicky about the shape of the function to be optimized and the quality of the initial guess. We chose one which seemed to provide sufficient robustness, a quasi-Newton method with a double-dogleg trust-region search [182] as implemented by Igor [171]. The whole trick of numerical optimization schemes, aside from making a good initial guess, is to select a good direction to search in parameter space and an appropriate step length along that direction. The simplest method along these lines, called gradient descent¹⁰, involves

¹⁰All optimization methods can search for either a minimum or a maximum; one need simply negate the function being optimized to turn its maxima into minima and vice versa. Much of the optimization literature

following the gradient some distance to choose the next point, the following the gradient from there, and so on. Gradient descent is almost never used, because in the case of flat, shallow minima it takes a very long time to converge and is thus computationally inefficient. Methods like gradient descent, which follow the derivative exactly, are referred to as Newton methods.

In the 1960s and 1970s, a class of methods called quasi-Newton methods arose. Quasi-Newton methods follow a modified prescription for finding the search direction, combining gradient information with information from the Hessian matrix, or an approximation to the Hessian matrix, which contains information on the second derivatives of the function. As a result, quasi-Newton methods have much more rapid convergence than simple gradient descent. One of the most popular quasi-Newton methods is the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method, which provides a technique for continuously updating an approximation to the Hessian matrix at each search point, making it computationally efficient. However, the BFGS method requires that the Hessian matrix always be positive-definite, which is equivalent to the statement that the concavity of the optimized function must have the same sign everywhere. If this is not the case, as appears to be the case with our optimization problem, the BFGS method fails. The double-dogleg trust-region method uses a hybrid approach, combining some of the robustness of gradient descent with the efficiency of quasi-Newton Hessian-based step direction selection, and employing trust-region techniques to restrict step size in a more robust way [182]. There are many other possible optimization routines available, and the interested reader is encouraged to consult references [182] and [183] to begin his or her explorations.

To perform the optimization, we need to give some initial guesses about the values of the parameters. We can provide rough estimates for the values of Γ_{up} and Γ_{dn} by assuming an exponential distribution and using equation (8.22) to determine $\tau_{\text{up}} = 1/\Gamma_{\text{up}}$ and $\tau_{\text{dn}} = 1/\Gamma_{\text{dn}}$. A reasonable initial guess for Γ_{det} might be κ , the cavity linewidth, which sets the rate at which information leaves the readout resonator.

There are a few other technical considerations for the numerical optimization related to floating point arithmetic. The log-likelihood function L has various terms of the form $\ln(\sinh(x))$ or $\ln(a \sinh(x) + b \cosh(x))$, as seen in equations (8.25) and (8.26). For sufficiently large values of x , $\sinh(x)$ and $\cosh(x)$ will return a NaN (“Not a Number”) in floating point arithmetic, so $\ln(\sinh(x))$ and $\ln(\sinh(x))$ will evaluate as NaNs as well. These NaNs turn the entire value of L to a NaN and thus are fatal to the optimization procedure. The solution is to recognize that for $x \gg 1$, $\ln(\sinh(x)) \approx x - \ln 2$ and $\ln(a \sinh(x) + b \cosh(x)) \approx x + \ln(a + b) - \ln 2$. When performing the sum to compute L , we simply check the value of x and use these alternative forms if $x \geq 20$ (a cutoff which gives fractional errors below 10^{-18} for any given term).

8.1.4 Tests of the extraction algorithm

To test the performance of the automated extraction algorithm, we made a series of simulated data traces with known parameters. By running these through the automated

is written for minimization problems, where one “descends” into a minimum, thus the nomenclature. We will use “descent”, even though it should strictly be “ascent” in our case, because they are made equivalent by just choosing to minimize $-L$ rather than maximizing L .

extraction algorithm, we could see how well the extracted qubit state tracked the nominal qubit state, and gauge the quality of the estimates both for ensemble qubit population across many time traces and for qubit transition rates.

In order for this test to be accurate, we need the simulated data traces to be as similar to real data as possible. This means modeling the limited measurement bandwidth, adding noise, and using a wide variety of realistic qubit transition rates. To generate simulated data, we specify qubit excitation and relaxation rates Γ_{\uparrow} and Γ_{\downarrow} , the cavity bandwidth κ , the noise bandwidth B , the initial excited state fraction, the signal-to-noise ratio, the duration of each trace, and the number of traces. We match the sampling rate in time (100 MS/s) to that of our real data.

First, we make a series of noiseless time traces, representing the state of the qubit. These are formed by stringing together random dwell times drawn from exponential distributions with decay rates Γ_{\uparrow} and Γ_{\downarrow} until we reach the time limit for the trace. The initial state is chosen randomly, with a specified probability for being in the ground state initially. After making these noiseless time traces, we convert the sharp transitions between states to rising or falling exponentials with a time constant $2/\kappa$, which models the behavior of the cavity output signal when the qubit state changes.

We generate Gaussian white noise at the data sampling frequency and then filter it to be in the bandwidth B , which is set to be equivalent to the bandwidth of the filters on the output of the demodulation circuit (see section 5.3.3). The filter rolloff is set to match that of the physical filters used in the experiment. We then histogram the filtered noise to determine its standard deviation.

With both signal and noise in hand, we add them together, scaling their relative values to give the desired signal-to-noise ratio (as defined in section 7.5.2). The absolute scaling of the combined trace is then set to match typical values of real data. The real data, in raw form, are 8-bit unsigned integers (integers in the range 0-255), with a difference in the mean high and low levels of around 70-100, so this is the scaling we apply to our simulated data. Finally, we turn the scaled simulated data into 8-bit unsigned integers, rounding off all data points to the nearest integer.

For each simulated data trace, we have access to the exact set of dwell times in the high and low excited states. Using these, we perform maximum likelihood estimation using equation (8.22) to give best-case estimates $\hat{\Gamma}_{\uparrow}$ and $\hat{\Gamma}_{\downarrow}$ if the qubit state is perfectly known. We also calculate the qubit ensemble population as a function of time, as well as the mean and median steady-state qubit population. These are our benchmarks for comparing the performance of the automated qubit extraction algorithm.

Figure 8.7 shows the results of state extraction algorithm tests for simulated data with either a nominal SNR of 3 (equivalent to the SNR of real data with $\bar{n} \approx 30$) or a nominal SNR of 1.6 (equivalent to the SNR of real data with $\bar{n} \approx 3$). These SNR values are quoted for the raw data, before any smoothing. The simulated data are generated for a variety of possible excitation and relaxation rates Γ_{\uparrow} and Γ_{\downarrow} corresponding to typical observed experimental values. We test our state extraction algorithm's estimates of the steady state qubit population (top two panels), the qubit relaxation time $1/\Gamma_{\downarrow}$ (middle two panels), and the qubit excitation time $1/\Gamma_{\uparrow}$ (bottom two panels). Each plotted point represents the mean percentage error of the extracted estimate (as compared to the true

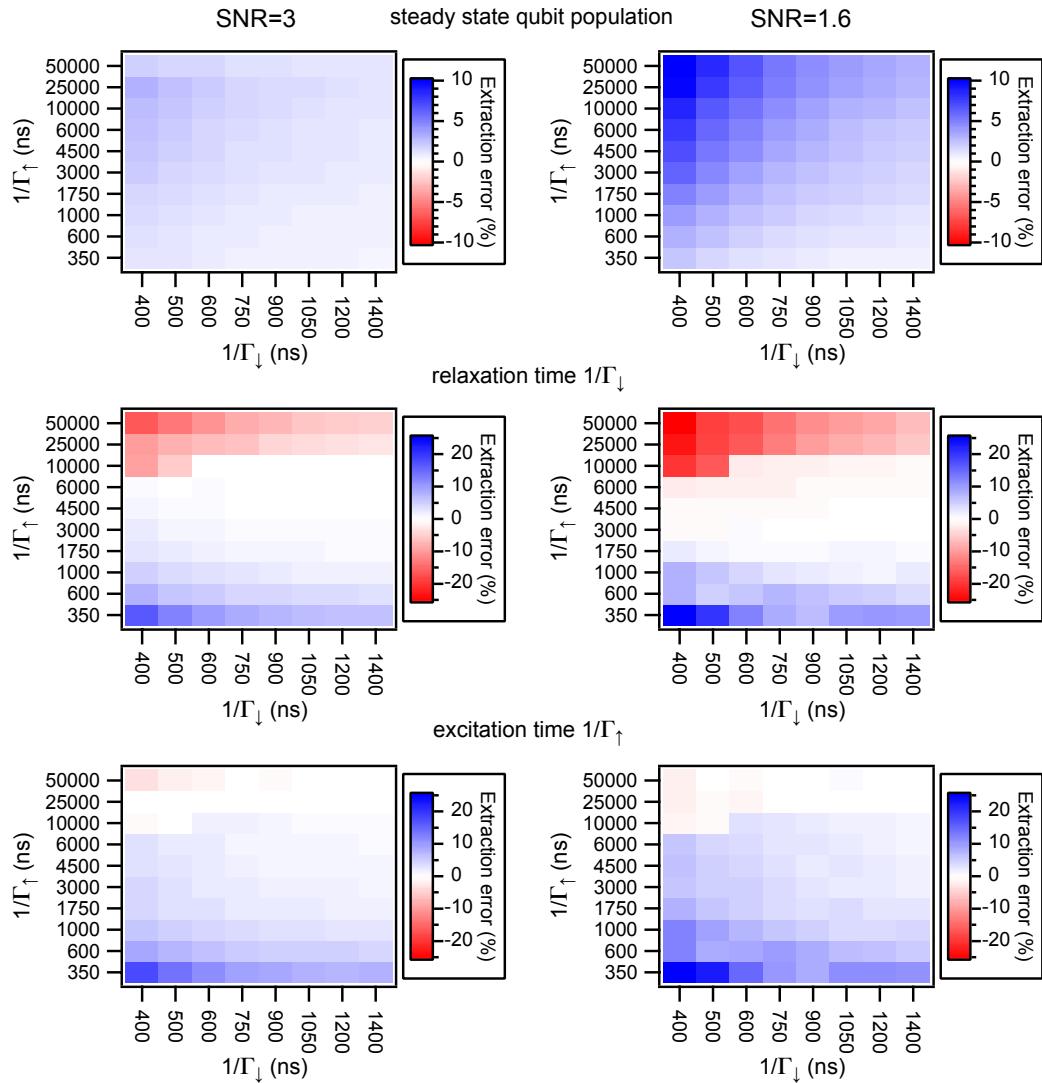


Figure 8.7: Performance of automated state extraction algorithm.

These six panels show the percentage error in the estimates provided by the automated state extraction algorithm for the steady state qubit population (top two panels), the qubit relaxation time (middle two panels), and the qubit excitation time (bottom two panels). The left column is from simulated data with nominal SNR=3, while the right column is for simulated data with nominal SNR=1.6; these SNRs are quoted before any smoothing by the state extraction algorithm.

value) for 10 different sets of simulated data generated with nominally identical values of Γ_{\uparrow} , Γ_{\downarrow} , and SNR. Each simulated data set consists of 10^4 traces of 1750 points each, the same size as the experimental data sets we use later in this chapter.

We find that the algorithm estimates the steady state qubit population quite accurately, with a typical fractional error less than 5% even for the low SNR data. The population estimates are least accurate for simulated data with $\Gamma_{\uparrow} \ll \Gamma_{\downarrow}$. However, in this regime the actual steady state excited population of the qubit is quite small ($\approx \Gamma_{\uparrow}/\Gamma_{\downarrow}$), so the absolute error in the qubit state population remains low even though the fractional error has increased.

The fractional error of the qubit transition rate estimates $\hat{\Gamma}_{\uparrow}$ and $\hat{\Gamma}_{\downarrow}$ is not as low as that for the population estimates, but is essentially always better than 20% and is typically better than 10%, even for the low SNR case. The performance is worst when one or both of the transition rates is very fast, (causing us to “miss” a greater fraction of events due to our finite measurement bandwidth and limiting our ability to reduce the noise bandwidth B to improve the SNR), or when the rates are very disparate. Our estimates of the excitation rate Γ_{\uparrow} are on the whole better than our estimates of Γ_{\downarrow} because Γ_{\uparrow} is typically much slower, and so we have fewer “missed” events. In all cases, improving the nominal SNR gives a more reliable estimate of the transition rates.

Overall, the performance of the automated qubit state extraction algorithm is good enough for use in our experimental analyses. Because the population estimates are more accurate than the rate estimates, we use population estimates when possible.

8.2 Measurement-induced state mixing

During the course of our quantum jump measurements, we saw that the qubit would occasionally jump up to the excited state during measurement without any applied qubit pulse. This phenomenon became more and more visible as we increased the readout cavity occupation \bar{n} . It has been known for some time through ensemble measurements [64] that something limits the fidelity and QND nature of circuit QED readout at high photon numbers. With our continuous monitoring capabilities and automated jump extraction techniques, we can quantify this behavior for the first time and attempt to discover its origin.

Figure 8.8 shows two typical sets of measurement time traces with the qubit initially in the ground state. As the cavity photon occupation \bar{n} increases, the qubit is more and more likely to be randomly excited during the measurement. This appears to be an effect caused during readout pulse itself, since the qubit is in the ground state at the start of readout, no matter the value of \bar{n} .

8.2.1 Dressed dephasing theory

One candidate theory to explain this phenomenon is the dressed dephasing theory [45, 184]. The dressed dephasing theory predicts that noise at the qubit-cavity detuning frequency $\Delta = \omega_q - \omega_{ro}$ will be mixed with readout photons to cause qubit transitions. The noise couples to the qubit through some control parameter (flux or charge, for example)

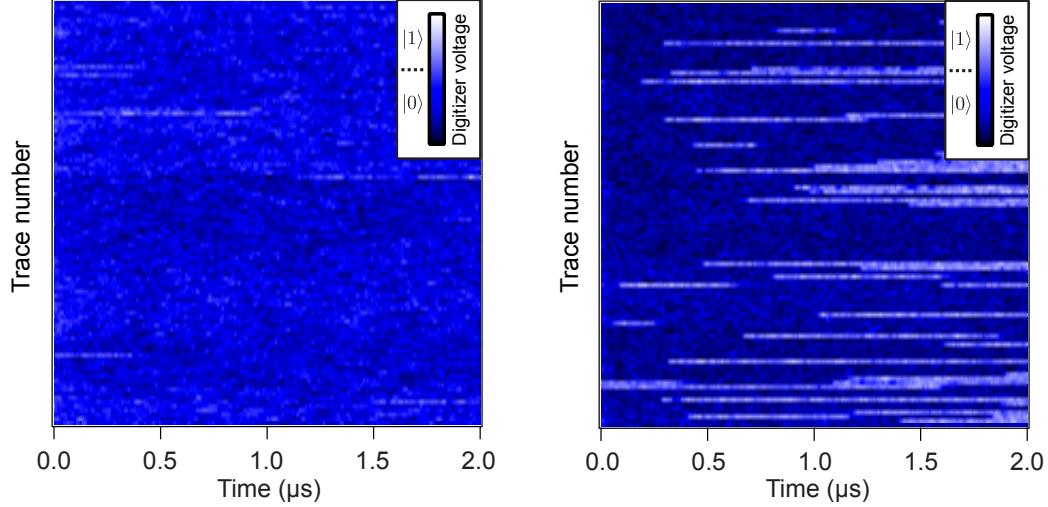


Figure 8.8: Spurious qubit state mixing.

We show two sets of measurement traces taken with the qubit nominally in the ground state for $\bar{n} = 5$ (left) and $\bar{n} = 46$ (right). Some of the traces are spuriously excited during the measurement. The effect appears to be measurement-induced because it becomes stronger with increasing \bar{n} , and because the spurious excited state population is not apparent when the measurement starts.

that enters in the qubit Hamiltonian. This is modeled as a perturbation

$$H_\varphi = \hbar\nu f_\varphi(t)\hat{\sigma}_z, \quad (8.27)$$

where ν characterizes the strength of the coupling between the noise and the qubit (e.g. if $f_\varphi(t)$ is a flux noise, as in the case of the transmon, ν is a flux-to-qubit-frequency transfer function). The low frequency components of $f_\varphi(t)$ are responsible for qubit dephasing. However, the frequency spectrum of $f_\varphi(t)$ can extend out to arbitrarily high frequencies, so that this dephasing noise can have components at the detuning frequency Δ . When this dephasing noise is “dressed” by the presence of readout photons (thus the name “dressed dephasing”), it can give rise to transitions between the qubit states. The rate for transitions up and down is given by [45, 184]

$$\Gamma_{\uparrow,\text{DD}} = 4 \frac{g^2}{\Delta^2} \nu^2 S(\mp\Delta) \bar{n}, \quad (8.28)$$

where g is the qubit-cavity coupling, \bar{n} is the cavity photon occupation, and $S(\Delta)$ is the power spectral density of $f_\varphi(t)$ at the detuning frequency Δ . Once the system has reached steady state, we expect a spurious excited state population to exist. Using the principle of detailed balance, we can express this in terms of $\langle \hat{\sigma}_z \rangle$ as

$$\langle \hat{\sigma}_z \rangle = -1 + \frac{\Gamma_{\uparrow,\text{DD}}}{\Gamma_1 + \Gamma_{\downarrow,\text{DD}}}. \quad (8.29)$$

Here $\Gamma_1 = 1/T_1$ is the intrinsic qubit decay rate, including the Purcell effect. In the case of a symmetric noise spectrum where $S(\Delta) = S(-\Delta)$, we have $\Gamma_{\downarrow,DD} = \Gamma_{\uparrow,DD}$. Furthermore, in most cases the spurious excitation and relaxation rates are much slower than Γ_1 . Using these two ideas, we can rewrite equation (8.29) as:

$$\langle \hat{\sigma}_z \rangle = -1 + \frac{\Gamma_{\uparrow,DD}}{\Gamma_1 + \Gamma_{\uparrow,DD}} \approx -1 + \frac{\Gamma_{\uparrow,DD}}{\Gamma_1}. \quad (8.30)$$

The fractional excited state population of the qubit is given by $(1 + \langle \hat{\sigma}_z \rangle)/2$, which means that it is directly proportional to $\Gamma_{\uparrow,DD}$. This in turn implies that the fractional excited state population we observe in steady state should be directly proportional to the cavity photon occupation \bar{n} and the spectral density of flux noise at the detuning frequency $S(\pm\Delta)$.

We now set out to perform experimental tests of the dressed dephasing theory by intentionally injecting signals and/or noise into the qubit loop. There are three main avenues we pursue. First, we apply flux signals to the qubit at different frequencies and see if the qubit responds to noise at all frequencies or just near Δ . This establishes whether or not the notion of mixing between flux noise at Δ and readout photons to generate qubit transitions is accurate. Second, we apply broadband noise with a known component at Δ and see if the steady state qubit population scales linearly with $S(\Delta)$ and \bar{n} as predicted. Finally, if these two tests support the dressed dephasing theory, we can use measurements of the qubit behavior with and without added noise to measure the power spectral density of flux noise at the detuning frequency Δ , which is of order ~ 1 GHz. Flux noise in superconducting devices has never been measured in this frequency range before.

