CONTROL AND MEASUREMENT OF MULTI-LEVEL STATES IN THE JOSEPHSON PHASE CIRCUIT

Thesis submitted for the degree of

"Doctor of Philosophy"

By

Yoni Prosper Shalibo

Submitted to the Senate of the Hebrew University of Jerusalem

December 2012

This work was carried out under the supervision of:

Dr. Nadav Katz

Acknowledgments

This work would not have been possible without the help and support of many people. First and foremost, I wish to thank my adviser, Dr. Nadav Katz for his exceptional guidance. Nadav has provided me with all the required tools for my work, and much more. He has been a great teacher, always seeking to impart his knowledge and views on a wide variety of topics, from the elements of experimental work to numerical methods, physics, scientific writing and scientific discussion. Whenever I needed someone to discuss current issues, I always found his office door open, even for long, sudden discussions. These discussions were always fruitful and enabled steady advances throughout my work. Lastly I wish to thank Nadav for encouraging me all along the way.

I wish to thank all the technical personnel at the university for helping me with everyday needs. George and Gabi, the electronics engineers for providing me with electronics parts whenever I needed them. In particular I want to thank George for teaching me important concepts in electronics, helping me out building stuff and solving noise issues in the lab. Avshalom and all the fine-mechanics shop staff for manufacturing my mechanical designs and teaching me how to draw properly. Amnon, for teaching me how to manufacture them by myself. The staff at the nanofabrication center, in particular Rami who has helped me for countless hours with wire-bonding.

I would like to thank the members of the Martinis group at UCSB for providing us with samples, hardware and software that enabled me to conduct all my experiments. In particular John, for collaborating with us on our experiments, Matthew, Erik and Markus for helping me with the electronics and software bringup and Radek and Matthew for designing and fabricating (with Nadav) the devices I measured.

For scientific discussions, I wish to thank the members of my PhD committee, Prof. Ronnie Kossloff and Prof. Oded Milo. In particular I wish to thank Ronnie

for discussions about simulation methods and Wigner tomography. For discussions on nonlinear dynamics, I wish to thank Prof. Lazar Friedland and Dr. Ido Barth.

I wish to thank all the past and current members of our group for making my work at the lab a pleasure. In particular, I wish to thank Dudy, for the countless talks and encouragements that kept my good spirit. I will never forget all our lunches, late night singing and wrestling and afternoon walks to buy snacks. I am grateful for his help in installing and removing the heavy μ-metal shield of the fridge for countless times, and for bearing the continuous acoustic noise in the lab for the sake of my experiments. Roy, whom with I had the pleasure to work with toward the end, and helped me with many of the projects described in this work. Felix, for enlighting discussions about materials and fabrications. Uri, for interesting discussions about multi-level control and measurement. Avraham, who helped me out with the programming and setting the necessary control software for our experiments and Ofer, for helping me out in the clean room and with chess problems.

Finally, I wish to thank my closest friends and family. My friends, David, Moosh, Dudy, Assaf, and Eli for putting up with me. My parents for giving me the opportunity to study and for their constant care. My sister Yael and her family and my brothers Meir and Daniel for their care, interest in my work and for being the best family I could ever ask for.

List of publications

1) **Y. Shalibo**, R. Resh, O. Fogel, D. Shwa, R. Bialczak, J. M. Martinis, N. Katz, "*Direct Wigner tomography of a superconducting anharmonic oscillator*", arXiv:1208.2441.

2) D. Shwa, E. Shtranvasser, **Y. Shalibo** and N. Katz, "*Controllable motion of optical vortex arrays using electromagnetically induced transparency*", Opt. Express 20, 22 (2012)

3) J. S. Kline, M. R. Vissers, F. C. S. da Silva, D. S. Wisbey, M. Weides, **Y. Shalibo**, N. Katz, B. R. Johnson, T. A. Ohki, D. P. Pappas, "*Sub-micrometer epitaxial Josephson junctions for quantum circuits*", Supercond. Sci. Technol. 25, 025005 (2012)

4) **Y. Shalibo**, Y. Rofe, I. Barth, L. Friedland, R. Bialczack, J. Martinis and N. Katz, "*Quantum and Classical Chirps in an Anharmonic Oscillator*", Phys. Rev. Lett. 108, 037701 (2012)

5) **Y. Shalibo**, Y. Rofe, D. Shwa, F. Zeides, M. Neeley, J. Martinis and N. Katz, "*Lifetime and Coherence of Two-Level Defects in a Josephson Junction*", Phys. Rev. Lett. 105, 177001 (2010)

Abstract

The quest for building a scalable quantum computer resulted in many research avenues, involving different disciplines of physics. Many implementations are continuously being tested and proposed in a wide variety of systems from the fields of solid state physics, atomic and molecular physics, and quantum optics. Fifteen years ago, a new technology based on superconducting circuits has emerged, which satisfies the necessary conditions for building a scalable quantum computer. Compared to other systems, superconducting circuits have the advantage of simple fabrication, control and scaling. Vital features of a superconducting quantum computer, such as universal gates and basic quantum algorithms have already been demonstrated, and coherence times are steadily improving due to novel designs and basic research. Until recently, all the effort of building quantum computers was focused on the conventional scheme, based on computational elements called quantum bits (qubits), which are two-level quantum states. However, each element in a superconducting quantum computer has many levels, several of which are accessible to control and therefore for computation. A quantum computer based on d-level quantum states (qudits) has the same computational space (Hilbert space) as a qubit quantum computer but with $\log_2 d$ fewer elements. It is thus not surprising that certain quantum algorithms and quantum communication schemes using qudits have been shown to speed-up computation and increase security respectively, compared to their qubit counterparts. In addition to their application as qudits, multi-level states are ideal for studying interesting physics that goes beyond two-levels. In this work we demonstrate the basic elements of control and measurements of a single qudit, formed by the quantized energy levels of an anharmonic system called the Josephson phase circuit. The Josephson phase circuit offers most of the advantages of superconducting circuits, but in addition a straightforward generalization of measurements to multi-levels and a wide-range tunability of the system's energy levels which enables one to

dramatically change the response of the system to control signals.

This work covers several different topics, which are described in detail in Chapter 3. In one work, we use the circuit's bias and control signals to change its dynamical response from being quantum to classical. We control the qudit with a nearly resonant, frequency-chirped signal and show that for some set of parameters, it phase-locks to the drive. Phase-locking by chirped excitation is a known classical phenomena, called autoresonance, which we observe here in a quantum system. However, it was shown theoretically that a similar phenomena should occur where the dynamics are purely quantum, and characterized by sequential excitation of energy levels, called "ladder climbing". We map the transition between the two regimes and find a good agreement with theory and simulations.

In a different work we use the lowest three levels of the anharmonic potential to create arbitrary three-level states and measure their density matrix via standard state tomography. Using these tools, we measure the decay of the coherence times and find that the coherence between the 1^{st} and the 3^{rd} level decays three times faster than the coherence between the 1^{st} and the 2^{nd}. This could be used to exclude certain types of noise sources, however additional coherences beyond the three level subspace must first be measured. Due to the short coherence times between distant levels, we find it increasingly difficult to measure larger subspaces with standard state tomography. We therefore developed a novel technique to implement phase-space tomography in our system which allows for a simple measurement of the density matrix of up to the lowest five levels. We use this method to explicitly show phase-locking in wavepackets, prepared with a chirp.

Contents

List of Figures

List of abbreviations

AWG	Arbitrary Waveform Generator
BCS	Bardeen Cooper and Schrieffer
DAC	Digital to Analog Converter
EPR	Einstein, Podolsky, Rosen
FPGA	Field Programmable Gated Array
LO	Local Oscillator
MQT	Macroscopic Quantum Tunneling
NMR	Nuclear Magnetic Resonance
QIP	Quantum Information Processing
QM	Quantum Mechanics
QTLS	Quantum Two-level system
qubit	Quantum Bit
qutrit	Quantum Trit
RCSJ	Resistively and Capacitively Shunted Junction
RWA	Rotating Wave Approximation
SNR	Signal-to-noise ratio
SQUID	Superconducting Quantum Interference Device
SST	standard state tomography
TLS	Two-Level defect State
WT	Wigner tomography

CHAPTER 1

Introduction

In this chapter we give a short overview on Quantum Information Processing (QIP), the physics of two-level systems and superconducting quantum devices with a special emphasis on the Josephson phase-circuit. In Sec. 1.1 we give a short historical overview on the field of QIP. In Sec. 1.2, we lay the foundation to the physics of a quantum two-level system under a coherent control, subjected to decoherence. In Sec. 1.3 we introduce superconducting quantum devices. We start with the basic and most crucial element, the Josephson junction, and explain its implementation in various configurations. We then discuss the limitations of these devices, due to undesired coupling to the environment.

1.1 Quantum information processing

Quantum mechanics (QM) is the most successful theory we have of the physical world. Since its inception in the early 20^{th} century, it has profoundly altered our understanding of the laws of nature, by introducing revolutionary concepts to science such as the wave-particle duality, the uncertainty principle and entanglement. Despite its remarkable precision in predicting measured phenomena, the very same new concepts it entailed have spurred much skepticism, mainly due to their disturbing philosophical implications. Perhaps the most difficult of all is the concept of entanglement; according to QM, two distinct objects can exist in a state, where despite their arbitrarily large separation, some of their physical observables are correlated. This property was soon invoked by Einstein, Podolsky and Rosen (EPR) [1] to show by logical reasoning that the *wave-function* (the mathematical

description of the state of a quantum system), cannot be a complete description of reality if QM were to be self-consistent. More specifically, they showed using entangled states that two physical quantities corresponding to two non-commuting operators, can have a simultaneous physical reality, in contradiction to the uncertainty principle. Later, alternative theories were formulated [2], where "hidden variables" are introduced to the description of a physical system, thereby offering a solution to the EPR paradox. In 1964 however, an experiment was proposed by John Bell to settle the debate [3]. In this experiment, the expectation values of two non-interacting spins is measured in different directions. Bell derived an inequality for the results of this experiment based on the hidden variables theories, which can be violated in a certain class of states if QM is correct. His assumptions were: 1) that all the measured quantities have definite values, independent of observation (*realism*). 2) The two spins do not interact (*locality*). Then, in 1982 the first violation of Bell's inequality was measured using photons by Aspect, Dalibard and Roger *[4].* Bell's inequality and other inequalities that followed have been violated numerous times since in various systems, ever strengthening the correctness of QM. Despite the current understanding of QM as being a non-local theory, it does not violate special relativity. For example, recent experiments use entanglement to transfer a quantum state between photons, without them ever interacting. This kind of protocol, called "quantum teleportation", cannot however be used to transfer information faster than the speed of light [5].

1.1.1 Quantum computation and communication

In the early 1980's, Richard Feynman argued that quantum systems would be ideal for simulating other quantum systems *efficiently* [6]. This is because, unlike classical computers, the number of states within the system used for simulation scales exponentially with its size, as does the computational complexity of the problem. Feynman's notion was more general and applied to other computational problems, not only quantum simulations, however it remained only a matter of

theoretical curiosity for more than a decade. In 1994, Peter Shor demonstrated theoretically a quantum algorithm that obtains the prime factors of an integer N in a *polynomial time* [7], that is, the algorithm's running time scales polynomially with $\log N$. This had a major implication on the field of cryptography, since the most powerful encryption method to date (RSA [8]) is based on the common understanding that the problem of factoring a number N into its primes scales *exponentially* with N. Shor's algorithm was soon followed by other important quantum algorithms [9], which initiated a global effort to build a quantum computer, capable of performing these tasks.

A quantum computer exploits a notorious property of quantum systems to perform a computation in *parallel* on all possible inputs. A quantum system can be in a *superposition* of states, which evolve simultaneously with time according to a linear equation (the Schrödinger equation). This makes it possible to engineer a time-evolution for the system based on a quantum algorithm, where the result of the computation is encoded in the final state of the system. The common quantum computation model uses *quantum bits* (qubits) as the elementary computational unit. Unlike a classical bit, a qubit can be simultaneously "0" and "1". In fact, there are infinite possible states in which a single qubit can be (see Sec. 1.2), however in a typical quantum algorithm, the input qubits are prepared in an equal superposition of "0" and "1", producing an input which is a superposition of all possible inputs. The programmed interaction between these qubits results in a final state which depends on all the initial input states.

In addition to computation, quantum states are also useful to securely transmit information. In one scheme, one makes use of the fact that it is impossible to duplicate an unknown state ("no-cloning theorem") [9] to detect an eavesdropper in a private communication channel, where information is encoded in single-photon qubits [10]. In other schemes, the communicating partners share pairs of entangled photons which are perfectly correlated unless there is an attempt to eavesdrop on the communication channel [11].

1.1.2 Implementations

Any physical system obeys the laws of QM, yet only a small number can be used to construct a practical quantum computer. The DiVincenzo [12] criteria list 5 crucial requirements for such a system: 1) it must be scalable and use well defined qubits. 2) one should be able to initialize it to an arbitrary state. 3) it must have long coherence times, 4) it must have universal quantum gates 5) one should be able to readout the state of individual qubits with high-fidelity. Later on, two additional criteria were added to include quantum communication and distributed quantum computing capabilities: 6) able to convert between stationary and "flying" qubits, 7) capable of transmitting flying qubits between distant locations. Despite the steady increase of experimentally demonstrated implementations of quantum bits and gates, and the vast effort to improve their characteristics, a scalable quantum computer is yet to be built. One of the main reasons is decoherence (see Sec. 1.2). Most of the implementations (for example quantum dots and superconducting circuits) suffer from sufficiently large, uncontrolled coupling to the environment that results in a decay of the quantum coherence with time, and consequently in errors in the computation. Quantum error correction schemes put a lower limit on the number of gates to be performed within the coherence time, which is currently $\sim 10^4$. Other schemes, which have much smaller decoherence rates (NMR, ion traps, optical photons) are difficult to scale up for many qubits. Photons are the most coherent systems of all, however for the same reason they are the most difficult to couple. It is becoming clearer, however that ultimately, a quantum computer will have to be built using multiple technologies, that combine long-lived quantum memories and efficient QIP qubits.

1.2 Quantum two-level system

The quantum two-level system (QTLS) is one of the most fundamental concepts in quantum mechanics. It arises in a wide variety of problems, belonging to different

fields, such as Nuclear Magnetic Resonance (NMR), atomic physics, quantum optics and many others. It is also the basic building block of Quantum Information Processing (QIP) schemes, where it is referred to as a qubit [9]. Understanding the physics of QTLSs is therefore essential for building QIP devices, and also for explaining many physical phenomena. While some physical systems, such as the electron spin, are true QTLSs, others can be considered only as an *effective* QTLS under some conditions. For example, an atom is a multi-level system, however in some cases two sub-levels are strongly coupled by an external perturbation such that the Schrödinger equation can be reduced to only two amplitudes[1].

A QTLS is described by a state vector $|\psi\rangle = a|0\rangle + b|1\rangle$, where a and b are complex amplitudes, satisfying the normalization $|a|^2 + |b|^2 = 1$. Since the global phase of the state cannot be detected, only two parameters are required to describe this system. It is convenient to map these parameters onto a sphere (Bloch sphere, see Fig. 1.1) using a polar angle θ and an azimuthal angle ϕ, and the state takes the form $|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$. The geometrical representation provides an intuitive visualization of the state: any state on the sphere can be represented by the Bloch vector,

$$\vec{r} \equiv [r_1, r_2, r_3] = [\sin(\theta/2)\cos\phi, \sin(\theta/2)\sin\phi, \cos(\theta/2)]\,.$$

For example, the states $|0\rangle$ and $|1\rangle$ are located on the poles, whereas equal superpositions span the equatorial.

Since any real QTLS is subject to energy relaxation and decoherence processes, we must generalize our description from *pure* states, where the system has a *definite* single quantum state, into a *statistical mixture* of quantum states (or *mixed* states). Mixed states are described by a density matrix ρ, rather than a state vector: $\rho = \sum_i p_i |i\rangle\langle i|$, where p_i is the probability of being in the state $|i\rangle$. The Bloch vector

[1] The external perturbation must usually satisfy certain characteristics, such as frequency matching, low bandwidth, specific polarization and small amplitude. The approximation may be valid only for short time scales, where energy relaxation and\or coupling to other levels can be neglected.

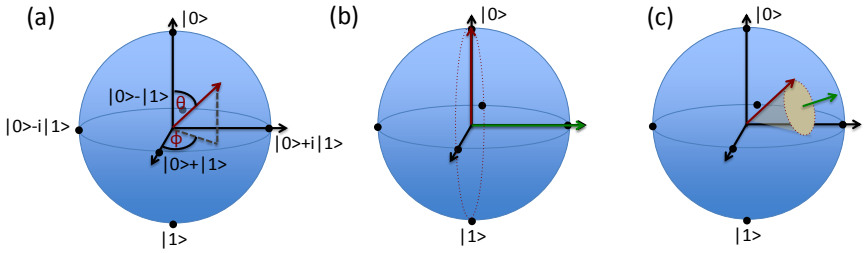

Figure 1.1: Dynamics on the Bloch sphere. (a) The Bloch vector (red) for the state $|\psi\rangle$. (b) and (c) show the time traces of the Bloch vector for resonant and off-resonant drives. The greed arrow represents the Rabi vector.

in this case is given by the expectation value of the Pauli matrices: $\vec{r} = Tr\left(\rho\vec{\sigma}\right)$. In contrast to pure states, mixed states have radii smaller than unity, which directly relates to their degree of purity: $\wp = Tr(\rho^2) = \frac{1+|\vec{r}|^2}{2}$. The quantity $\sqrt{r_1^2 + r_2^2}$ represents the degree of coherence and relates to the absolute value of the off-diagonal elements of the density matrix.

1.2.1 State dynamics

The dynamics of a density matrix state are dictated by the quantum Liouville equation:

$$\dot{\rho} = [H, \rho] / i\hbar. \tag{1.1}$$

In the case of a QTLS with an energy difference $\hbar\omega_a$, the Hamiltonian takes the form $H_a = \frac{\hbar\omega_a}{2}\sigma_z$, and the corresponding Liouville equation has a simple solution: $\rho_{01}(t) = \rho_{10}^*(t) = \rho_{01}(0)e^{i\omega_a t}$, and $\rho_{00}(t) = \rho_{00}(0) = 1 - \rho_{11}$. When a nearly resonant time-dependent perturbation is added, having the form $H_t = \hbar\Omega(t)\sigma_x \cos(\omega t)$[2], where $\Omega(t)$ is a slowly varying envelope, it is convenient to transform the Hamiltonian to the rotating frame: $H_R = V^\dagger H V - i\hbar V^\dagger \dot{V}$, where $V = e^{i\omega t\sigma_z}$. This leads to $H_R/\hbar = -\frac{\Delta}{2}\sigma_z + \frac{\Omega}{2}\sigma_x + e^{2i\omega t}\sigma_+ + e^{-2i\omega t}\sigma_-$, where we define the frequency detuning $\Delta = \omega - \omega_a$ and $\sigma_\pm = \frac{1}{2}\left(\sigma_x \pm \sigma_y\right)$. In the limit $\Delta \ll 2\omega$, the rapidly varying terms average to zero at the time scale relevant to

[2]This is not the most general form of a resonant perturbation. In the case where the drive has a different phase, one must add a term having the σ_y operator. A σ_z term may also exist, however it does not lead to transitions between energy levels, only to a change in the system's energy.

changes in the state and therefore can be neglected. This is called the Rotating Wave Approximation (RWA). It leads to simplified solutions in the case of a constant drive and greatly reduces the time required to solve the equations numerically in the case of a slowly varying envelope [3]. Using the RWA, the dynamics in the rotating frame are determined by the equation:

$$\dot{\rho} = [H_{\mathrm{RWA}}, \rho] \,/ i\hbar, \tag{1.2}$$

where:

$$H_{\mathrm{RWA}} = -\frac{\hbar\Delta}{2}\sigma_z + \frac{\hbar\Omega}{2}\sigma_x. \tag{1.3}$$

The eigenenergies of H_{RWA} are $\tilde{E} = \pm\frac{\hbar}{2}\sqrt{\Delta^2 + \Omega^2}$ and the eigenstates are $\left|\tilde{0}\right\rangle = \sin(\theta/2)\left|0\right\rangle - \cos(\theta/2)\left|1\right\rangle$ and $\left|\tilde{1}\right\rangle = \cos(\theta/2)\left|0\right\rangle + \sin(\theta/2)\left|1\right\rangle$, where $\theta = \arctan(\Omega/\Delta)$. For zero detuning, the energy difference reduces to $\hbar\Omega$ and the eigenstates reduce to an equal superpositions of the original eigenstates. Starting at the original ground state $\left|0\right\rangle = \left(\left|\tilde{0}\right\rangle + \left|\tilde{1}\right\rangle\right)/\sqrt{2}$, the state then evolves according to Eq. 1.2, namely $\left|\psi(t)\right\rangle = \left(\left|\tilde{0}\right\rangle + e^{-i\Omega t}\left|\tilde{1}\right\rangle\right)/\sqrt{2}$. Writing the state in the original eigenstates basis, and taking the absolute square of the amplitudes we get sinusoidal oscillations of the probability between the states $\left|0\right\rangle$ and $\left|1\right\rangle$, with a frequency $\Omega/2\pi$. This result can be visualized geometrically on the Bloch sphere as a precession of the Bloch vector about the x axis.

Equation 1.2 can be transformed into an equation for the Bloch vector [13]:

$$\frac{d\vec{r}}{dt} = \vec{\Omega} \times \vec{r}, \tag{1.4}$$

where we have introduced the generalized Rabi vector, $\vec{\tilde{\Omega}} = [\Omega, 0, \Delta]$[4]. The dynamics of a QTLS under a nearly resonant drive in the rotating frame are therefore described by a precession of the Bloch vector about the Rabi vector at an angular

[3] A recent work gives an exact analytical solution to the Rabi problem (no RWA required)

[4] In the most general case, the drive can take an arbitrary phase ϕ_d and $\vec{\tilde{\Omega}} = (\Omega\cos\phi_d, \Omega\sin\phi_d, \Delta)$.

frequency $\tilde{\Omega} = \sqrt{\Omega^2 + \Delta^2}$ (see Fig. 1.1c).

1.2.2 Decoherence

Energy relaxation and decoherence can be accounted for by adding the term,

$$\left[-r_1/T_2, -r_2/T_2, -\left(r_3 - r_3^{eq}\right)/T_1\right]$$

into Eq. 1.4, where T_1 is the energy (or longitudinal) relaxation time, T_2 is the phase (or transverse) relaxation time and r_3^{eq} is the equilibrium value of r_3 which is usually determined by system's temperature. The resulting set of equations is often called the Maxwell-Bloch equations. The transverse relaxation determines the characteristic time for the decay of coherent superpositions into a statistical mixture. Note that even in the absence of processes beyond energy relaxation, T_2 must be finite. By requiring the norm of the Bloch vector to be maximally bounded by 1, one finds that $T_2 \leq 2T_1$[5].

The general solution to the equation in the presence of decoherence and energy relaxation is complex, however it is instructive to write the solutions in some special cases. For example, in steady state ($\dot{\vec{r}} = 0$), the solution is given by:

$$\begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \frac{r_{eq}}{1 + (\Delta T_2)^2 + T_1 T_2 \Omega^2} \begin{pmatrix} -\Delta\Omega T_2 \\ \Omega T_2 \\ 1 + (\Delta T_2)^2 \end{pmatrix}. \tag{1.5}$$

The excited state population takes the shape of a Lorenzian as a function of detuning. The width of this function is given by $\Delta_{\text{FWHM}} = \sqrt{\frac{1+T_1T_2\Omega^2}{T_2}}$. In the limit of weak drives $\Omega \ll 1/\sqrt{T_1T_2}$, the width becomes $\Delta_{\text{FWHM}} = 2/T_2$. This allows

[5] Assuming a QTLS is excited to a superposition state $\vec{r} = [1, 0, 0]$ and the drive is turned off, we find from the Maxwell-Bloch equations $\vec{r}(t) = \left[e^{-t/T_2}, 0, 1 - e^{-t/T_1}\right]$. Since $|\vec{r}(t)|^2 \leq 1$, one finds $e^{-2t/T_2} - 2e^{-t/T_1} + e^{-2t/T_1} \leq 0$. Dividing the last relation by e^{-t/T_1}, one gets $e^{-(2t/T_2 - t/T_1)} + e^{-t/T_1} \leq 2$. Since this inequality must be satisfied for any time t, and for any T_1 we conclude that the first term on the left hand side must be smaller than unity, and therefore T_2 must be smaller or equal to $2T_1$.

for a direct measurement of the qubit's total coherence time (see 2.9). Note that in this case there is no coherence left in the qubit, and therefore we call such a measurement "incoherent spectroscopy". At large drive amplitudes the linewidth becomes power broadened, with $\Delta_{\mathrm{FWHM}} = 2\Omega\sqrt{T_1/T_2}$. In this case the transition is said to be saturated.

1.3 Superconducting quantum circuits

1.3.1 Superconductivity and the Josephson effect

The microscopic theory of superconductivity (BCS) predicts that below the transition temperature (T_c), electrons in a superconductor couple together in pairs (Cooper pairs) that condense into a ground state [14]. This low energy configuration can be described by a macroscopic wavefunction which gives rise to a lossless flow of electric current. Yet, the coherent character of the condensate remained hidden until the discovery of the Josephson effect and the demonstration of quantum interference of macroscopic currents in the Superconducting Quantum Interference Device (SQUID) [15].

The Josephson effect occurs when two superconductors are sufficiently close so that their wavefunctions overlap (see Fig. 1.2). Each superconductor acquires a random phase below the transition temperature due to spontaneous symmetry breaking and a phase difference can emerge. In 1962, Josephson predicted [16] that any phase difference across the junction results in a constant macroscopic current of Cooper pairs flowing without any voltage. Alternatively, when the junction is biased below a critical current, a corresponding phase difference develops. Above the critical current, a voltage $2\Delta/e$ drops across the junction and the current becomes normal, following Ohm's law $V = IR_n$, where R_n is the junction's normal state resistance. When a voltage is applied, the phase difference will vary in time, leading to an AC current and vice versa - when the junction is biased with an AC current, the phase difference will oscillate, and a non-zero voltage will

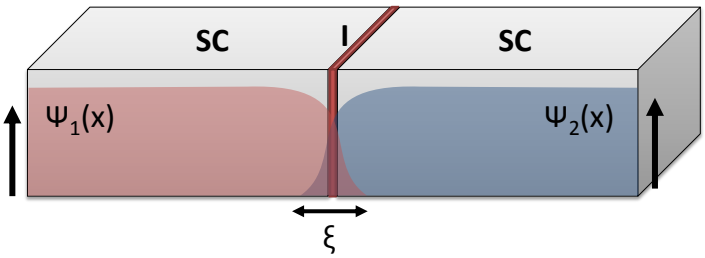

Figure 1.2: The Josephson junction

develop. These characteristics are formulated in the Josephson relations:

$$I = I_c \sin \delta \tag{1.6}$$

and,

$$\frac{d\delta}{dt} = \frac{2eV}{\hbar}, \tag{1.7}$$

where δ is the phase difference, I_c is the critical current, e is the electron charge, and $\hbar = h/2\pi$ is the reduced Planck constant. The critical current of the junction is determined by the normal state resistance and is calculated using the Ambegaokar-Baratoff relation [17]:

$$I_c R_n = \frac{\pi\Delta}{2e} \tanh(\Delta/2k_B T), \tag{1.8}$$

where Δ is the superconducting gap, k_B is Boltzmann's constant and T is the temperature.

The Josephson junction has a wide range of applications, from single-photon detectors and extremely sensitive magnetic field sensors, to high-frequency transistors and recently quantum circuits.

To process quantum signals using an electric circuit one must use a non-linear element. In addition, this element must be non-dissipative. The Josephson junction is the only known element which satisfies both conditions, since it supports a supercurrent that behaves electrically as a non-linear inductor. Using the standard definition for the inductance $L = V/\left(\frac{dI}{dt}\right)$ and the Josephson relations, one finds

$L_J = \frac{\hbar}{2eI_c \cos\delta}$. The inductance depends on the phase δ and therefore on the current, which means the inductance is non-linear. Unlike a regular inductor, there is no magnetic field involved in this effect and it is a direct manifestation of the kinetic energy of the Cooper pair current.

1.3.2 Macroscopic quantum effects

Despite the success of the theories of superconductivity and superfluidity, and the understanding that macroscopic objects can exhibit quantum effects [15, 18], it was unclear until the mid 1980s whether macroscopic variables can behave quantum mechanically. As proposed by A. Leggett [19], a distinction must be made between macroscopic quantum effects arising from objects that are composed of a macroscopic number of microscopic quantum systems, and macroscopic degrees of freedom such as the position and momentum of a macroscopic particle. In this respect, flux quantization and the Josephson effect belong to the first kind.

In a series of experiments by J. Clarke and co-workers [20, 21, 22], conducted on a current-biased Josephson junction, it was established that the phase difference δ is in fact a macroscopic quantum variable. In these experiments, the escape rates of the macroscopic phase out of an energy well were measured as a function of temperature and current bias, showing a crossover from thermal activation to quantum tunneling at decreasing temperatures (or increased critical currents). In addition, evidence of quantized levels within the well was found in spectroscopy measurements, which is inconsistent with classical theory. The measured transitions and their lineshape strongly agree with the prediction of quantum mechanics. It is important to note that these experiments were made possible due to the non-linear properties of the Josephson junction.

Observations of macroscopic quantum variables were recently extended to the position and momentum of a macroscopic mechanical oscillator [23] using a superconducting qubit as a non-linear interface to the classical control signals. Additional manifestations of quantum behavior in superconducting circuits such as en-

tanglement, and violation of Bell's inequality are continually being demonstrated [24, 25, 26].

1.3.3 The current-biased Josephson junction

To understand the dynamics of the phase in a current-biased Josephson junction we must account for the junction's capacitance and internal\external dissipation. This is done by modeling the circuit as a Resistively and Capacitively Shunted Junction (RCSJ). Using Kirchhoff's laws and the Josephson relations, we can write an equation of motion for the phase:

$$I = I_c \sin\delta + V/R + C\frac{dV}{dt} = I_c \sin\delta + \frac{\Phi_0}{2\pi R}\frac{d\delta}{dt} + \frac{C\Phi_0}{2\pi}\frac{d^2\delta}{dt^2}, \tag{1.9}$$

where $\Phi_0 = h/2e$ is the quantum of flux. This equation can be brought to a canonical form,

$$m\frac{d^2\delta}{dt^2} + \frac{m}{RC}\frac{d\delta}{dt} + \frac{dU(\delta)}{d\delta} = 0, \tag{1.10}$$

where we have defined an effective mass $m = C\left(\frac{\Phi_0}{2\pi}\right)^2$, and potential energy $U(\delta) = -\frac{I_c\Phi_0}{2\pi}\cos\delta - \frac{I\Phi_0}{2\pi}\delta$. Therefore, the dynamics of the phase are mapped out into that of a fictitious particle of mass m, moving in a one-dimensional potential $U(\delta)$ of the position coordinate δ and a friction force. The potential has a shape of a tilted washboard, with a tilt slope determined by the current. When the junction is not biased, the equation of motion is that of a pendulum, with a small-amplitude resonance frequency of $\omega_p = \sqrt{\left(\frac{d^2U}{d\delta^2}\right)_{\delta=\delta_0}/m} = \sqrt{\frac{2eI_c}{\hbar C}}$. For bias currents $I < I_c$, the potential wells become asymmetric with an increasing anharmonicity. The resonance frequency at their bottom is in this case $\omega_0 = \sqrt{\left(\frac{d^2U}{d\delta^2}\right)_{\delta=\delta_0'}/m} = 1/\sqrt{L_JC} = \omega_p\left(1-\gamma^2\right)^{1/4}$, where $\gamma = I/I_c$ is the normalized current. When $I > I_c$, there are no longer potential wells, and the phase increases with time, leading to a non-zero average voltage across the junction according to Eq. 1.7.

At low temperatures $k_BT \ll \hbar\omega_0$ and weak coupling to the environment, δ and

N can no longer be treated as classical variables, but rather as operators which satisfy the uncertainty relation $[\delta, N] = -i$. This leads to a number of observable quantum phenomena, such as energy quantization and quantum tunneling from metastable wells for $I < I_c$. By solving the Hamiltonian for the case of shallow wells ($I \lesssim I_c$), we find that the lowest two transition energies are given by [20]:

$$\omega_{01} = \omega_p \left(1 - \frac{5}{36}\frac{\hbar\omega_p}{\Delta U}\right) \tag{1.11}$$

and

$$\omega_{12} = \omega_p \left(1 - \frac{10}{36}\frac{\hbar\omega_p}{\Delta U}\right), \tag{1.12}$$

where $\Delta U = 2\sqrt{2}\frac{I_0\Phi_0}{3\pi}\left(1-\gamma\right)^{3/2}$ is the current dependent potential barrier. For temperatures much smaller than the thermal escape temperature $\hbar\omega_p/2\pi k_B$, escape from the well is dominated by quantum tunneling. For high-Q systems, the tunneling rates are given by [27]:

$$\Gamma_q = \frac{a_t\omega_p}{2\pi}\exp\left(-7.2\frac{\Delta U}{\hbar\omega_p}\right), \tag{1.13}$$

where $a_t = \left(120\pi\left(7.2\frac{\Delta U}{\hbar\omega_p}\right)\right)^{1/2}$. Note that tunneling from the potential well involves a macroscopic number of particles and does not occur in real space. This is to be distinguished from tunneling of single Cooper pairs across the junction, and therefore the former is coined Macroscopic Quantum Tunneling (MQT).

1.3.4 Types of superconducting circuits

Superconducting quantum circuits suffer from decoherence processes. These include both energy relaxation and pure dephasing which result in the system's quantum coherence vanishing at time-scales large compared to the characteristic decoherence time. The dominant sources of dephasing in these systems is fluctuations in the macroscopic, classical parameters which set the energy of the system (see 1.3.6), such as offset-charge, capacitance, critical current and flux. To reduce

the state's sensitivity to these fluctuations, a number of circuit architectures have been developed. There are currently three main types of such circuits: the charge, flux, phase qubits [28], with two relatively new ones: the transmon [29] and fluxonium [30] qubits. In each, the energy terms are designed to have different relative sizes. There are three relevant energy scales: the electrostatic energy $E_C = 2e^2/C$ (associated with an excess number of one Cooper pair across the junction), the Josephson energy $E_J = I_c\Phi_0/2\pi$ (the maximal tunnel coupling) and the magnetic energy $E_L = \Phi_0^2/2L$ (due to one flux quantum threading the circuit loop).

The charge qubit (or Cooper-pair box) is the first superconducting quantum circuit [31]. It features a superconducting island, coupled to the circuit on one side through a Josephson junction and on the other side through a small capacitor or an additional junction. In this circuit, the ratio E_C/E_J is chosen to be larger than unity. The circuit is sensitive to offset-charge noise since the latter is large compared to quantum fluctuations in charge. The offset charge on the capacitor is tuned via a voltage source, allowing for a control over the system's energy levels. Using this control, it is possible to bias the offset charge to a point where the energies are to first order insensitive to offset-charge fluctuations.

The flux qubit (or RF-SQUID) is a superconducting loop, interrupted by a Josephson junction [32]. The loop is inductively coupled to an external flux bias that enables control over the energy. Here, $E_C < E_J$ so offset-charge fluctuations are made small compared to quantum fluctuations in charge, thereby making this circuit immune to charge noise. On the other hand, this system is sensitive to flux noise, however in a similar manner to the charge qubit, the external flux can be tuned to roughly $\Phi_0/2$, where the potential shape is symmetric and therefore insensitive to first order to fluctuations in the external flux. The circuit is designed to work in the regime $L \lesssim L_{J0}$ ($L_{J0} \equiv L_J(\delta = 0)$), where typically only one level is trapped within the potential wells. This level is split into two due to the tunnel coupling between the wells, and serves as the qubit.

The phase qubit is a current-biased Josephson junction with $E_J \approx 10^4 E_C$ [33]. Sim-

ilar to the flux qubit, this system is insensitive to charge fluctuations. Since E_J is much larger than E_C, quantum fluctuations in δ are small and it behaves like a classical variable. Since the inductance does not vary over the wavefunction's span, the inductance can be considered linear and the junction behaves like an harmonic oscillator. To take advantage of the junction's non-linearity, it must be carefully biased close to the critical current, where it is anharmonic. This leads to a strong sensitivity of the energy to the bias current and consequently to current noise. Unlike the flux and charge qubits, the phase qubit does not have an optimal point ("sweet spot"), however it has two main advantages: a) It has a large tunability. In this work we utilize this tunability to perform our experiments. b) It facilitates device fabrication. While the charge and flux qubits utilize small junctions (with a typical size smaller than 1 μm) that require fabrication using e-beam lithography, phase qubits have large junctions and can be fabricated using only standard optical lithography.

The transmon is a charge-flux hybrid. It is a split Cooper pair box (having two junctions) with a $E_J \backsimeq 50E_C$, that reduces sensitivity to charge noise [29]. While sensitivity to charge noise reduces exponentially with E_J/E_C, anharmonicity decreases only as $\sqrt{E_C/E_J}$, thereby allowing for an optimal point in the number of gates during the coherence time. The transmon can be biased with a voltage or an external flux which changes the amplitude of the cosine potential. The fluxonium is a charge qubit, shunted with a large non-geometric inductance, obtained with a series of large critical-current Josephson junctions [30]. This ensures large impedance $\sqrt{L/C}$ of the junction array at high-frequency, keeping the circuit in the charging regime, while eliminating charge offset noise at low-frequency.

1.3.5 The flux-biased Josephson phase circuit

In the original phase-qubit circuit, the energy levels are controlled via a current bias, leading to a tilted-washboard potential. A significant limitation of this design is its large coupling to the bias source, causing an increased decay rate and

sensitivity to bias noise. This property is characterized by the electrical impedance to the bias source, which for a current bias is limited to $\sim 100\,\Omega$ by the wire impedance at microwave frequencies. For a given impedance $Z(\omega)$ between the junction and the bias (see Fig. 1.3a), the resulting lifetime is $T_1 = Z(\omega)C$, where $C \sim 1\,\mathrm{pF}$ is the typical parallel capacitance. By using a flux transformer, it is possible to achieve a broadband, high-impedance current source. Assuming a geometric inductance L and a mutual inductance M to the current source, we get $Z(\omega) = Z_0 \left(L/M\right)^2$, where Z_0 is typically $50\,\Omega$. For our devices, L/M is typically chosen to be ~ 300, leading to a few $\mathrm{M}\Omega$ of impedance isolation. This limits T_1 to a few μs, which is much longer than our typical device lifetime. An additional advantage of this design is that the classical state of the circuit is easily read out via a magnetic field measurement. In an experiment, state occupation probabilities are mapped onto the probability of escaping the qubit well after a measurement pulse (see Sec. 2.3 for a detailed explanation of the measurement process). The latter probability is related to the total current circulating in the qubit loop and therefore can be measured accurately via a SQUID magnetometer. This brings an improvement over the current-biased phase circuit, in which the escaped phase freely runs, creating a voltage on the junction and heating which leads to quasiparticle generation.

The schematics of the flux-biased phase circuit are depicted in Fig. 1.3b. The total current circulating in the qubit loop I_q with geometric inductance L adds a magnetic energy term to the Hamiltonian: $E_M = \frac{1}{2}LI_q^2$. This current results from both the phase difference across the junction and the external flux. We can write the total flux threading the qubit loop as $\Phi_q = \Phi_{\mathrm{ext}} + LI_q$, where Φ_{ext} is the external flux bias. Since the gauge invariant phase difference across the junction δ is related to the total flux $\delta = 2\pi\Phi_q/\Phi_0$, we get $I_q = \delta\Phi_0/2\pi - \Phi_{\mathrm{ext}}$, and therefore $E_M = \frac{1}{2L}\left(\Phi_{\mathrm{ext}} - \frac{\delta\Phi_0}{2\pi}\right)^2$. This term replaces the current bias term in the current biased Josephson junction Hamiltonian to give:

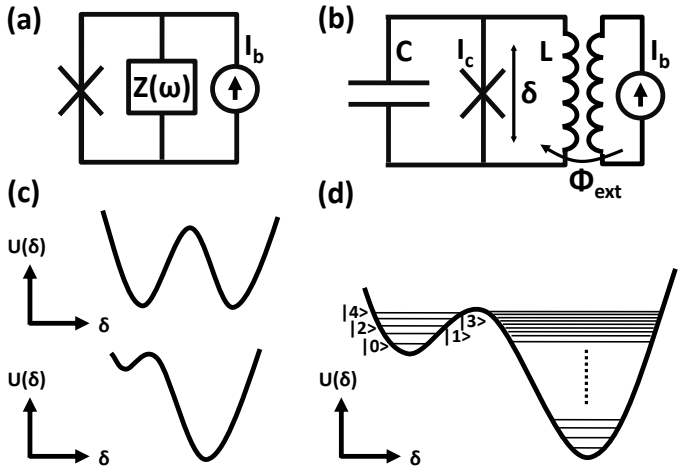

Figure 1.3: The flux-biased phase circuit. (a) Effective circuit of a current biased Josephson junction. (b) A flux-biased junction using a current transformer (c) Potential shape at away from working point (upper) and at the typical working point (lower) (d) Energy levels of the circuit at the working point. The right well contains hundreds of levels, while the left only a few.

$$\hat{H} = -\frac{2e^2}{C}\frac{d^2}{d\hat{\delta}^2} - \frac{I_c\Phi_0}{2\pi}\cos\hat{\delta} + \frac{1}{2L}\left(\Phi_{\text{ext}} - \frac{\hat{\delta}\Phi_0}{2\pi}\right)^2. \tag{1.14}$$

We identify a potential energy term $U(\delta) = -\frac{I_c\Phi_0}{2\pi}\cos\hat{\delta} + \frac{1}{2L}\left(\Phi_{\text{ext}} - \frac{\hat{\delta}\Phi_0}{2\pi}\right)^2$ which is a sum of a cosine and a shifted parabola. Depending on the parameters I_c, L and Φ_{ext}, we can have either one potential well or multiple wells (see Fig. 1.3c). Typically the device is designed to have $L \backsimeq 4L_{J0}$, so that two wells exist at the most[6]. After reset, the state of the system is completely determined by the probability amplitudes of the trapped states within a single well. The lowest two eigenstates are typically used for qubit experiments (see Fig. 1.3d), however in this work we are interested in many of the trapped levels (see Part 3). Note that the potential *form* is periodic in Φ_{ext} with a period Φ_0, since $U(\Phi_{\text{ext}} + n\Phi_0, \delta) = U(\Phi_{\text{ext}}, \delta - 2\pi n)$. At some critical flux values Φ_c[7], the size of one of the potential wells decreases to zero and the trapped phase escapes from this well. Sufficiently close to Φ_c, the

[6] A double well potential facilitates resetting of the phase in one well for qubit experiments (see Sec. 2.3 for more information)

[7] due to periodicity, Φ_c is defined up to an integer multiple of Φ_0

potential is well approximated by a cubic function, similar to the current-bias potential. We can therefore map between the flux-bias and current bias parameters and use it to directly obtain the energies of the lowest transitions using equations 1.11 and 1.12. By expanding the potential to third order around the critical phase δ_0 (where $[U'(\delta)]_{\delta_0} = [U''(\delta)]_{\delta_0} = 0$) we find $U(\tilde{\delta}) \approx (I_c \sin\delta_0)\sin\tilde{\delta} - (\Phi - \delta_0\Phi_0/2\pi)\,\tilde{\delta}/L$, where $\tilde{\delta} = \delta - \delta_0$ and $\cos\delta_0 = -L_{J0}/L$. We can therefore identify an effective critical current $\tilde{I}_c = I_c \sin\delta_0$ and an effective bias current $\tilde{I} = (\Phi - \delta_0\Phi_0/2\pi)/L$. The critical flux is obtained by requiring $\tilde{I} = \tilde{I}_c$, and gives: $\Phi_c = \frac{\Phi_0}{2\pi}(\delta_0 - \tan\delta_0)$.

1.3.6 Decoherence in superconducting circuits

There are a number of decoherence processes that are relevant to superconducting circuits in general and to the flux-biased phase circuit in particular. We divide these processes into two - those that cause energy relaxation and those that cause only decoherence. As stated earlier (see Sec. 1.2), even when no pure decoherence process occur, energy relaxation limits the coherence time to be smaller than $2T_1$, where T_1 is the energy relaxation time.

1.3.6.1 Energy relaxation

Dielectric loss. This is currently the dominant relaxation mechanism in the phase-circuit [34] and other superconducting systems [35]. It is caused by energy relaxation of electric dipoles within the insulating material used for capacitors, the Josephson junction or even the native oxide layer of the superconducting wires. As discussed in Sec. 3.1, some of the loss can be attributed to two-level defect states within amorphous dielectrics. The physical parameter of the material describing the losses is its loss tangent $\tan\delta = \epsilon''/\epsilon'$, where ϵ' and ϵ'' are the real and imaginary parts of the dielectric function. The loss tangent of the dielectric constant of a capacitor directly determines its quality factor $Q_d = 1/\tan\delta$. For cases where the electric field is only partially distributed in the lossy material, one gets

an effective Q_d factor which is higher than the intrinsic Q_d of the material. The low-temperature loss tangent of materials is to large extent unknown. Typically, the superconducting material is chosen to be aluminum, mainly since its oxide is used to obtain the barrier of the Josephson junction and has a relatively low-loss oxide (AlO_x, where $x \approx 1.5$). However, this amorphous material still limits the Q factor of most of the devices. In its crystalline form (sapphire), this material has a much lower loss tangent and it is currently an active research avenue growing crystalline junctions [36, 37] or even use microbridges [38] instead of tunnel junctions, where no dielectric is required. The phase qubit in particular suffers from dielectric losses due to its large capacitance. This capacitance is achieved by an external capacitor using low-loss amorphous dielectrics (SiN_x or H:a-Si), or using interdigitated capacitors.

Quasi-particles. At any finite temperature bellow the critical point T_c, only some of the electrons are paired. The two populations (paired and unpaired) can be effectively described within the two-fluid model as parallel resistors. At zero frequency the superconducting fluid (condensate of paired-electrons) has infinite admittance and therefore acts as a short for electric current. At finite frequency the resistance of normal electrons (quasi-particle excitations) has to be taken into account. The dominant contribution to the resistance comes from quasiparticles tunneling through the junction. At $T < T_c$ this resistance becomes exponentially larger than the normal state resistance, $R_{qp} = R_n\,(1 + \exp(1.76T_c/T))$, due to the exponentially smaller quasiparticle density. This leads to a quality factor $Q = \omega_0 RC$ which for $T = 50\,\mathrm{mK}$ is in excess of 10^{17}. Our working temperature is even lower ($\sim 10\,\mathrm{mK}$), however the actual temperature of the electrons is probably higher. Despite the low operation temperature of these devices, significantly higher quasi-particle densities may exist due to non-equilibrium processes. In the Cooper-pair box, densities of $\sim 10\,\mu\mathrm{m}^{-3}$ were reported[39], leading to a calculated lifetime $T_1 \sim 2\,\mu\mathrm{s}$ for the phase qubit [40]. A number of possible sources for these quasiparticles exists: thermal radiation propagating along the coax cables

from higher temperature stages, blackbody radiation at 1 K loading the sample mount, local heating from SQUID switching and more.