8.2.2 Experimental scheme

The experiment is similar to that presented in the previous chapter. We use a transmon qubit with $E_J(\Phi = 0) = 20.3$ GHz and $E_C = 231$ MHz, capacitively coupled to a superconducting quasi-lumped-element resonator consisting of a meander inductor ($L = 3.68$ nH) in parallel with an interdigital capacitor ($C = 175$ fF). The qubit/cavity sample is shown in Figure 8.9(a). Unlike the sample in the previous chapter, this resonator has asymmetric coupling and is operated in transmission, with excitations entering from the weakly coupled port via a heavily attenuated injection line and leaving through the strongly coupled port towards the amplifiers¹¹. The qubit state is manipulated by applying microwave signals through the weakly coupled port of the resonator. In addition, while the sample in the previous chapter was on a Bollywood launch, this sample was in an enclosed, shielded type “B” sample box with bulkhead SMA connectors (see section 5.1 for more details). Qubit coherence times are better in this sample, varying monotonically depending on the qubit frequency from $T_1 = 290$ ns and $T_2 = 550$ ns at $f_q = 5.705$ GHz to $T_1 = 910$ ns and $T_2 = 1.35$ μ s at $f_q = 5.075$ GHz. These numbers represent T_1 values about a factor of two below the Purcell limit, and pure dephasing times T_φ much longer than T_1 . The readout cavity has

¹¹This design has the additional advantage that the cavity serves as a filter for the qubit pulses, substantially reducing their power at the paramp relative to the reflection geometry. The paramp does not appear to be affected by the presence of qubit pulses in either geometry since they are very far detuned from the pump, but it’s nice to control this variable nonetheless.

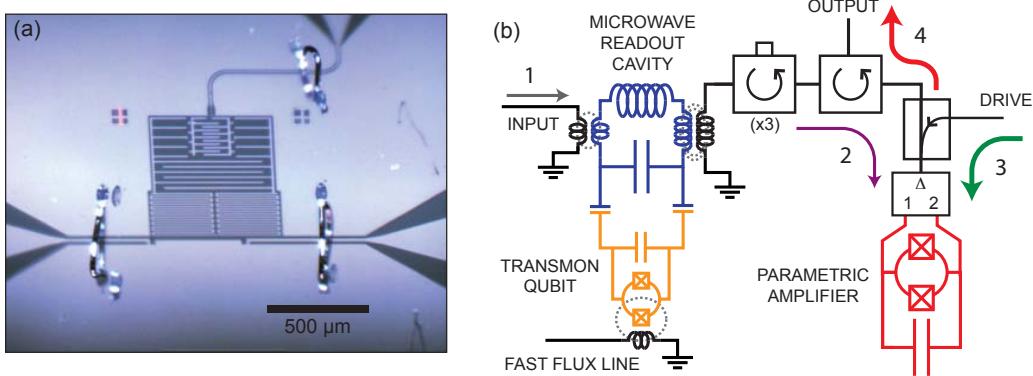


Figure 8.9: Transmon qubit and resonator with fast flux line.

Part (a) shows a photograph of the qubit/cavity sample with input port (left), output port (right), and fast flux port (top). Part (b) depicts the experimental setup, similar to the previous chapter. Readout photons enter the cavity via the weakly coupled input port (grey arrow, “1”), interact with the qubit, and acquire a phase shift which depends on the qubit state. The readout signal then leaves the cavity via the strongly coupled port (purple arrow, “2”), passes through several circulators, and combines with the paramamp pump tone (green arrow, “3”). The readout and pump interact in the paramamp and the amplified readout signal (red arrow, “4”) is reflected to the output port. The readout signal then is further amplified, mixed down to zero frequency (converting the phase shift signal into a voltage signal), and digitized. Noise or coherent tones can be injected into the qubit loop via a weakly coupled fast flux line with a bandwidth of 2.2 GHz. Grey dashed lines indicate flux coupling.

a bare frequency of 6.2724 GHz with a linewidth $\kappa/2\pi = 7$ MHz and a coupling to the qubit $g/2\pi = 106$ MHz. The cavity frequency is Lamb-shifted up by roughly 10-20 MHz depending on the qubit bias point.

For each qubit bias point, the dispersive shift 2χ is calculated using an expression that accounts for the higher excited states of the transmon qubit and includes corrections for the Kerr nonlinearity [93], as described in section 2.3.3. We use the dispersive shift information to calibrate the number of photons in the readout cavity using the ac Stark shift [42] (see section 6.4.3) and to set the frequency $\omega_{\text{ro}}/2\pi$ to use for readout photons. We select a readout frequency halfway between the cavity resonant frequencies corresponding to the qubit in the ground and first excited states $\omega_{\text{ro}} = \frac{1}{2}[\tilde{\omega}_{\text{cav}}(|0\rangle) + \tilde{\omega}_{\text{cav}}(|1\rangle)] = \tilde{\omega}_{\text{cav}}(|0\rangle) + \chi$. This choice of readout frequency means that the cavity photon occupation \bar{n} is the same whether the qubit is in state $|0\rangle$ or $|1\rangle$, simplifying the analysis of the dependence of transition rates on \bar{n} .

The experimental setup, shown in Figure 8.9(b), is similar to that presented in the previous chapter. Photons enter the readout cavity, where they acquire a phase shift that depends on the state of the qubit. These readout photons leave the cavity through the

strongly coupled port and pass through four microwave circulators, which isolate the qubit from the paramp pump. The paramp amplifies the signal, which is then further amplified by cryogenic and room temperature amplifiers (not shown) and finally mixed down to zero frequency and digitized. Detailed experimental schematics are shown in Appendix C.

This sample also has a weakly coupled fast flux line, which allows modulation of the qubit Hamiltonian by noise or coherent signals. The fast flux line has a bandwidth of 2.2 GHz, defined by a reactive VLFX-1350 filter at 700 mK and three lossy impedance-matched roach filters (see section 5.2.3 and Appendix A for more details) at 4 K, 100 mK, and 50 mK. The roach filters thermalize the line and prevent any high frequency signals in the re-entrant part of the VLFX filter’s passband from reaching the qubit. Because the roach filters have no loss at dc, we can pass substantial (\sim mA) dc currents without causing too much dissipation on the cold stages and heating the fridge. A 20 dB attenuator at the 4 K stage provides some low-frequency thermalization.

We calibrate the coupling of the fast flux line to the qubit loop by extracting the flux-to-qubit-frequency transfer function at the qubit operating point from spectroscopy¹². We then measure the qubit frequency using Ramsey fringes and watch the frequency change as a function of applied dc current through the fast flux line. Combining the frequency-to-flux and current-to-frequency transfer functions gives us a calibration of the flux in the qubit loop as a function of the current in the fast flux line at base temperature. We can then use the measured frequency-dependent attenuation of the fast flux line to convert room-temperature power into a flux in the qubit loop. The coupling is sufficiently weak ($180 \text{ mA}/\Phi_0$) that noise from the 50Ω impedance of the fast flux line should not be the dominant source of flux noise in the qubit loop for frequencies at or below $\omega_q/2\pi^{13}$.

The measurement protocol consists of readout pulses lasting $17.5 \mu\text{s}$ occurring every $100 \mu\text{s}$. The long delay ensures that the qubit will fully relax to its thermal ground state (1.5% excited state population, corresponding to a qubit temperature of 62 mK) between measurement runs. We take 10^4 individual time traces for each combination of experimental parameters. The measurement traces are then analyzed using the automated state extraction algorithm described in section 8.1. All fast flux excitations, both coherent tones and noise, are applied continuously (not pulsed) independent of whether the readout is on or off.

8.2.3 State mixing with added flux tones

To test the dressed dephasing theory, we first would like to determine if signals at the detuning frequency can be mixed with readout photons to cause state transitions, as postulated. We do this by injecting a continuous microwave tone into the fast flux line of frequency $\omega_{\text{ff}} \sim \Delta(\bar{n})$, where we note that Δ depends on \bar{n} due to the ac Stark shift [42]. This tone produces a small flux oscillation in the qubit loop; the power P_{ff} of the tone was

¹²In this particular sample, we were unable to tune the qubit through a full flux quantum due to weak coupling to our external coil, so we took spectroscopy over a partial range and fit it to the theoretical expression for qubit frequency versus flux to get the desired transfer function.

¹³The equivalent flux noise in the qubit loop due to quantum fluctuations in the fast flux line is $3 \times 10^{-12} \Phi_0/\sqrt{\text{Hz}}$ at the qubit frequency, several orders of magnitude below the value of the intrinsic flux noise we extract in section 8.2.4.

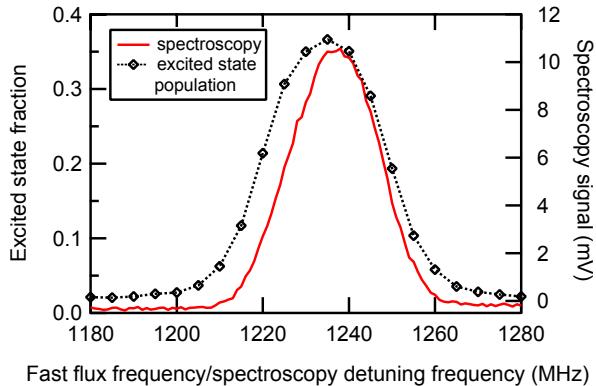


Figure 8.10: Spurious excitation and qubit spectroscopy.

We plot the qubit excited state population (black diamonds) versus the frequency of fast flux excitation for a fast flux signal of $625 \mu\Phi_0$ RMS and a readout cavity occupation $\bar{n} = 12$. For comparison, we plot independently measured qubit spectroscopy (red line), with the spectroscopic frequency given in terms of the detuning from the cavity frequency $\Delta(\bar{n})$. The spectroscopic qubit line has been shifted and broadened by the ac Stark effect. Note that spurious excitation of the qubit occurs primarily in regions where the frequency of the fast-flux excitation signal is commensurate with Δ .

varied to produce RMS flux excitations of up to $625 \mu\Phi_0$, corresponding to qubit frequency fluctuations of 2-4 MHz RMS (depending on qubit bias point). These fluctuations are much smaller than the ac-Stark-broadened qubit linewidth $\gamma_q(\bar{n})/2\pi$ for the values of \bar{n} studied. For each value of \bar{n} , we stepped the frequency of the coherent tone through a range from $\omega_{\text{ff}} < |\Delta(\bar{n})| - \gamma_q(\bar{n})$ to $\omega_{\text{ff}} > |\Delta(\bar{n})| + \gamma_q(\bar{n})$ and stepped P_{ff} through ten values linearly spaced over a factor of 40 in power, as well as a control where $P_{\text{ff}}=0$. The qubit is detuned by about 1.2 GHz at our operating point, with $\omega_q/2\pi = 5.075$ GHz.

Figure 8.10 shows the extracted steady state qubit population as a function of fast flux frequency ω_{ff} for a cavity readout occupation \bar{n} of 12 photons and an RMS flux excitation of $650 \mu\Phi_0$. We observe that spurious excitation occurs primarily in regions where the frequency of the fast flux tone is commensurate with the detuning frequency $\Delta(\bar{n})$ between the cavity frequency and the ac-Stark-shifted qubit line, which is measured independently by spectroscopy as described in section 6.4.3.

A larger plot of experimental results is shown in Figure 8.11. When $P_{\text{ff}} \neq 0$, qubit state mixing occurs as long as ω_{ff} is within roughly a qubit linewidth of the detuning frequency, and is most noticeable when $\omega_{\text{ff}} = |\Delta(\bar{n})|$. The effect grows stronger with increasing P_{ff} . Figure 8.11(f) shows the same bias parameters as Figure 8.11(e), except it displays the qubit population at the start of the readout as opposed to the steady state population. This initial excited population is much lower than the steady state excited population and is essentially independent of ω_{ff} and \bar{n} , indicating that the spurious excitation observed in (a)-(e) is caused by the presence of readout photons and does not occur when the readout

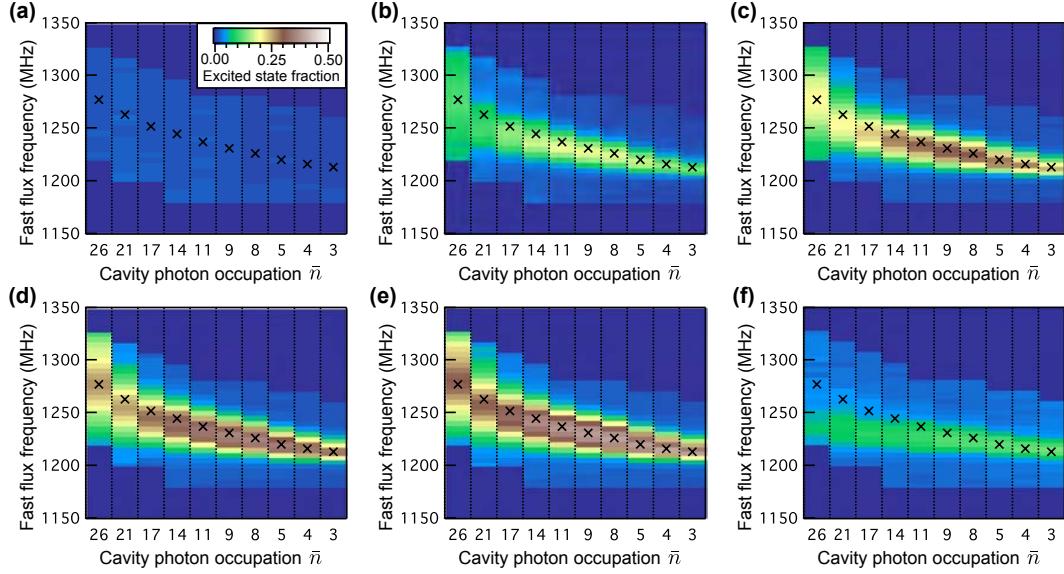


Figure 8.11: Spurious excitation with coherent fast flux tone.

Parts (a)-(e) show the qubit steady state population during measurement with a coherent microwave tone applied to the fast flux line. Each panel is for a different value of P_{ff} , corresponding to RMS added fluxes of $0 \mu\Phi_0$ (a), $230 \mu\Phi_0$ (b), $370 \mu\Phi_0$ (c), $470 \mu\Phi_0$ (d), and $550 \mu\Phi_0$ (e). The population data are shown as a function of ω_{ff} and of the cavity photon occupation \bar{n} . The black 'x's denote the qubit-cavity detuning frequency $\Delta(\bar{n})$ as determined independently by qubit spectroscopy. The qubit remains primarily in the ground state, even for large fast flux excitations, except when $\omega_{ff} \approx \Delta(\bar{n})$ (within about one qubit linewidth γ_q ; note that γ_q grows with \bar{n}). The excited state fraction also grows with increasing P_{ff} . Part (f) shows the qubit population measured 40 ns after the readout has turned on (the earliest time point at which we could reliably extract qubit population) with the same P_{ff} as part (e). The qubit is primarily in the ground state even when $\omega_{ff} \approx \Delta(\bar{n})$, showing that the excited state population in (e) developed during the measurement and did not exist before the measurement turned on. There is a small ($\approx 10\%$) excited state population visible, which does not track with the measured $\Delta(\bar{n})$ as in the other panels. We theorize this may be the result of mixing between the fast flux tone and the cavity photons in the very early stages of readout, or potentially of mixing between a residual thermal cavity population and the fast flux signal when the measurement is off. The vertical color scale is the same for all panels.

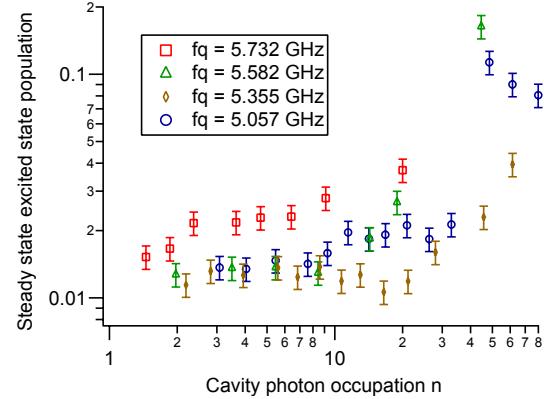


Figure 8.12: Qubit population with no fast flux tone.

This figure plots the steady-state excited state population of the qubit in the absence of fast flux tones as a function of \bar{n} for four different qubit frequencies. The higher qubit frequencies correspond to smaller detunings Δ but also reduced qubit sensitivity to flux ν . The excited state population has a roughly linear dependence on \bar{n} until $\bar{n} \approx 20$, at which point it increases rapidly. The data are plotted on log-log axes to show the large range of qubit populations and \bar{n} .

is off. This fact suggests that the applied flux signals at the detuning frequency are mixing with readout photons to cause qubit transitions. The small excited state population visible in Figure 8.11(f) does not track with the measured $\Delta(\bar{n})$ as in the other panels, and may arise from mixing between the fast flux tone and the cavity photons in the very early stages of readout (since \bar{n} is lower during cavity ring-up than in steady state, the qubit frequency is shifted less and Δ is smaller), or possibly from mixing between a residual thermal cavity population and the fast flux signal when the measurement was off.

In the absence of a fast flux tone, we can still observe some spurious qubit excitation with increasing \bar{n} , as seen in Figure 8.12. The effect is about 1% additional excited population per 10 photons cavity occupation, increasing more rapidly at higher photon numbers (up to $\approx 10\%$ total for $\bar{n} \approx 40$). Since there is no applied fast flux tone, the dressed dephasing theory postulates that the spurious excitation is due to intrinsic flux noise at $\Delta(\bar{n})$ in the qubit being upconverted by measurement photons. If so, this represents a limit on measurement fidelity—even with quantum-limited post-amplification—by giving a penalty for increasing \bar{n} . It would also suggest a route to optimizing measurement fidelity by addressing sources of low-frequency noise in the qubit.

8.2.4 Flux noise measurements

Having determined that signals at $\Delta(\bar{n})$ cause spurious excitation, we can use the qubit as a spectrometer to measure flux noise at Δ ¹⁴. To test the qubit's response to noise, we applied flux noise to the qubit loop through the fast flux line and examining spurious excitation during measurement. The noise was generated by amplifying the Johnson noise of a room-temperature $50\ \Omega$ termination, which was then sent to the fast flux line through a computer controlled attenuator, allowing the strength of the added noise to be varied (detailed diagrams of the noise generation setup are in Appendix C). The experiment was performed both with full-spectrum white noise from 10 MHz to 2.2 GHz and with noise filtered to lie in the range 180 MHz - 2.2 GHz; the steady state qubit populations were essentially identical between these two types of applied flux noise, suggesting again that only noise components near $\Delta(\bar{n})$ are responsible for spurious excitation.

Figure 8.13(a) shows the steady state qubit excited population as a function of cavity photon occupation \bar{n} and the spectral density of added flux noise at the detuning frequency $S_{\Phi}^{1/2}(\Delta(\bar{n}))$. As anticipated, the qubit excited state population increases with both \bar{n} and $S_{\Phi}^{1/2}(\Delta(\bar{n}))$. The scaling is roughly linear in \bar{n} and quadratic in $S_{\Phi}^{1/2}(\Delta(\bar{n}))$, as predicted by the dressed dephasing theory. To ensure that the added noise is not causing excitation in the absence of measurement, we also examined the initial qubit population at the start of measurement from the same data set, shown in Figure 8.13(b). As expected, the excited state population at the start of the measurement corresponds to the thermal population, and is independent of \bar{n} and $S_{\Phi}^{1/2}(\Delta(\bar{n}))$.

If we attribute all spurious excitation to upconverted flux noise, we can extrapolate these data back to a value of the added flux noise where there is no excitation above the thermal population, yielding an estimate of the intrinsic flux noise at the detuning frequency. Using this method, we extract a flux noise spectral density $S_{\Phi}^{1/2} = 0.005 \pm 0.002\ \mu\Phi_0/\sqrt{\text{Hz}}$ at 690 MHz. We can test whether this number agrees with the dressed dephasing expressions (8.28) and (8.30) by inserting our experimental parameters g , Δ , Γ_1 , ν , and the slope of the excited state population versus \bar{n} . This latter quantity is determined by fitting a line to the data in Figure 8.12. Using the extracted numbers, we find that the dressed dephasing theory predicts a flux noise spectral density $S_{\Phi}^{1/2} = 0.004 \pm 0.001\ \mu\Phi_0/\sqrt{\text{Hz}}$, in good agreement with our experimentally extracted number.