External circuitry. To be able to control and measure the system, there must be a finite coupling between the system and the external circuitry, such as depicted in Fig. 1.3a. For a dissipationless coupler, the admittance of the external circuit $Y_1(\omega)$ is transformed into an admittance $\mathrm{Re}Y_2(\omega) = |dI_2/dI_1|^2 \,\mathrm{Re}Y_1(\omega)$ [41]. Each coupling element in the phase-circuit is specifically designed to minimize $\mathrm{Re}Y_2(\omega)$. As shown in 1.3.5, the flux biased phase-circuit implements a high-impedance transformer with $1/\mathrm{Re}Y_2(\omega) > 1\,\mathrm{M}\Omega$, leading to $T_1 > 1\,\mu$s. To couple microwave currents to drive transitions in the circuit an external microwave source is coupled through a small coupling capacitor with $C_c \backsimeq 1$ fF. Since the source admittance is $Y_0 = 1/Z_0$, we find that the transformed admittance, as seen by the phase circuit is $Y'(\omega) = Z_0/\left|Z_{C_c}\right|^2 = Z_0\left(\omega C_c\right)^2$. At the typical qubit frequency, this leads to $T_1 > 10\,\mu$s. A useful feature of the SQUID circuitry is the SQUID-bias-dependent coupling to the phase-circuit, which enables one to dynamically turn on the coupling only during the classical measurement [41].

Two-level defects. A qubit can couple to defect states coherently (see Sec. 3.1). Even for a weak interaction with the defects, given a large enough bath, the qubit can decay with a relatively short time. The energy can then decay within this bath by means of a second coupling mechanism.

Additional decay mechanisms include electric and magnetic dipole radiations and spontaneous emission which are negligible (by several orders of magnitude) relative to the above mechanisms.

1.3.6.2 Dephasing and decoherence

A distinction must be made between two kinds of mechanisms which both lead to the decay of off-diagonal elements of the density matrix[8] ("state coherence"): noise, and coherent coupling to external degrees of freedom. Let's consider a

[8] When written in the eigenstate basis.

qubit for example. Its state is fully described by an excitation probability and a relative phase. The relative phase varies with time according to $\phi = \int \omega_{01}(t)dt$, where $\hbar\omega_{01}$ is the energy difference between the qubit states. When this energy fluctuates in time unpredictably, a measurement device operating at a fixed frequency *dephases* from the system's frequency over time. Since one cannot determine the relative phase with a single measurement but rather by averaging over many instances of the system, at sufficiently long times one measures an average state which is indistinguishable from a statistical mixture. This mechanism is usually called dephasing, and it can result from classical noise sources, and to some extent can be eliminated (see Sec. 2.3). The same measurable effect can result from quantum interactions. Interaction between two quantum systems can lead to a non-separable state in which the two are correlated. This correlation can only be seen by a joint measurement of both systems. For example, when two qubits are in a Bell state: $|\Psi^+\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$, a measurement on only one of the them reveals no coherence at all despite each qubit being in a coherent superposition of $|0\rangle$ and $|1\rangle$. In other cases, the qubit may form such superpositions with other spurious qubits which are subjected to other decoherence processes, thereby reducing the coherence of the controlled qubit. Unlike dephasing, these kind of processes cannot be reversed and are usually called decoherence.

Low-frequency noise. This is the principal source of dephasing in our system. In general, the phase qubit is sensitive to noise in any of the macroscopic parameters I_c, Φ_{ext}[9]. In addition, noise can exist in the control electronics, such as in the local oscillator (LO) and bias electronics (see Sec. 2.2 for more details) which may also lead to dephasing. Our measurement apparatus is specifically designed to eliminate these effects to a minimum. In the phase-qubit, noise in Φ_{ext} is found to be the dominant source of noise. This was confirmed by measuring the qubit

[9]The loop inductance is much more stable than the other parameters. It is determined by the loop geometry which has negligible relative fluctuations in time due to its large size. I_c is also a macroscopic parameter (typical junction sizes are 1 μm^2), however, due to the exponential dependence on tunnel barrier, fluctuation of defects within the junction may cause appreciable fluctuations in I_c by changing the number of micro-conductance channels.

energy fluctuations at symmetric flux bias points[10] [42]. The *energy* fluctuations at both points were found to be anti-correlated. That is, as the energy of the qubit decreased due to some random process, the qubit energy at the symmetric flux point has increased. From the potential form it is straightforward to check that for a noise dominated by critical current fluctuations, the distance between the symmetric points should change, while for flux-dominated noise, the distance remains constant, causing anti-correlated energy fluctuations. The noise was found to have a $1/f^{\alpha}$ power spectrum with $\alpha = 0.95$ and an extrapolated amplitude $S_{\Phi}(1\,\mathrm{Hz}) = 4\,\mu\Phi_0/\sqrt{\mathrm{Hz}}$, consistent with flux noise measurements on SQUIDs. The scaling and its amplitude are found to be universal, that is, only weakly dependent on a wide variety of parameters, such as geometry and materials. By measuring the flux of a field-cooled a DC-SQUID as a function of temperature below T_c, Sendelbach *et al.* [43] have been able to show presence of magnetic spin defects on the surface of the superconducting electrodes, and to extract their densities. At sufficiently small magnetic fields, where no vortices are created, a relatively large magnetic field is measured at low temperatures, indicating a high degree of ordering of these spins below $\sim 100\,\mathrm{mK}$, suggestive of the typical coupling energy scale between these spins. It is possible that a surface treatment will reduce the defect densities and thereby the observed noise, however such an experiment has not yet been reported.

Two-level defects. Defect states can cause decoherence through quantum interactions with the qubit. As described in Sec. 3.1, these are reduced by using smaller junctions (having a smaller number of these defects) and biasing the qubit far from these defects, or by fabricating epitaxially grown junctions[37, 36].

[10] The phase circuit energies within a specific well are symmetric with respect to a certain flux bias (see Sec. 2.3 for more details). The initial bias points at which the qubit was measured were selected to yield the same qubit energies

CHAPTER 2

Materials and Methods

In this chapter, the implementation of a Josephson phase-circuit and its measurement are described in detail. We discuss the design aspects of a phase-circuit device and its fabrication process. Then, we go over the measurement apparatus, including the dilution fridge and wiring, our dedicated electronics for control and measurement, and corrections we implement to eliminate imperfections in the control signal. Observation of quantum coherent phenomena in a superconducting device requires a relatively complex protocol for control and measurement. In the following we explain the basic operation of the Josephson phase circuit and how to obtain its characteristics. This can be considered as an "operation manual" for the device.

2.1 Device design and fabrication

The starting point for designing a device is its phase-circuit parameters: I_c, C and L. These will determine the qubit frequency and anharmonicity at its working point, as well as its sensitivity to flux noise. In practice, the values obtained in fabrication are only within 10 % of the designed ones, however since the above requirements allow for a large range of parameters, this does not pose a problem. Once the phase-circuit parameters are set, the flux bias, microwave control and SQUID circuits are chosen so as to minimize external dissipation and noise while maintaining easy control.

2.1.1 Device parameters

The qubit frequency has a lower bound, set by thermal excitation. Dilution refrigerators offer the most convenient method of cooling devices to the lowest technologically attainable temperatures ($\sim 30\,\mathrm{mK}$)[1]. To eliminate thermal population in the excited state we require the frequency to be at least 10 times higher than the thermal frequency $k_B T/h \approx 500\,\mathrm{MHz}$, above 5 GHz. The higher bound is set by current microwave technology to be $\sim 20\,\mathrm{GHz}$. Above this frequency, it is currently much more difficult to shape pulses and avoid significant signal loss in the cables.

The anharmonicity ($f_{01} - f_{12}$) must be large enough to allow a large number of single-qubit gates to be applied within the coherence time. Since the present coherence time is approximately 100 ns, it is desirable to be able to perform high-fidelity gates within $\sim 10\,\mathrm{ns}$. This restricts the anharmonicity to be larger than $\sim 200\,\mathrm{MHz}$. Although much shorter gates can in principle be achieved with much larger anharmonicities, commercially available arbitrary waveform generators (AWGs) are currently much too expensive above a $\sim 500\,\mathrm{MHz}$ bandwidth and at sampling rates greater than 1 GS/s.

As the sensitivity of the circuit to flux noise decreases with increasing inductance, it is generally desirable to have a large inductance. However, as pointed out in 1.3.5, a larger inductance (relative to the critical-current-dependent Josephson inductance) leads to multiple potential wells, which require a longer reset time. The capacitance itself must be sufficiently large so as to be in the phase-qubit regime ($E_J \approx 10^4 E_C$) . Using an eigenenergies solver, one finds that parameters in the vicinity of $I_c = 1\,\mu\mathrm{A}$, $C = 1\,\mathrm{pF}$ and $L = 800\,\mathrm{pH}$ (within $\sim 10\,\%$ of variation) are a good choice to satisfy the above requirements.

The bias inductor must have a small mutual inductance with the phase-circuit

[1] Dilution refrigerators can reach lower temperatures, however it is much more difficult to thermalize the electrons in the devices to these temperatures. Adiabatic demagnetization refrigerators are a possible alternative, however their cooling periods are much shorter making it more difficult to perform long measurements.

loop, so that the impedance coupling is sufficiently large to protect the qubit from decohering through this channel. On the other hand, it must not be too small, otherwise a too large current will be required to tilt the potential sufficiently ($M\Delta I \sim 2\,\Phi_0$). A value of several pH for the mutual inductance and $\sim 100\,\mathrm{pH}$ for its self induce allows for reasonable bias currents (smaller than $\sim 1\,\mathrm{mA}$) to fully shift the potential and obtain a lifetime which is not be limited by the bias coupling.

2.1.2 The SQUID circuit

The SQUID circuit employs a unique design which reduces the coupling to the qubit while its critical current is not measured. In a standard dc-SQUID, a superconducting loop is interrupted by two identical Josephson junctions. Since the phase difference across the junctions is related to the magnetic flux threading the loop, one can use the Josephson relations to find that this configuration has an effective flux-dependent critical current, $I_c' = 2I_c \sin(\pi\Phi/\Phi_0)$.

The SQUID used in this work is schematically shown in Fig. 2.1a. Its loop consists of three Josephson junctions and it is inductively coupled to the phase-circuit loop. On one branch we have a single junction with a critical current I_c^{sq}, while on the other branch we have two junctions, each with a critical current αI_c^{sq}. The loop is biased directly with a current bias I_{sq}, which flows symmetrically through the two branches into the ground, and is also biased inductively by the qubit loop with a current I_q which flows circularly in the loop. Using the Josephson relations for these currents we get:

$$I_{sq} = I_c^{sq} \sin\delta + \alpha I_c^{sq} \sin\frac{\delta}{2}, \tag{2.1}$$

and

$$I_q = I_c^{sq} \sin\delta - \alpha I_c^{sq} \sin\frac{\delta}{2}. \tag{2.2}$$

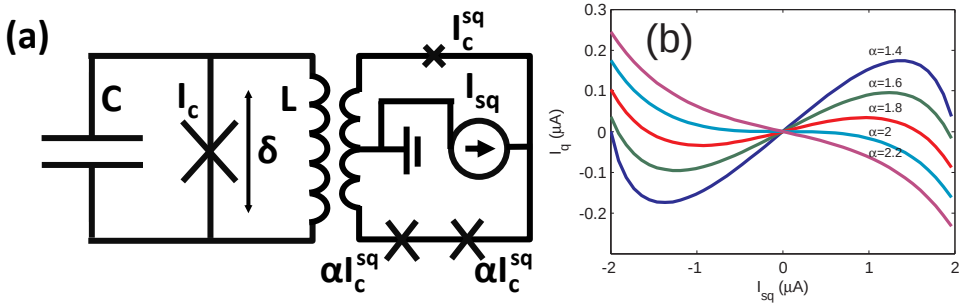

Figure 2.1: On-chip SQUID. (a) Equivalent circuit (b) SQUID sensitivity to qubit current.

Using these two equations, one finds a relation between I_q and I_{sq}, plotted in Fig. 2.1b, for several values of α. For α smaller than 2, there exist bias values of I_{sq}, where the qubit current is to first order insensitive to the SQUID current, $dI_q/dI_{sq} = 0$. The value of α is typically chosen to be ~ 1.7 so that the optimal point is not too close to the effective critical current of the SQUID $3\sqrt{3}I_c^{sq}/2$ and not too close to 2.0, given the typical uncertainty in I_c^{sq} in fabrication. The value of I_c^{sq} is usually chosen to be identical to I_c, the SQUID's critical current.

At each switching event the SQUID heats up, creating quasi-particles in the qubit loop which cause energy decay. To minimize heat dissipation on the chip, a shunt resistor is placed on the SQUID bias line, outside the sample mount. The resistance is chosen to be smaller than the effective SQUID resistance in its normal state so that some of the bias current is shunted to the ground through the resistor. A too small resistor reduces the voltage in the switched state, reducing the signal-to-noise ratio (SNR). We typically use a 30 Ω-50 Ω resistor.

2.1.3 Circuit geometry and materials

The junctions have a large fabrication uncertainty. This is due to the exponential sensitivity of the critical current on barrier thickness, controlled by oxidation time and pressure. Variations of only a few angstroms in the thickness can lead to a useless device, and therefore each die is designed with a different junction area to increase the yield. The upper electrode of the junction in each die has a triangular shape (see Fig. 2.2a) but a different offset relative to the base electrode so as to

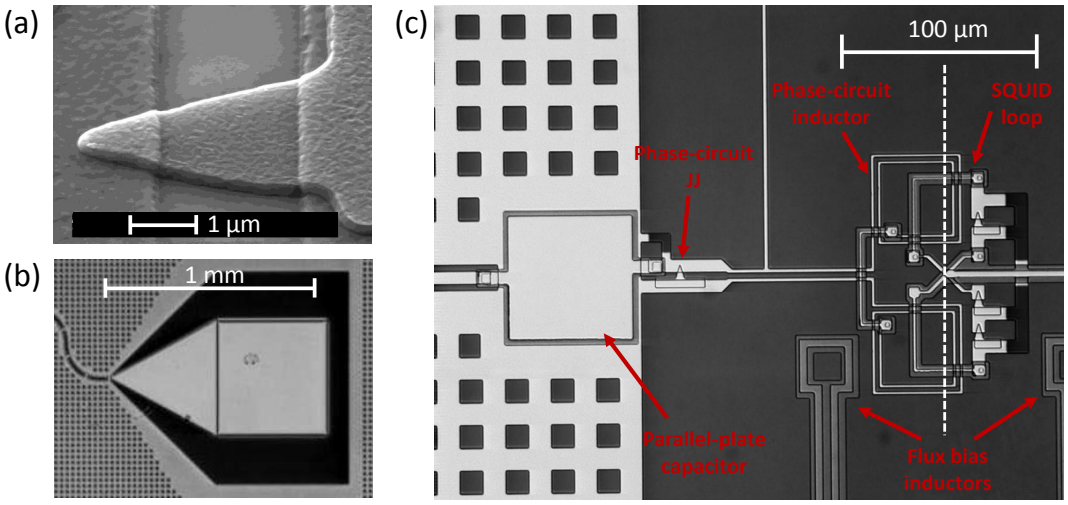

Figure 2.2: Phase circuit device images. (a) Scanning electron microscope image of the Josephson junction. (b) Microwave launcher. (c) The phase circuit, SQUID and flux bias. Dashed line shows the flux bias symmetry axis where $B = 0$.

vary the area from die to die. The area is chosen to be about 1 μm^2 so as to reduce the number of TLSs (see sections 1.3.6 and 3.1). This results in additional uncertainty in the critical current since it is close to the standard optical lithography resolution. This problem is relieved by the area variation technique as well.

To keep the large capacitance of the circuit with the smaller junction, an external capacitor is added. The capacitors of the devices used in this work have either a parallel plate (see Fig..2.2c) or an interdigitated geometry. The interdigitated geometry enables a dielectric-free (and therefore loss-free) capacitor. It consists of two fork-shaped electrodes placed one against the other. This geometry increases their mutual surface and therefore the capacitance. Devices with parallel-plate capacitors have either a-SiN$_x$ or a-Si:H dielectrics which have a relatively low loss at microwave frequencies and low temperatures[34, 44].

The qubit inductor is designed with a figure of 8 geometry. This makes it insensitive to spatially homogeneous stray magnetic fields and only sensitive to field gradients. It is therefore called a gradiometric inductor. The flux-bias inductor is composed of two counter-wound loops, displaced from each other. The magnetic field due to these loops cancels out along a straight line perpendicular to the

vector connecting these loops and passing midway between them. To minimize coupling of the SQUID to the qubit bias, the SQUID loop is placed along this line. The qubit inductor is placed parallel to the SQUID loop, but offset from the symmetry axis to have non-zero mutual inductance with the flux-bias. To make sure that the coupling between the qubit and the SQUID is zero at zero SQUID bias, the SQUID loop is placed parallel to the qubit loop but its inductor is split symmetrically with respect to the ground, so that the mutual inductance is canceled out (the equivalent circuit is shown in Fig. 2.1a). When the SQUID bias is close to the critical current, the inductance becomes non-linear and different for the two branches. The mutual inductance with the qubit loop is no longer canceled and the qubit can be measured.

The microwave line is a $50\,\Omega$ transmission line with a coplanar geometry that couples to the circuit through a small capacitor (interdigitated geometry). To provide an impedance matched interface between the large connection pads and the narrow transmission line, the coplanar waveguide is linearly scaled down in a launcher (see Fig..2.2b).

A second identical circuit (including SQUID, flux bias and microwave line) is placed symmetrically on the chip, with an optional small coupling capacitor to enable resonant interaction between the qubits.

The material used for the superconducting electrodes is Aluminum. This is mostly due to its relatively high-quality oxide that can be used for the creation of the Josephson junction. In addition, its critical temperature is well above the working temperature, reducing equilibrium quasi-particles to a minimum, and it is easy to etch. The whole circuit is surrounded by a ground grid with thickness of $\sim 1\,\mu$m, to make it less favorable for vortices to form[2]. For cross-over dielectrics (to be used for multilayer inductors), a-Si:H or a-SiN_x are used. The substrate is made of sapphire, which has low dielectric losses.

[2] Although aluminum is a type-I superconductor, thin films show type-II behavior [45]

2.1.4 Fabrication

The devices used in this work were fabricated at the nanofabrication facility of the University of California, Santa-Barbara (UCSB) by the Martinis group. The main steps of fabrication are outlined here, however this outline contains only partial information about the process. More detailed information is available in a thesis by Markus Ansmann[3].

The device includes 5 layers, fabricated in 7 steps. The parallel-plate capacitor and inductor require two layers with an insulator between ("base wiring" - insulator - "top wiring"). The additional top layers consist of the thin junction oxide and the junctions upper layer ("junction layer").

(a) First, a thin aluminum film (150 nm) is deposited on a 2" sapphire (crystalline Al_2O_3) wafer via sputtering.

(b) The base wiring structure is then patterned into the aluminum layer via optical lithography, followed by inductively coupled plasma (ICP) etching.

(c) After the photo-resist layer is removed, the insulating layer of a-Si:H or SiN_x (250 nm) is deposited using plasma enhanced chemical vapor deposition (PECVD).

(d) Holes are drilled through the insulating layer (vias) down to the base wiring using ICP to allow for an electric short between the base and top wiring.

(e) The top wiring aluminum layer (200 nm) is deposited via sputtering.

(f) The top wiring layer is milled with Argon ions to remove the native oxide layer. This is an uncontrolled ~ 2 nm thick layer that forms quickly when exposing aluminum to air. After milling, a controlled amount of oxygen is allowed into the sputtering chamber for a precisely calibrated time, to form the oxide junction barrier. An additional aluminum layer (150 nm) is then deposited.

(g) The junction layer is patterned using optical lithography. It is then etched using Argon-Chlorine ICP etch which gives a fine control over the etching depth, which has to be precisely controlled here due to the oxide layer which is too thin (~ 1 nm) to prevent etching into the top wiring layer. The top electrode layer is

[3] http://web.physics.ucsb.edu/~martinisgroup/theses/Ansmann2009.pdf

then patterned and all the unnecessary insulation is removed (to reduce dielectric loss to a minimum). Finally, shorts between the junction electrodes that prevent charging effects from blowing the junctions during plasma etching, are removed via wet chemical etching.
A photo-resist protective layer is deposited on the wafer. The wafer is diced into 6 mm $\times$ 6 mm dies, which are ready for measurement.

Three phase-circuit devices were used in this work: Device_A, Device_B, Device_C. Device_A (1^a_b, coupled, parallel plate capacitor) was fabricated by Matthew Neeley using a-Si:H for the capacitor and cross-over wiring, while Device_B (9, 4, uncoupled, interdigitated capacitor) and Device_C (7, 5 coupled, interdigitated capacitor) were fabricated by Radoslaw Bialczak using a-SiN$_x$ for cross-over wiring. For measurements described in Sec. 3.1 and for standard state tomography on the qubit and qutrit (Sec. 3.4) we used Device_A. For measurements described in Sec. 3.2 we used Device_B (except for the holding measurement, where we used Device_C). For measurements described in Sec. 3.3 and on Wigner tomography measurements (Sec. 3.4) we used Device_C.

2.2 Measurement apparatus

The device is mounted in an hexagonal aluminum box, designed to provide further radiation isolation, and an electrical interface between the device and the signal-carrying coaxial cables. The box supports six coaxial feeds to the device through SMA connectors, which are impedance matched to 50 Ω to allow signals of up to 20 GHz to be transmitted. This design is typically used for controlling and measuring two qubits (each having its own flux bias, microwave and SQUID lines). The sample mount is connected to the mixing chamber plate of a dilution refrigerator, cooled down to 10 mK.

2.2.1 The dilution fridge

We cool down our devices with an Oxford-Vericold DR-200, cryo-free dilution refrigerator (see Fig. 2.3b). A dilution refrigerator, exploits the phase-separation of a ^{4}He-^{3}He mixture below $\sim 1\,\mathrm{K}$ to provide cooling power down to a few millikelvins. At these temperatures, ^{4}He is already in its superfluid phase, while ^{3}He is a normal fluid. Below the tri-critical point $T \lesssim 0.86\,\mathrm{K}$, the mixture separates into a ^{3}He rich phase (top) and a ^{4}He rich phase (bottom). The concentration of ^{3}He in the ^{4}He phase decreases with decreasing temperature. In the ^{4}He phase, the enthalpy of ^{3}He atoms is higher, and therefore an excess energy is required for a ^{3}He atom to cross the phase boundary. By diluting the ^{4}He phase of ^{3}He atoms, the system is driven out of equilibrium, and will attempt to restore it by allowing ^{3}He atoms from the ^{3}He rich phase to cross the phase boundary. The excess energy required for that is provided by the walls of the mixing chamber, where phase separation occurs, thereby cooling the mixing chamber and the metal plate that is in thermal contact with. The dilution process stops at a $\sim 3\,\mathrm{mK}$, where ^{3}He atoms pair-up to form a superfluid. The dilution of ^{3}He atoms from the ^{4}He phase is done by pumping on the ^{4}He side of the mixing chamber and heating it in a still to preferentially evaporate ^{3}He atoms (at these temperatures ^{3}He atoms evaporate at a much faster rate).