We can perform cross-checks on this noise level by using Ramsey and Rabi experiments (detailed in section 6.4.2) to measure flux noise in different frequency ranges. First, we measured Ramsey fringes at a rate of one per second for an hour, then fit each fringe to extract the qubit frequency. We then translated noise in the qubit frequency into an effective flux noise using the flux-to-qubit-frequency transfer coefficient ν , allowing us to extract a flux noise spectral density for frequencies below 0.5 Hz. Second, we performed Rabi oscillations with a zero-detuning Rabi drive tone for several Rabi drive amplitudes. Each drive amplitude yielded to a different Rabi frequency Ω_{Rabi} , and we extracted the

¹⁴Since the transmon qubit does not couple to charge noise, flux noise should be the main source of qubit dephasing in our experiment. In this section, we will quote noise on the qubit frequency in terms of an equivalent flux noise. In principle, the observed spurious excitation could also be due to critical current noise. Recent experiments suggest, however, that critical current noise in transmon junctions is much weaker than had previously been supposed [17], and flux noise should be the dominant source of dephasing.

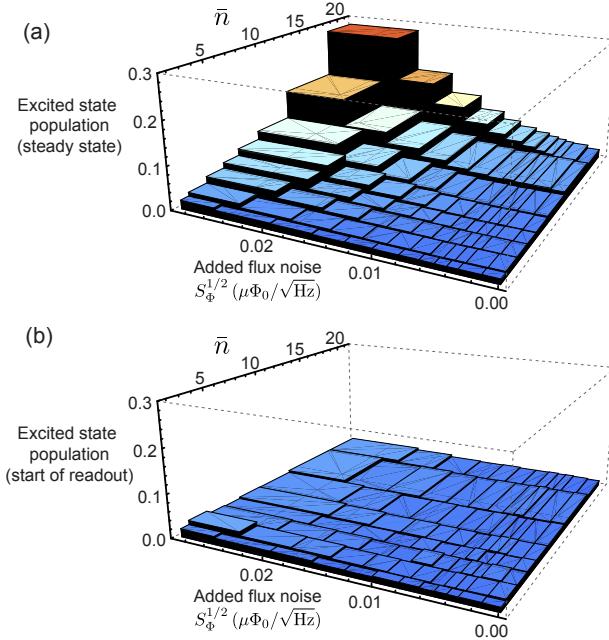


Figure 8.13: Qubit population with added noise.

Part (a) shows the qubit steady state population as a function of added flux noise and the number of measurement photons \bar{n} in the readout cavity. Part (b) shows the qubit population immediately after the readout has fully energized. A thermal population of 1.5% is visible, and is equivalent for all values of added flux noise and \bar{n} .

decay rate Γ_{Rabi} for each Ω_{Rabi} . We used this decay rate to obtain a spectral density of qubit frequency fluctuations $S_{\delta\omega_q}$ at Ω_{Rabi} using the relation

$$\Gamma_{\text{Rabi}} = \frac{3}{4}\Gamma_1 + \frac{1}{2}\Gamma_\nu, \quad (8.31)$$

where $\Gamma_1 = 1/T_1$ and $\Gamma_\nu = \pi S_{\delta\omega_q}(\Omega_{\text{Rabi}})$ [159]. Converting $S_{\delta\omega_q}(\Omega_{\text{Rabi}})$ to an effective flux noise gives us data for frequencies between 1 and 20 MHz. These methods are also discussed in section 6.4.2.

The flux noise values extracted by these methods, as well as those calculated from our spurious excitation data, are shown in Figure 8.14. We fit the Ramsey data to a $1/f^\alpha$ power law [160, 185] to yield the red trend line shown. This fit agrees well with the extracted values from both Rabi decay and from the spurious excitation data, representing a power law for flux noise that appears to hold over 11 orders of magnitude in frequency. The fit coefficients give $\alpha = 0.58 \pm 0.04$ and $S_\Phi^{1/2}(1\text{Hz}) = 1.4 \pm 0.1 \mu\Phi_0/\sqrt{\text{Hz}}$, both of which agree with typical values reported in the literature [185, 186, 187, 188, 189]. We note that other recent work produced a similar $S_\Phi^{1/2}(1\text{Hz})$ but found $\alpha = 0.9 - 1.0$ [160, 190]; sample-to-sample variation in α of this magnitude has been noted elsewhere [185, 189].

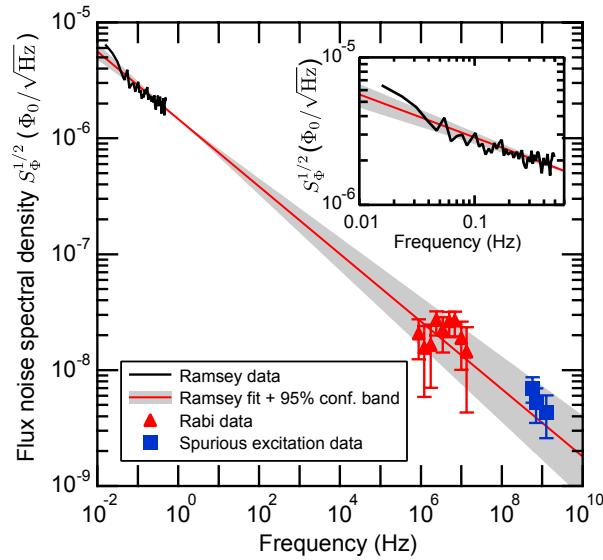


Figure 8.14: Spectral density of flux noise vs frequency.

We plot effective flux noise extracted from Ramsey fringes (black), Rabi oscillation decay (red triangles), and our measurements of spurious excitation (blue squares). The red line is a fit to the Ramsey fringe data using a $1/f^\alpha$ power law, with the grey shaded area representing the 95% confidence interval on the fitted value of $\alpha = 0.58$. The inset view shows a detail of the Ramsey fringe data and fit line.

The correspondence of the low-frequency fit to the extracted flux noise at $\Delta(\bar{n})$ suggests that the “universal” low-frequency flux noise [191] causes not only qubit dephasing but also reduces the QND character of circuit QED measurement, in quantitative agreement with the proposed dressed dephasing theory [45, 184]. Further work to improve low-frequency dephasing noise, such as is being carried out by many groups including ours, may therefore also provide a route to improved qubit readout fidelity. As T_1 values grow increasingly longer, the effects of measurement non-idealities like dressed dephasing become more and more important.

8.3 Quantum Zeno effect

The process of quantum measurement forces a quantum system into an eigenstate of the measured observable. This measurement backaction gives rise to a number of unusual phenomena, including the quantum jumps described in the previous chapter. One of the most famous examples of measurement backaction is the quantum Zeno effect. The quantum Zeno effect [169] pertains to the competition between quantum state evolution and projective measurement; the first tries to change the state of the quantum system of interest, while the second forces it into an eigenstate. The presence of measurement actually

slows the evolution of the quantum system.

The quantum Zeno effect was first observed experimentally in 1990 using an ensemble of trapped ions [192]. Since that time, it has also been seen in single trapped ions [193], molecules [194], cold atoms [195], photons [196], and the polarization of light [197, 198]. In 2000, the anti-Zeno effect, where a quantum system can decay faster in the presence of measurement than without, was proposed [199]. A recent theoretical paper looked at the Zeno effect in the classical limit, as $\hbar \rightarrow 0$, and found that it vanishes to all orders in \hbar [200]. This means that the Zeno effect is truly quantum mechanical, and provides evidence for the quantum mechanical nature of systems in which it is observed.

The basic theory of the quantum Zeno effect is rather straightforward, the classic example being a qubit undergoing Rabi oscillations while being periodically measured. We can write a simple model for the qubit's behavior, after Cook [163], using density matrix formalism. For a qubit undergoing Rabi oscillations at frequency Ω from an on-resonance drive, we can write the time evolution of the density matrix ρ in the rotating frame as follows:

$$\dot{\rho}_{00} = -\frac{1}{2}i\Omega(\rho_{01} - \rho_{10}) \quad (8.32)$$

$$\dot{\rho}_{11} = \frac{1}{2}i\Omega(\rho_{01} - \rho_{10}) \quad (8.33)$$

$$\dot{\rho}_{01} = \frac{1}{2}i\Omega(\rho_{11} - \rho_{00}). \quad (8.34)$$

Let's say that at time $t = 0$ we make a measurement of the qubit, projecting it into the state $|1\rangle$. This sets boundary conditions for equations (8.32)-(8.34) at $t = 0$, namely $\rho_{00} = \rho_{01} = \rho_{10} = 0$ and $\rho_{11} = 1$. We can then write down explicit solutions for the components of the density matrix as a function of time.

$$\rho_{00}(t) = \sin^2(\Omega t/2) \quad (8.35)$$

$$\rho_{11}(t) = \cos^2(\Omega t/2) \quad (8.36)$$

$$\rho_{01}(t) = \frac{i}{2} \sin(\Omega t). \quad (8.37)$$

The probability $P_{|0\rangle}(\tau)$ or $P_{|1\rangle}(\tau)$ for a subsequent measurement at time τ to find the qubit in either state $|0\rangle$ or $|1\rangle$ is simply given by $\rho_{00}(\tau)$ and $\rho_{11}(\tau)$ respectively. Let's consider the case when another measurement is made a very short time (a small fraction of a Rabi period) after the first one, so that the qubit will not have strayed far from being in the state $|1\rangle$. Mathematically, we can write this by saying $\Omega\tau \ll 1$, allowing us to Taylor expand equations (8.35)-(8.37). We then write the density matrix components as:

$$\rho_{00}(\tau) \approx (\Omega\tau/2)^2 = \frac{\Omega^2\tau^2}{4} \quad (8.38)$$

$$\rho_{11}(\tau) \approx \left(1 - \frac{1}{2}(\Omega\tau/2)^2\right)^2 \approx 1 - (\Omega\tau/2)^2 = 1 - \frac{\Omega^2\tau^2}{4} \quad (8.39)$$

$$\rho_{01}(\tau) \approx \frac{i\Omega\tau}{2}. \quad (8.40)$$

The probability for us to observe a qubit state change $P_{\text{sc}}(\tau)$ when we measure at time τ is $P_{|0\rangle}(\tau) = \rho_{00}(\tau)$, or alternatively $1 - P_{|1\rangle}(\tau) = 1 - \rho_{11}(\tau)$. We note that in the case where we do not measure between time 0 and τ , this probability is just given by

$$P_{\text{sc}}(\tau) = \Omega^2\tau^2/4, \quad (8.41)$$

which increases quadratically with the time τ . Now we can ask what happens if we perform multiple measurements within the time τ . Let's imagine we perform N measurements evenly distributed in time; the time between measurements will be τ/N . To find the probability $P_{\text{sc}}(\tau)$ of a state change, we need to determine the probability that the qubit will have remained in state $|1\rangle$ throughout all the measurements. We can write down the probability for the qubit to be in state $|1\rangle$ after one such measurement using equation (8.39), which gives:

$$P_{|1\rangle}(\tau/N) = 1 - \frac{\Omega^2(\tau/N)^2}{4} \quad (8.42)$$

Every time a measurement occurs, the qubit is projected back into either $|0\rangle$ or $|1\rangle$. As long as the qubit is projected back into $|1\rangle$ each time, the probability in equation (8.42) only depends on the time τ/N since the last measurement and not the overall time since $t = 0$. The case where the qubit is projected back to $|1\rangle$ each time is the one we are interested in, because we would like to know the probability that the qubit has remained in state $|1\rangle$ throughout all the measurements. Because the probability for each interval is the same, we can write that the total probability to remain in $|1\rangle$ given N measurements in time τ is just the product of the individual probabilities to remain in $|1\rangle$ at each interval:

$$P_{|1\rangle}(\tau, N) = (P_{|1\rangle}(\tau/N))^N = \left(1 - \frac{\Omega^2(\tau/N)^2}{4}\right)^N \quad (8.43)$$

Recalling that $\Omega\tau \ll 1$, we can expand this form out to give:

$$P_{|1\rangle}(\tau, N) \approx 1 - \frac{\Omega^2\tau^2}{4N} = 1 - \frac{\Omega^2\tau}{4f_m}, \quad (8.44)$$

where $f_m = N/\tau$ is the frequency at which measurements are made. We can then find the probability of state change in the time τ , which is just $1 - P_{|1\rangle}(\tau, N)$:

$$P_{\text{sc}}(\tau, f_m) = \frac{\Omega^2\tau}{4f_m} \quad (8.45)$$

Let's compare this result with equation (8.41). In the absence of measurement, $P_{\text{sc}}(\tau) \propto \tau^2$, showing a quadratic dependence of the qubit evolution on time. By contrast, $P_{\text{sc}}(\tau, f_m) \propto \tau$ in the presence of repeated measurements, exhibiting linear rather than quadratic scaling behavior in time. Furthermore, we notice that the act of repeated measurement actually decreases the probability of changing states, effectively slowing down the qubit evolution.

The stronger (i.e. more frequent) the measurement, the more the qubit evolution is inhibited; this is most readily apparent from the form in equation (8.44). This phenomenon of inhibition of qubit evolution through repeated measurement is the quantum Zeno effect. In the limit where $\tau \rightarrow \delta\tau$ becomes infinitesimal, we can talk about the probability of changing states per unit time $\delta\tau$, in other words a rate Γ_{jump} of quantum jumps between states. Using (8.45) we find

$$\Gamma_{\text{jump}} = \Omega^2 / 4f_m \quad (8.46)$$

The model derived above is fairly simplistic, and we must take a little bit of care interpreting its results. If we carry equation (8.45) out to the limit of infinitely fast measurement, the transition probability would drop to zero and we would expect to see the qubit “frozen” in state $|1\rangle$. This is referred to in the literature as the quantum Zeno paradox, as distinct from the quantum Zeno effect (which is merely a slowing of qubit evolution). The quantum Zeno paradox is not realizable in real systems which exhibit radiative or radioactive decay because at the very short time scales required for the measurement, the energy uncertainty is so large that the system couples to channels outside the subspace of interest [199, 201].

In our experiments, we have continuous rather than pulsed measurement. However, the Zeno effect should still hold theoretically in this instance [202], with the time between individual pulsed measurements replaced by an effective measurement time defined by the measurement rate Γ_m . If we take $\delta\tau=1/\Gamma_m$, we have $\Gamma_{\text{jump}} = \Omega^2 / 4\Gamma_m$. The model above also does not account for qubit relaxation processes.

Gambetta and co-workers [91] were able to derive approximate results for the Zeno effect in the case of a continuous circuit QED measurement. Their analytical expression only holds in the weak measurement regime, and takes the form

$$\Gamma_{\text{jump}} = \frac{\Omega^2}{2(\Gamma_2 + \Gamma_d)}, \quad (8.47)$$

where $\Gamma_2 = \frac{1}{2}\Gamma_1 + \Gamma_\varphi$ ($\Gamma_1 = 1/T_1$ and Γ_φ is the low-frequency dephasing rate) and Γ_d is the measurement-induced dephasing rate. The measurement rate Γ_m is given by $\frac{1}{2}\Gamma_d$ in the ideal case¹⁵, so in the limit of low intrinsic decoherence ($\Gamma_2 \rightarrow 0$) this expression reduces to the same one given in the previous paragraph derived from the simplistic theory.

In the strong measurement regime, one cannot rid the qubit stochastic master equation used to derive equation (8.47) of cavity effects, and so there is no analytical formula presented for Γ_{jump} with strong measurement in [91]. However, the results of full numerical simulations of the qubit/cavity system seems to suggest a scaling $\Gamma_{\text{jump}} \propto \Omega$, rather than the $\Gamma_{\text{jump}} \propto \Omega^2$ of the simple analytical form.

While the quantum Zeno effect is traditionally described in the time domain, the expression in (8.47) gives a hint that we can think about it in the frequency domain as well for our experiment. The qubit linewidth is set by the dephasing; normally, in the absence of measurement, the qubit linewidth is given by Γ_2 . As can be seen in circuit QED ac Stark shift data, the presence of measurement photons broadens the qubit line [42]. One can think of this in terms of the action of the measurement-induced dephasing Γ_d , which

¹⁵We remind the reader that there are two conventions for the relationship between Γ_m and Γ_d which differ by a factor of two. We use the convention of [91] and [92] here because the expression in (8.47) uses this convention.

with strong measurement is considerably larger than the intrinsic dephasing Γ_φ and so sets the qubit linewidth. As described above, the measurement-induced dephasing is intimately tied to the measurement rate.

When a qubit undergoes Rabi oscillations, a set of sidebands appear on the qubit line in the frequency domain, with the distance between the central peak and the sidebands given by the Rabi frequency. The qubit line and two Rabi sidebands are sometimes referred to as the “Mollow triplet” in atomic and optical physics [203]. The presence of the sidebands is an indication in the frequency domain of the qubit’s time domain oscillations. Qualitatively, the quantum Zeno effect occurs when the qubit linewidth becomes so broadened by measurement that one can no longer resolve the Rabi sidebands separately from the main peak. Mathematically, this can be expressed roughly as $\Gamma_d \gg 2\Omega$. If we use the relations $\Gamma_d = 2\Gamma_m \sim 1/\tau$, this gives us the rough criterion $\Omega\tau \ll 1$, which we note is the essentially same criterion of “fast” measurement repeat rate used in equation (8.39) for our derivation of the simple case of the quantum Zeno effect.

The slowing and eventual disappearance of coherent Rabi oscillations under increasingly strong measurement, an indication of the quantum Zeno effect, has been observed experimentally by the Saclay group [23]. We have duplicated these results and found that Rabi oscillations in our qubit/cavity system tend to be suppressed by measurement for cavity photon occupations $\bar{n} \gtrsim 1$. Above this value of \bar{n} , we should be in the Zeno regime.

For $\bar{n} \gtrsim 3$, we are able to use our automated jump extraction algorithm to determine the excitation and relaxation rates of the qubit directly from the individual measurement traces. Our protocol consists of turning on the readout, waiting 3 μs , and then turning on a qubit drive tone at the center frequency of the ac-Stark-shifted qubit line $\omega_q(\bar{n})$. We leave the qubit drive tone on for 10 μs . This experiment is repeated 10^4 times each (at a 100 μs repetition rate) for a range of values of both \bar{n} and qubit drive amplitude. The measurement setup and qubit sample are the same as were presented in the previous section on measurement-induced state mixing. We use our automated jump extraction algorithm to analyze the data. We only analyze data taken while both readout and qubit drive tones were on, because the rates will differ from when the qubit drive is off. The time in the data traces before the qubit drive turns on can be separately analyzed to calibrate the amount of measurement-induced state mixing if we would like to subtract off this effect.