In our dilution unit, a Cryomech pulse-tube system cools down a ^{4}He reservoir down to $\sim 4\,\mathrm{K}$ via a compression-expansion cycle, controlled with a rotary valve. The ^{4}He gas is in thermal contact with the 70 K and 4 K stages of the cryostat. All temperature stages (plates) are thermally decoupled from each other to reduce heat load at low temperatures. To provide the initial cooling to 4 K of the lower plates (still plate, cold plate and mixing chamber plate), the ^{4}He-^{3}He mixture is cycled through a precool circuit, which allows heat exchange with all stages, eventually bringing all the lower plates down to 4 K ($\sim 20\,\mathrm{hrs}$). At this point, the ^{4}He-^{3}He mixture is pumped out from the precool circuit and run through the condensing line. A compressor forces the mixture through a narrow tube ($< 1\,\mathrm{mm}$) at the

still stage, which condenses the mixture via a Joule-Thomson process, eventually bringing it down to $\sim 800\,\mathrm{mK}$. Then, the dilution cycle is initiated by heating the still and pumping out ^{3}He from the still with a turbo pump.

The mixing chamber plate temperature is monitored with a Ruthenium-Oxide resistor, calibrated down to a few mK with ^{60}Co nuclear thermometry. The base temperature of our fridge is $10\,\mathrm{mK}$[4], with a specified cooling power $> 200\,\mu\mathrm{W}$ at the cold plate ($100\,\mathrm{mK}$). All temperatures in the fridge are measured using a 4-wire Lakeshore resistance bridge, which also provides control over mixing-chamber and still heaters, giving control over mixing chamber plate temperature and cooling power.

To provide the necessary thermal isolation from stage to stage, we also shield from thermal radiation with metal cans at the still, $4\,\mathrm{K}$ and $70\,\mathrm{K}$ stages. The coax lines used to transmit electrical signals are also thermally anchored at all stages to reduce heat load. To provide a thermal contact with the center conductor of the coax, we use filters or attenuators at the $4\,\mathrm{K}$ and mixing chamber stages[5]. This reduces both the heat load at the mixing chamber stage and Johnson noise in the signals that may cause qubit dephasing. The filter boxes are filled with thermal grease between the ground and center pin, while attenuators provide a $\sim 100\,\Omega$ resistance between the ground and center conductor which give a sufficient heat conductance. The outer can serves as a vacuum seal, and allows for a $10^{-6}\,\mathrm{mbar}$ of pressure during operation, achieved with an external turbo pump and cooling induced cryo-pumping. The outer can is surrounded by three additional cans, made of a high magnetic permeability (a few 10^4) iron-nickel-copper alloy, called μ-metal. These are used to screen low-frequency stray magnetic fields, which may cause dephasing in our flux-sensitive device. We measure a $> 10^3$ attenuation inside the fridge with the mounted cans, however this should be measured period-

[4]The whole fridge is electrically connected to ground only through the still heaters, to reduce ground noise from affecting temperature measurement. However, we find that temperature measurement of the mixing-chamber's RuO_x is affected by heavy electrical equipment in the lab (such as air-conditioning).

[5]The center conductor is electrically and thermally insulated from the outer conductor with Teflon.

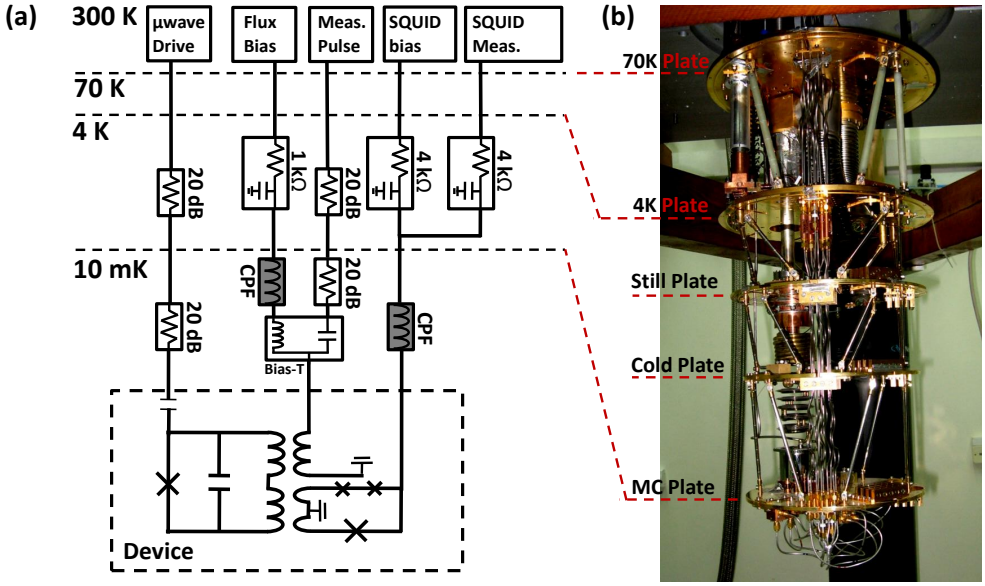

Figure 2.3: Fridge wiring. (a) schematics of coaxial lines connecting the device at 10 mK and control\measurement electronics at room temperature. (b) A photo of the dilution fridge, showing the different temperature stages.

ically as the magnetic properties of these materials are known to deteriorate with time.

2.2.2 Fridge wiring

A schematic representation of the fridge wiring is shown in Fig. 2.3. Measurement and control signals are transmitted through semi-rigid coax lines between room temperature and the mixing chamber plates. The coax lines support signals of up to 20 GHz and are terminated by SMA connectors. To reduce thermal coupling between the stages, the coax lines between room temperature and 4 K are made of stainless-steel which is a poor conductor relative to other metals, and the coax lines between 4 K and the mixing chamber are made of Niobium. Niobium becomes superconducting bellow 9.3 K, and its electronic heat conductivity exponentially vanishes at low temperatures $T < T_c$. Standard tin-coated coaxes are used to connect between the sample mount and the Niobium coaxes at the mixing chamber stage.

High-frequency signal cables (used for microwave control and measurement pulses) are interrupted by 50 Ω, 20 dB attenuators at the 4 K and mixing chamber plates. These provide cold attenuation to the required signal level at the device. Cold attenuation is used to keep the large SNR at low signals levels, by cooling down the resistive elements that provide the attenuation to reduce their thermal (Johnson) noise. The attenuators are distributed among the two plates to both reduce the heat load on the mixing chamber alone from the large initial signal dissipated in the attenuators and to enable thermal contact for the center conductor of the coax at 4 K.

Low-frequency signals (SQUID bias and measurement, and flux-bias) are interrupted at 4 K by low-pass RC filters. These include a split series resistance (4 kΩ for the SQUID lines and 1 kΩ for the flux bias line) and 300 pF split capacitance to ground which block noise above $\sim$ 1 MHz. The series resistance serves also as a bias, limiting the applicable currents to a few mA. The RC filter stops attenuating at larger frequencies ($\gtrsim$ 200 MHz) and therefore we add a copper-powder filter in series. This is made of a coil, immersed in an non-conductive epoxy-copper-powder mix. RF radiation (above $\gtrsim$ 100 MHz) is absorbed in these filters by plasmons on the surface of the fine copper grains ($\sim 10\,\mu$m), and provide large attenuation ($\sim$ 60 dB) up to $\sim$ 15 GHz[6].

The flux bias coil in the device must be fed with the low-frequency flux-bias signal (flux channel) and the high bandwidth flux-bias (measurement pulse channel), which are generated by separate sources, and go through different lines inside the fridge. We combine the two in the mixing chamber stage using a bias-T. This is a 3-port component, which includes two input ports for low-frequency and high-frequency signals and an output port. The high-frequency input has a capacitor in series to block low-frequency signals from leaking out, and the low-frequency input has an inductor in series to block high-frequency signals from leaking out. Our bias-T does not require a capacitor, since the high-frequency sig-

[6]More details on the copper-powder filters is available in a thesis by Ya'ara Rofe

nals are shunted to the ground by the bias coil in the device[7]. For an inductor we use a Mini-Circuits ADCH-80 RF-choke with $\sim 7\,\mu$H inductance. This gives a $\sim 7\,$GHz bandwidth between the RF-input and the output port, which is more than enough for the measurement pulse channel.

2.2.3 Technical noise

Electric noise has negative effects on our measurements: Intermediate frequency noise causes dephasing of the qubit, while low frequency noise causes drifts in our calibrations (such as mixer offsets, SQUID cutoff offsets and qubit frequency) and makes it difficult to distinguish between circuit flux states via the SQUID. While some of the noise sources are fundamental and difficult to eliminate (thermal noise, shot noise, two-level fluctuators), other sources exist in any measurement apparatus which are technical and can be reduced with proper design. One of the main sources of technical noise is ground noise, resulting from nearby instruments connected with the same conductor to the ground, and from fluctuating magnetic fields that induce emf on wire loops. In our system, these are supplemented by acoustic noise generated by the fridge[8], coupled capacitively to the coax lines. We characterize these noise sources using a low-noise amplifier, specifically designed for measuring thermal noise at the low-frequency regime.
We use a 40 dB low noise amplifier (using two Texas Instruments OPA602AP op-amps), having a $12\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ voltage noise. On Fig. 2.4 we show several noise spectra measured using an oscilloscope on the fridge signal ports for various configurations. In Fig. 2.4a, we measure noise on a $1\,\mathrm{M}\Omega$ resistor while the fridge is off and the scope is connected to the same electric outlet. The noise is dominated by Johnson noise from the resistor $V_{\mathrm{JN}} = \sqrt{4Rk_BT} \approx 10^{-7}\,\mathrm{V}/\sqrt{\mathrm{Hz}}$ at low frequencies, however this noise is not entirely transmitted due to low-pass filtering from

[7] This coil has a low impedance relative to the $1\,\mathrm{k}\Omega$ bias resistors at GHz frequencies.
[8] The fridge generates a broadband acoustic noise from pumps, compressors and the pulse tube (the closed-cycle ^{4}He cooling system).

the coax[9]. At higher frequencies, the noise drops to the background level of the amplifier. In Fig. 2.4b we perform the same measurement, while the fridge is on and the resistor is at 4 K. While the thermal noise is now comparable to the background, we get a strong noise in the $10 - 10^4$ Hz range. Acoustic noise from the pulse tube causes microphonic noise on the resistor - $V_n = RVdC/dt$, where R is the resistance (1 MΩ) and V is a pre-existing voltage on the cables (due to ground noise for example). This noise is much stronger while the fridge is not running (2.4b), probably due to the lack of vacuum, which provides acoustic isolation. In Fig. 2.4d we plot the noise on a the SQUID line (having a 4 KΩ resistor at 4 K). Despite the pulse tube being on, the microphonic noise is much less pronounced since the resistor is 250 times smaller. This noise is further reduced in Fig. 2.4e where we turn off the pulse tube. In both of these measurements we see strong narrow peaks at integer multiples of 50 Hz. These are the result of ground loops, however they are absent in Fig. 2.4a and 2.4b where we have connected oscilloscope on the same electric outlet as the fridge. Finally, in Fig. 2.4f we measure the spectra on a 50 Ω, 20 dB attenuator at room temperature (similar to the microwave control lines) while the fridge is off.

While fridge noise appears to be fairly large, we expect it to have only a minor effect on qubit dephasing. To calculate qubit dephasing due to low frequency noise, we extract the current noise in the flux channel from the noise measurements[10]. In the simulation we start off with white noise having a spectral noise density of $1/\sqrt{\text{Hz}}$, and multiply it in the spectral domain by the measured spectral noise density $\tilde{I}_n(f)$. Our noise has a floor value of $40\,\text{pA}/\sqrt{\text{Hz}}$, with several peaks, the largest of which has an amplitude of $20\,\text{nA}/\sqrt{\text{Hz}}$. We then transform back to the time domain to obtain $I_n(t)$. The induced noise in the qubit loop is then given by $I_n^q = I_n/\sqrt{Z_0\,(L/M)^2}$, where $Z_0\,(L/M)^2 \approx 4.5\,\text{M}\Omega$, and the correspond-

[9] Our coaxes have roughly 30 pF/m capacitance to the ground. Together with the large resistance, we get a few kHz of bandwidth

[10] We use the measured voltage noise from the SQUID line (see fig.2.4d), divided by the resistor value to find the current noise I_n. As expected, separate measurements confirm similar noise in the flux bias line. Noise in the wide-bandwidth line is ignored in our calculation since its floor level is smaller than the amplifier noise, as seen in fig. 2.4f.

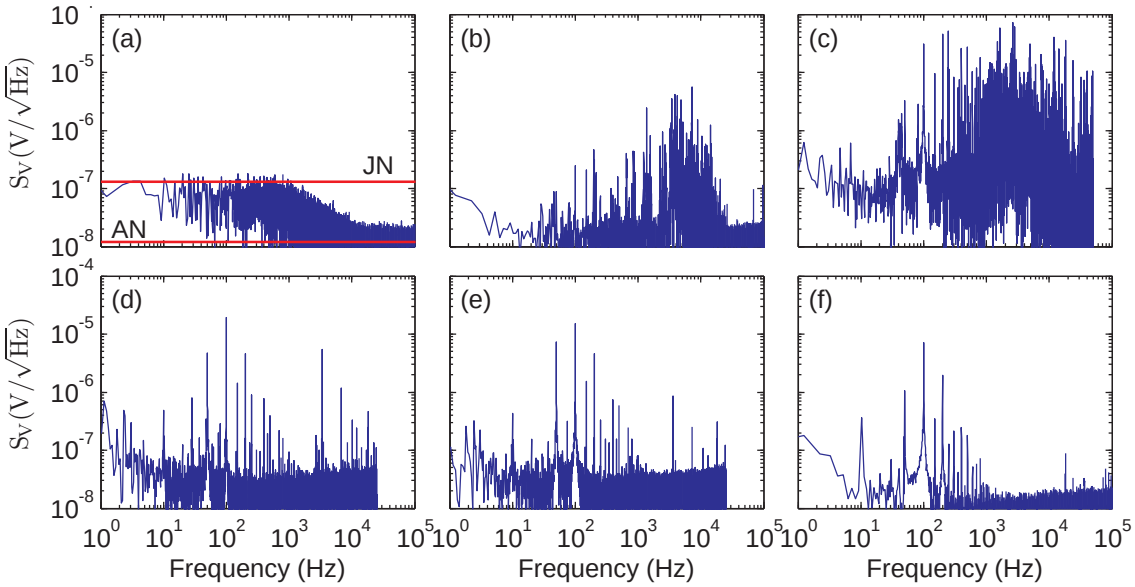

Figure 2.4: Noise spectra for various configurations. Measurement on (a) 1 MΩ resistor at room temperature located on the 4 K plate, while pulse tube is off. The horizontal lines show the calculated Johnson noise (JN) level for this temperature and the added amplifier noise (AN). (b) the same resistor at 4 K (actual temperature), with the fridge running and cans mounted. (c) the same resistor at room temperature, while only the pulse tube is on and the cans not mounted. (d) SQUID bias line (with 4 kΩ resistor) on the 4 K plate, at 4 K while fridge is on and pulse tube is on. (e) SQUID bias line (with 4 kΩ resistor) on the 4 K plate, at 4 K while fridge is on and pulse tube is off. (f) 20 dB attenuator located on the 4 K plate at room temperature, while fridge is off. (a) and (b) are measured while the oscilloscope is connected on the same power outlet as the fridge.

ing noise in the qubit transition angular frequency is $\omega_{01}^n(t) = L\frac{\partial\omega_{01}}{\partial\Phi}I_n^q(t)$. From qubit spectroscopy (see Sec. 2.3), we extract $\frac{\partial\omega_{01}}{\partial\Phi} \approx 2\pi \times 25\,\mathrm{GHz}/\Phi_0$ at the typical working point (large anharmonicity). The qubit phase noise is then given by $\phi_n(t) = \int^t \omega_{01}^n(t)dt$, and the qubit coherence signal is proportional to $\cos\phi_n$. Since we are interested in the low frequency noise ($f < 10^5$ Hz) we assume a constant angular frequency during a single experiment $t < 10\,\mu$s (ϕ_n changes linearly with time), and average the signal from multiple experiments, repeated at 10 kHz up to 1 s. This leads to a Gaussian decay envelope for both white noise and $1/f$ noise and is considered as *inhomogeneous dephasing*. Our simulation yields a typical decay time larger than 1 μs, whereas a similar simulation for $1/f$ noise using the well known measured flux noise[42, 34], yields $T_2^* \sim 300$ ns. We find that most of the contribution to qubit dephasing comes from the large peaks in the spectrum

which are believed to be caused by ground loop noise (they come in multiples of 50 Hz).

Despite the large noise from the pulse tube, it has no measurable effect on qubit dephasing. We performed several measurements of qubit dephasing time via Ramsey sequences (see Sec. 2.3), where the pulse tube, pumps and compressor have been turned off and found no effect on T_2^*. However, when we disconnected the still heater from the ground[11] we found an increased T_2^* (from 80 ns to 100 ns). This increase is consistent with the simulation, giving a $\sim 1\,\mu$s contribution coming from the flux line. This suggests a large contribution of ground noise to qubit dephasing, and should be further checked.

2.2.4 Control and measurement electronics

Our experiments require the generation of arbitrary pulses with large bandwidth at gigahertz frequencies. We must also be able to apply short measurement pulses, bias the device and measure the SQUID's critical current at the end of each experiment. All signal channels must be synchronized and support a high repetition rate (> 1 kHz) to obtain state probabilities with a reasonable SNR at a rate of ~ 1 Hz. Satisfying these requirements currently requires an expensive equipment $>$ \$100k, which probably lacks the flexibility and scalability that will surely be needed in future experiments. During the last decade, there has been a surge of custom built data acquisition and control electronics by labs around the world, based on FPGA technology. FPGA (Field Programmable Gated Array) is a programmable integrated circuit, which allows a user to quickly design and implement (virtually) arbitrary digital data processing at high speeds. While traditional laboratory experiments rely only on desktop computers to perform an experiment specific computation and to control an essentially analog experimental apparatus, communication times with the electronic equipment would normally limit the rate at which these experiments can be repeated. FPGA-based electronics solves

[11] The fridge is grounded only through the still heater. All the other connections are isolated

this limit by enabling an on-device, user-defined data processing.
We use dedicated FPGA-based electronics, designed by the Martinis Group[12] at UCSB. Our control and measurement apparatus (see Fig. 2.5) includes two GHz DAC boards, a microwave pulse shaping setup, a multichannel biasing board (Fast-Bias) and pre-amplifier board (Preamp-Trigger).

2.2.4.1 GHz DACs

The main electronic boards are the GHz DACs. Each GHz DAC includes a 2-channel Arbitrary Waveform Generator (AWG) which is used to shape microwave pulses and generate measurement pulses. In addition, these boards send the biasing sequences to the Fast-Bias board and acquire switching data from the Preamp-Trigger board. Multiple GHz DAC boards are connected together in a daisy-chain configuration and operate in synchronization via a trigger which is provided by a primary board. The boards use an internal 1 GHz clock, locked to an external 10 MHz ultra low-phase noise reference clock.
AWG. Each of the two AWG channels in the board includes a 14-bit, 1 GS/s DAC with a programmable sequence of length 8192. This translates to 8.192 μs long microwave pulses, which is currently much longer than the coherence time and lifetime of the qubit. The bandwidth of arbitrary waveforms is 500 MHz (Nyquist frequency), however, it is further reduced by filters[13] to about 300 MHz to remove clock feedthroughs and aliasing problems. Each DAC outputs a 0 V-0.5 V signal at 50 Ω in two channels V^+ and V^- related by $V^- = 0.5 - V^+$. These are later transformed into two identical signals (up to a "-" sign) in the range -0.25 V-0.25 V using a differential amplifier. The DACs have a typical jitter smaller than ~ 1 ps, which is crucial for coherent control experiments at < 20 GHz frequencies.
Control and acquisition. Each GHz DAC communicates directly with a computer to receive arbitrary waveforms and an experiment sequence and to send

[12] http://web.physics.ucsb.edu/~martinisgroup/
[13] We use custom built Gaussian filters (having a Gaussian transmission profile in the frequency domain), with a 3 dB point of 275 MHz.

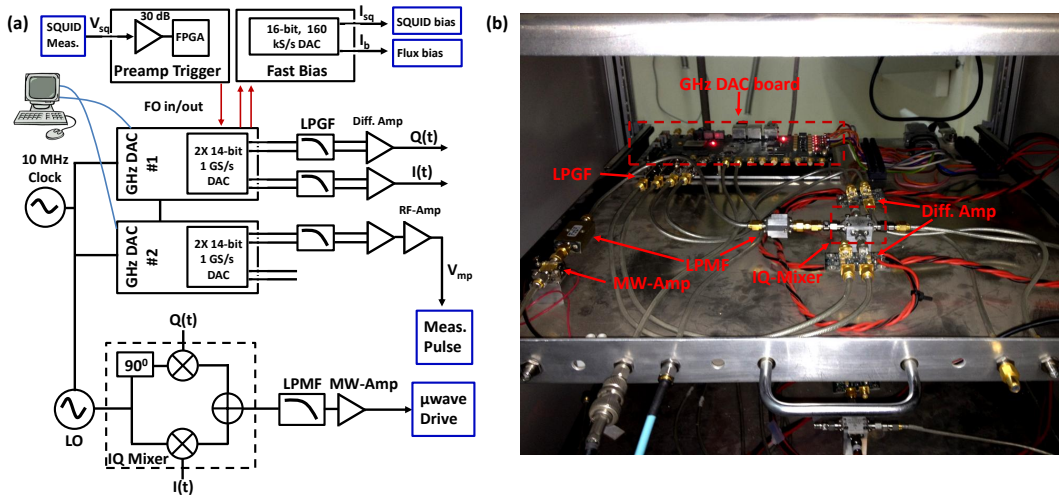

Figure 2.5: Control and measurement electronics. (a) Schematics of the electronics setup. Up: low-frequency bias card and pre-amplifier. Down: microwave pulse shaping setup. (b) A photo of pulse-shaping drawer, including the primary board for control and data acquisition.

back switching data using an Ethernet port. The experiment sequence includes a time-ordered set of commands that tell the biasing board which voltages to output, when and for how long (these are sent to the Fast-Bias board via fiber optics cables). It also includes a timer command that counts an internal 250 MHz clock cycles until an incoming input pulse comes from the Preamp-Trigger board (via a fiber-optic cable), indicating a SQUID switching event. The experiment sequences are executed sequentially and the switching times are registered and sent back to the computer. The repetition rate is only limited by the required length of each experiment[14].

2.2.4.2 Biasing and measurement

Flux bias and SQUID bias signals are generated by the Fast-Bias. This board has 4 low-noise channels, each capable of outputting voltages in the range -2.5 V-2.5 V with 16 bits of resolution. The board has an FPGA chip which is used to decode

[14]In our case it is limited to $\sim$ 10 kHz, mostly due to the low-pass filtering on the Fast-Bias channels.

the incoming GHz DAC commands and perform additional tasks. The FPGA contains a programmable internal clock that enables reading programming instructions at a rate of 1 command per $\sim 6\,\mu$s. Each channel is equipped with two DACs: one DAC for coarse low-noise signals (16 bits of resolution over the range -2.5 V-2.5 V, and $25\,\text{nV}/\sqrt{\text{Hz}}$ DAC noise) with fast or slow slew rates (RC=1 μs and 10 μs) and another DAC for fine lower-noise signals (16 bits of resolution over the range 0 V-2.5 V, and $7\,\text{nV}/\sqrt{\text{Hz}}$ DAC noise) with a slow slew rate (RC=10 μs). In all of our experiments, we use the coarse DAC for both the SQUID and Flux bias[15].