The extracted jump rates are plotted in Figure 8.15. Because the actual Γ_{jump} values vary over several orders of magnitude depending on the experimental bias parameters, we plot scaled values rather than absolute values. We use Ω , the Rabi frequency in the absence of measurement, as our scaling factor. In part (a), we plot the rate of excitation from the ground state Γ_\uparrow (the total $\Gamma_{\text{jump}} = \Gamma_\uparrow + \Gamma_\downarrow$) as a function of Ω and $\sqrt{\bar{n}}$. Because there is no T_1 contribution to Γ_\uparrow , the rate should be just that given by the naïve Zeno theory, with $\Gamma_\uparrow = \Omega^2/\Gamma_m$. We plot the scaled quantity Ω^2/Γ_\uparrow on the vertical axis; the use of the scaling factor Ω^2 allows us to readily compare qubit excitation rates which vary over more than two orders of magnitude as we change Ω and \bar{n} , and show that these rates obey the scaling in (8.47). The data show the quantity Ω^2/Γ_\uparrow to be more or less independent of Ω , as suggested by the simple Zeno theory. In addition, we see that Ω^2/Γ_\uparrow increases with increasing photon number \bar{n} . This demonstrates the quantum Zeno effect; for a given Rabi drive strength, increasing the measurement strength decreases the rate at which the qubit

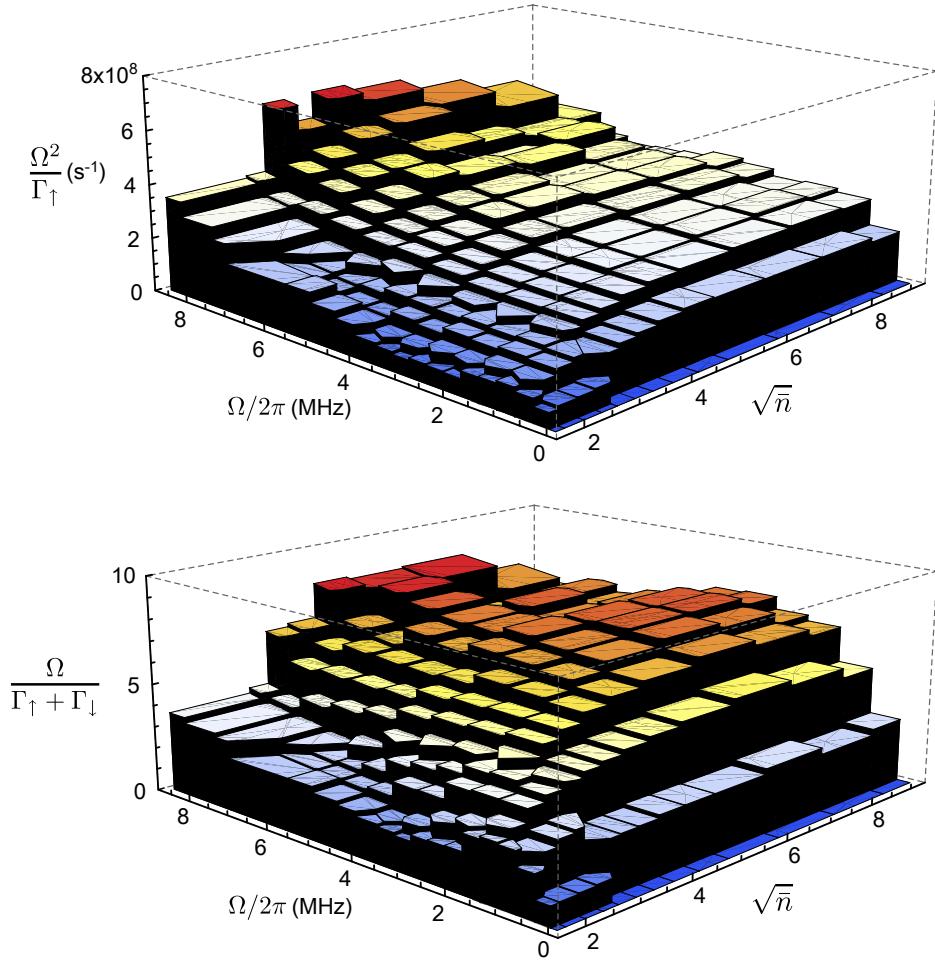


Figure 8.15: Quantum Zeno effect.

Part (a) plots the scaled quantity Ω^2/Γ_\uparrow . For a given value of Ω , this quantity increases with increasing \bar{n} , showing that Γ_\uparrow decreases with \bar{n} as expected for the quantum Zeno effect. The use of the scaling factor Ω^2 lets us readily compare the values of Γ_\uparrow , which vary by more than two orders of magnitude in the plotted data. We expect Ω^2/Γ_\uparrow to be independent of Ω , and this appears to be more or less the case. Part (b) shows plots the scaled quantity $\Omega/(\Gamma_\uparrow + \Gamma_\downarrow)$ for same data set. Since these rates now include the effects of T_1 decay, we choose the scaling factor Ω (rather than Ω^2) as suggested by numerical simulations of the Zeno effect including decoherence effects given in [91]. Again, for a given Ω the rate $\Gamma_{\text{jump}} = \Gamma_\uparrow + \Gamma_\downarrow$ decreases for increasing \bar{n} , showing a slowing of the qubit transition rates with increasing measurement strength.

changes state. Based on the simple theory we expect to see a linear dependence on Γ_m and thus on \bar{n} , but we instead find a sublinear trend. In addition, the behavior appears to saturate at very high measurement strengths.

Part (b) shows the same data set, but now plots $\Gamma_{\text{jump}} = \Gamma_{\uparrow} + \Gamma_{\downarrow}$. We scale the vertical axis as $\Omega/(\Gamma_{\uparrow} + \Gamma_{\downarrow})$, based on the roughly linear (and not quadratic) dependence of Γ_{jump} on Ω observed in the full numerical simulations in ref. [91], which account for decay processes as well as the applied qubit drive. Again, we see that the scaled quantity $\Omega/(\Gamma_{\uparrow} + \Gamma_{\downarrow})$ appears to be essentially independent of Ω , but increases with \bar{n} , showing that increased measurement strength slows the qubit evolution as expected from the Zeno effect.

Our data clearly show that continuous measurement inhibits qubit state evolution, demonstrating the quantum Zeno effect. Because the Zeno effect has no classical analogue, this observation also serves as further confirmation of the quantum mechanical nature of superconducting qubits.

Chapter 9

Conclusions and outlook

9.1 Future work

The ability to perform continuous qubit state monitoring should enable us to implement quantum feedback and error correction and to create heralded single microwave photon sources and detectors. These efforts may also be furthered by improving the performance of the qubit and the paramp in subsequent designs.

9.1.1 Quantum feedback and error correction

One potential application of the qubit measurement technology demonstrated in this thesis is to quantum feedback and error correction. All quantum systems eventually decohere, albeit at widely varying rates. If we seek to combine these in quantum computers or quantum simulators where the quantum state of each constituent element cannot be allowed to decohere, we quickly run into difficulties.

One method for addressing the problem of decoherence is quantum error correction, where a decoherence-free logical qubit is constructed from an assembly of imperfect physical qubits [1, 204, 205]. The logical quantum state is stored in a subspace of the multi-qubit Hilbert space which is protected against decoherence. By making appropriately designed measurements of the multi-qubit state, one can check for and fix state errors in any of the physical qubits without actually measuring (and thus affecting) the logical quantum state. As long as the error rate for each physical qubit is not too high, this technique could in principle be used to store a logical quantum state indefinitely.

In a simple example, one can entangle three qubits, with two of the qubits serving as ancillae, and correct for a phase flip or bit flip error in any of the qubits by performing state-conditional gate operations. At the end, the error has been moved to the states of the ancilla qubits and the primary qubit recovers to its initially prepared state. One can then reset the ancillae, re-entangle them with the primary qubit, and repeat. This error-correcting three-qubit gate, known as the Toffoli gate, has recently been implemented with trapped ions [206] and superconducting qubits [24, 85, 86].

One could also perform a similar three-qubit error correction protocol based on measurement, where several fast, high-fidelity projective QND measurements of the pairwise

parity of the three-qubit state are performed (these measurements do not affect the encoded logical qubit state). The resulting information is used to send a sequence of appropriate pulses to correct any errors in the physical qubits. We could potentially implement the necessary fast, high-fidelity, QND parity measurements using a scheme like the one described in this thesis.

Our fast, high-fidelity measurement capabilities could also be used for quantum feedback, performing weak measurements on the quantum system and using the resulting information to stabilize a desired quantum state. In one such proposal, periodic weak measurement of the state of a microwave photon field in a Fabry-Perot cavity is used to feed back and stabilize Fock states of the field [207]. This proposal was recently demonstrated in the one of the first experiments showing stabilization of a quantum state through feedback [208]. Other proposals have shown the possibility of using continuous weak measurement and feedback to stabilize different states on the Bloch sphere [209]. Because the circuit QED/LJPA combination detailed in this thesis has excellent noise performance, particularly in the small-signal regime, it should be possible to implement this type of state stabilization feedback in a superconducting qubit.

One might also use continuous weak measurement techniques to stabilize qubit dynamics, for example Rabi oscillations. A continuous weak measurement on a qubit undergoing Rabi oscillations can be used to feed back and lock the Rabi oscillations to a desired frequency [210]. This could allow Rabi oscillations of a qubit to persist for times much longer than those set by the intrinsic dephasing of the qubit.

9.1.2 Single microwave photon source and detector

The qubit/cavity system with following paramp can also function as a heralded single photon source and/or detector. The decay of the qubit corresponds to the emission of a photon at the qubit frequency, which could potentially be directed into a particular channel and used in quantum optics experiments. The timing of the photon emission is accessible to the experimenter because we observe a quantum jump at that point. The time of absorption of a single resonant microwave photon by a qubit could be determined in a similar manner. Such heralded single-photon emitters and detectors could be used to demonstrate quantum effects such as antibunching of microwave photons [211].

9.1.3 Improved qubits

Great strides have been made recently in the effort to improve the coherence times of superconducting qubits, with the 3D transmon design demonstrating coherence times reliably longer than $20\ \mu s$ [17]. Such long-lived qubits would be useful for extensions of the quantum jump experiments presented in this thesis for a variety of reasons. First, the longer coherence time would allow for improved single-shot fidelity. Secondly, the longer coherence times would allow more subtle measurement backaction effects to become apparent. We could harness this to make further measurements of flux noise and explore the behavior of flux noise at high frequencies for a variety of qubit parameters and designs.

In addition, the ability to make single-junction versus SQUID-loop transmon qubits in the 3D transmon configuration could potentially be used to reduce the effects

of spurious qubit state mixing by preventing flux noise from causing noise on the qubit frequency. This could point the way to improving the QND nature of circuit QED readout at high readout power.

Although the circuit QED techniques used in this thesis were applied to the readout of a superconducting qubit, circuit QED coupling and readout have also been successfully used with other solid state quantum systems such as lateral quantum dots and electron spin ensembles [212, 213, 214]. The addition of a following paramp, as described in this thesis, could enable fast, high-fidelity QND state monitoring for these and other solid state quantum systems coupled to a superconducting readout cavity.

9.1.4 Parametric amplifier development

The LJPA used to amplify the quantum jump signal has excellent gain and noise properties but suffers from limited dynamic range. Improvements in the dynamic range of the amplifier could pave the way for use in a much broader range of quantum information experiments. One potential method for improving dynamic range would be to use arrays of SQUIDs rather than just a single SQUID; the array would have the same inductance as our current single-SQUID design but would have a considerably higher critical current, which in turn would increase the pump power and the dynamic range. However, there are potential difficulties with such an arrangement because the array would have multiple normal modes and might not be stable to phase slips when driven into the paramp regime.

Another possibility for improving dynamic range would be to use nulling feedback to keep the pump power in the linear regime of the transfer function. Such feedback would work in a similar way to the classic flux-locked loop used to extend the dynamic range of the dc SQUID amplifier. The tradeoff in such a system would be reduced amplifier bandwidth, but this issue could potentially be ameliorated by the use of fast electronics or cryogenic components to make the feedback loop. Modern dc SQUID amplifiers can operate in flux-locked loop configurations with 20 MHz of bandwidth using fast feedback electronics [215].

A different approach to the dynamic range and bandwidth problem would be to develop a new type of superconducting parametric amplifier. One potential design, called the traveling wave parametric amplifier or TWPA, uses a nonlinear transmission line loaded with Josephson junctions to create a transmission-mode microwave amplifier analogous to an optical fiber parametric amplifier. Because the TWPA does not have resonant structures, it should be inherently broadband and tunable over a wide frequency range simply by changing the pump frequency. Initial theoretical calculations suggest that the TWPA could achieve close to 1 GHz of instantaneous bandwidth with a 6 GHz pump and gains on the order of 15-20 dB. The dynamic range of the TWPA should be considerably larger than that of the LJPA, with saturation occurring at powers at least 20 dB higher than for the LJPA. These features might allow for the frequency multiplexing of a number of different qubit readout signals on a single amplifier channel, aiding in the scalability of quantum computing architectures based on superconducting qubits read out with parametric amplifiers. In addition, the fact that the TWPA is a transmission amplifier, and that the pump can be very far detuned from the amplified signals, may reduce the need for bulky, expensive microwave circulators between the readout cavity and the amplifier.

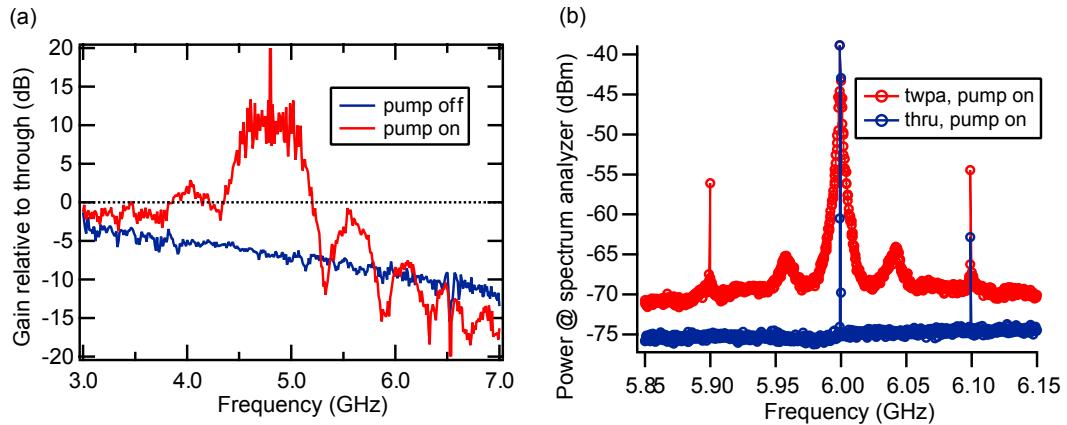


Figure 9.1: Preliminary TWPA performance.

Part (a) shows the gain of the TWPA relative to a microwave through with the pump on and off. The pump frequency is 4.8 GHz. Part (b) shows the output spectrum of a TWPA and a microwave through with a pump at 6 GHz and small signal detuned by 100 MHz. The TWPA shows both signal and idler gain, with a signal gain of 9 dB and an output signal-to-noise improvement of 4 dB relative to the through.

Some preliminary gain and noise data from a TWPA prototype are shown in Figure 9.1. The TWPA has about 10 dB of gain relative to a microwave through, with more than 500 MHz of instantaneous 3 dB bandwidth. The noise performance is analyzed in part (b); the TWPA shows both direct and trans gain and gives a signal-to-noise ratio improvement of about 4 dB over a microwave through. The TWPA fabrication process is complex, and is currently under development.

9.2 Conclusions

This thesis showed the first observation of quantum jumps in a superconducting qubit, a macroscopic quantum system. The work described harnessed the combination of projective QND readout offered by the circuit QED architecture and fast, quantum-limited amplification enabled by a superconducting parametric amplifier, the LJPA. The resulting observation of quantum jumps brings quantum measurement capabilities to solid state quantum systems which had previously been the exclusive domain of atomic and optical physics. The ability to perform continuous high-fidelity qubit monitoring points the way to quantum error correction and feedback in solid state systems.

The continuous high-fidelity monitoring capabilities demonstrated here allow the quantification of measurement backaction in superconducting qubits. They show the perhaps surprising result that low-frequency noise is the culprit in measurement-induced qubit state mixing at high readout powers in circuit QED, a problem that has limited the fidelity of circuit QED readout. This points the way for to the unification of several disparate

research directions in the superconducting qubit community, namely the quest to maximize readout fidelity and the quest to understand and reduce the sources of universal flux noise. We also observe the quantum Zeno effect.

One way forward is to use the techniques described in this thesis to enable quantum error correction. This pursuit is one of the major goals in contemporary experimental quantum information research. The successful implementation of quantum error correction to create long-lived logical qubits would be an important breakthrough in the quest to build quantum computers and quantum simulators. We hope that some of the work in this thesis can contribute to this and other research efforts in the future, helping to advance our capabilities and knowledge as we explore the fascinating quantum world.

Bibliography

- [1] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press (2000).
- [2] R. P. Feynman. “Simulating physics with computers.” *Int. J. Theor. Phys.* **21**, 467 (1982).
- [3] W. H. Zurek. “Decoherence, einselection, and the quantum origins of the classical.” *Rev. Mod. Phys.* **75**, 715 (2003).
- [4] I. Siddiqi. “Superconducting qubits: poised for computing?” *Supercond. Sci. Tech.* **24**, 091002 (2011).
- [5] J. Clarke and F. K. Wilhelm. “Superconducting quantum bits.” *Nature* **453**, 1031 (2008).
- [6] A. J. Leggett. “Macroscopic Quantum Systems and the Quantum Theory of Measurement.” *Prog. Theor. Phys. Suppl.* **69**, 80 (1980).
- [7] J. M. Martinis, M. H. Devoret, and J. Clarke. “Experimental tests for the quantum behavior of a macroscopic degree of freedom: The phase difference across a Josephson junction.” *Phys. Rev. B* **35**, 4682 (1987).
- [8] Y. Nakamura, Y. A. Pashkin, and J. S. Tsai. “Coherent control of macroscopic quantum states in a single-Cooper-pair box.” *Nature* **398**, 786 (1999).
- [9] V. Bouchiat, D. Vion, P. Joyez, D. Esteve, and M. H. Devoret. “Quantum Coherence with a Single Cooper Pair.” *Phys. Scr.* **T76**, 165 (1998).
- [10] J. M. Martinis, S. Nam, J. Aumentado, and C. Urbina. “Rabi Oscillations in a Large Josephson-Junction Qubit.” *Phys. Rev. Lett.* **89**, 117901 (2002).
- [11] I. Chiorescu, Y. Nakamura, C. J. P. M. Harmans, and J. E. Mooij. “Coherent quantum dynamics of a superconducting flux qubit.” *Science* **299**, 1869 (2003).
- [12] 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 (2002).