Once the SQUID is switched, a voltage develops on the junction (100 μV), which is amplified by a factor $\sim 10^3$ by low-noise amplifiers, located on the Preamp-Trigger board. The board has 4 identical input channels (only one is needed per qubit), each in line with programmable low-pass and high-pass filters to remove noise from the measurement[16]. The board is equipped with an FPGA chip which sends a pulse through a fiber-optic cable once the amplified signal crosses a user defined threshold. Programming instructions of threshold and filtering are carried out through a separate card, connected to the computer through a fiber-optic cable. This card also provides direct monitoring of all channels on the Fast-Bias and Preamp-Trigger on a standard oscilloscope.

Both the Fast-Bias and Preamp-Trigger have a floating power source (+6 V and -6 V batteries[17]) to eliminate direct electrical connection to the ground that may result in excess noise in the bias control and measured signals.

[15]The flux bias DAC is usually set to fast, while the SQUID is set to either slow or fast. Slower SQUID biases may provide a better distinguishability between qubit wells due to the corresponding larger time separation between switching events, but lengthens the experiment sequence. Despite working with the faster DAC, we find that it takes $\sim 24\,\mu$s for the bias to stabilize on the voltage set-point with an error smaller than 100 μV (or $\Phi_{\text{ext}} \approx 100\,\mu\Phi_0$), corresponding to ~ 1 MHz change in qubit frequency.

[16]This noise is caused mostly by ground loops and coupled acoustic noise from the fridge

[17]The batteries are capable of providing a continuous operation of the two cards for ~ 12 hrs. Longer experiments are possible via an additional set of batteries that can be switched from charging to the boards and vice versa.

2.2.4.3 Control software

All of our experiments are designed and executed via a desktop computer. We use LabRAD[18] to communicate with multiple instruments via a scripting language (Python), log data and monitor measurements in real-time. LabRAD is an open source project, developed by the Martinis group at UCSB. It is a software environment that enables a user to design experiments at a chosen language (Delphi Python and Labview) with all the required functionalities for running experiments built-in, and provides network based sharing of lab resources. Its modular architecture enables efficient control over complex apparatuses, specifically designed for multi-qubit experiments.

2.2.4.4 Microwave pulse generation

Arbitrary pulse shaped microwave signals are generated by modulating a microwave oscillator with lower frequency ($0 - 500\,\mathrm{MHz}$) signals. This is achieved with a mixer, which is a nonlinear device (typically a diode or a field effect transistor) that outputs a signal which is the product of its input signals. To control both the envelope and the phase of our pulses, we use an IQ-mixer (see Fig. 2.5). In an IQ mixer, a high-frequency oscillator (LO) $A\cos(2\pi f_{\mathrm{LO}}t)$ is split into two identical signals, one is multiplied with a signal $I(t)$ and the other is phase shifted by 90° and multiplied with a signal $Q(t)$, where both $I(t)$ and $Q(t)$ change much slower than the oscillator. The two multiplied signals are then combined to yield a signal proportional to $I(t)\cos(2\pi f_{\mathrm{LO}}t) + Q(t)\sin(2\pi f_{\mathrm{LO}}t) = \sqrt{I^2+Q^2}\cos(2\pi f_{\mathrm{LO}}t+\phi)$ at the output, where $\phi = \arctan(Q/I)$. I and Q provide complete control over the amplitude and phase as a function of time, within the bandwidth limits. A useful feature we use often is side-band mixing. Note that when $I(t) = A_{\mathrm{SBM}}\cos(2\pi f_{\mathrm{IF}}t)$ and $Q(t) = A_{\mathrm{SBM}}\sin(2\pi f_{\mathrm{IF}}t)$, we get an output $A_{\mathrm{SBM}}\cos\left(2\pi\left(f_{\mathrm{LO}}+f_{\mathrm{IF}}\right)t\right)$ (see Fig. 2.6). This allows us to vary the frequency of the oscillator rapidly (at a rate

[18] http://en.wikipedia.org/wiki/LabRAD

limited by the bandwidth of I and Q). This ability is essential since communication times with the oscillator are much longer (~ 1 ms).

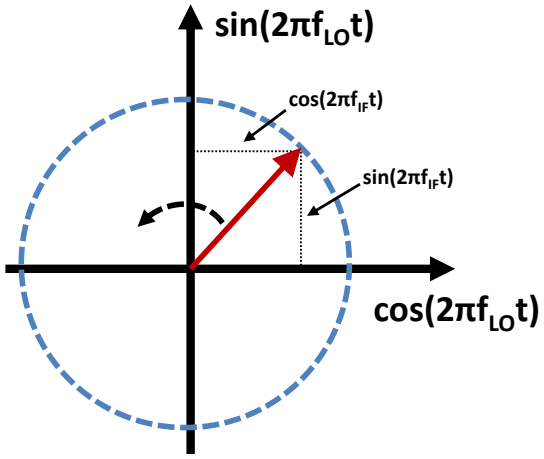

Figure 2.6: Side-band mixing. The modulation signal is easily understood when viewed in the rotating frame (the I and Q channels are 90^o out of phase, and therefore can be mapped to perpendicular axes in the complex plane). We choose the I and Q channels to be the horizontal and vertical axes, respectively. In this picture, a clock-wise rotating vector (whose components are the I and Q signals) at a frequency f_{IF} results in a frequency shift $-f_{\mathrm{IF}}$, relative to the local oscillator and vice versa for a counter clock-wise rotation.

2.2.4.5 IQ mixer imperfections

We use an ultra-low phase noise LO (Anritsu MG3692B) as a microwave source, and modulate our mixers[19] with signals generated by a GHz DAC board. IQ mixers have some nonidealities that must be corrected, the most critical to our experiments being: 1) the generation of LO harmonics, 2) constant LO leakage at zero modulation amplitudes, 3) cross-talk between the I and Q channels and 4) ringing of the pulses at the output.

1) Although transitions close to twice the qubit frequency are negligible, for sufficiently large anharmonicity and at sufficiently strong drives these cannot be neglected. LO harmonics are typically not modulated by I and Q and therefore have a constant and large amplitude. This could be deleterious especially during measurement pulses, where the anharmonicity is greatly increased. To remove these

[19] We use Marki IQ0307 (for 3-7 GHz signals) or IQ0714 (for 7-14 GHz signals) IQ mixers

signals, we use a low-pass filter designed for microwave frequencies. We custom built a set of filters based on width modulated microstrip transmission line, however due to their poor attenuation above the cutoff ($\sim -30\,\mathrm{dB}$ max) we replaced them with commercial filters in our more recent experiments[20].

2) Leakage of the LO frequency may result in a continuous, undesired qubit excitation. The leakage is found to be a result of voltage offset of the I and Q ports: $V_{\mathrm{out}} = (I(t) + I_{\mathrm{offset}})\cos(2\pi f_{\mathrm{LO}} t) + (I(t) + Q_{\mathrm{offset}})\sin(2\pi f_{\mathrm{LO}} t)$. The offsets are easily found by searching for a minimum of V_{out} as a function of I and Q. Since the offsets in I and Q are frequency dependent we sweep the relevant frequency range and find V_{out} at fixed intervals[21]. The measured offsets are used in all of our experiments to correct for the IQ-mixer offsets. Our calibration measurement uses a spectrum analyzer to accurately measure V_{out} down to -110 dBm. We typically achieve an extinction ratio of 80 dB between the maximal signal power and the leakage level power, limited by the voltage resolution of our search. This calibration tends to deteriorate with time to $\sim 60\,\mathrm{dB}$ due to temperature drifts of the IQ-mixer. In some of our experiment we integrate a periodic (automatic) measurement and correction of the offsets using the qubit excitation as a sensitive detector to account for these drifts.

3,4) A pulse envelope in the I (or Q) channel can leak to the Q (or I) quadrature of the IQ-mixer, emerging out with the wrong phase. This may lead to undesired effects, such as a lower extinction of the opposite band in side-band mixing[22]. In addition, impedance mismatch at the I and Q pathways (cables, filters and amplifiers) causes a partial reflection of the signal. In some cases the transmitted signal may encounter a second impedance mismatch at the I and Q ports and partially reflect on it again. This gives rise to a series of reflections (ringing), separated by a

[20] A detailed write-up on our custom built filters by Yoni Shalibo is available at the group internal directory.

[21] We can then choose arbitrary frequencies within the range with this calibration by interpolating between neighboring points. We typically use 20 MHz intervals to get smooth offset-frequency curves.

[22] Side-band modulation suppresses one band, while strengthens the other. In an ideal modulation, their ratio is infinite, however in practice it is limited by cross-talk between the I and Q channels to ~ 30 dB.

time interval $t_r = x_r/v$, where v is the propagation velocity[23], and x_r the distance between the reflection points. As a result, the output signal may not decrease to zero at its end, but rather show additional peaks or a slowly decaying envelope. A ringing microwave signal causes errors in state preparation and gates due to pulse concatenation and in measurement during measurement pulses. We solve both problems (cross-talk and ringing) by a generalized two-channel deconvolution procedure we developed, explained in detail in Appendix C.

2.3 Qubit measurement

Here, we describe basic measurements on a phase circuit device: potential mapping and macroscopic quantum tunneling, qubit spectroscopy, coherent control (Rabi oscillations) and T_1, T_2 characterization. These measurements are the first to be performed in each device since they provide crucial information for more complex experiments.

2.3.1 Mapping of the phase-circuit potential

To create and measure quantum states in our devices we must first reset the phase φ in one of the potential wells, and then let the energy relax to the ground state. By measuring the phase distribution as a function of the external flux Φ_{ext}, we can map out the potential minima and use it to construct a control sequence that adiabatically transfers the population to a specific well.

The time sequence of such mapping is depicted in Fig. 2.7a. For a given flux Φ_{ext}, we start at the maximal\minimal possible flux value, wait a time T_R for relaxation and then move adiabatically to Φ_{ext}. This ensures a more homogenous population distribution among the wells. The flux sequence is followed by a ramp on the SQUID, after which we record the time it takes for the SQUID to switch to the resistive state. This time is directly related to the SQUID's critical current, which

[23] In our coaxes, $v \approx 0.7c$, where c is the speed of light

measures the total flux through the SQUID loop Φ_s. The total current through the qubit loop is then given by $I_q = \Phi_s / M_{qs}$, where M_{qs} is the mutual inductance between the qubit and the SQUID. From the Josephson relations, $I_q = I_0 \sin(\varphi)$ and therefore the switching time directly relates to the average phase φ.

For each flux, we repeat the measurement $\sim 10^3$ times to obtain the distribution $P(\varphi)$. The result of this measurement is shown in Fig. 2.7b. The measurement reveals a series of branches, called SQUID steps, having a period Φ_0 in flux[24]. The peaks of the distribution at any given flux correspond to the potential minima. Each branch in the plot follows a specific potential minima, and therefore corresponds to a single meta-stable well. Note that it is essential to start from both large positive and large negative Φ_{ext} values at the beginning of each sequence to make sure that all the wells are occupied (on average) at any given flux.

To reset the population at a given potential well, we periodically change the flux from Φ_{ext}^{min} to Φ_{ext}^{max}, between which *only one* potential well remains stable (see Fig. 2.7c). At the end of every period, each occupied well (except for one) will have poured its population into neighboring wells, one of which is always stable and therefore monotonically increases its population. The number of periods required to reset the population with a given fidelity depends on the number of simultaneous wells. This number is uniquely determined by the product $I_0 L$, which also determines the length of each branch. For devices with only 2 simultaneous wells, we can always find a flux in which there is only one stable well and therefore no "potential shaking" reset sequence is required. We find that for the case of 5 simultaneous wells (as in Fig. 2.7), the residual population after 25 reset periods becomes smaller than 10^{-3}.

[24]This results from the flux-phase periodicity of the potential $U(\Phi_{ext} + n\Phi_0, \varphi) = U(\Phi_{ext}, \varphi - 2\pi n)$ and the periodicity of the current I_q with φ

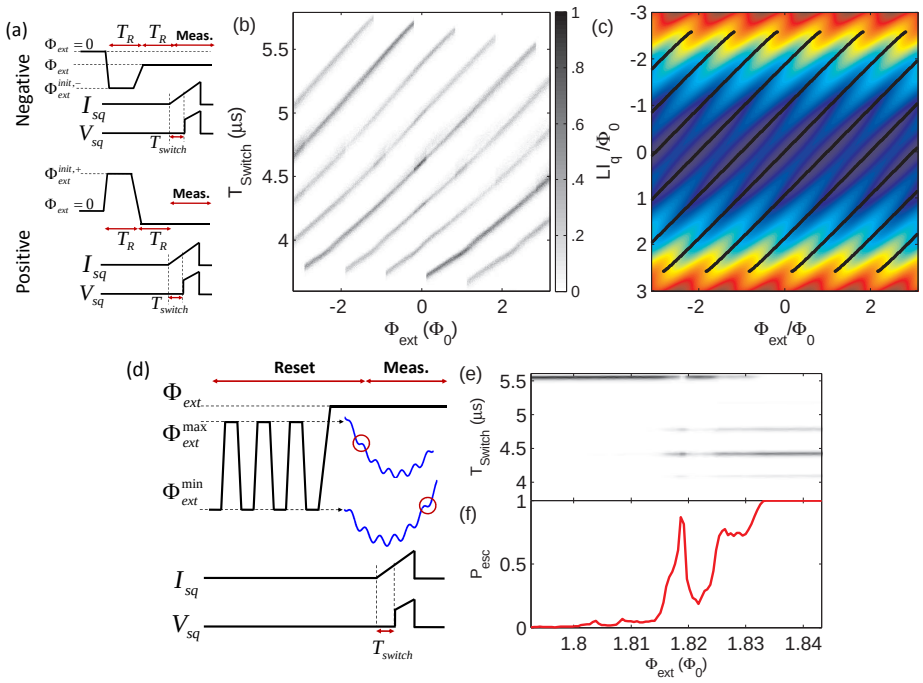

Figure 2.7: Potential mapping. (a) Time sequence of a SQUID-steps measurement. (b) Normalized histogram (grayscale) of switching probability as a function of external flux. (c) Calculated potential (colormap) as a function of flux. The black lines track the potential minima. (d) Experimental sequence for phase reset by "potential shaking", followed by step edge measurement. (e) Step edge histogram and corresponding escape probability (f). deviceA

2.3.2 Macroscopic quantum tunneling measurements

In Fig. 2.7d we plot a histogram of switching times as a function of flux after reset, and near the edge of a branch. One can see an overlap between two branches that results from the population escaping from the higher, shallow well to one of the lower, deeper wells during the waiting time. The escape from the well is dominated by quantum tunneling (MQT), rather than thermal activation due to our low electron temperature $k_B T \ll \hbar\omega$ [27]. We define a cutoff between the two branches and compute the probability of ending up in one of the bottom wells (see Fig. 2.7e). We observe distinct peaks in the escape probability near the edge that result from resonant tunneling [46].

Tunneling rates are exponentially sensitive to the potential barrier height, and therefore can be used to distinguish between the populations of each level in the

well. Due to the relatively short lifetime of excited states ($\sim 10^{-7}$ s), we use much higher tunneling rates in order to avoid population decay during this process. Tunneling is achieved by applying a short pulse (1-5 ns duration) in the fast flux-bias channel I_{bf}, which adiabatically reduces the potential barrier down to a point where tunneling occurs only from above a certain energy level (see Fig. 2.8a). In principle, one can measure all the occupation probabilities inside the well by performing a series of experiments with different measurement amplitudes, each measuring the cumulative probability from above a different level (see Sec. 3.2.2 and 3.4.4 for examples). To calibrate the required measurement amplitude I_{mp} to distinguish between consecutive levels, we perform measurement of the escape probability as a function of I_{mp} for different initial states.

In Fig. 2.8 we plot the escape curves for the lowest levels. The escape probability increases sharply from 0 to 1 around a certain measurement amplitude. This amplitude decreases with increasing level number due to the smaller potential barrier of increasing levels. The measurement visibility, defined as the maximal difference between the escape probabilities of consecutive levels typically ranges between 80 % and 90 % for the first two levels, and is limited by the device parameters. In most cases we calibrate the measurement amplitude I_{mp}^{k} (of pulses that distinguish populations of level $k-1$ and k) to be in the maximal visibility point. We then account for the smaller than 100 % visibility by a proper normalization matrix [25]. In cases where one wishes to increase sensitivity (for example, to calibrate drive pulses), it is advised to use a smaller amplitude pulse in which there is virtually no tunneling from levels smaller than $k-1$ but a measurable tunneling from level k. This constitutes a weak measurement that increases SNR [47].

For higher levels, the maximal visibility decreases due to the decreasing energy difference. The apparent visibility of excited levels in the data is even smaller, due to the decreasing fidelity of the prepared initial states of higher levels (see Sec. 2.3.4). However, this can be corrected by iteratively optimizing the shape of the escape curves (see Sec. 3.2.2) and using optimized control sequences for state

[25] See Sec. 3.2.2 for further information

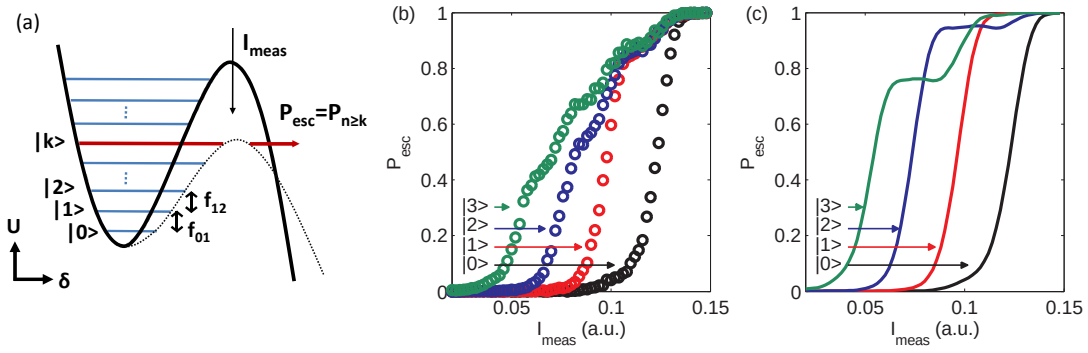

Figure 2.8: Occupation probability measurement. (a) cartoon illustration of the level-selective tunneling. (b) Escape curves for the initial states $|\psi\rangle = |0\rangle, |1\rangle, |2\rangle, |3\rangle$, shown in black, red, blue and green respectively. (c) Smoothed and normalized escape curves. The steps in the escape curves of states $|2\rangle$ and $|3\rangle$ are believed emerge from trapping into TLSs during the measurement pulse. These data were obtained from Device_C

preparation (see Sec. 3.3). Additional features in the single level escape curves can reduce the visibility due to the capturing of excitation by TLSs during the measurement pulse [48].

By applying measurement pulses, we map a qubit state into a classically separable state. In this state, the qubit level occupation probabilities are encoded in the probability of being in the right\left potential well, which is measured with the SQUID. Due to the large measurement fidelity (distinguishability of potential wells in SQUID measurement), our precision in measuring the probability is mostly limited by shot noise.

2.3.3 One-photon and two-photon spectroscopy

The qubit frequency is found using *incoherent* spectroscopy measurements. In these measurements, a long microwave square pulse ($\tau_{\mu w} \gg T_2^*$) is applied in the microwave channel $I_{\mu w}$, followed by measurement of the excited state probability P_1. The microwave frequency is swept over a large range around the expected

qubit frequency[26]. A typical spectroscopy measurement at low microwave power is shown in Fig. 2.9 in blue. A single peak, corresponding to the microwave drive being resonant with the qubit is seen at $\nu_{\mu w} = 8.637\,\text{GHz}$. The resonance curve is a Lorenzian, having a width $\Delta\nu_{FWHM} = 2.7\,\text{MHz}$ which is related to the inhomogeneous coherence time $T_2^* = 1/\pi\Delta\nu_{FWHM} = 118\,\text{ns}$. At larger drive amplitudes, a second peak is revealed, corresponding to the two-photon transition from the ground state to the 2^{nd} excited level at $\nu_{\mu w} = f_{20}/2 = 8.622\,\text{GHz}$. This is approved by repeating the measurement with a smaller measurement amplitude, showing only peaks corresponding to excitations from level two and above. The peak occurs at a lower frequency than the qubit's, as expected from a negative anharmonicity $f_{21} < f_{10}$, and implies an anharmonicity $\beta/2\pi = 30\,\text{MHz}$ since $\beta/2\pi = f_{10} - f_{21} = 2(f_{10} - f_{20}/2)$. Depending on the anharmonicity, additional multi-photon transitions can be seen at still larger drive amplitudes. Larger anharmonicity gives an advantage when working with qubits because pulse durations required to suppress leakage to level 2 scale as $1/\beta$ [49], while decoherence restricts pulses to be much shorter than T_2.

At certain flux values, the qubit frequency is split into two or more frequencies (see upper inset in Fig. 2.9b), due to coupling to TLSs (see Sec. 1.3.6 and 3.1). TLSs have several undesired effects on the qubit; they can reduce the fidelity of gates, cause the qubit population amplitudes to vary with time, shorten its lifetime through off-resonant coupling and reduce measurement visibility. To avoid these effects, we map the qubit frequency as a function of flux and find optimal regions where TLSs do not couple to the qubit. In Fig. 2.9b we plot the escape probability as a function of drive frequency and flux. The plot reveals two curves corresponding to one-photon (qubit) and two-photon excitations, in which the fre-

[26]The qubit frequency can be guessed, knowing the SQUID's measured critical current: while there may be large uncertainties in the junctions' critical current due their large sensitivity to the oxidation process, the qubit and SQUID junctions are designed to have a fixed area ratio, and therefore one can estimate the qubit junction's critical current by knowing the SQUID's. In the devices used in this work, the ratio of SQUID and qubit critical currents is $3\sqrt{3}/22.1$. Using the expected values for the capacitance (dominated by the external capacitor) and inductance, one can calculate roughly the qubit frequency close to the edge of the branch1.3.5.

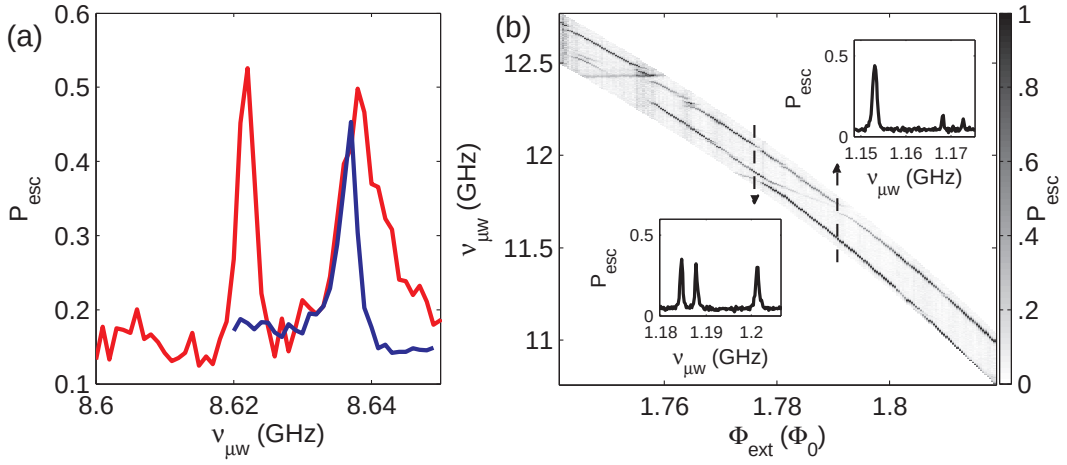

Figure 2.9: Incoherent spectroscopy. (a) Low-power ($\Omega/2\pi = 1.4$ MHz) in blue and high-power ($\Omega/2\pi = 5.7$ MHz) in red, showing a two-photon transition peak (measured on Device_B). (d) Spectroscopy vs. flux, showing avoided level crossings in the qubit and two-photon transitions (measured on Device_A)

quency monotonically decreases with flux. In fact, the frequency curves closely follow $f_{10} \propto \left(a + b\Phi_{ext} + b\Phi_{ext}^2\right)^{1/4}$, as expected from theory (see Sec. 1.3.5). In this measurement, the potential well becomes shallower with increasing flux, and the anharmonicity increases correspondingly. This is readily seen in the plot where the separation between the two lines increases with flux. We also see multiple avoided-level crossings, each resulting from coupling between the qubit and a single TLS close to resonance. To avoid undesired coupling to TLSs during our experiments we make sure that the closest avoided-level crossing is at least $\Omega_{\max}$ away from the qubit frequency, where $\Omega_{\max}$ is the maximal Rabi frequency in our experiment.