- [13] 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.” *Phys. Rev. A* **76**, 042319 (2007).
- [14] V. E. Manucharyan, J. Koch, L. I. Glazman, and M. H. Devoret. “Fluxonium: single cooper-pair circuit free of charge offsets.” *Science* **326**, 113 (2009).
- [15] M. Steffen, S. Kumar, D. P. DiVincenzo, J. R. Rozen, G. A. Keefe, M. B. Rothwell, and M. B. Ketchen. “High-Coherence Hybrid Superconducting Qubit.” *Phys. Rev. Lett.* **105**, 100502 (2010).
- [16] S. J. Srinivasan, A. J. Hoffman, J. M. Gambetta, and A. A. Houck. “Tunable Coupling in Circuit Quantum Electrodynamics Using a Superconducting Charge Qubit with a V-Shaped Energy Level Diagram.” *Phys. Rev. Lett.* **106**, 083601 (2011).
- [17] H. Paik, D. I. Schuster, L. S. Bishop, G. Kirchmair, G. Catelani, A. P. Sears, B. R. Johnson, M. J. Reagor, L. Frunzio, L. 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.” *ArXiv* 1105.4652v3 (2011).
- [18] M. Steffen, M. Ansmann, R. C. Bialczak, N. Katz, E. Lucero, 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 (2006).
- [19] 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 (2009).
- [20] M. Neeley, R. C. Bialczak, M. Lenander, E. Lucero, M. Mariantoni, A. D. O’Connell, D. Sank, H. Wang, M. Weides, J. Wenner, Y. Yin, T. Yamamoto, A. N. Cleland, and J. M. Martinis. “Generation of three-qubit entangled states using superconducting phase qubits.” *Nature* **467**, 570 (2010).
- [21] M. Ansmann, H. Wang, R. C. Bialczak, M. Hofheinz, E. Lucero, M. Neeley, A. D. O’Connell, D. Sank, M. Weides, J. Wenner, A. N. Cleland, and J. M. Martinis. “Violation of Bell’s inequality in Josephson phase qubits.” *Nature* **461**, 504 (2009).
- [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**, 574 (2010).
- [23] A. Palacios-Laloy, F. Mallet, F. Nguyen, P. Bertet, D. Vion, D. Esteve, and A. N. Korotkov. “Experimental violation of a Bells inequality in time with weak measurement.” *Nature Phys.* **6**, 442 (2010).
- [24] M. Mariantoni, H. Wang, T. Yamamoto, M. Neeley, R. C. Bialczak, Y. Chen, M. Lenander, E. Lucero, A. D. O’Connell, D. Sank, M. Weides, J. Wenner, Y. Yin,

- J. Zhao, A. N. Korotkov, A. N. Cleland, and J. M. Martinis. “Implementing the Quantum von Neumann Architecture with Superconducting Circuits.” *Science* **334**, 61 (2011).
- [25] A. Dewes, R. Lauro, F. R. Ong, V. Schmitt, P. Milman, P. Bertet, D. Vion, and D. Esteve. “Demonstrating quantum speed-up in a superconducting two-qubit processor.” *ArXiv* 1110.5170 (2011).
- [26] N. Bohr. “I. On the constitution of atoms and molecules.” *Philos. Mag.* **26**, 1 (1913).
- [27] W. Neuhauser, M. Hohenstatt, P. E. Toschek, and H. Dehmelt. “Localized visible Ba+ mono-ion oscillator.” *Phys. Rev. A* **22**, 1137 (1980).
- [28] D. J. Wineland and W. M. Itano. “Spectroscopy of a Single Mg+ Ion.” *Phys. Lett. A* **82**, 75 (1981).
- [29] W. Nagourney, J. Sandberg, and H. Dehmelt. “Shelved optical electron amplifier: observation of quantum jumps.” *Phys. Rev. Lett.* **56**, 2797 (1986).
- [30] T. Sauter, W. Neuhauser, R. Blatt, and P. E. Toschek. “Observation of quantum jumps.” *Phys. Rev. Lett.* **57**, 1696 (1986).
- [31] J. C. Bergquist, R. G. Hulet, W. M. Itano, and D. J. Wineland. “Observation of quantum jumps in a single atom.” *Phys. Rev. Lett.* **57**, 1699 (1986).
- [32] T. Basché, S. Kummer, and C. Bräuchle. “Direct spectroscopic observation of quantum jumps of a single molecule.” *Nature* **373**, 132 (1995).
- [33] S. Peil and G. Gabrielse. “Observing the Quantum Limit of an Electron Cyclotron: QND Measurements of Quantum Jumps between Fock States.” *Phys. Rev. Lett.* **83**, 1287 (1999).
- [34] S. Gleyzes, S. Kuhr, C. Guerlin, J. Bernu, S. Deléglise, U. B. Hoff, M. Brune, J.-M. Raimond, and S. Haroche. “Quantum jumps of light recording the birth and death of a photon in a cavity.” *Nature* **446**, 297 (2007).
- [35] Y. Yu, S.-L. Zhu, G. Sun, X. Wen, N. Dong, J. Chen, P. Wu, and S. Han. “Quantum Jumps between Macroscopic Quantum States of a Superconducting Qubit Coupled to a Microscopic Two-Level System.” *Phys. Rev. Lett.* **101**, 157001 (2008).
- [36] P. Neumann, J. Beck, M. Steiner, F. Rempp, H. Fedder, P. R. Hemmer, J. Wrachtrup, and F. Jelezko. “Single-shot readout of a single nuclear spin.” *Science* **329**, 542 (2010).
- [37] A. N. Vamivakas, C.-Y. Lu, C. Matthiesen, Y. Zhao, S. Fält, A. Badolato, and M. Atatüre. “Observation of spin-dependent quantum jumps via quantum dot resonance fluorescence.” *Nature* **467**, 297 (2010).
- [38] S. Ulmer, C. C. Rodegheri, K. Blaum, H. Kracke, A. Mooser, W. Quint, and J. Walz. “Observation of Spin Flips with a Single Trapped Proton.” *Phys. Rev. Lett.* **106**, 253001 (2011).

- [39] V. B. Braginsky and F. Y. Khalili. *Quantum Measurement*. New York: Cambridge University Press (1992).
- [40] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf. “Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation.” *Phys. Rev. A* **69**, 062320 (2004).
- [41] 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 (2004).
- [42] 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.” *Phys. Rev. Lett.* **94**, 123602 (2005).
- [43] A. A. Houck, J. A. Schreier, B. R. Johnson, J. M. Chow, J. Koch, J. M. Gambetta, D. I. Schuster, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. “Controlling the Spontaneous Emission of a Superconducting Transmon Qubit.” *Phys. Rev. Lett.* **101**, 080502 (2008).
- [44] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. “Approaching Unit Visibility for Control of a Superconducting Qubit with Dispersive Readout.” *Phys. Rev. Lett.* **95**, 060501 (2005).
- [45] M. Boissonneault, J. M. Gambetta, and A. Blais. “Dispersive regime of circuit QED: Photon-dependent qubit dephasing and relaxation rates.” *Phys. Rev. A* **79**, 013819 (2009).
- [46] R. Vijay, M. H. Devoret, and I. Siddiqi. “Invited review article: The Josephson bifurcation amplifier.” *Rev. Sci. Instrum.* **80**, 111101 (2009).
- [47] F. Mallet, F. R. Ong, A. Palacios-Laloy, F. Nguyen, P. Bertet, D. Vion, and D. Esteve. “Single-shot qubit readout in circuit quantum electrodynamics.” *Nature Phys.* **5**, 791 (2009).
- [48] J. M. Martinis. “Superconducting phase qubits.” *Quant. Inf. Proc.* **8**, 81 (2009).
- [49] T. Picot, R. Schouten, C. J. P. M. Harmans, and J. E. Mooij. “Quantum Nondemolition Measurement of a Superconducting Qubit in the Weakly Projective Regime.” *Phys. Rev. Lett.* **105**, 040506 (2010).
- [50] A. A. Clerk, M. H. Devoret, S. M. Girvin, F. Marquardt, and R. J. Schoelkopf. “Introduction to quantum noise, measurement, and amplification.” *Rev. Mod. Phys.* **82**, 1155 (2010).
- [51] M. J. Feldman, P. T. Parrish, and R. Y. Chiao. “Parametric amplification by unbiased Josephson junctions.” *J. Appl. Phys.* **46**, 4031 (1975).
- [52] A. H. Silver, D. C. Pridmore-Brown, R. D. Sandell, and J. P. Hurrell. “Parametric properties of SQUID lattice arrays.” *IEEE Trans. Magn.* **17**, 412 (1981).

- [53] N. Calander, T. Claeson, and S. Rudner. “Shunted Josephson tunnel junctions: High-frequency, self-pumped low noise amplifiers.” *J. Appl. Phys.* **53**, 5093 (1982).
- [54] A. D. Smith, R. D. Sandell, J. F. Burch, and A. H. Silver. “Low noise microwave parametric amplifier.” *IEEE Trans. Magn.* **21**, 1022 (1985).
- [55] B. Yurke, L. R. Corruccini, P. G. Kaminsky, L. W. Rupp, A. D. Smith, A. H. Silver, R. W. Simon, and E. A. Whittaker. “Observation of parametric amplification and deamplification in a Josephson parametric amplifier.” *Phys. Rev. A* **39**, 2519 (1989).
- [56] B. Yurke, M. L. Roukes, R. Movshovich, and A. N. Pargellis. “A low-noise series-array Josephson junction parametric amplifier.” *Appl. Phys. Lett.* **69**, 3078 (1996).
- [57] M. A. Castellanos-Beltran, K. D. Irwin, G. C. Hilton, L. R. Vale, and K. W. Lehnert. “Amplification and squeezing of quantum noise with a tunable Josephson metamaterial.” *Nature Phys.* **4**, 929 (2008).
- [58] N. Bergeal, F. Schackert, M. Metcalfe, R. Vijay, V. E. Manucharyan, L. Frunzio, D. E. Prober, R. J. Schoelkopf, S. M. Girvin, and M. H. Devoret. “Phase-preserving amplification near the quantum limit with a Josephson ring modulator.” *Nature* **465**, 64 (2010).
- [59] E. M. Levenson-Falk, R. Vijay, and I. Siddiqi. “Nonlinear microwave response of aluminum weak-link Josephson oscillators.” *Appl. Phys. Lett.* **98**, 123115 (2011).
- [60] B. Abdo, F. Schackert, M. Hatridge, C. Rigetti, and M. Devoret. “Josephson amplifier for qubit readout.” *Appl. Phys. Lett.* **99**, 162506 (2011).
- [61] M. Hatridge, R. Vijay, D. H. Slichter, J. Clarke, and I. Siddiqi. “Dispersive magnetometry with a quantum limited SQUID parametric amplifier.” *Phys. Rev. B* **83**, 134501 (2011).
- [62] R. Vijay, D. H. Slichter, and I. Siddiqi. “Observation of Quantum Jumps in a Superconducting Artificial Atom.” *Phys. Rev. Lett.* **106**, 110502 (2011).
- [63] J. Q. You and F. Nori. “Atomic physics and quantum optics using superconducting circuits.” *Nature* **474**, 589 (2011).
- [64] D. I. Schuster. *Circuit Quantum Electrodynamics*. Ph.D. thesis, Yale University (2007).
- [65] P. W. Shor. “Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer.” *SIAM Rev.* **41**, 303 (1999).
- [66] L. Grover. “Quantum Mechanics Helps in Searching for a Needle in a Haystack.” *Phys. Rev. Lett.* **79**, 325 (1997).
- [67] B. Yurke and J. Denker. “Quantum network theory.” *Phys. Rev. A* **29**, 1419 (1984).

- [68] M. H. Devoret. “Quantum Fluctuations in Electrical Circuits.” In *Quantum Fluctuations, Les Houches LXIII*, S. Reynaud, E. Giacobino, and J. Zinn-Justin, eds., chap. 10. New York: Elsevier (1995).
- [69] B. Yurke. “Input-Output Theory.” In *Quantum Squeezing*, P. D. Drummond and Z. Ficek, eds., chap. 3. New York: Springer (2004).
- [70] D. F. Walls and G. J. Milburn. *Quantum Optics*. New York: Springer, 2nd ed. (2008).
- [71] M. H. Devoret, A. Wallraff, and J. M. Martinis. “Superconducting qubits: A short review.” *ArXiv cond-mat/0411174* (2004).
- [72] M. Tinkham. *Introduction to Superconductivity*. Mineola, NY: Dover, 2nd ed. (1996).
- [73] J. Clarke and A. I. Braginsky (Eds.). *The SQUID Handbook: Fundamentals and Technology of SQUIDs and SQUID Systems, Vol. I*. Weinheim, Germany: Wiley-VCH Verlag (2004).
- [74] A. Cottet. *Implementation of quantum bit in a superconducting circuit*. Ph.D. thesis, Université Paris VI (Pierre et Marie Curie) (2002).
- [75] Z. Kim, B. Suri, V. Zaretsky, S. Novikov, K. D. Osborn, A. Mizel, F. C. Wellstood, and B. S. Palmer. “Decoupling a Cooper-Pair Box to Enhance the Lifetime to 0.2 ms.” *Phys. Rev. Lett.* **106**, 120501 (2011).
- [76] J. M. Chow, L. DiCarlo, J. M. Gambetta, F. Motzoi, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf. “Optimized driving of superconducting artificial atoms for improved single-qubit gates.” *Phys. Rev. A* **82**, 040305 (2010).
- [77] S. Haroche and J.-M. Raimond. *Exploring the quantum: atoms, cavities, and photons*. Oxford: Oxford University Press (2006).
- [78] C. Monroe. “Demolishing quantum nondemolition.” *Physics Today* **64**, 8 (2011).
- [79] J. Majer, J. M. Chow, J. M. Gambetta, J. Koch, B. R. Johnson, J. A. Schreier, L. Frunzio, D. I. Schuster, A. A. Houck, A. Wallraff, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. “Coupling superconducting qubits via a cavity bus.” *Nature* **449**, 443 (2007).
- [80] M. A. Sillanpää, J. I. Park, and R. W. Simmonds. “Coherent quantum state storage and transfer between two phase qubits via a resonant cavity.” *Nature* **449**, 438 (2007).
- [81] M. Mariantoni, H. Wang, R. C. Bialczak, M. Lenander, E. Lucero, M. Neeley, A. D. O’Connell, D. Sank, M. Weides, J. Wenner, T. Yamamoto, Y. Yin, J. Zhao, J. M. Martinis, and A. N. Cleland. “Photon shell game in three-resonator circuit quantum electrodynamics.” *Nature Phys.* **7**, 287 (2011).
- [82] A. A. Houck, D. I. Schuster, J. M. Gambetta, J. A. Schreier, B. R. Johnson, J. M. Chow, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. “Generating single microwave photons in a circuit.” *Nature* **449**, 328 (2007).

- [83] M. Hofheinz, H. Wang, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, a. D. O'Connell, D. Sank, J. Wenner, J. M. Martinis, and a. N. Cleland. "Synthesizing arbitrary quantum states in a superconducting resonator." *Nature* **459**, 546 (2009).
- [84] B. R. Johnson, M. D. Reed, A. A. Houck, D. I. Schuster, L. S. Bishop, E. Ginossar, J. M. Gambetta, L. DiCarlo, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf. "Quantum non-demolition detection of single microwave photons in a circuit." *Nature Phys.* **6**, 663 (2010).
- [85] 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." *ArXiv* 1109.4948 (2011).
- [86] A. Fedorov, L. Steffen, M. Baur, and A. Wallraff. "Implementation of a Toffoli Gate with Superconducting Circuits." *ArXiv* 1108.3966 (2011).
- [87] L. S. Bishop. *Circuit Quantum Electrodynamics II*. Ph.D. thesis, Yale University (2010).
- [88] B. R. Johnson. *Controlling Photons in Superconducting Electrical Circuits*. Ph.D. thesis, Yale University (2010).
- [89] J. M. Chow. *Quantum Information Processing with Superconducting Qubits*. Ph.D. thesis, Yale University (2010).
- [90] D. Leibfried, R. Blatt, C. Monroe, and D. Wineland. "Quantum dynamics of single trapped ions." *Rev. Mod. Phys.* **75**, 281 (2003).
- [91] J. Gambetta, A. Blais, M. Boissonneault, A. A. Houck, D. I. Schuster, and S. M. Girvin. "Quantum trajectory approach to circuit QED: Quantum jumps and the Zeno effect." *Phys. Rev. A* **77**, 012112 (2008).
- [92] 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." *Phys. Rev. A* **74**, 042318 (2006).
- [93] M. Boissonneault, J. M. Gambetta, and A. Blais. "Improved Superconducting Qubit Readout by Qubit-Induced Nonlinearities." *Phys. Rev. Lett.* **105**, 100504 (2010).
- [94] J. Gambetta, W. A. Bratt, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf. "Protocols for optimal readout of qubits using a continuous quantum nondemolition measurement." *Phys. Rev. A* **76**, 012325 (2007).
- [95] C. D. Motchenbacher and J. A. Connell. *Low-Noise Electronic System Design*. New York: John Wiley (1993).
- [96] C. M. Caves. "Quantum limits on noise in linear amplifiers." *Phys. Rev. D* **26**, 1817 (1982).

- [97] J. B. Johnson. “Thermal Agitation of Electricity in Conductors.” *Phys* **32**, 97 (1928).
- [98] H. Nyquist. “Thermal Agitation of Electric Charge in Conductors.” *Phys. Rev.* **32**, 110 (1928).
- [99] H. B. Callen and T. A. Welton. “Irreversibility and Generalized Noise.” *Phys. Rev.* **83**, 34 (1951).
- [100] H. A. Haus and J. A. Mullen. “Quantum Noise in Linear Amplifiers.” *Phys. Rev.* **128**, 2407 (1962).
- [101] W. H. Louisell. *Coupled Mode and Parametric Electronics*. New York: John Wiley (1960).
- [102] D. P. Howson and R. B. Smith. *Parametric Amplifiers*. Maidenhead, England: McGraw-Hill (1970).
- [103] W. W. Mumford. “Some Notes on the History of Parametric Transducers.” *Proc. IRE* **48**, 848 (1960).
- [104] E. F. W. Alexanderson and S. P. Nixdorff. “A Magnetic Amplifier for Radio Telephony.” *Proc. IRE* **4**, 101 (1916).
- [105] L. Sivan. *Microwave tube transmitters*. New York: Chapman & Hall (1994).
- [106] M. E. Marhic. *Fiber Optical Parametric Amplifiers, Oscillators, and Related Devices*. New York: Cambridge University Press (2008).
- [107] T. Yamamoto, K. Inomata, M. Watanabe, K. Matsuba, T. Miyazaki, W. D. Oliver, Y. Nakamura, and J. S. Tsai. “Flux-driven Josephson parametric amplifier.” *Appl. Phys. Lett.* **93**, 042510 (2008).
- [108] R. Vijay. *Josephson Bifurcation Amplifier: Amplifying quantum signals using a dynamical bifurcation*. Ph.D. thesis, Yale University (2008).
- [109] M. A. Castellanos-Beltran. *Development of a Josephson Parametric Amplifier for the Preparation and Detection of Nonclassical States of Microwave Fields*. Ph.D. thesis, University of Colorado Boulder (2010).
- [110] N. Bergeal, R. Vijay, V. E. Manucharyan, I. Siddiqi, R. J. Schoelkopf, S. M. Girvin, and M. H. Devoret. “Analog information processing at the quantum limit with a Josephson ring modulator.” *Nature Phys.* **6**, 296 (2010).
- [111] F. Mallet, M. Castellanos-Beltran, H. Ku, S. Glancy, E. Knill, K. Irwin, G. Hilton, L. Vale, and K. Lehnert. “Quantum State Tomography of an Itinerant Squeezed Microwave Field.” *Phys. Rev. Lett.* **106**, 220502 (2011).
- [112] M. I. Dykman and M. A. Krivoglaz. “Fluctuations in nonlinear systems near bifurcations corresponding to the appearance of new stable states.” *Physica A* **104**, 480 (1980).

- [113] B. Cord, J. Yang, H. Duan, D. C. Joy, J. Klingfus, and K. K. Berggren. “Limiting factors in sub-10nm scanning-electron-beam lithography.” *J. Vac. Sci. Technol. B* **27**, 2616 (2009).
- [114] M. J. Rooks, E. Kratschmer, R. Viswanathan, J. Katine, R. E. Fontana, and S. A. MacDonald. “Low stress development of poly(methylmethacrylate) for high aspect ratio structures.” *J. Vac. Sci. Technol. B* **20**, 2937 (2002).
- [115] B. Cord, C. Dames, K. K. Berggren, and J. Aumentado. “Robust shadow-mask evaporation via lithographically controlled undercut.” *J. Vac. Sci. Technol. B* **24**, 3139 (2006).
- [116] R. W. Johnstone, I. G. Foulds, M. V. Pallapa, and A. M. Parameswaran. “Isopropanol/water as a developer for poly(dimethylglutarimide).” *J. Micro/Nanolithogr. MEMS* **7**, 043006 (2008).
- [117] K. Koshelev, M. A. Mohammad, T. Fito, K. L. Westra, S. K. Dew, and M. Stepanova. “Comparison between ZEP and PMMA resists for nanoscale electron beam lithography experimentally and by numerical modeling.” *J. Vac. Sci. Technol. B* **29**, 06F306 (2011).
- [118] F. J. Pantenburg. “Influence of developer temperature and resist material on the structure quality in deep x-ray lithography.” *J. Vac. Sci. Technol. B* **16**, 3547 (1998).
- [119] W. W. Hu, K. Sarveswaran, M. Lieberman, and G. H. Bernstein. “Sub-10 nm electron beam lithography using cold development of poly(methylmethacrylate).” *J. Vac. Sci. Technol. B* **22**, 1711 (2004).
- [120] B. Cord, J. Lutkenhaus, and K. K. Berggren. “Optimal temperature for development of poly(methylmethacrylate).” *J. Vac. Sci. Technol. B* **25**, 2013 (2007).
- [121] J. S. Greeneich. “Solubility Rate of Poly-(Methyl Methacrylate), PMMA, Electron-Resist.” *J. Electrochem. Soc.* **121**, 1669 (1974).
- [122] J. S. Greeneich. “Developer Characteristics of Poly-(Methyl Methacrylate) Electron Resist.” *J. Electrochem. Soc.* **122**, 970 (1975).
- [123] M. Yan, S. Choi, K. R. V. Subramanian, and I. Adesida. “The effects of molecular weight on the exposure characteristics of poly(methylmethacrylate) developed at low temperatures.” *J. Vac. Sci. Technol. B* **26**, 2306 (2008).
- [124] L. E. Ocola and A. Stein. “Effect of cold development on improvement in electron-beam nanopatterning resolution and line roughness.” *J. Vac. Sci. Technol. B* **24**, 3061 (2006).
- [125] R. Vijay, E. M. Levenson-Falk, D. H. Slichter, and I. Siddiqi. “Approaching ideal weak link behavior with three dimensional aluminum nanobridges.” *Appl. Phys. Lett.* **96**, 223112 (2010).

- [126] H. Wang, G. M. Laws, S. Milicic, P. Boland, A. Handugan, M. Pratt, T. Eschrich, S. Myhajlenko, J. A. Allgair, and B. Bunday. “Low temperature ZEP-520A development process for enhanced critical dimension realization in reactive ion etch etched polysilicon.” *J. Vac. Sci. Technol. B* **25**, 102 (2007).
- [127] G. J. Dolan. “Offset masks for lift-off photoprocessing.” *Appl. Phys. Lett.* **31**, 337 (1977).
- [128] J. Mooij, T. Orlando, L. Levitov, L. Tian, C. H. van der Wal, and S. Lloyd. “Josephson persistent-current qubit.” *Science* **285**, 1036 (1999).
- [129] P. J. Koppinen, L. M. Vaisto, and I. J. Maasilta. “Complete stabilization and improvement of the characteristics of tunnel junctions by thermal annealing.” *Appl. Phys. Lett.* **90**, 053503 (2007).
- [130] I. M. Pop, T. Fournier, T. Crozes, F. Lecocq, I. Matei, B. Pannetier, O. Buisson, and W. Guichard. “Fabrication of stable and reproducible sub-micron tunnel junctions.” *ArXiv* 1105.6204 (2011).
- [131] J. M. Martinis, K. B. Cooper, R. McDermott, M. Steffen, M. Ansmann, K. D. Osborn, K. Cicak, S. Oh, D. P. Pappas, R. W. Simmonds, and C. C. Yu. “Decoherence in Josephson Qubits from Dielectric Loss.” *Phys. Rev. Lett.* **95**, 210503 (2005).
- [132] A. D. OConnell, M. Ansmann, R. C. Bialczak, M. Hofheinz, N. Katz, E. Lucero, C. McKenney, M. Neeley, H. Wang, E. M. Weig, a. N. Cleland, and J. M. Martinis. “Microwave dielectric loss at single photon energies and millikelvin temperatures.” *Appl. Phys. Lett.* **92**, 112903 (2008).
- [133] H. Paik and K. D. Osborn. “Reducing quantum-regime dielectric loss of silicon nitride for superconducting quantum circuits.” *Appl. Phys. Lett.* **96**, 072505 (2010).
- [134] S. J. Weber, K. W. Murch, D. H. Slichter, R. Vijay, and I. Siddiqi. “Single crystal silicon capacitors with low microwave loss in the single photon regime.” *Appl. Phys. Lett.* **98**, 172510 (2011).
- [135] D. M. Pozar. *Microwave Engineering*. Wiley, 3rd ed. (2005).
- [136] M. J. Hatridge. *SQUID magnetometry from nanometer to centimeter length scales*. Ph.D. thesis, Univ. of California, Berkeley (2010).
- [137] R. Barends, J. Wenner, M. Lenander, Y. Chen, R. C. Bialczak, J. Kelly, E. Lucero, P. O’Malley, M. Mariantoni, D. Sank, H. Wang, T. C. White, Y. Yin, J. Zhao, A. N. Cleland, J. M. Martinis, and J. J. A. Baselmans. “Loss and decoherence due to stray infrared light in superconducting quantum circuits.” *ArXiv* 1105.4642 (2011).
- [138] A. D. Córcoles, J. M. Chow, J. M. Gambetta, C. Rigetti, J. R. Rozen, G. A. Keefe, M. B. Rothwell, M. B. Ketchen, and M. Steffen. “Protecting superconducting qubits from external sources of loss and heat.” *ArXiv* 1108.1383 (2011).

- [139] F. Pobell. *Matter and Methods at Low Temperatures*. Berlin: Springer, 3rd ed. (2007).
- [140] K. Bladh, D. Gunnarsson, E. Hurfeld, S. Devi, C. Kristoffersson, B. Smalander, S. Pehrson, T. Claeson, P. Delsing, and M. Taslakov. “Comparison of cryogenic filters for use in single electronics experiments.” *Rev. Sci. Instrum.* **74**, 1323 (2003).
- [141] P. Schiffres. “A Dissipative Coaxial RFI Filter.” *IEEE Trans. Electromagn. Compat.* **6**, 55 (1964).
- [142] H. W. Denny and W. B. Warren. “Lossy Transmission Line Filters.” *IEEE Trans. Electromagn. Compat.* **10**, 363 (1968).
- [143] A. B. Zorin. “The thermocoax cable as the microwave frequency filter for single electron circuits.” *Rev. Sci. Instrum.* **66**, 4296 (1995).
- [144] F. P. Milliken, J. R. Rozen, G. A. Keefe, and R. H. Koch. “50 Ω Characteristic Impedance Low-Pass Metal Powder Filters.” *Rev. Sci. Instrum.* **78**, 024701 (2007).
- [145] A. Lukashenko and A. V. Ustinov. “Improved powder filters for qubit measurements.” *Rev. Sci. Instrum.* **79**, 014701 (2008).
- [146] H. A. Wheeler. “Transmission-Line Properties of a Strip Line Between Parallel Planes.” *IEEE Trans. Microwave Theory Tech.* **26**, 866 (1978).
- [147] M. Metcalfe. *A new microwave resonator readout scheme for superconducting qubits*. Ph.D. thesis, Yale University (2008).
- [148] D. F. Santavicca and D. E. Prober. “Impedance-matched low-pass stripline filters.” *Meas. Sci. Technol.* **19**, 087001 (2008).
- [149] J. B. Peterson and P. L. Richards. “A cryogenic blackbody for millimeter wavelengths.” *Int. J. Infrared Millimeter Waves* **5**, 1507 (1984).
- [150] H. Hemmati, J. C. Mather, and W. L. Eichhorn. “Submillimeter and millimeter wave characterization of absorbing materials.” *Appl. Opt.* **24**, 4489 (1985).
- [151] M. Halpern, H. P. Gush, E. Wishnow, and V. De Cosmo. “Far infrared transmission of dielectrics at cryogenic and room temperatures: glass, Fluorogold, Eccosorb, Stycast, and various plastics.” *Appl. Opt.* **25**, 565 (1986).
- [152] H. P. Gush, M. Halpern, and E. H. Wishnow. “Rocket measurement of the cosmic-background-radiation mm-wave spectrum.” *Phys. Rev. Lett.* **65**, 537 (1990).
- [153] J. C. Mather, D. J. Fixsen, R. A. Shafer, C. Mosier, and D. T. Wilkinson. “Calibrator design for the COBE Far Infrared Absolute Spectrophotometer (FIRAS).” *Astrophys. J.* **512**, 511 (1999).
- [154] A. Kogut, E. Wollack, D. J. Fixsen, M. Limon, P. Mirel, S. Levin, M. Seiffert, and P. M. Lubin. “Design and calibration of a cryogenic blackbody calibrator at centimeter wavelengths.” *Rev. Sci. Instrum.* **75**, 5079 (2004).

- [155] D. H. Slichter, O. Naaman, and I. Siddiqi. “Millikelvin thermal and electrical performance of lossy transmission line filters.” *Appl. Phys. Lett.* **94**, 192508 (2009).
- [156] V. E. Manucharyan, E. Boaknin, M. Metcalfe, R. Vijay, I. Siddiqi, and M. Devoret. “Microwave bifurcation of a Josephson junction: Embedding-circuit requirements.” *Phys. Rev. B* **76**, 014524 (2007).
- [157] J. M. Fink, L. Steffen, P. Studer, L. S. Bishop, M. Baur, R. Bianchetti, D. Bozyigit, C. Lang, S. Filipp, P. J. Leek, and A. Wallraff. “Quantum-To-Classical Transition in Cavity Quantum Electrodynamics.” *Phys. Rev. Lett.* **105**, 163601 (2010).
- [158] M. D. Reed, L. DiCarlo, B. R. Johnson, L. Sun, D. I. Schuster, L. Frunzio, and R. J. Schoelkopf. “High-Fidelity Readout in Circuit Quantum Electrodynamics Using the Jaynes-Cummings Nonlinearity.” *Phys. Rev. Lett.* **105**, 173601 (2010).
- [159] 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).
- [160] J. Bylander, S. Gustavsson, F. Yan, F. Yoshihara, K. Harrabi, G. Fitch, D. G. Cory, Y. Nakamura, J.-S. Tsai, and W. D. Oliver. “Noise spectroscopy through dynamical decoupling with a superconducting flux qubit.” *Nature Phys.* **7**, 565 (2011).
- [161] M. Steffen, M. Ansmann, R. McDermott, N. Katz, R. C. Bialczak, E. Lucero, M. Neeley, E. M. Weig, A. N. Cleland, and J. M. Martinis. “State Tomography of Capacitively Shunted Phase Qubits with High Fidelity.” *Phys. Rev. Lett.* **97**, 050502 (2006).
- [162] M. Steffen, M. Ansmann, R. C. Bialczak, N. Katz, E. Lucero, 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 (2006).
- [163] R. J. Cook. “What are Quantum Jumps?” *Phys. Scr.* **T21**, 49 (1988).
- [164] E. Schrödinger. “Are there quantum jumps? Part I.” *Brit. J. Philos. Sci.* **III**, 109 (1952).
- [165] E. Schrödinger. “Are there quantum jumps? Part II.” *Brit. J. Philos. Sci.* **III**, 233 (1952).
- [166] J. M. Chow, L. DiCarlo, J. M. Gambetta, A. Nunnenkamp, L. S. Bishop, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. “Detecting highly entangled states with a joint qubit readout.” *Phys. Rev. A* **81**, 062325 (2010).
- [167] M. S. Khalil, F. C. Wellstood, and K. D. Osborn. “Loss Dependence on Geometry and Applied Power in Superconducting Coplanar Resonators.” *IEEE Trans. Appl. Supercond.* **21**, 879 (2011).
- [168] R. J. Cook and H. J. Kimble. “Possibility of direct observation of quantum jumps.” *Phys. Rev. Lett.* **54**, 1023 (1985).

- [169] B. Misra and E. C. G. Sudarshan. “The Zenos paradox in quantum theory.” *J. Math. Phys.* **18**, 756 (1977).
- [170] J. M. Chow, J. M. Gambetta, L. Tornberg, J. Koch, L. S. Bishop, A. A. Houck, B. R. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf. “Randomized Benchmarking and Process Tomography for Gate Errors in a Solid-State Qubit.” *Phys. Rev. Lett.* **102**, 090502 (2009).
- [171] WaveMetrics. “The Igor Pro Reference Manual.” (2011).
- [172] Y. Yuzhelevski, M. Yuzhelevski, and G. Jung. “Random telegraph noise analysis in time domain.” *Rev. Sci. Instrum.* **71**, 1681 (2000).
- [173] J. P. Klein and M. L. Moeschberger. *Survival analysis: techniques for censored and truncated data*. New York: Springer, 2nd ed. (2003).
- [174] L. R. Rabiner. “A tutorial on hidden Markov models and selected applications in speech recognition.” *Proc. IEEE* **77**, 257 (1989).
- [175] M. G. House, H. W. Jiang, and X. C. Zhang. “Analysis of electron tunneling events with the hidden Markov model.” *Phys. Rev. B* **80**, 113308 (2009).
- [176] D. V. Hinkley. “Inference about the change-point from cumulative sum tests.” *Biometrika* **58**, 509 (1971).
- [177] J. Canny. “A Computational Approach to Edge Detection.” *IEEE Trans. Pattern Anal. Mach. Intel.* **PAMI-8**, 679 (1986).
- [178] O. Naaman and J. Aumentado. “Poisson Transition Rates from Time-Domain Measurements with a Finite Bandwidth.” *Phys. Rev. Lett.* **96**, 100201 (2006).
- [179] P. J. Bickel and K. A. Doksum. *Mathematical Statistics: Basic Ideas and Selected Topics*. Upper Saddle River, N.J.: Prentice-Hall, 2nd ed. (2001).
- [180] G. Casella and R. L. Berger. *Statistical Inference*. Pacific Grove, CA: Duxbury, 2nd ed. (2002).
- [181] G. Roussas. *An Introduction to Probability and Statistical Inference*. San Diego: Academic Press (2003).
- [182] J. E. Dennis and R. B. Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*. Englewood Cliffs, N.J.: Prentice-Hall (1983).
- [183] J. Nocedal and S. J. Wright. *Numerical optimization*. New York: Springer, 2nd ed. (2006).
- [184] M. Boissonneault. “Multilevel dressed dephasing and sidebands driving: working notes.” (2011).
- [185] F. C. Wellstood, C. Urbina, and J. Clarke. “Low-frequency noise in dc superconducting quantum interference devices below 1 K.” *Appl. Phys. Lett.* **50**, 772 (1987).

- [186] F. C. Wellstood, C. Urbina, and J. Clarke. “Low-frequency noise in dc superconducting quantum interference devices below 1 K.” *Appl. Phys. Lett.* **50**, 772 (1987).
- [187] D. D. Awschalom, J. R. Rozen, M. B. Ketchen, W. J. Gallagher, A. W. Kleinsasser, R. L. Sandstrom, and B. Bumble. “Low-noise modular microsusceptometer using nearly quantum limited dc SQUIDs.” *Appl. Phys. Lett.* **53**, 2108 (1988).
- [188] F. Yoshihara, K. Harrabi, A. O. Niskanen, Y. Nakamura, and J. S. Tsai. “Decoherence of Flux Qubits due to 1/f Flux Noise.” *Phys. Rev. Lett.* **97**, 167001 (2006).
- [189] D. Drung, J. Beyer, J.-H. Storm, M. Peters, and T. Schurig. “Investigation of Low-Frequency Excess Flux Noise in DC SQUIDs at mK Temperatures.” *IEEE Trans. Appl. Supercond.* **21**, 340 (2011).
- [190] R. C. Bialczak, R. McDermott, M. Ansmann, M. Hofheinz, N. Katz, E. Lucero, M. Neeley, A. D. O’Connell, H. Wang, A. N. Cleland, and J. M. Martinis. “1/f Flux Noise in Josephson Phase Qubits.” *Phys. Rev. Lett.* **99**, 187006 (2007).
- [191] R. H. Koch, D. P. DiVincenzo, and J. Clarke. “Model for 1/f Flux Noise in SQUIDs and Qubits.” *Phys. Rev. Lett.* **98**, 267003 (2007).
- [192] W. M. Itano, D. J. Heinzen, J. J. Bollinger, and D. J. Wineland. “Quantum Zeno effect.” *Phys. Rev. A* **41**, 2295 (1990).
- [193] C. Balzer. “The quantum Zeno effect evolution of an atom impeded by measurement.” *Opt. Commun.* **180**, 115 (2000).
- [194] B. Nagels, L. Hermans, and P. Chapovsky. “Quantum Zeno Effect Induced by Collisions.” *Phys. Rev. Lett.* **79**, 3097 (1997).
- [195] E. Streed, J. Mun, M. Boyd, G. Campbell, P. Medley, W. Ketterle, and D. Pritchard. “Continuous and Pulsed Quantum Zeno Effect.” *Phys. Rev. Lett.* **97**, 260402 (2006).
- [196] J. Bernu, S. Deléglise, C. Sayrin, S. Kuhr, I. Dotsenko, M. Brune, J. Raimond, and S. Haroche. “Freezing Coherent Field Growth in a Cavity by the Quantum Zeno Effect.” *Phys. Rev. Lett.* **101**, 180402 (2008).
- [197] P. G. Kwiat, A. G. White, J. R. Mitchell, O. Nairz, G. Weihs, H. Weinfurter, and A. Zeilinger. “High-Efficiency Quantum Interrogation Measurements via the Quantum Zeno Effect.” *Phys. Rev. Lett.* **83**, 4725 (1999).
- [198] O. Hosten, M. T. Rakher, J. T. Barreiro, N. A. Peters, and P. G. Kwiat. “Counterfactual quantum computation through quantum interrogation.” *Nature* **439**, 949 (2006).
- [199] A. Kofman and G. Kurizki. “Acceleration of quantum decay processes by frequent observations.” *Nature* **405**, 546 (2000).
- [200] P. Facchi, S. Graffi, and M. Ligabò. “The classical limit of the quantum Zeno effect.” *J. Phys. A: Math. Theor.* **43**, 032001 (2010).

- [201] H. Nakazato, M. Namiki, and S. Pascazio. “Temporal behavior of quantum mechanical systems.” *Int. J. Mod. B* **10**, 247 (1996).
- [202] L. Schulman. “Continuous and pulsed observations in the quantum Zeno effect.” *Phys. Rev. A* **57**, 1509 (1998).
- [203] M. Fox. *Quantum Optics: An Introduction*. New York: Oxford University Press (2006).
- [204] P. W. Shor. “Scheme for reducing decoherence in quantum computer memory.” *Phys. Rev. A* **52**, R2493 (1995).
- [205] C. Ahn, A. Doherty, and A. Landahl. “Continuous quantum error correction via quantum feedback control.” *Phys. Rev. A* **65**, 042301 (2002).
- [206] P. Schindler, J. T. Barreiro, T. Monz, V. Nebendahl, D. Nigg, M. Chwalla, M. Henrich, and R. Blatt. “Experimental repetitive quantum error correction.” *Science* **332**, 1059 (2011).
- [207] I. Dotsenko, J. Bernu, S. Deléglise, C. Sayrin, M. Brune, J.-M. Raimond, S. Haroche, M. Mirrahimi, and P. Rouchon. “The quantum Zeno effect and quantum feedback in cavity QED.” *Phys. Scr.* **T140**, 014004 (2010).
- [208] C. Sayrin, I. Dotsenko, X. Zhou, B. Peaudecerf, T. Rybarczyk, S. Gleyzes, P. Rouchon, M. Mirrahimi, H. Amini, M. Brune, J.-M. Raimond, and S. Haroche. “Real-time quantum feedback prepares and stabilizes photon number states.” *Nature* **477**, 73 (2011).
- [209] J. Wang and H. Wiseman. “Feedback-stabilization of an arbitrary pure state of a two-level atom.” *Phys. Rev. A* **64**, 063810 (2001).
- [210] A. Korotkov. “Simple quantum feedback of a solid-state qubit.” *Phys. Rev. B* **71**, 201305 (2005).
- [211] D. Bozyigit, C. Lang, L. Steffen, J. M. Fink, C. Eichler, M. Baur, R. Bianchetti, P. J. Leek, S. Filipp, M. P. da Silva, A. Blais, and A. Wallraff. “Antibunching of microwave-frequency photons observed in correlation measurements using linear detectors.” *Nature Phys.* **7**, 154 (2010).
- [212] D. I. Schuster, A. P. Sears, E. Ginossar, L. DiCarlo, L. Frunzio, J. J. L. Morton, H. Wu, G. A. D. Briggs, B. B. Buckley, D. D. Awschalom, and R. J. Schoelkopf. “High-Cooperativity Coupling of Electron-Spin Ensembles to Superconducting Cavities.” *Phys. Rev. Lett.* **105**, 140501 (2010).
- [213] T. Frey, P. J. Leek, M. Beck, A. Blais, T. Ihn, K. Ensslin, and A. Wallraff. “Dipole coupling of a double quantum dot to a microwave resonator.” *ArXiv* 1108.5378 (2011).
- [214] Y. Kubo, C. Grezes, A. Dewes, T. Umeda, J. Isoya, H. Sumiya, N. Morishita, H. Abe, S. Onoda, T. Ohshima, V. Jacques, A. Dréau, J. F. Roch, I. Diniz, A. Auffeves,

- D. Vion, D. Esteve, and P. Bertet. “Hybrid quantum circuit with a superconducting qubit coupled to a spin ensemble.” *ArXiv* 1110.2978 (2011).
- [215] D. Drung, C. Hinnrichs, and H.-J. Barthelmess. “Low-noise ultra-high-speed dc SQUID readout electronics.” *Supercond. Sci. Tech.* **19**, S235 (2006).
- [216] “Eccosorb MFS.” Tech. rep., Emerson & Cuming Microwave Products, Randolph, MA (2011).
- [217] R. H. Giles, A. J. Gatesman, A. P. Ferdinand, and J. Waldman. “Design and fabrication of narrow band radar absorbing materials at terahertz frequencies.” In *IEEE Proceedings from the 15th International Conference on Infrared and Millimeter Waves*, 1–3. Orlando, FL (1990).
- [218] R. H. Giles, T. M. Horgan, and J. Waldman. “Silicon-Based Anechoics at Terahertz Frequencies.” In *Proc. of the 17th Int. Conf. on Infrared and Millimeter Waves*. Los Angeles, CA (1992).