2.3.4 Time-domain characterization

One of the most important factors in assessing the quality of a qubit device are its energy relaxation (T_1) and dephasing times (T_2^*). These are usually compared with the shortest gate time, which is usually taken to be ~ 10 ns for single qubit gates and a few 10 ns for two-qubit gates.

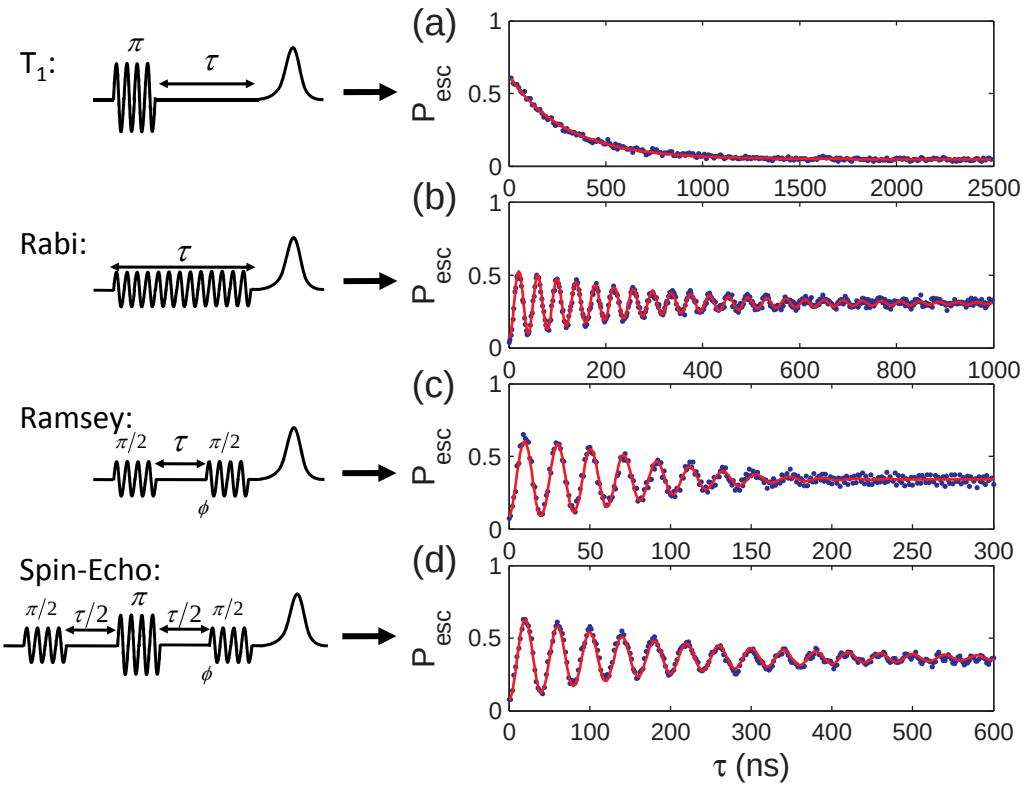

Figure 2.10: Time domain characterization. Escape probability as a function of delay, in T_1 (a), Rabi (b), Ramsey (c) and spin echo (d) measurements. Pulse sequences used in these measurements are shown to the left of each. Red curves: fits to a decaying exponent (a), oscillatory decaying exponent (b), oscillatory Gaussian (c) and oscillatory decaying exponent (d). These data were measured with Device_B.

In a T_1 measurement we start at the ground state, and apply a short π-pulse[27] to rotate the qubit to the excited state $|1\rangle$. We then wait a variable time τ and measure the excited state population with a calibrated measurement pulse. This measurement (see Fig. 2.10a) yields an exponential decay curve $P_1 = P_{10}\exp(-t/T_1)$, with and extracted lifetime $T_1 = 296\,\text{ns}$. The maximal lifetime we measure in our devices is in the range $T_1 \approx 150 - 350\,\text{ns}$. This time is limited by dielectric relaxation in the junction oxide and external capacitor, however we find it to deteriorate after a few cooldowns[28]. In some cases, where we wish to eliminate

[27] see Sec. 3.4 for details about pulse calibration procedures

[28] This deterioration is not gradual, but relatively sudden. We notice a large change in T_1 at some

excitation of other modes (for example multi-photon transitions or TLSs), we use incoherent excitation in the T_1 sequence. The maximal excitation in this measurement is limited to $P_1 = 0.5$, however due to the long time and small amplitude excitation pulses, Fourier and power broadening are minimized, giving more reliable T_1 values.

The coherence time T_2 is measured using a Ramsey sequence (see Fig.2.10b). In this measurement, the qubit is rotated with a $\pi/2$-pulse to the coherent superposition state $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. We then wait for a variable time τ, followed by an identical $\pi/2$-pulse which rotates the remaining horizontal polarization to the energy basis, where it is measured. During free evolution, the Bloch vector precesses about the z-axis[29], contracts in the horizontal direction (dephasing) and acquires magnitude in the vertical direction (decay), however only the horizontal polarization is measured. We find that the envelope of the fringes decays like a Gaussian $P_\psi = \exp\left(-(t/T_2)^2\right)$[30], as expected from dephasing dominated by $1/f$ noise [50, 51] and inhomogeneous dephasing (see Sec. 2.2). We extract a dephasing time $T_2 = 100\,\text{ns}$ at large anharmonicity ($\beta/2\pi = 150\,\text{MHz}$), consistent with the qubit's spectral linewidth (not shown). At small anharmonicity ($\beta/2\pi \approx 20\,\text{MHz}$), $\pi/2$-pulse lengths are comparable to the dephasing time, and therefore it is difficult to use time-domain methods to measure it. In this case we use spectroscopy, which confirms coherence times in excess of 200 ns. This result is in agreement with dephasing dominated by flux noise, since the sensitivity of the qubit frequency to fluctuations in flux decreases with decreasing anharmonic-

point in time (with a change period maximally bounded by a few days), which then remains stable for months. We suspect that these changes occur as a result of thermal expansion\contraction cycles that induce defects in the metal-dielectric interface. This point is not well understood and remains to be explored.

[29]Since we work in the rotating frame, we do not see a precession, however in most cases the qubit is slightly detuned from the drive, causing low-frequency oscillations that may induce errors in the extracted T_2. We therefore deliberately detune the drive from the qubit by a large frequency or vary the phase of second pulse linearly in time to emulate precession.

[30]The coherence time is affected by both pure dephasing and by energy relaxation. Energy relaxation results in exponential decay of the coherence, and therefore the actual envelope has a more complex form. However, since pure dephasing is the dominant decoherence process, our Ramsey envelope curves are well fitted by a Gaussian.

ity [49].

Low frequency fluctuations can be canceled out using dynamical decoupling methods. In a spin-echo for example (see inset of Fig. 2.10c), an additional π-pulse is applied in the x or y directions - halfway through the free evolution period. The π-pulse at time $\tau/2$ reflects the current Bloch vector about its initial direction, causing it to precess back to its original direction at time τ, regardless of its precession frequency. Therefore, an ensemble of qubits with different energies will all acquire the same phase at the end of the free evolution time. In our case, where there is only one physical qubit, dephasing emerges from averaging an ensemble of identical experiments occurring at different times. Since the spin-echo sequence is only effective in cases where the qubit frequency is fixed during the free precession time, it acts as a high-pass filter for noise, with a frequency cutoff $1/\tau$. A spin-echo measurement (see Fig. 2.10c) yields a homogeneous decay time $T_2' = 205\,\text{ns}$ which is still not T_1-limited, implying either high-frequency noise (at $f > 1\,\text{MHz}$) or a decoherence mechanism that is not related to noise.

CHAPTER 3

Results

This chapter describes all the achievements made throughout this work. The first section stands apart from the other sections, since it is not directly related to the control and measurement of multi-level states, the main topic of this work. In this section we discuss two-level defect states (TLSs) in the Josephson junction and their contribution to decoherence in our device. We then explain our measurements, which directly probe their lifetime and coherence time, and extract from these data important characteristics of the TLSs. In Sec. 3.2 we measure the nonlinear response of our anharmonic system to a controlled drive in the quantum and classical regimes. We map out the transition between them using 3 control parameters and compare the results to theory. In Sec. 3.3 we demonstrate a method to prepare quantum states with relatively high fidelity, which does not require any calibration effort. This is done with an optimization algorithm, applied directly on our system. The results shown in these section are only preliminary, but provide an essential tool for testing some of the state tomography methods explained in Sec. 3.4. In Sec. 3.4, we perform full standard quantum state tomography of a 3-level state in the phase circuit for the first time. We also demonstrate arbitrary state preparation within this subspace and discuss the limitations of this method. We then show a novel method of obtaining the density matrix within this system which is much more simple to implement and currently allows for an accurate state tomography of up to 5 levels.

3.1 Two-level defects

Two-level defect states (TLSs) in amorphous insulators are of fundamental interest due to their impact on many low temperature properties, such as the heat conductivity [52] and the generation of $1/f$ noise [53, 54]. On the practical side, these defects limit the operation of solid state devices, for example amplifiers [55] and CCD detectors [56], and increase the dielectric loss of insulators [34].

While the thermal properties of amorphous materials at low temperatures were already measured decades ago [52], their dielectric properties such as the loss tangent at low temperatures and power are still largely unknown. It was recently proposed that two-level charge defects within the dielectrics used in phase-qubit devices limit their decay and decoherence times [34]. Amorphous dielectric materials are used in qubit devices in the external capacitor, between metal layers at wiring cross-overs and in the insulating layer of the Josephson junction. Martinis et al. [34] found that for large Al-AlO_x-Al junctions ($A \sim 200\,\mu$m), loss from the junction dominates with a typical decay time of ~ 10 ns, and many resonances appear in the qubit spectrum, scattered uniformly in energy. The number of defects that couple to the qubit scales as the junction area, and becomes on the order of 1 for $\sim 1\,\mu$m junctions. In this regime, loss from the cross-over insulator (amorphous SiO_2) dominates. Measurements of the loss-tangent of several dielectrics using resonators show an increase at increased power and temperature, consistent with coupling to a TLS bath. It was concluded that while these two loss mechanisms (sparse TLSs in junctions and a macroscopic defect bath in thick insulating layers) are different, they are fundamentally related, since they both originate from coupling to a spin bath [34]. Since the phase-qubit operates in the large capacitance limit ($E_J \gg E_C$), to increase its decay time while maintaining the large capacitance, the large area junction has been replaced by a small one ($\sim 1\,\mu$m) and the cross-over insulator and external capacitor with lower-loss dielectrics (SiN_x). The notion of coupling of the phase qubit to a strongly anharmonic microscopic system was further fortified through careful analysis of the multilevel spectrum

of the phase and flux qubit near resonance with a defect [57, 58]. Additional paths taken to eliminate loss from TLSs include epitaxial growth of the junction insulator [36, 59, 37], and using micro-bridges instead of a tunnel barrier [38]. In other works, TLSs were used as a controlled decay channel [60] or as a quantum memory [61].

Additional models for the origin of defects and defect-qubit coupling mechanisms have been proposed, including modulation of the junction capacitance or critical current by microscopic charged states [58, 62, 63], magnetic moments originating from metal-induced gap states at the metal-oxide interface [64], delocalized oxygen atoms [65], a cooper pair tunneling from the superconductor into localized electronic states [66] and electron trapping in shallow sub-gap states in the metal insulator interface [67]. The true mechanism for TLSs is still unclear [68]. Some of the proposed models indicate different noise spectra from TLSs with respect to other mechanisms. This can be used to rule out specific TLS mechanisms via future measurements.

Additional understanding of the nature of defects is still needed, in order to better understand their origin and possibly improve the dielectric materials in qubit devices. In this work, we use the phase qubit to study the properties of the defects themselves. We measure the energy relaxation and dephasing times of an ensemble of defects. We find the energy relaxation rates to correlate with the coupling to the qubit, suggesting a phonon relaxation mechanism for the TLSs. The dephasing times are optimized at intermediate couplings. We find a possible explanation for this result from optimized sensitivity to fluctuations in this regime.

3.1.1 Defect states in glasses

TLSs were first suggested to explain low temperature thermal properties of amorphous insulators [52]. Measurements of the heat capacity below 4 K show a linear increase with temperature, compared with the cubic dependence in the crystalline phase that is expected from phonons. In addition, the magnitude of the heat

capacity in this regime is larger in the amorphous phase by nearly two orders of magnitude. Since at low temperatures, thermal phonons have low frequency, and correspondingly long wavelengths ($\sim 100\,\text{nm}$), it was argued that phonons should be oblivious to the microscopic arrangement of atoms, and therefore their contribution to the heat capacity in the amorphous phase should be similar to that of the crystalline phase. The heat conductivity in amorphous insulators behaves differently as well, showing a quadratic dependence instead of a cubic dependence that is expected from phonon heat transport at low temperatures.

To explain these results, several models were suggested, the most successful of which was the TLS model. According to the model, two-level states result from atomic impurities, tunneling between two (or more) configurational states, similar to those in the ammonia molecule. Each configuration is associated with a potential minimum, however occasionally, an additional configuration(s) exists which is close in energy and has a non-negligible tunneling coupling to it. This arrangement can be effectively described by a double-well potential, with two parameters: the energy asymmetry between the two wells Δ, and the tunneling energy Δ_0 associated with the potential barrier between the two wells (see Figure 3.1a). Due to the random nature of defects, the model uses a distribution function $f(\Delta_0, \Delta)$ to determine their statistical properties. The distribution is assumed to be broad and symmetric in Δ, with a width on the order of $k_B T_g$, where $T_g \sim 1000\,\text{K}$ is the glass transition temperature. At low temperatures ($T < 1\,\text{K}$), only TLSs at the center of the distribution are relevant, and therefore it is taken to be independent of Δ. The distribution of tunneling energies $\Delta_0 \sim \exp(-\lambda)$ is determined by the distribution of the tunneling parameter λ[1]. Since Δ_0 is exponentially sensitive to λ, only a small range of λ is sampled for a broad range of Δ_0 and therefore the distribution of λ is taken as constant over this range. This translates to a distribution $\sim 1/\Delta_0$ for Δ_0, and therefore $f(\Delta_0, \Delta) = P/\Delta_0$, where P is some material dependent constant. The Hamiltonian of a given defect can be

[1]The tunneling parameter is typically taken as $\lambda = \sqrt{2mV_0/\hbar^2}d$ [52], where m is the mass of the tunneling atom, V_0 is the height of the potential barrier and d is its width.

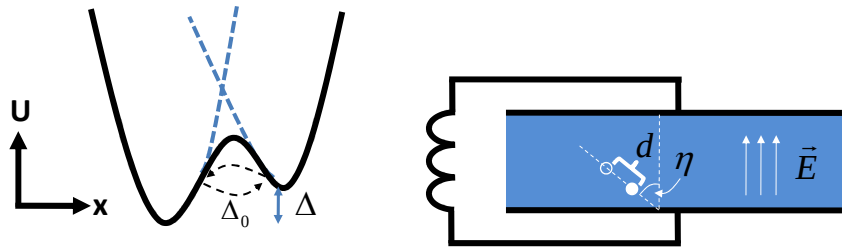

Figure 3.1: TLS model and Qubit-TLS interaction. (a) Atomic impurities tunneling between two configurational states, associated with minima of an effective potential. (b) The junction is modeled as a parallel plate capacitor, and the TLS as an electric dipole.

written in the position basis ($\{|L\rangle, |R\rangle\}$):

$$H_{TLS} = \frac{1}{2}\begin{pmatrix} \Delta & \Delta_0 \\ \Delta_0 & -\Delta \end{pmatrix}. \tag{3.1}$$

The eigenenergies are then given by $E = \pm\frac{1}{2}\sqrt{\Delta^2 + \Delta_0^2}$ and the corresponding ground and excited states are $|g\rangle = \sin(\theta/2)|L\rangle - \cos(\theta/2)|R\rangle$ and $|e\rangle = \cos(\theta/2)|L\rangle + \sin(\theta/2)|R\rangle$, where $\theta = \arctan(\Delta_0/\Delta)$.

3.1.2 Defect states in a Josephson junction

The TLS model can be applied to the Josephson phase qubit to calculate important properties, such as the qubit decay rate and splitting distribution in the qubit spectrum. The coupling mechanism between the qubit and a single defect is assumed to be that of an electric dipole with an electric field. The electric dipole associated with a single charged defect emerges from an electric charge that can switch (via tunneling) between two spatial configurations inside the electric field of the junction (see Fig. 3.1b). The dipole moment can be computed using the eigenstates of the TLS Hamiltonian (Eq. 3.1) : $D_e = \langle e | q\hat{X} | e \rangle$, where the position operator can be written as $\hat{X} = d(|L\rangle\langle L| - |R\rangle\langle R|)$ assuming similar wavefunctions for the uncoupled wells. We can then compute, $D_e = qd\Delta/E_{ge}$, and similarly for the ground state. The Hamiltonian of such an interaction can be written as:

$$H_{int} = \overrightarrow{D} \cdot \overrightarrow{E} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} (qVd/x)\cos\eta, \tag{3.2}$$

where V is the voltage across the junction, x is the thickness of the barrier, d is the spatial size of the dipole, η is the dipole angle relative to the electric field of the junction. The dipole is static when $\Delta_0 = 0$ and oscillating when $\Delta = 0$. In the case of a non-classical electric field, such as that produced in a qubit state, we must replace the voltage by an operator $\hat{V} = \hat{q}/C$, where C is the qubit capacitance, and $\hat{q}$ is the excess charge across the junction. Writing the voltage state in the eigenstate basis, we get: $\hat{V} = \sqrt{\frac{E_{10}}{2C}} \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$ which implies zero average voltage in the qubit eigenstates. We can now write the coupling Hamiltonian between the qubit and the TLS in operator form (after neglecting non-resonant terms):

$$H_{int} = \overrightarrow{\mathbf{D}} \otimes \overrightarrow{\mathbf{E}} = i\,(S/2)\,(|0,e\rangle\langle 1,g| - |1,g\rangle\langle 0,e|)\,, \tag{3.3}$$

where $S = S_{max}\frac{\Delta_0}{E}\cos\eta$ is the coupling strength and $S_{max} = 2\,(d/x)\,\sqrt{E_{10}e^2/2C}$ is its maximal value. The probability distribution can now be transformed from the variables Δ, Δ_0 which cannot be measured directly, to the physically observable parameters E_{ge}, S. The density of states can be written after some mathematical manipulations as: $\frac{d^2N}{dE_{ge}dS} = \sigma A/S\sqrt{1 - \left(\frac{S}{S_{max}\cos\eta}\right)^2}$, where A is the junction area and σ is the number of TLSs per unit area. This is independent of the TLS energy E_{ge}, consistent with experiments, showing constant density of TLSs up to 20 GHz. By averaging over η (assuming an isotropic distribution), we get:

$$\frac{d^2N}{dE_{ge}dS} = \frac{\sigma A}{S}\sqrt{1 - \left(\frac{S}{S_{max}}\right)^2} \tag{3.4}$$

3.1.3 Lifetime and coherence of defect states in a Josephson junction

Several mechanisms have been proposed for relaxation and dephasing of the dielectric defects themselves. Energy relaxation is caused by coupling to phonon states, while dephasing could be caused by spectral diffusion [52]. However, to date, only limited measurements were carried out on TLSs to characterize these processes at the single defect level [69]. Such measurements could be used to better understand the nature of the defects and their decay mechanisms, and possibly engineer long-lived quantum memories in future devices. In [69], the coherence times of several TLSs were measured spectroscopically and were found to distribute as $P \sim 1/T_2$, where T_2 is the spectroscopic coherence time. Neeley's method [61] of probing the TLS adds the capability of measuring the coherence time more accurately and also measure their lifetime separately.

In this work, we present a measurement of the decay of energy and coherence for a large ensemble of TLSs in a small area junction using the phase qubit. We find that on average, the energy relaxation time (T_1), follows a power law dependence on the coupling parameter to the phase qubit. The exponent of this power law is in fair agreement with what is expected from phonon radiation by a dipole (proportional to the coupling strength) inside the junction. The average dephasing time ($T_\phi = (1/T_2 - 1/2T_1)^{-1}$) is coupling dependent as well, peaking at intermediate couplings. We interpret this optimum coupling to be caused by anti-correlated fluctuations in the physical parameters which determine the TLS energy.

For small area junctions ($\sim 1\,\mu m^2$), the typical measurement bandwidth allows us to detect and measure about 10 TLSs in a particular cooldown. Instead of using many different samples to acquire sufficient statistics, we use the fact that heating resets the TLS characteristics. The device is thermally anchored to the mixing chamber of a dilution refrigerator during measurement. We find that the TLS distribution is reset upon raising the temperature above 20 K and cooling down to the base temperature (10 mK). Some memory of the TLS distribution remains if

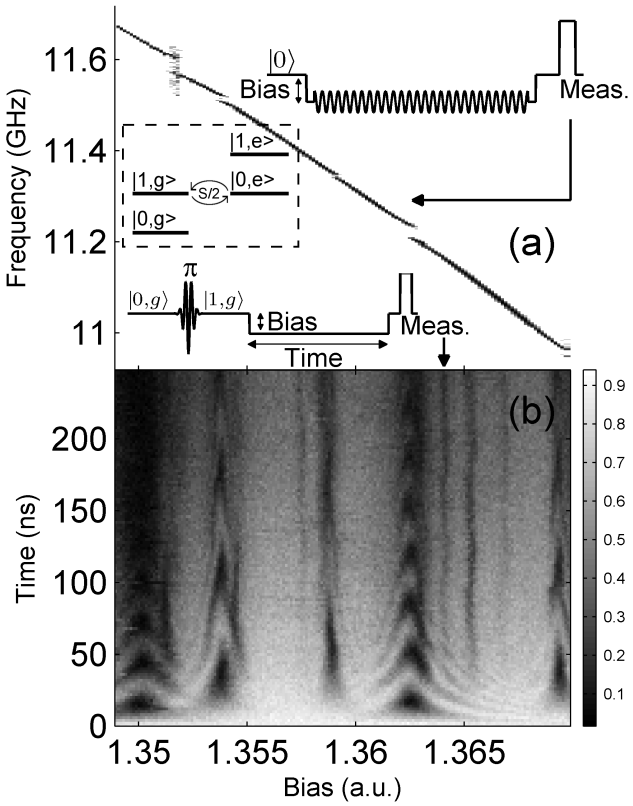

Figure 3.2: Frequency-domain and time-domain signatures of TLSs in qubit measurement. (a) The probability of the qubit being excited (P_1, normalized) after a long microwave tone at different frequencies and biases. On top of the expected smooth change of qubit frequency with bias, we observe randomly scattered splittings due to coupling to TLSs. (b) P_1 after qubit excitation followed by a "free" evolution at different bias values. P_1 oscillates near resonance with TLSs, consistent with the position of splittings in the spectrum. Upper and lower insets: control sequences used to produce (a) and (b) respectively. $|0,g\rangle$ and $|1,g\rangle$ stand for states where the qubit is at its ground and excited state respectively while the TLS is at its ground state. Dashed inset: level diagram for the combined qubit-TLS system on resonance.

the temperature is increased to only 1.5 K (see 3.1.4). We utilize this feature to produce a new set of TLSs and generate an ensemble. The data was taken over 82 different TLSs, obtained from 8 different cooldowns.