Appendix A

Microwave ‘‘roach motel’’ filters

A.1 Design and modeling of roach filters

The roach filters are implemented as a length of lossy stripline transmission line in a copper box. The center trace of the stripline is a piece of copper tape cut to the appropriate width, while the dielectric is a magnetically loaded silicone rubber microwave absorber, Emerson & Cuming Eccosorb MFS-117. We seek to impedance-match the filters to the $50\ \Omega$ microwave environment, which is accomplished by choosing the width of the stripline center conductor. The impedance Z is related to the capacitance C_ℓ and inductance L_ℓ per unit length of the transmission line as [135]

$$Z = \sqrt{\frac{L_\ell}{C_\ell}} = \frac{\sqrt{L_\ell C_\ell}}{C_\ell} = \frac{1}{v_p C_\ell}, \quad (\text{A.1})$$

where v_p is the phase velocity of the TEM mode propagating down the stripline. We can express this phase velocity in terms of the permittivity ε and permeability μ of the dielectric as:

$$v_p = \frac{1}{\sqrt{\mu\varepsilon}}. \quad (\text{A.2})$$

We note that these are the absolute (not the relative) permittivity and permeability, and that in general they are complex and frequency-dependent. We can write them in terms of their real parts ε' and μ' and their imaginary parts ε'' and μ'' as

$$\varepsilon(\omega) = \varepsilon'(\omega) + i\varepsilon''(\omega) \quad (\text{A.3})$$

$$\mu(\omega) = \mu'(\omega) + i\mu''(\omega). \quad (\text{A.4})$$

To determine the impedance, all that remains is to calculate the capacitance per unit length C_ℓ . We can solve for this in the electrostatic approximation, giving [135]

$$C_\ell = W_{\text{eff}} \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{(n\pi)^2 \varepsilon \cosh(n\pi b/2a)}{2a \sin(n\pi W_{\text{eff}}/2a) \sinh(n\pi b/2a)}, \quad (\text{A.5})$$

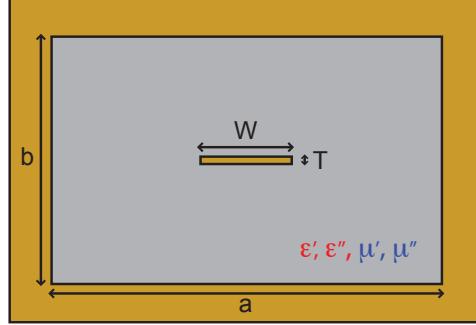


Figure A.1: Roach filter stripline geometry.

This figure shows a cross-section of a roach filter stripline with box width a and height b and center strip width W and thickness T . The Eccosorb dielectric is characterized by the real and imaginary parts of the permittivity and permeability ϵ' , ϵ'' , μ' , and μ'' .

where a is the width of the box, b is its depth, and W_{eff} is the effective width of the center strip. The sum in (A.5) is over odd values of n only. Although it is nominally a sum over all odd values of n , in practice one can truncate the sum at $n \approx 10^3 - 10^4$ without impacting the result. The effective strip width W_{eff} is given by [146]

$$W_{\text{eff}} = W + \frac{T}{\pi} [\ln(2b/T) + 1], \quad (\text{A.6})$$

where W and T are the width and thickness of the strip, respectively. Figure A.1 shows the cross-section of a roach filter stripline with dimensions labeled.

With these expressions, we can write down an expression for the impedance of the roach filter in terms of its cross-sectional geometry and the permittivity and permeability of the Eccosorb dielectric:

$$Z = \frac{\sqrt{\mu\varepsilon}}{W_{\text{eff}}} \times \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{2a \sin(n\pi W_{\text{eff}}/2a) \sinh(n\pi b/2a)}{(n\pi)^2 \varepsilon \cosh(n\pi b/2a)}. \quad (\text{A.7})$$

The values of ε and μ for Eccosorb MFS-117 have been measured at various frequencies by the manufacturer and are presented in the data sheet [216]. These values change substantially with frequency, so the impedance of the filter will vary with frequency as well. However, by choosing the geometry appropriately it is possible to keep the impedance in the range $50 \pm 7 \Omega$ out to 18 GHz.

Figure A.2 plots the real and imaginary parts of ε and μ (in units of the vacuum permittivity and permeability ε_0 and μ_0 , respectively) as a function of frequency. We see that the loss is a combination of dielectric loss and magnetic loss at low frequencies, but that magnetic loss dominates as we move to higher frequencies. The magnetic loss tangent μ''/μ' is greater than 1 above 8 GHz.

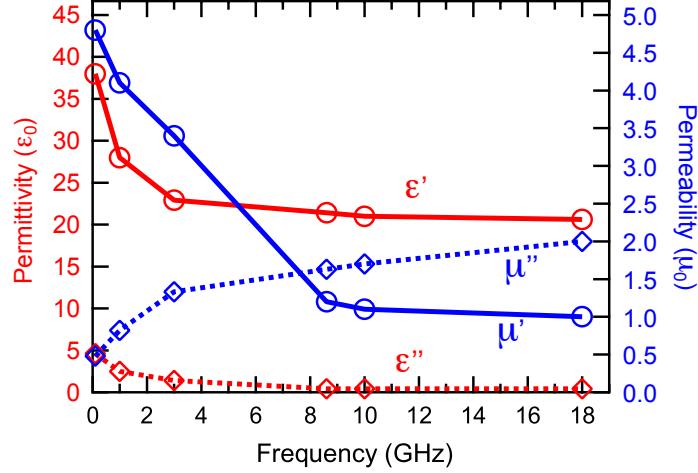


Figure A.2: Permittivity and permeability of Eccosorb MFS-117.

This figure plots the real and imaginary parts of the permittivity and permeability ϵ' , ϵ'' , μ' , and μ'' as a function of frequency in units of ϵ_0 and μ_0 . The magnetic loss tangent μ''/μ' increases with frequency; this is the primary loss mechanism for Eccosorb.

The filter boxes are milled from oxygen-free high-conductivity copper, and consist of a cavity with holes at the ends for connecting to SMA bulkhead connectors designed for launching to microstrip (AEP/Radiall 9308-9113-001). Matching copper lids screw down on top to seal the cavity and finish the filter. The impedance depends on the cross-sectional dimensions of the cavity and the center trace as described above, while the filter cutoff frequency $f_{3\text{dB}}$ depends on the length of the cavity [148]. Our standard roach filter has $f_{3\text{dB}} = 1.3$ GHz using a cavity 9.5 mm long, 8.9 mm wide, and 4.1 mm deep, with a stripline center conductor 1.35 mm wide and 50 μm thick. One can decrease the cutoff frequency by making a longer cavity, however making the cavity shorter to raise the cutoff frequency tends to give rise to re-entrant behavior above 30 GHz due to direct coupling between the pins of the SMA launches. We are working to develop roach filters using FIRAM microwave absorber material [217, 218] instead of Eccosorb; the FIRAM has lower loss, allowing higher cutoff frequencies. FIRAM is a carbon-loaded silicone rubber which shows dielectric loss instead of magnetic loss, and has been shown to be lossy to THz frequencies. However, its performance at millikelvin temperatures has not yet been tested.

The stripline is fabricated by cutting a piece of Eccosorb with a razor blade which will fit snugly in the copper box with no gaps. This Eccosorb is then cut in half and one half is placed in the box. A piece of copper tape cut to the correct dimensions with a razor blade is used for the center conductor; we remove the sticky backing of the tape with acetone to facilitate filter assembly. The tape is soldered to the SMA launch pins and the second half of the Eccosorb is placed on top. The lid of the box is then screwed on and the filter is complete.

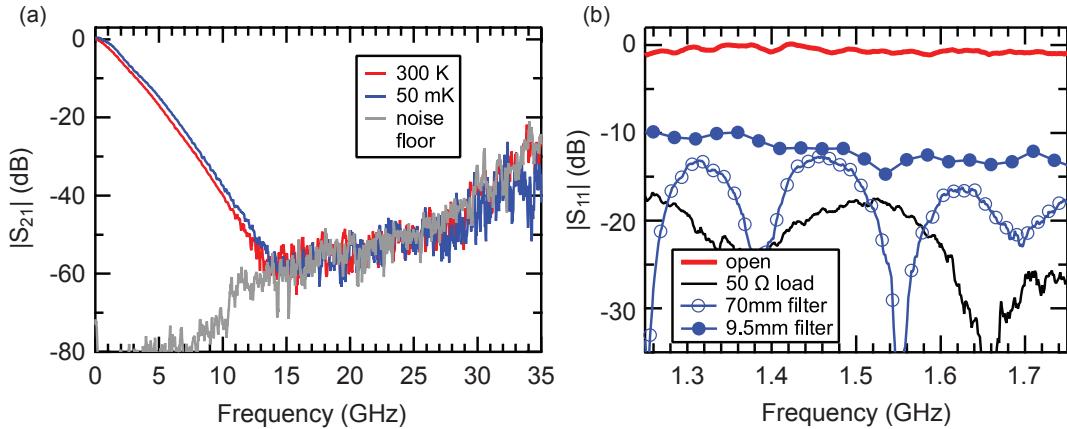


Figure A.3: Millikelvin S-parameters of roach filters.

Part (a) shows the filter attenuation $|S_{21}|$ at 300 K and 50 mK, along with the noise floor of the measurement. The noise floor rises with frequency due to attenuation in the cryogenic coaxial lines. Part (b) shows the reflection $|S_{11}|$ at 25 mK from an open circuit (bold red), a $50\ \Omega$ termination (black), a 70-mm-long roach filter with the opposite end open-circuited (empty blue circles), and a standard 9.5-mm-long roach (at 50 mK) connected to a $50\ \Omega$ load (filled blue circles).

A.2 Millikelvin filter performance

We measured the characteristics of roach filters at dilution fridge temperatures, which we reported in ref. [155]. Measurements of the transmission and reflection properties of several roach filters are shown in Figure A.3. We found that the frequency-dependent loss of the roach filters was essentially unchanged between room temperature and 50 mK, and that in both cases the attenuation of the filter was below the noise floor of our measurement at frequencies above 15 GHz. In addition, we found that the filters remained impedance-matched at low temperatures, with return loss better than 10 dB.

We also examined the thermalization of a 70-mm-long roach filter at millikelvin temperatures by measuring the noise power emitted in the filter stopband. A direct measurement of the Johnson noise emitted from the filter is hampered by $1/f$ drifts in the gain and system noise temperature of the measurement chain, which limit the maximum integration time. These drifts limit our maximum measurement resolution to approximately 150 mK for a direct noise measurement.

To circumvent this problem, we used a differential measurement technique analogous to a Dicke radiometer, shown schematically in Figure A.4. We used a Hittite switch (see section 5.2.2) located on the mixing chamber stage of our refrigerator to chop between our filter and a 40 dB NiCr attenuator, both anchored to the mixing chamber plate. The noise from both sources was amplified, bandpass filtered, and then detected with a zero-bias Schottky diode. The difference in noise power emission between the filter and the attenua-

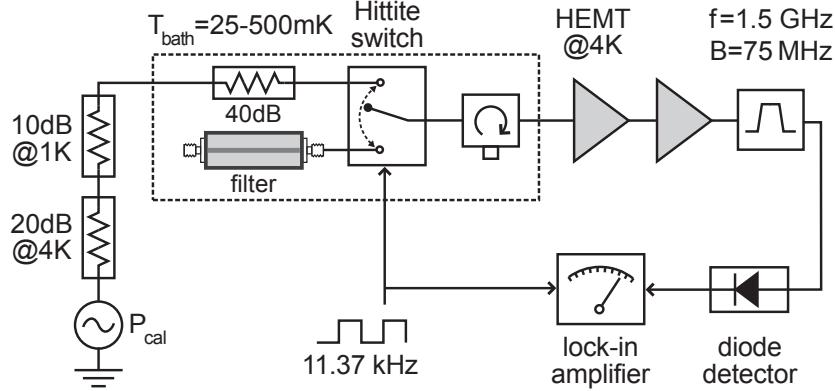


Figure A.4: Thermalization measurement setup.

We measure the noise power emitted by a 70-mm-long roach filter and a NiCr attenuator anchored to the mixing chamber stage of a dilution fridge. A Hittite switch is used to chop back and forth between the filter and the attenuator in a Dicke radiometer configuration. The noise emitted by each is amplified, bandpass filtered, and detected with a zero-bias Schottky diode used as a power detector. The noise power signal is measured synchronously with a lock-in amplifier. An external generator is used to provide a calibration signal

tor was measured synchronously with a lock-in amplifier. A signal generator connected to the attenuator was used for calibrating the absolute noise power.

This radiometric technique allows us to compare the noise radiated from the filter to that of a NiCr attenuator, which is believed to thermalize well at millikelvin temperatures. To quantify the temperature difference between the filter and the attenuator, we set the calibration generator power $P_{\text{cal}} = -135$ dBm and measured the in-phase quadrature of the lock-in signal

$$V_{\text{lockin}} = \eta G k_B B (T_{\text{atten}} - T_{\text{filter}}), \quad (\text{A.8})$$

where η is the power to voltage rectification ratio of the diode, G is the gain of the measurement chain, $B = 75$ MHz is the bandwidth of the bandpass filter on the output, and T_{atten} and T_{filter} are the noise temperatures of the attenuator and filter respectively. To extract the temperature difference, we used two independent calibration procedures. First, after each data point acquisition, five calibration points with different values of P_{cal} ranging from -63 dBm to -78 dBm were recorded. In this case, the lock-in signal is

$$V_{\text{lockin}} = \eta G [A P_{\text{cal}} + k_B B (T_{\text{atten}} - T_{\text{filter}})] \quad (\text{A.9})$$

where $A = -76$ dB is the attenuation of the stainless steel coaxial input line and associated attenuators (measured independently at room temperature). This room temperature value of A is accurate to within a few dB when the fridge is cold. Given A and P_{cal} , we can find $\Delta T \equiv T_{\text{atten}} - T_{\text{filter}}$ from the $V_{\text{lockin}} = 0$ intercept of (A.9), independent of the system gain. As a crosscheck, we connected the Hittite switch to the NiCr attenuator, varied the attenuator temperature from 200 mK to 600 mK, and recorded the diode output voltage.

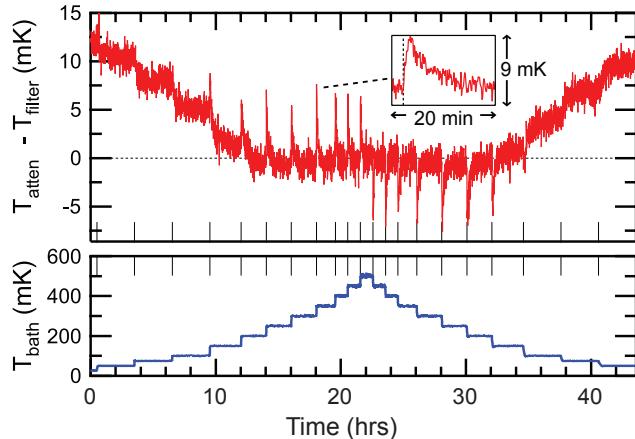


Figure A.5: Thermalization measurement.

The noise temperature difference ΔT between the filter and the attenuator as a function of time is shown in the top panel. The bottom panel shows the temperature of the mixing chamber stage on the same time axis. The long tick marks on the horizontal axis denote the times when the refrigerator temperature was changed. At temperatures above 150 mK, each temperature step is accompanied by a spike in ΔT ; the inset shows an expanded view of one such ΔT spike. The time of the temperature change is shown with a dotted line.

This yielded the product ηG which can also be used to obtain ΔT from V_{lockin} using (A.8). These two methods yield values of ΔT that agree to within 30%. The first calibration procedure is carried out for every data point, while the latter was performed only once for the entire measurement. However, the close agreement of these two calibrations indicates a high degree of immunity to $1/f$ noise. Additionally, we measured the insertion loss of both channels of the Hittite switch and calculate the effects of transmission imbalance and noise emission to be negligible.

In Figure A.5, we plot the measured ΔT as a function of time, together with the temperature of the thermal bath as it is varied between 25 mK and 500 mK. The data indicate that the attenuator and the filter have temperatures within a few millikelvin of each other down to 25 mK. At high temperatures (> 150 mK), the measured temperature difference between the attenuator and the filter is within the experimental uncertainty, and each temperature step is accompanied by a spike in ΔT : a sharp increase followed by a slower decrease, shown in the figure inset. We interpret this as the attenuator rapidly thermalizing to the new bath temperature, followed by the filter equilibrating with a longer time constant. Note that the spike polarity is reversed when we step the temperature downwards. Below 150 mK, the filter appears to be slightly colder than the attenuator, even after several hours of measurement.

These data show that the roach filters maintain both their loss and reflection characteristics at base temperature, and that they thermalize as well as a NiCr attenuator even down to millikelvin temperatures. This makes the roach filters very useful for low-

temperature applications requiring low loss at dc but high loss at microwave frequencies (such as the thermalization of a fast flux line), and for applications requiring a microwave stopband that extends to very high frequencies (such as measurements of loss in high- Q resonators).

Appendix B

Fabrication recipes

B.1 Resist Spinning

This recipe gives the standard MMA/ZEP bilayer, with about 250-300 nm of ZEP and 1 μm of MMA. Other work has used longer bake times for the MMA layer, but we found that 5 minutes is sufficient.

1. Procure substrate (wafer or chip) and place on spinner chuck, checking alignment carefully.
2. Coat substrate with MMA/MAA EL-13 resist.
3. Spin at 3000 rpm for 60 sec.
4. Remove substrate and bake on 180° C hot plate for 5 min.
5. Return substrate to spinner chuck, aligning carefully.
6. Coat substrate with ZEP-520A7 resist.
7. Spin at 3000 rpm for 60 sec.
8. Remove substrate and bake on 200° C hot plate for 3 min.
9. Remove from hot plate and heat sink to cool substrate. Put resist-coated substrate in a dark place until ready for use.

B.2 E-beam lithography

Typical doses for the ZEP/MMA bilayer with ice-water-bath development are 210 $\mu\text{C}/\text{cm}^2$ for fine features written at small currents (e.g. junctions), 240-270 $\mu\text{C}/\text{cm}^2$ for coarser features written with larger currents (e.g. resonators), and 50-60 $\mu\text{C}/\text{cm}^2$ for resist undercut boxes. ZEP resist is available from the Nippon Zeon Corp.; all other resists in this thesis were purchased from Micro-Chem Corp. Gold colloid and spring clips for lithography can be purchased from Ted Pella. Faraday cup parts and gold standards can be purchased

from EB Sciences. The NPGS website (<http://www.jcnabity.com/>) has a listing of vendors for electron beam lithography supplies.