The initial identification of TLSs and their coupling parameters are carried out as follows. First, the qubit spectrum is swept over the bias to locate the frequencies of TLSs from the positions of the avoided level crossing structures (see Fig. 3.2a). A complementary picture of the interacting qubit-TLS system in the time domain is shown in Fig. 3.2b, where we excite the qubit with a short resonant pulse (π-

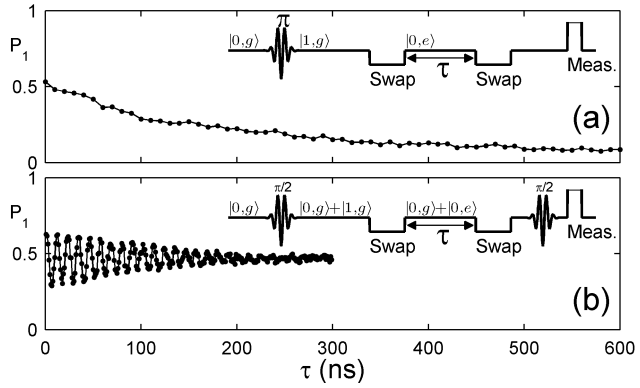

Figure 3.3: Representative T_1 and T_2 measurements of a TLS. (a) T_1 measurement of the TLS, along with its experimental sequence (inset). The qubit is first excited with a π-pulse, then brought into resonance with a TLS for a "swap time" (the time to fully transfer an excitation between the qubit and the TLS. It is found for each TLS by locating the first minimum in the oscillations in Fig. 3.2b). After a free evolution of the TLS of time τ, a swap gate is again applied, after which the qubit excitation probability P_1 is measured. (b) T_2 measurement of a TLS, along with its experimental sequence (inset). This sequence is similar to a T_1 measurement, only that superposition states are produced in the TLS and their decay is measured vs. time (Ramsey sequence [61]). The amplitude of the oscillations is proportional to the degree of coherence in the TLS.

pulse) far away from any observable TLS and then apply a bias pulse of varying amplitude and length. As seen in the figure, for bias values where the qubit is resonant with a TLS we observe oscillations that have the same frequency as the splitting size in the spectrum.

Following Neeley *et al.* [61], the characteristic energy relaxation and decoherence time scales were extracted from T_1 and Ramsey experiments on the TLS, with sequences schematically represented in the insets of Fig. 3.3. Figure 3.3a and 3.3b show representative T_1 and T_2 decay curves of the same TLS with characteristic times of 187 ns and 148 ns respectively, obtained from a fit to a decaying exponent and an oscillatory decaying exponent.

The size distribution of the observed splittings (see Fig. 3.4a) follows Eq. 3.4 and agrees with previous results on similar junctions (generated by measuring different samples) [34]. The maximal splitting size is found to be 45 MHz. From the measured $S_{\max}$ and known junction parameters we compute a dipole size $d \simeq 1$ Å. The minimal observable splitting size is ~ 3 MHz, and is mainly limited by the

f_{ge} (GHz)	S (MHz)	T_1 (ns)	T_2 (ns)	T_ϕ (ns)	CD No.	f_{ge} (GHz)	S (MHz)	T_1 (ns)	T_2 (ns)	T_ϕ (ns)	CD No.
12.8	3.3	365	-	-	6	11.59	16.0	550	-	-	7
11.43	4.5	220	-	-	3	11.78	16.1	51	-	-	5
11.35	4.7	6000	-	-	7	11.62	16.1	70	100	350	5
12.45	4.9	26	-	-	7	10.95	16.1	600	-	-	1
11.55	5.0	4700	-	-	4	12.642	16.4	1080	70	72	2
11.31	5.0	132	-	-	7	11.85	16.4	37	-	-	7
11.23	5.3	590	-	-	4	12.41	17.5	165	110	165	5
11.462	5.6	650	-	-	3	12.37	17.5	150	-	-	7
13.08	5.6	117	-	-	4	12.542	18.0	106	81	131	8
11.33	5.6	160	-	-	4	11.42	18.5	91	144	690	4
12.89	5.9	4400	-	-	6	11.915	18.9	208	-	-	2
11.4	5.9	623	-	-	6	12.47	19.2	72	36	48	6
12.79	6.3	335	-	-	4	11.21	20.0	127	100	165	5
13.11	7.4	40	50	133	4	12.73	20.8	41	-	-	7
12.85	8.1	291	66	74	3	11.86	21.7	156	120	195	2
12.13	8.2	190	50	58	5	12.15	22.2	85	-	-	7
11.57	8.3	400	-	-	4	11.72	23.8	60	120	-	2
11.45	9.4	373	68	75	5	12.66	23.8	32	-	-	7
11.835	9.5	170	-	-	2	12.44	25.0	44	-	-	6
11.342	9.5	185	85	110	2	11.86	25.0	355	-	-	1
12.26	9.6	453	47	50	2	11.61	25.6	53	100	1767	3
12.32	10.0	147	76	102	5	12.34	25.6	32	-	-	6
11.85	10.4	45	95	-	4	11.73	25.6	224	48	54	7
11.6	10.6	12	-	-	6	13.04	26.3	217	50	57	6
12.25	11.1	275	68	78	7	12.675	27.0	110	-	-	1
12	11.6	1000	-	-	4	11.77	27.8	62	130	2687	4
11.86	11.8	240	-	-	2	12.2	28.0	50	-	-	1
11.22	12.0	2400	-	-	1	12.613	29.4	56	52	97	2
11.88	12.2	1000	90	94	5	11.22	29.4	89	41	53	2
12.57	12.5	243	50	56	6	12.67	29.4	193	50	57	3
11.96	12.8	70	80	187	3	12.147	30.0	150	212	723	8
11.515	13.7	187	148	245	2	12.34	30.3	82	58	90	4
11.96	14.0	300	-	-	4	11.89	31.3	31	-	-	6
10.96	14.1	209	115	159	2	10.91	31.3	300	120	150	7
10.8	14.3	1600	-	-	7	12.78	34.5	125	114	210	5
13.22	14.7	55	95	697	6	12.12	34.5	48	52	113	6
11.57	14.9	210	63	74	3	11.59	35.7	26	27	56	2
11.27	15.0	158	65	82	5	11.17	37.0	243	68	79	6
12.24	15.2	165	100	143	3	11.772	41.0	220	156	242	1
11.7	15.6	62	-	-	4	12.05	45.5	476	950	-	4
11.38	15.9	123	40	48	7	11.99	47.0	60	-	-	1

Table 3.1: Full measurement data of TLSs: TLS energy (f_{ge}), splitting (S), lifetime (T_1), coherence time (T_2) calculated dephasing time (T_ϕ) and cooldown number. The last column indicates which TLSs were measured on the same cooldown (i.e., temperature was not raised to more than 20 mK between measurements of TLSs belonging to the same cooldown).

coherence time of the qubit. In addition, we find the distribution of TLS energies E_{ge} to be constant throughout our qubit measurement bandwidth (see Fig. 3.4b), consistent with theory.

Although most of the T_1 decay curves of the TLSs are similar in their shape (i.e. a simple exponential decay), their decay times range almost 3 orders of magnitude, from 12 ns to more than 6000 ns. Coherence times on the other hand range from 30 ns to only 150 ns (excluding a single anomalous TLS which will be discussed later). For comparison, when the qubit is biased far from any observable splitting, its lifetime is 270 ns, and its coherence time is 90 ns. TLS energy relaxation times at a given splitting are not random. We find that they are shorter for larger splittings (stronger interaction with the qubit), although short lifetimes are measured for the smallest splittings as well. This trend is apparent in Fig. 3.4c where we plot average T_1 values as a function of splitting. In this plot we divide the ensemble into groups of TLSs having similar splitting values, in a 7 MHz window size[2]. We find the average values $\langle T_1(S) \rangle$, excluding two points, to be best fitted by a power law $T_1 \propto S^{\alpha}$, where $\alpha = -1.44 \pm 0.15$. We would like to point out that two points in 3.4c have been excluded from the fit to a power law. These are the points with the largest average splittings (38 MHz and 46 MHz), where the statistics within each window is low (3 data points and 2 data points respectively). If we include these points in the fit we obtain an exponent $\alpha = -0.93$. We believe that the data points at the largest splittings may result from anomalous TLSs.

Figure 3.4d (black circles) shows the processed T_2 data, obtained similarly from only 43 different TLSs. In this case we observe a weak dependence on the coupling with a peak at $S \approx 25$ MHz. This feature is more pronounced in the dephasing time T_ϕ, represented by red triangles in Fig. 3.4d.

The full data measured for 82 TLSs is presented in Tab. 3.1 and the $T_1(S)$, $T_2(S)$ and $T_\phi(S)$ values are plotted in figure 3.5. Figure 3.5a contains 82 data points

[2]The errors in Fig. 3.4c and and 3.4d represent the statistical spread of the data within the 7 MHz window. They are calculated by normalizing the standard deviation by $\sqrt{N}$, where N is the number of points within the window.

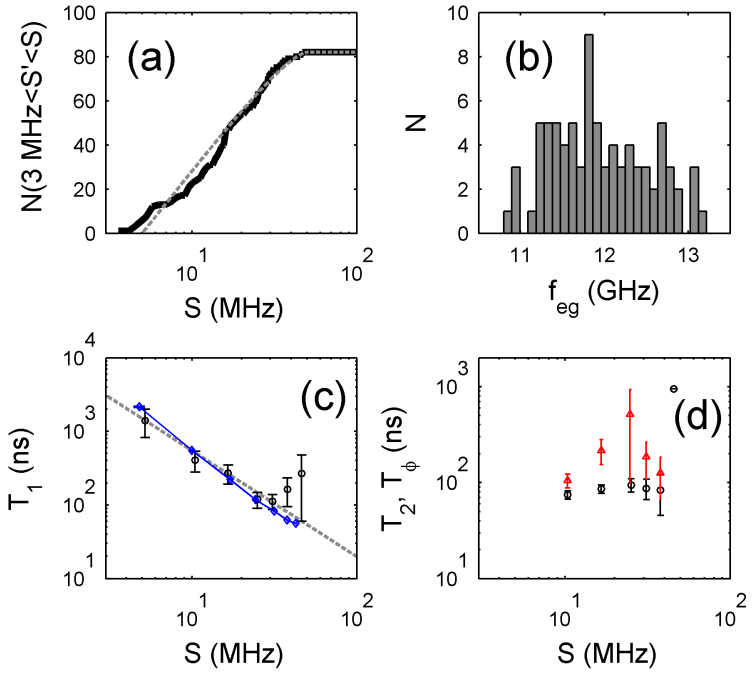

Figure 3.4: TLS survey results. (a) splitting-size and (b) energy distribution of 82 TLSs. The curve in (a) is the best-fitted log-normal distribution of the standard TLS model (see Eq. 3.4). (c) Average T_1 values (black circles) as a function of average splittings taken for TLSs inside a 7 MHz splitting window, and the best fitted power law in dashed gray. Average T_1 from stochastic simulation (arbitrary amplitude is shown) in blue diamonds. We attribute the deviation at the largest splittings to low statistics within these windows (d) Average T_2 (black circles) and T_ϕ (red triangles) value as a function of average splitting size within a 7 MHz window.

while 3.5b contains only 42. For about 50 % of the cases the dephasing time could not be determined, due to both low visibility and short dephasing time. This happens mostly for small splittings, which do not lie in the range of points presented in 3.4d. For points that lie within this range and are omitted from the analysis we separately checked that the shorter coherence time does not affect the trend. The number of points in Fig. 3.5c is further reduced because for some T_1 limited TLSs we measured T_2 which is slightly longer than $2T_1$ due to measurement error. T_ϕ is excluded from the figure and table for these cases (three TLSs).

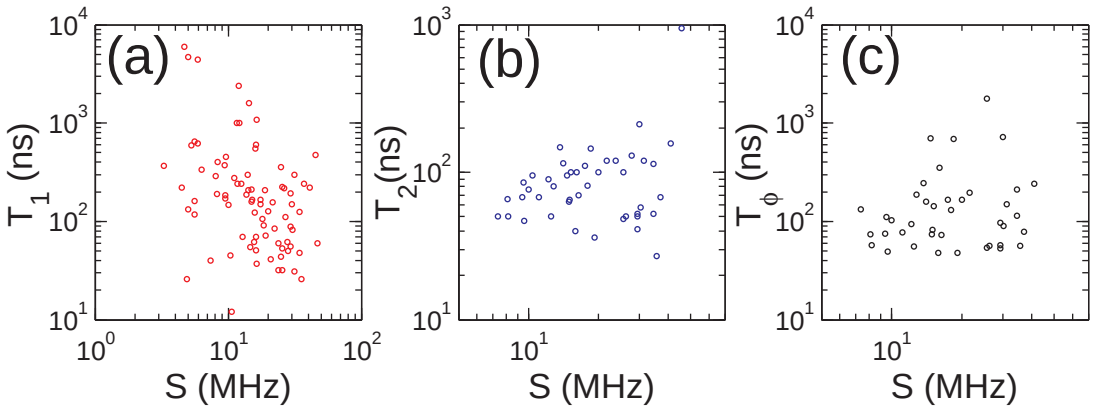

Figure 3.5: (a) Measured T_1 *vs.* splitting, (b) T_2 *vs.* splitting and (c) T_ϕ *vs.* splitting for all the measured TLSs. The spread in the T_1 data at a particular splitting value results from the random distribution of TLS orientation in the junction. It is apparent that the maximal lifetime shortens at larger splittings, consistent with dipole radiation. The spread of the T_2 data appears independent of the splitting, however the average values show some dependence which becomes more prominent in the dephasing times.

The T_1 results can be understood within the standard TLS model. The excited state of the TLS involves a local deformation of the insulator. This deformation couples to phonon modes, leading to the decay of the TLS excitation. The expected lifetime for such a process [52], is given by

$$T_1^{-1} = \frac{E_{ge}\Delta_0^2\gamma^2}{2\pi\rho\hbar^4}\left(\frac{1}{v_l^5} + \frac{2}{v_t^5}\right), \tag{3.5}$$

where γ is the deformation potential, v_l and v_t are the speeds of sound for the longitudinal and transverse modes respectively, Δ_0 is the energy splitting due to tunneling and ρ is the mass density. This is consistent with a power law dependence on S, since the interaction strength with the qubit satisfies $S \propto S_{max}\Delta_0/E_{ge}$ [34].

According to the model described earlier (see 3.1.2), the qubit-TLS interaction depends on the dipole orientation relative to the electric field of the junction. This feature explains the large spread in the data at a given splitting: both large dipoles (large Δ_0/E_{ge}) perpendicular to the junction's electric field and small dipoles aligned with the field can have the same S but different lifetimes. To more rigorously com-

pare experiment with this theory, we simulate an ensemble of TLSs with uniform distribution of dipole orientation and a log distribution of dipole moment sizes (see 3.1.2), from which we calculate the average lifetimes as a function of splitting size. The simulation data (see Fig. 3.4c) yields an average exponent $\alpha = -1.63$, in agreement with our measurement (a detailed description of the simulation is provided in 3.1.6).

The magnitude of the times that we extract from the experiment at a given splitting can be compared to the expected values for defects inside an AlO_x dielectric using Eq. 3.5. We approximate the deformation potential by $\gamma = \frac{1}{2}\rho v^2 \Delta V$ [70], where v is the average speed of sound of the transverse and longitudinal modes and ΔV is the local difference in volumes. For Al_2O_3 values with the difference in volumes taken as $\Delta V \simeq a^3$, where $a \simeq 1\,\text{Å}$ is the extracted dipole size, we get $\gamma \approx 1\,\text{eV}$, consistent with defects in other dielectrics [54]. Since the dielectric layer thickness is much smaller than the relevant phonon wavelength, the speed of sound in Eq. 3.5 is set by the aluminum layers of the electrodes. Using the speed of sound for thin aluminum films [71], we get $T_1(S_{\max}) \simeq 30\,\text{ns}$ which is very similar to what we measure. A more specific estimation should take into account the size of the junction and the layered structure.

Assuming the dephasing process is caused by fluctuations in energy, we note that the maximum observed in Fig. 3.4d can be explained by an anti-correlated dependence of the charging energy and tunneling energy on fluctuations in the TLS environment. According to the TLS model, $E_{\text{ge}} = \sqrt{\Delta^2 + \Delta_0^2}$ where Δ is the energy difference between the bare states of two spatial configurations $|L\rangle$ and $|R\rangle$ and Δ_0 is the tunneling interaction energy. Both Δ and Δ_0 are dependent on a set of environmental parameters $\overrightarrow{P}$, which fluctuate in time. As is standard for the TLS model, we assume a linear sensitivity for Δ on $\overrightarrow{P}$ and an exponential sensitivity for Δ_0: $\Delta_0(\overrightarrow{P}) = N_1 e^{-\sum_i P_i/P_{0i}}$, $\Delta(\overrightarrow{P}) = N_2 \sum_i P_i/P_{1i}$, with overall dimensional normalization constants N_1 and N_2 and parameter specific constants $\overrightarrow{P_0}$ and $\overrightarrow{P_1}$.

By expanding the TLS energy to first order in the position parameters $\overrightarrow{P}$, we get:

$$\delta E_{\mathrm{ge}} \approx \frac{\partial E_{\mathrm{ge}}}{\partial \Delta}\frac{\partial \Delta}{\partial \overrightarrow{P}} + \frac{\partial E_{\mathrm{ge}}}{\partial \Delta_0}\frac{\partial \Delta_0}{\partial \overrightarrow{P}} = \frac{1}{E_{\mathrm{ge}}}\sum_i \delta P_i \left(\Delta^2/P_i - \Delta_0^2/P_{0i}\right).$$

Assuming that $\delta P_i \ll P_i$, the terms inside the parenthesis can be assumed to be time independent. This expansion becomes interesting for the situation where $\Delta_0 \sim \Delta$ as there is a possibility for the contribution in δE_{ge} to pairwise cancel[3]. Since $\Delta_0 = S\left(E_{\mathrm{ge}}/S_{\max}\cos\eta\right)$, where η is the dipole orientation relative to the junction's electric field, we expect to find such a cancellation at a particular splitting value S. Note that the dependence on $\cos\eta$ smears this somewhat but we still expect a significant effect, as is observed in Fig. 3.4d.

As seen in the figures, the power-law describing $T_1(S)$ cannot explain all the measured TLSs. We find that three out of 82 TLSs with large splittings (37 MHz, 41 MHz and 45 MHz) have much longer lifetimes than expected (220 ns, 243 ns and 476 ns - respectively). In addition, one TLS out of 41 has much longer coherence than all the others (about a factor of 6 longer than the longest T_2 of all the others), associated with a splitting of size 30 MHz. Other anomalies we discovered are related to the stability of a particular TLS in time. We find that the energy E_{ge} of some TLSs (about 5 %) changes spontaneously at varying time scales, from seconds to days. All the rest were remarkably stable.

Some of these changing TLSs have long lifetime (a few microseconds), which is consistent with the power-law trend we discussed above. We also find that the instability of some TLSs increases at elevated temperatures (i.e., the change in TLS energy becomes more frequent). We conclude that some of the TLSs we measure have a different nature, perhaps related to their internal structure or position inside the junction.

[3]Note that the sign of P_{0i} could be negative, so this is not always the case

3.1.4 Temperature effects

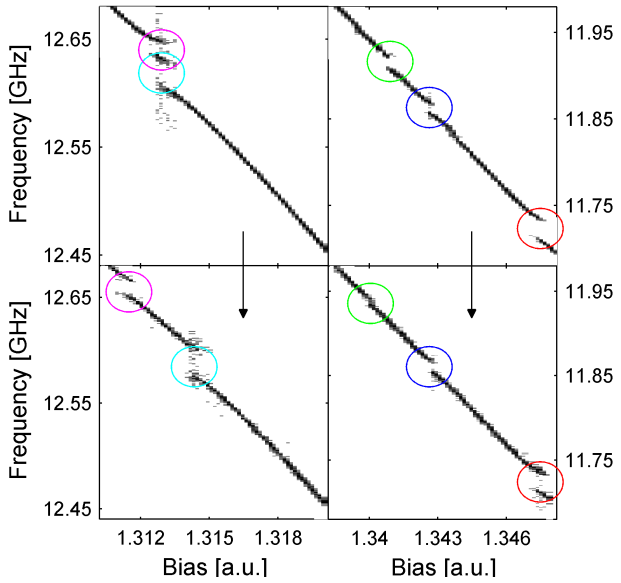

Figure 3.6: Qubit spectrum as a function of bias, before and after warmup to 1.5 K. Upper: two sections of the spectrum before warmup. Lower: the same sections, taken after heating the sample to 1.5 K and cooling down back to 10 mK. Circles with the same color indicate a splitting that we attribute to the same TLS. Both the position in frequency and splitting size of the TLSs are similar, indicating that the TLSs are not fully reset at this temperature.

New sets of TLSs are produced from the same device by warming up to 20 K. We believe that some TLSs are not fully reset upon warming up to only 1.5 K. Some splittings in the spectrum are similar in their frequency and splitting size to those before a partial warmup to 1.5 K, as indicated in Fig. 3.6. Both the reset of TLSs at high temperatures, and the fact that some of the TLSs are changing in time indicate the true nature of the TLS as an approximation of a multilevel state, resulting from a multi-well energy structure.

Furthermore, we also measure a few representative TLSs as a function of temperature. We find no significant change in T_1 and T_2 below 100 mK for TLS-B, consistent with the expected $\tanh(E_{ge}/2k_BT)$ dependence [52] (see Fig. 3.7). TLS-A shows an anomalous T_1 behavior, with optimized lifetime around 100 mK, while TLS-B follows the $\tanh(E_{ge}/2k_BT)$ dependence within the temperature range. The coherence time T_2 and the pure dephasing time T_ϕ decrease significantly with in-

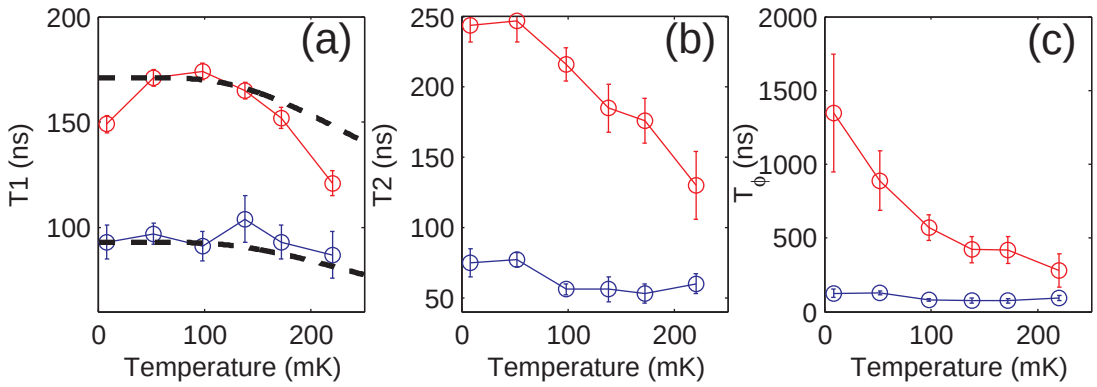

Figure 3.7: Temperature dependent data. Lifetime (a), dephasing (b) and calculated pure dephasing as a function of temperature for two representative TLSs. The red and blue curves are associated with measurements of TLS-A and TLS-B, having a splitting 30 MHz and 18 MHz respectively. Dashed lines in (a) mark the $\tanh(E_{ge}/2k_BT)$ trend.

creasing temperature in TLS-A, but only slightly for TLS-B. Above 220 mK we observe non-negligible thermal population in the qubit which reduces the visibility. In addition, the TLSs become unstable (frequent changes in TLS energy) which further reduces the visibility, thereby making the measurement increasingly difficult.