1. Spin resist on a full wafer or smaller sample, then dice as appropriate.
2. Mount the sample on an SEM stub. The wafer should be held down with spring clips.
3. Put a toothpick in the bottle of gold colloid and let it sit there for a minute or two. Remove the toothpick and put a small drop of colloid at each corner of the sample. Allow the drops to dry completely.
4. Place the SEM stub in the SEM, pump down, and turn on the beam.
5. Using the computer controls, move around until you find the Faraday cup. Focus on the edge of the Faraday cup, link the stage Z to the focal distance, and zoom in to the Faraday cup hole so the picture is black.
6. Make a notation of the beam current on all spot sizes to be used for lithography.
7. Zoom out and move to the gold standard. Focus, link the stage Z, and adjust the stage height until the working distance is 10 mm.
8. Adjust the focus and astigmatism until the image is as clean as you can make it. Adjust the lens alignment using the modulator. Link the Z of the stage to the focal distance.
9. In NPGS, select “Direct Stage Control” and skip rotation correction. This prepares the system to accept dynamic focus correction points.
10. Zoom back out to low magnification and find a corner of the sample. Find where the gold colloid was deposited and zoom in on that area. Manually adjust the stage Z until the image comes into good focus.
11. Find an isolated gold ball sitting on the resist (not on a piece of debris) and focus on it. Optimize the focus with the computer controls.
12. Make a new point on the stage map for your location, then add the focus point to NPGS dynamic focus correction list.
13. Move to another corner, then repeat steps 11 and 12. Do this for all drops of colloid placed on the sample.
14. Accept the dynamic focus points in NPGS, then hit “Beam Off” and ensure the electrostatic beam blanker is set to “External”.
15. Check to make sure the doses, beam currents, magnifications, and spot sizes are correct in the NPGS run file.
16. Set the SEM scan mode to “External”.
17. Using the stage map, move to the desired location to begin lithography.
18. Use the NPGS program to perform lithography.

B.3 Resist Development

1. Fill the bath beaker with ice, then add water to make slush. Place the beaker inside its thermal insulation jacket.
2. Pour n-amyl acetate into the small developer beaker and place this beaker in its holder inside the bath beaker. Ensure that the developer is steady and no water gets in the beaker. Place the lid on the bath beaker to reduce condensation in the developer.
3. Wait 5 minutes for developer to cool.
4. Using tweezers, plunge the sample into the cold n-amyl acetate and agitate gently for 60 seconds.
5. Remove sample and blow dry immediately. DO NOT rinse (this may cause overdeveloping). Blow dry for at least 15 seconds to warm the sample so no condensation forms on its surface.
6. Using tweezers, plunge the sample into a beaker of room-temperature 1:3 MIBK:IPA and agitate gently for 4 minutes (large undercut features) or 1 minute (small undercut features).
7. Remove the sample and plunge it into a beaker of isopropanol, agitating gently for 15 seconds.
8. Remove sample and blow dry.

B.4 Thin film deposition

Typical evaporation parameters for qubit samples are layers of 20-30 nm for the first evaporation and 70-90 nm for the second evaporation, with angles of $\pm 18^\circ$, $\pm 30^\circ$, or $-5^\circ / + 30^\circ$, depending on the design. Typical oxidation parameters are 10 Torr of 5% O₂ in Ar for 10 minutes. Paramp samples have evaporation thicknesses of 60 and 80 nm for the two layers at angles of $\pm 30^\circ$. Oxidation parameters are typically 8 Torr of 5% O₂ in Ar for 8 minutes.

1. Place the sample in the evaporator, taking care to align the sample axis appropriately to the stage rotation axis.
2. Close the load lock and pump down. Once the load lock pressure is below 1×10^6 Torr, open the load lock gate valve and allow the cryo pump to pump on the load lock as well.
3. Allow at least 4 hours, or preferably overnight, for the load lock to be pumped by the cryo pump.
4. Fill and flush the oxidation volume with ArO₂ three times, with a fill pressure of 100 Torr.

5. Fill the oxidation volume to the correct pressure of ArO₂.
6. Allow the chamber to return to base pressure.
7. Set the stage to the first angle and turn on the e-gun.
8. Slowly turn up the e-gun current until a rate appears, then continue until it settles. Pre-evaporate 50-60 nm of aluminum.
9. Open the shutter and deposit the desired thickness of aluminum.
10. Turn the e-gun down slowly until it is off. Close the load-lock gate valve and the turbo gate valve.
11. Open the oxidation volume to the load lock and allow the oxidation to proceed for the desired time.
12. Turn on the turbo pump and its backing pump 25 seconds before oxidation is to complete, then open the turbo gate valve at the completion time. Allow the load lock to pump out to below 5×10^7 Torr.
13. Open the load lock gate and allow the load lock to pump down.
14. Set the stage to the second angle for evaporation and turn on the e-gun.
15. Slowly turn up the e-gun current until a rate appears, then continue until it settles. Pre-evaporate 50-60 nm of aluminum.
16. Open the shutter and deposit the desired thickness of aluminum.
17. Turn the e-gun down slowly until it is off. Return the sample stage to an angle of 0°.
18. Close gate valves, vent load lock, and remove sample.

B.5 Silicon nitride deposition

The silicon nitride for the paramp capacitors is deposited on top of the Nb ground planes and Si substrate using plasma-enhanced chemical vapor deposition (PECVD). We use the oxford2 machine in the UC Berkeley Nanolab for this purpose. Deposition is performed at 350° C substrate temperature with 100 W of RF power, using 100 sccm of 10% SiH₄/90% Ar and 30 sccm NH₃ as precursor gases. The deposition rate is roughly 15-20 nm/minute, and we grow between 120 and 180 nm of nitride depending on the desired specific capacitance.

Appendix C

Experimental schematics

These schematic diagrams show the arrangement of components used for our experiments. Filters and splitters not labeled with the name of a manufacturer are from Mini-circuits. Figure C.1 shows the setup used to generate broadband white noise for the experiments in section 8.2.4. Figure C.2 shows a typical room temperature experimental setup, including qubit and readout pulse generation, paramp pump generation, demodulation, and digitization. Coherent fast flux tones, if desired, are applied with a different generator (not shown).

Figures C.3 and C.4 show the wiring arrangement inside the fridge for the experiments performed in Chapters 7 and 8, respectively. These figures do not show the low-frequency wiring. The current for flux biasing of the qubit and paramp samples is sent to the flux bias coils via filtered low-frequency superconducting lines (copper above 4 K). The voltage signals used to operate the Hittite switch are sent via filtered low-frequency manganin lines. The signals to operate the Radiall switches are sent on unshielded superconducting looms (copper above 4K).

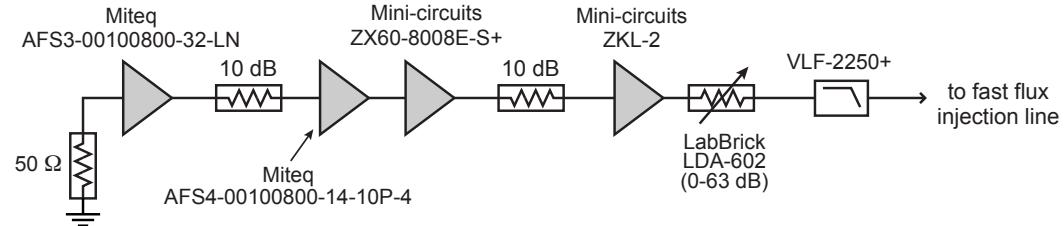


Figure C.1: Fast flux noise generation.

Circuit for generating broadband white noise 10 MHz - 2.5 GHz from amplified Johnson noise of a $50\ \Omega$ termination. The strength and frequency spectrum of noise at the output can be characterized using a spectrum analyzer. This circuit was used to generate flux noise for dressed dephasing experiments in Chapter 8.

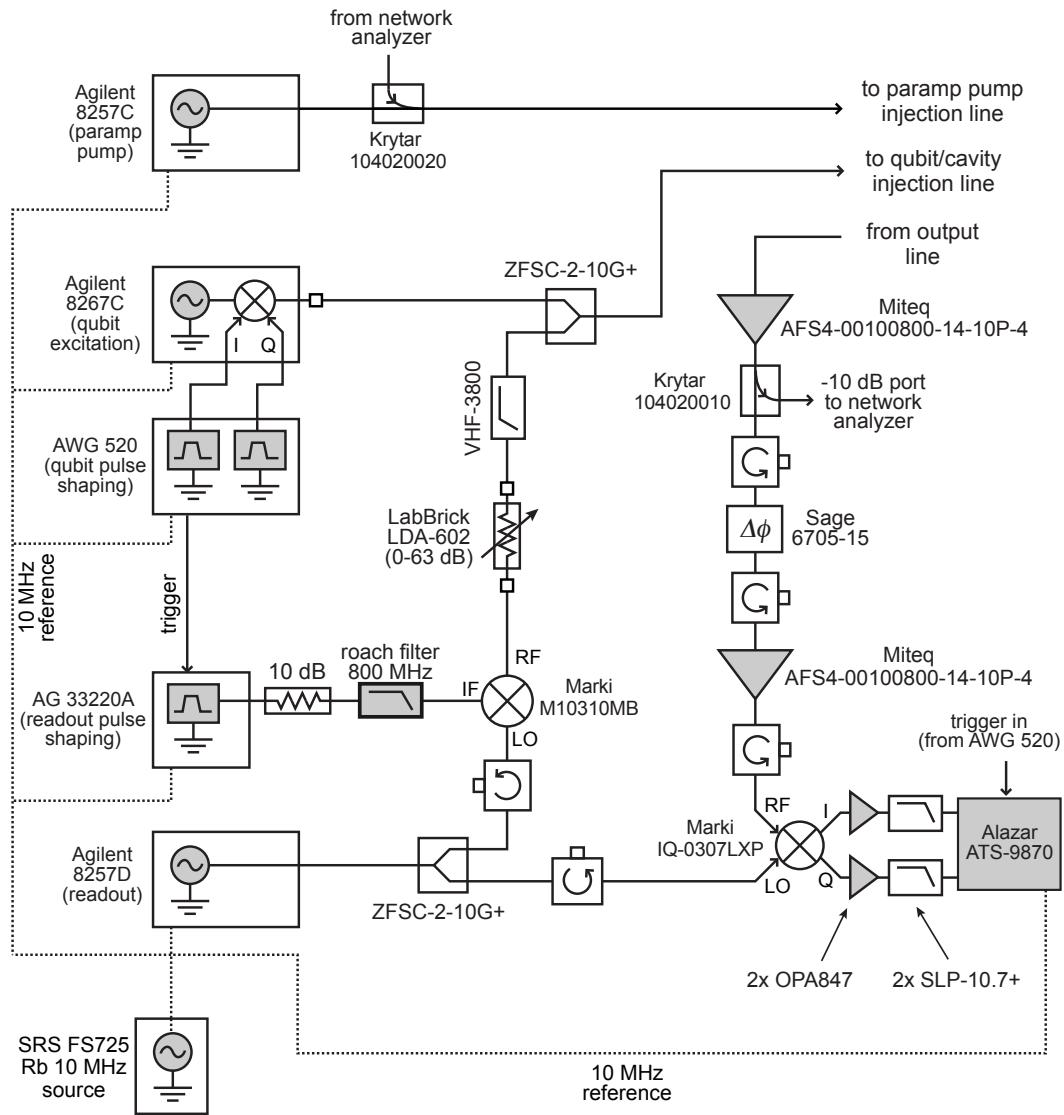


Figure C.2: Typical room temperature wiring.

Typical room temperature wiring for quantum jump experiments. A network analyzer (not shown) is connected to the directional couplers as labeled to perform paramp tune-up without needing to change the circuit. After tune-up, the network analyzer is removed and replaced with $50\ \Omega$ terminations. The room temperature circulators are all from DiTom, model D3I4080. The small unlabeled squares represent inner-outer DC blocks (Inmet 8039).

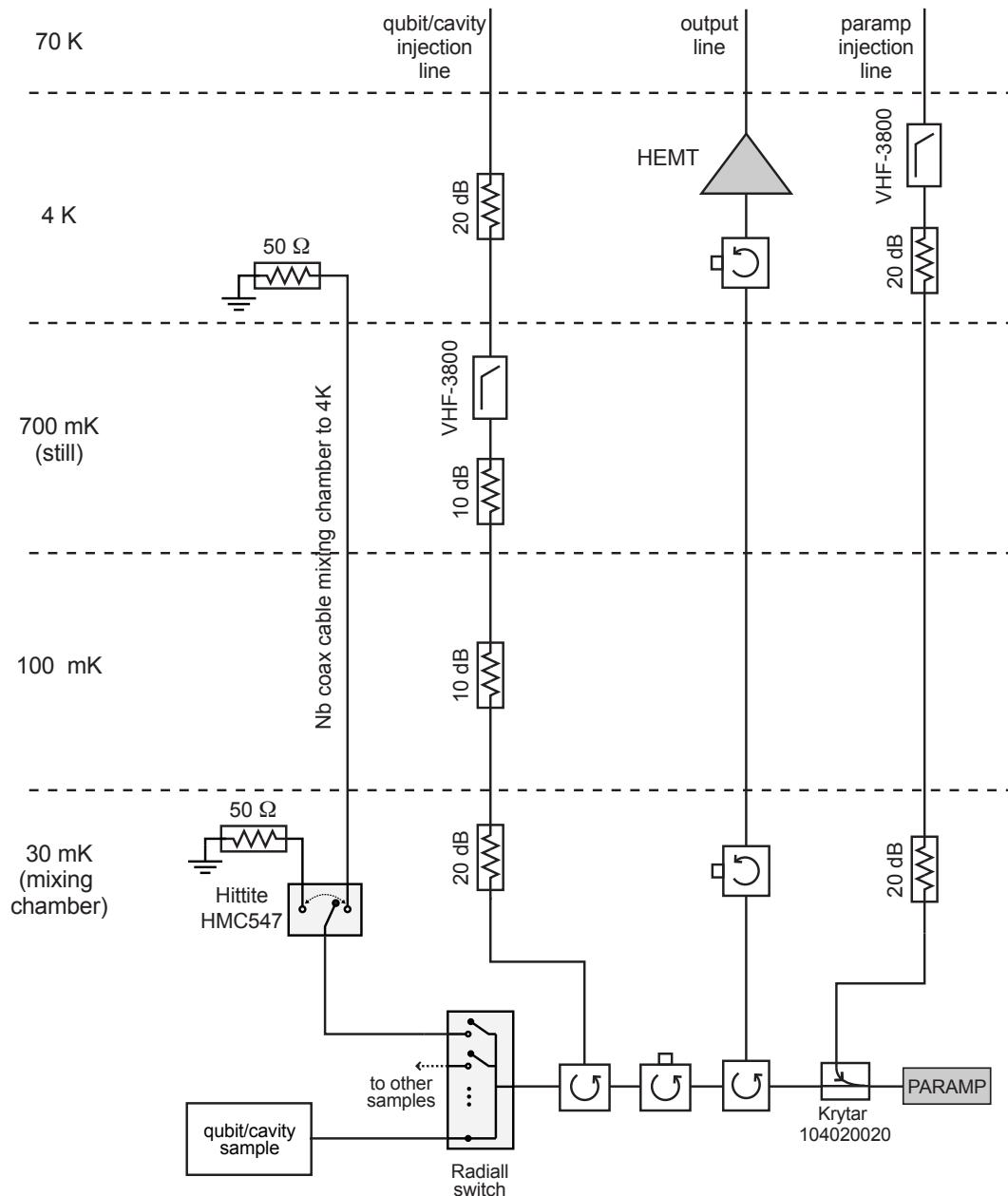


Figure C.3: Fridge wiring for Chapter 7 experiments.

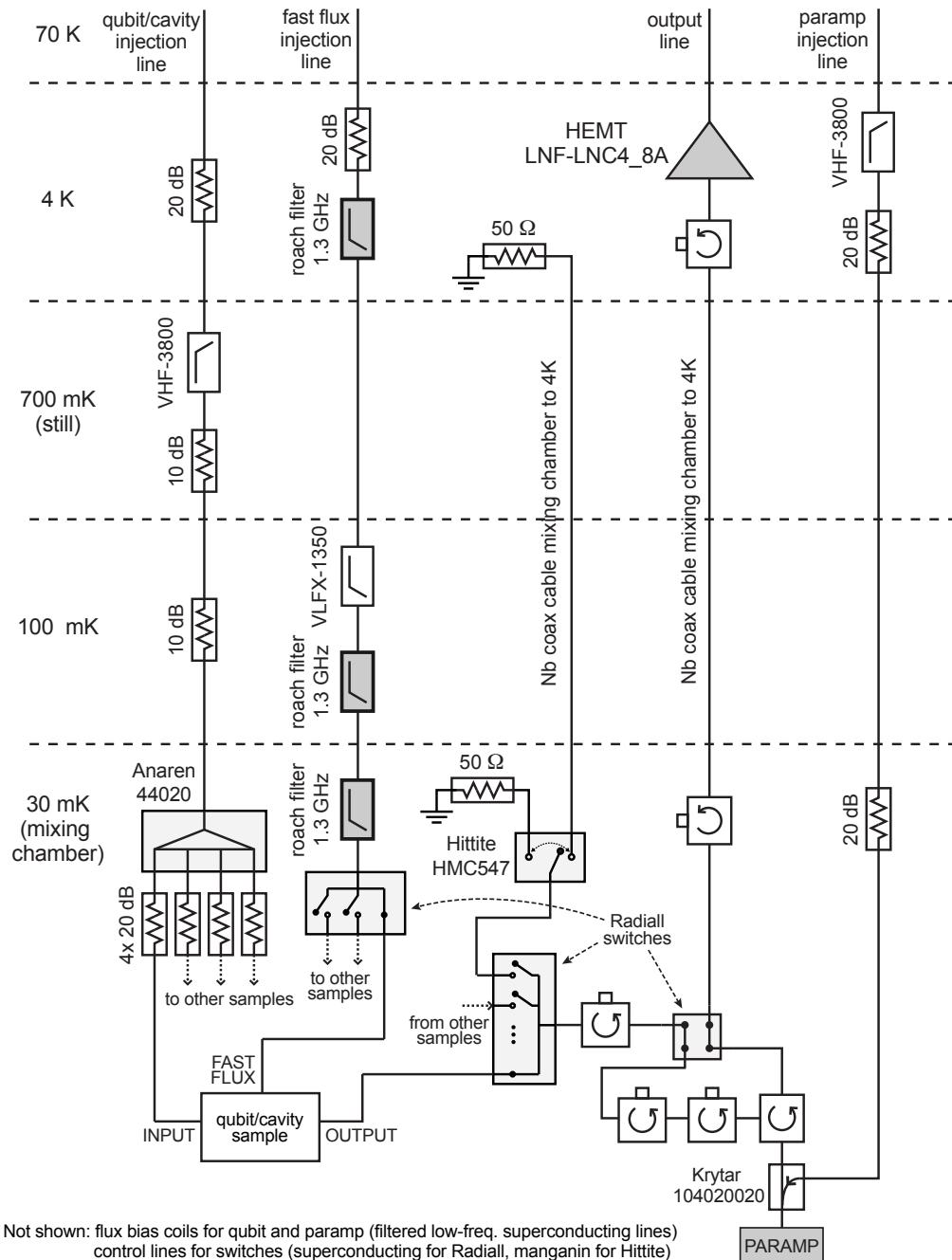


Figure C.4: Fridge wiring for Chapter 8 experiments.