3.1.5 Related work

A recent experiment [72] measured the lifetime and coherence time of two TLSs in a similar device using a different technique. In their work, they used a second-order Raman process via the third level of the phase qubit [73], to excite the TLS directly while the qubit and the TLS are detuned. Instead of exciting the qubit separately and then transferring the excitation to the TLS, they applied microwaves to the qubit at the TLS frequency while the qubit remained in the ground state. Using this method, the authors have measured the lifetime and coherence time of two TLSs as a function of detuning between the qubit and the TLS. At positive detuning ($\Delta_{ge,01} = f_{ge} - f_{01} > 0$) the measured values were found to be unaffected by the qubit (T_1, T_2 independent of detuning), however for negative values, the qubit shortens the measured values significantly down to $\Delta_{ge,01} = -300$ MHz.

In addition, the authors measured the temperature dependence of two representative TLSs up to 300 mK. In their measurement, the temperature dependence of T_1 is found to be parabolic, and decreases faster than the $\tanh(E_{ge}/2k_BT)$ trend above 100 mK. In addition, they report T_1-limited dephasing, which is not the case in our TLSs. We believe that the differences between our measurements and the ones reported in [72] are a result of the different nature of the TLSs. A different experiment has been able to directly control the characteristics of TLSs by the use of mechanical strain, applied with a piezoelectric crystal on the substrate of the device [74]. The authors have been able to show a control over the asymmetry parameter Δ by spectroscopically tracking the energy of individual TLSs.

3.1.6 Stochastic simulation

We reproduced the lifetime distribution of an ensemble of TLSs according to the TLS model in the following way. For each TLS we assume a uniform distribution of dipole orientations ($\cos\eta$ distributes uniformly, where η is the angle relative to the electric field inside the junction), a potential asymmetry Δ from a uniform distribution ($\Delta \propto z$, where z is the effective distance between position states of the TLS inside the junction) and a tunneling energy Δ_0 from a log distribution ($\ln \Delta_0 \propto z$). These distributions are consistent with the TLS model, which assumes linear sensitivity of Δ on z and exponential sensitivity of Δ_0 on z. We excluded TLSs having smaller couplings than we can measure. For each TLS we compute the lifetime according to Eq. 3.5: $T_1(\sin(\theta)) = a/\sin(\theta)^2$, where $\theta = \arctan(\Delta_0/\Delta)$ and a is some constant. The simulation data points were then averaged over a 7 MHz window size, as done for the experimental data.

The resulting $\langle T_1(S)\rangle$ behavior (see Fig. 3c in the main paper, blue diamonds) resembles a sum of two power laws. At smaller splittings $S \lesssim S_{max}/2$, the points fit a power law with an exponent $\alpha_1 \approx -1.9$, while for larger splittings they fit a power law with an exponent $\alpha_2 \approx -1$. A similar trend is observed in our data as well.

3.1.7 Discussion

Our results clearly show a trend of decreasing TLS lifetime with increasing coupling. In light of the recent experiment by Lisenfeld *et al.* [72], we wish to stress that our results cannot be explained by frequency proximity of the TLS and the qubit. According to the cited paper, a stronger coupling with the qubit results in an enhanced decay through the qubit. As shown in their work, this effect is relevant mostly for negative detunings. To exclude such an effect on our results, we note that our measurements were performed mostly at large detunings (typically a few 100 MHz, and always larger than 100 MHz), where we believe that the coupling with the qubit has negligible effect on the measured lifetimes. In addition, the qubit lifetime was about 2.5 times longer than that reported in [72], and therefore we expect a smaller effect on the TLS. We also checked that our results are unaffected by small detuning measurements, by using only TLSs from the ensemble for which $\Delta_{ge,01} < -300\,\mathrm{MHz}$ or $\Delta_{ge,01} > 100\,\mathrm{MHz}$. This does not change the results, and yields the same power-law exponent.

The T_2 data is less conclusive. The ensemble in this case was taken from only *half* of the TLSs used for the T_1 measurements since the Ramsey signal decayed too quickly and\or had a too small visibility. In this case there is no correlation between detuning from the qubit and absence of Ramsey fringes. Further measurements are required with a larger ensemble to conclude whether the average dephasing time peaks at intermediate couplings. In addition, spin-echo measurements on the same range may determine whether the dephasing is induced by energy fluctuations. We wish to add that such measurements were attempted by us, however the signal was too weak to measure, probably due to the additional π pulse with the longer interaction time with the qubit. Using the off-resonant method performed by Lisenfeld *et al.*, we believe such a measurement will be possible.

3.2 Nonlinear Dynamics in the Josephson phase circuit

The Josephson phase circuit is typically used as a qubit, which is encoded in the lowest two eigenstates of the anharmonic potential well. Indeed, when the anharmonicity is sufficiently large relative to the excitation bandwidth, such an approximation is valid and allows accurate single and multi-qubit gates to be performed [75]. However, at large amplitudes and\or large drive bandwidth, the true nature of the circuit as an anharmonic oscillator is revealed. The non-linearity of superconducting oscillators has been used previously to build low-noise amplifiers by biasing the Josephson junction with a strong microwave drive [76], which in principle can reach the standard quantum limit of noise added by an amplifier [77]. This method, called bifurcation amplification, has been utilized to readout the state of a qubit that is dispersively coupled to the amplifier [78]. Recently, a different method was proposed to readout the qubit state with high sensitivity and relatively small amplitudes. In this method, an anharmonic oscillator is driven with a frequency-chirp near a threshold, above which the oscillator phase-locks to the drive and the oscillation amplitude can become very large [79]. In a recent experiment, the width of the phase-locking threshold was measured as a function of temperature to illustrate the quantum-limit of such a method on detection sensitivity [80]. This classical nonlinear phenomenon, called autoresonance, has a wide range of relevance - from plasma physics [81] to planetary dynamics [82] and the generation of trapped anti-hydrogen atoms[83] but only recently has been measured in an electrical circuit.

Autoresonance has a universal scaling law for the phase-locking threshold, however, for any quantum system it changes at a certain parameter range that marks the transition from classical to quantum response [84]. In this work, we measure this transition in the Josephson phase circuit. The unique combination of our drive and device parameters makes it possible to map the phase-locking thresh-

old around the quantum to classical transition, and also measure the dynamics at any given point in parameter space. In the quantum regime, we show that using chirps the system can be easily prepared in an energy eigenstate with relatively high fidelity, and in the classical regime we demonstrate exponentially enhanced lifetimes of the phase-locked state, held by the drive. Our measurements demonstrate excitations up to the 18^{th} excited level for the first time in this system.

3.2.1 Classical autoresonance

An anharmonic oscillator cannot be excited to arbitrary large amplitudes with a fixed-frequency drive. In any such system, an initially resonant drive will at some point become detuned, as the system's oscillation frequency changes with its increasing oscillation amplitude. A classic example is the pendulum, where the oscillation frequency *decreases* as it is driven to higher amplitudes. At intermediate drive amplitudes, the resonance shape changes from a Lorenzian to an asymmetric shape, with the maximal response shifting to lower frequency. Above a critical amplitude, the steady state response becomes bi-stable (see Fig. 3.8a). The solution to the equations of motion at some frequency range below the linear resonance, reveals two stable modes corresponding to small and large-amplitude oscillations, and an additional unstable one between them [85]. In the bi-stable regime, a fixed-frequency drive will excite the pendulum only to the lower branch if initially at rest. The upper branch can be reached from rest either adiabatically, or non-deterministically via fluctuations. By sweeping the drive frequency (chirping) through the linear resonance sufficiently slowly, the upper branch can be reached deterministically, and ideally, with arbitrary small amplitudes. This results from phase-locking between the system and the drive, where the system automatically adjust its amplitude so that its instantaneous frequency matches that of the drive. The phenomena is thus called autoresonance. We briefly derive its main properties using the pendulum as a toy model. The potential energy in our system has a different shape, however the results of this analysis are similar

in our system.

The equation of motion of a driven pendulum is given by:

$$\ddot{\theta} + \omega_0 \sin\theta = \epsilon \cos\gamma t, \tag{3.6}$$

where ω_0 is the linear resonance frequency and the drive frequency γ is chirped down with a constant rate, $\gamma(t) = \omega_0 - \alpha t/2$, where α is the chirp rate. At negative times, the pendulum will not phase-lock to the drive because its driven mode of oscillation is comparable to the undriven one. Near $t = 0$, the driven mode becomes dominant due to the linear resonance condition and the system phase-locks to the drive. To track the dynamics at positive times, we first approximate Eq. 3.6 to third order (keeping the first nonlinear terms), assuming the drive frequency is not too far from the linear resonance:

$$\ddot{\theta} + \omega_0 \left(1 - \theta^2/6\right) \theta = \epsilon \cos\gamma t. \tag{3.7}$$

Assuming a solution of the form $\theta(t) = \Re(a(t)e^{i\phi(t)})$ we get two equations, for the real and imaginary parts. Neglecting harmonics of the drive and under a slowly varying amplitude approximation, Eq. 3.7 can be written using action-angle variables as:

$$\dot{I} = -2\tilde{\epsilon}\sqrt{I}\sin\Phi \tag{3.8}$$

$$\dot{\Phi} = \omega(I) - \gamma(t) - \tilde{\epsilon}\sqrt{I}\cos\Phi, \tag{3.9}$$

where we define the action $I = a^2/2$ and phase difference between the system and drive $\Phi = \phi - \gamma t$, the amplitude-dependent nonlinear frequency $\omega(I) = \omega_0(1 - \beta I)$ and the normalized drive amplitude $\tilde{\epsilon} = \epsilon/\sqrt{8}\omega_0$. To maintain phase-locking at positive times, $\dot{\Phi}$ must be close to zero at all times. We thus define an equilibrium action I_0 for which $\dot{\Phi} = 0$. Since I_0 varies slowly in time it is more

convenient to write $I = I_0 - \Delta$, where Δ is the small deviation of the action from its equilibrium value.

Note that equations 3.8 and 3.9 can be derived from the following Hamiltonian:

$$H(\Phi, \Delta) = S\Delta^2/2 + V_{\text{pseudo}}(\Phi), \tag{3.10}$$

where $V_{\text{pseudo}}(\Phi) = 2\epsilon\sqrt{I_0}\cos\Phi - \alpha\Phi/S$ and $S = \omega_0/8 + \epsilon/2I_0^{3/2}$ is only weakly dependent on time. We end up with a pseudo-particle moving in a time-dependent potential, of a tilted-washboard form (see Fig. 3.8b). For small tilts $\kappa = \alpha/2\tilde{\epsilon}S\sqrt{I_0} < 1$ the pseudo-particle is trapped within one of the metastable wells where there cannot be a cumulative change of the phase difference, and therefore phase-locking is maintained. In this case, the system is in autoresonance. If $\kappa > 1$, there is no trapping, and therefore the phase difference can accumulate. By requiring $\kappa < 1$ for any possible action, we arrive at the following criterion for autoresonance:

$$\epsilon_c > \sqrt{8\omega_0/\beta}\,(\alpha/3)^{3/4}\,. \tag{3.11}$$

Computer simulations show that this condition is sufficient to maintain phase-locking even in the strong-anharmonic regime (when the cubic approximation is no longer valid). In addition, it turns out that after phase-locking is achieved, the drive amplitude can even be reduced without losing phase-locking. Thus, for a non-dissipative system, the system can become phase-locked for an arbitrary small-amplitude drive, as long as condition 3.11 is met. As will be seen shortly, this condition breaks down below a certain amplitude due to quantum effects.

3.2.2 Phase-locking in superconducting anharmonic oscillators

Phase-locking in superconducting anharmonic oscillators was first demonstrated by Naaman *et al.* [79] on a system composed of a Josephson junction, embedded in a superconducting transmission line resonator. In their experiment, a frequency-chirped voltage was applied and the current oscillations in the system were mon-

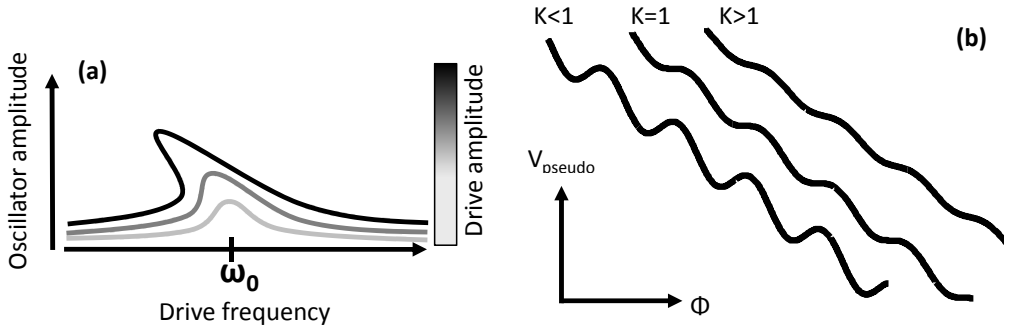

Figure 3.8: Autoresonance. (a) steady state response of the pendulum to a constant frequency drive at small, intermediate and large amplitudes. (b) The pseudo potential of the chirped pendulum.

itored as a function of drive amplitude. The amplitude of current oscillations (averaged over many chirp trials) was increased from zero to a significant fraction of the critical current over a finite range of drive amplitudes, and the threshold amplitude (defined as the voltage where the oscillation amplitude increases to $1/2$ of its maximum value) was found to scale with $\alpha^{3/4}$, consistent with the theoretical prediction. As we show in 3.2.5, the experiment described above operates in the *classical* regime [86], due to the small anharmonicity of the oscillator, where autoresonance theory holds.

Ideally, the threshold for autoresonance in the classical regime should be infinitely sharp, however due to its sensitivity to initial conditions, it gets broadened by fluctuations [86]. The effect of thermal fluctuations on threshold width was measured recently by Murch *et al.* [80]. In their work, the authors show that the transition width $\Delta\epsilon$ (the range of drive amplitudes where the phase-locking probability increases from 0.1 to 0.9) scales as the square-root of the effective temperature $T_{\text{eff}} = \frac{\hbar\omega_0}{2k_B}\coth(\frac{\hbar\omega_0}{2k_BT})$. At high-temperatures ($\hbar\omega_0 \ll k_BT$) the noise is dominated by thermal fluctuations and $T_{\text{eff}} \approx T$, while at low temperatures ($k_BT \ll \hbar\omega_0$) it is dominated by quantum fluctuations. Thus, quantum fluctuations impose an upper boundary on the sensitivity of an autoresonance-based measurement. Note that while the width of the threshold is dominated by quantum effects at low temperatures, the dynamical response is still classical. However, this is not the case at higher anharmonicity.

3.2.3 Phase-locking in the quantum regime

The dynamics of a quantum anharmonic oscillator in response to a nearly resonant drive are easily understood in the dressed basis. In Fig. 3.9a we show the dressed energies as a function of qubit-drive detuning $\Delta = f - f_{01}$, obtained by diagonalizing the system Hamiltonian in the rotating frame:

$$H_{\mathrm{N}} = \hbar \begin{pmatrix} 0 & \Omega/2 & 0 & 0 & \dots & 0 \\ \Omega/2 & -\Delta & \sqrt{2}\Omega/2 & 0 & & 0 \\ 0 & \sqrt{2}\Omega/2 & \epsilon_{02} - 2\Delta & \sqrt{3}\Omega/2 & & 0 \\ 0 & 0 & \sqrt{3}\Omega/2 & \epsilon_{03} - 3\Delta & & \vdots \\ \vdots & & & & \ddots & \sqrt{N}\Omega \\ 0 & 0 & 0 & \dots & \sqrt{N}\Omega & \epsilon_{0,N} - (N-1)\Delta \end{pmatrix}, \tag{3.12}$$

where $\epsilon_{0,\mathrm{n}} = 2\pi\left(f_{0,\mathrm{n}} - n f_{0,1}\right)$ is the cumulative anharmonicity[4] and $\Delta = \Delta(t) = 2\pi\left(f(t) - f_{01}\right)$ is the frequency detuning of the drive. In Eq. 3.12, we use the rotating wave approximation for a weakly anharmonic oscillator, and assume a harmonic drive coupling (see Appendix 5). Starting in the ground state at $t = 0$ and with a positive detuned drive (upper branch, left side of the figure), we change the detuning at a constant rate $\alpha < 0$. At $\Delta = 0$, the first transition is reached ($f = f_{01}$), where the ground and excited eigenstates anti-cross leading to a splitting of size $\hbar\Omega$ in energy. From the adiabatic theorem [87] it follows that if the chirp rate is sufficiently slow, the population stays on the branch, protected by the splitting. By remaining on the adiabatic branch the system will shift to the 1^{st} excited state. At a later time, the next transition (f_{12}) is reached and the system can similarly transit adiabatically to the 2^{nd} excited state, and so on. The Landau-Zener theory predicts that in the case of a two-level system, the probability of remaining in an eigenstate of the system at the end of the chirp is given by $P_{\mathrm{LZ}} = 1 - \exp(-2\pi\Gamma_{\mathrm{LZ}})$, where $\Gamma_{\mathrm{LZ}} = S^2/4\alpha$ and S is the full size of the splitting in units

[4] In Fig. 3.9a we use $\epsilon_{0,\mathrm{n}} = \frac{\beta n(n-1)}{2}$, where $\beta/2\pi = 158\,\mathrm{MHz}$.

of angular frequency. In the case of an anharmonic energy ladder (such as that of Eq. 3.12), we can find a generalized probability of remaining in an eigenstate:

$$P_{\mathrm{LZ}}^{\mathrm{n}} = \prod_{k}^{N} \left(1 - \exp(-2\pi\Gamma_{\mathrm{LZ}})^{k}\right), \tag{3.13}$$

where in this case $S = \Omega$ [86]. By solving for $P_{\mathrm{LZ}}^{\mathrm{n}} = 0.5$, we find that the threshold for adiabatically climbing the energy ladder is when $\Omega/\sqrt{\alpha} \approx 0.8$. Note that the Landau-Zener formula assumes that the chirp extends in time from plus to minus infinity. However, for sufficiently slow chirps $\beta \gg \sqrt{\alpha}$ [86, 84] one can treat the problem as that of successive Landau-Zener transitions and in this case Eq. 3.13 is correct.

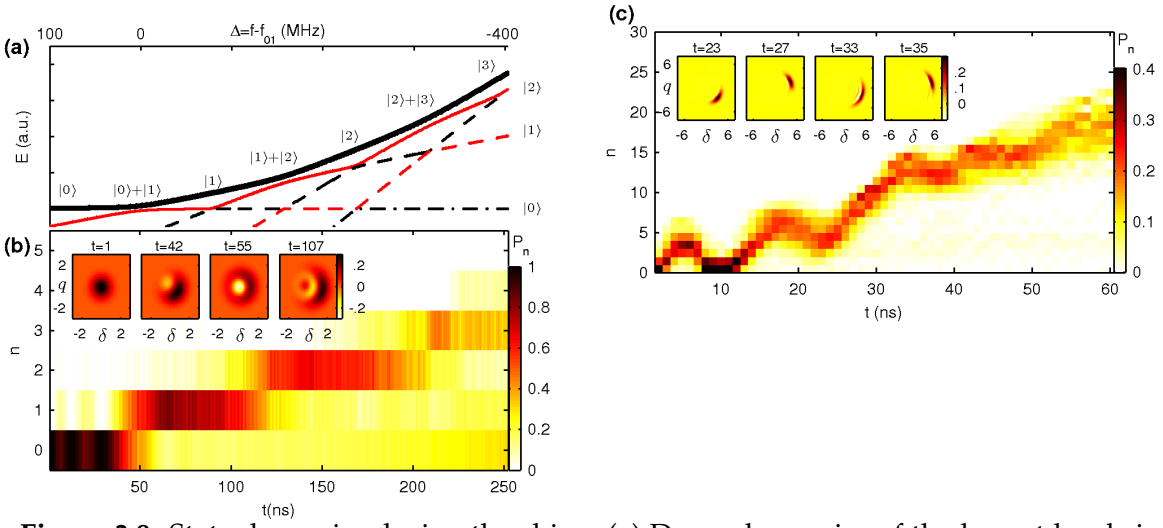

Figure 3.9: State dynamics during the chirp. (a) Dressed energies of the lowest levels in the rotating frame as a function of the drive frequency detuning Δ from the first transition f_{01}. As the chirp progresses (decreasing Δ), for a sufficiently small chirp rate the state remains on the adiabatic branch (solid black line). (b) Measured occupation probability (color-scale) as a function of time and level number in the ladder climbing regime ($\beta/2\pi = 158\,\mathrm{MHz}$, $\alpha/2\pi = 2\,\mathrm{MHz/ns}$, $\Omega/2\pi = 27\,\mathrm{MHz}$) and (c) autoresonance regime ($\beta/2\pi = 20\,\mathrm{MHz}$, $\alpha/2\pi = 10\,\mathrm{MHz/ns}$, $\Omega/2\pi = 190\,\mathrm{MHz}$). The detuning scale in (a) and the time scale in (b) are bound by the start and the end of the chirp. Insets: simulated Wigner distribution at different times along the chirp.

We start measuring the dynamics at a large anharmonicity $\beta = 158\,\mathrm{MHz}$. Figure 3.9b shows the processed data of P_{n} vs. time along the chirp for the relevant states n (see data processing details in 3.2.6). We clearly observe steps in the occupation,

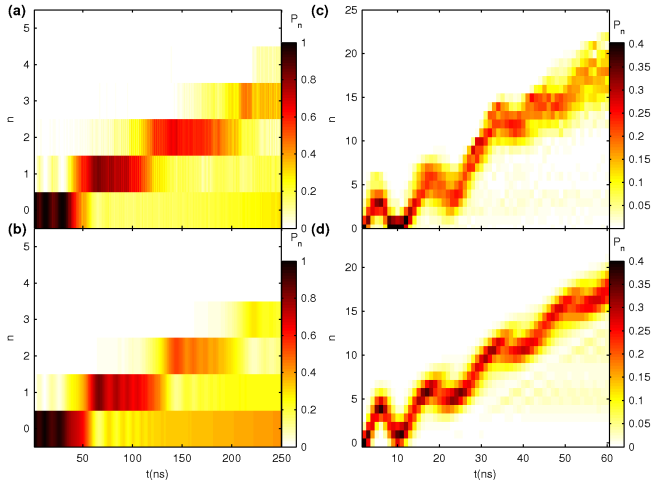

Figure 3.10: (a) Experimental data and (b) simulation of the dynamics experiment at large anharmonicity ($\beta/2\pi$=158 MHz) shown in Fig. 2b. (c) Experimental data and (d) simulation of the dynamics experiment at small anharmonicity ($\beta/2\pi$=18 MHz) shown in Fig. 2b.

corresponding to the ladder climbing effect. In phase space (see insets of Fig. 3.9b for Wigner distribution calculated from simulation), the phase is delocalized during each step, as expected from a Fock-type state ($|\psi\rangle = |n\rangle$). In between the steps, there is a partial localization of the phase due to the interference of two such states. The fidelity of each step in the experiment (the degree of correspondence with a Fock-type state) decreases as the state number n is increased, as a result of the chirp time being comparable to the energy decay time (T_1) of the first excited state.

3.2.4 Phase-locking in the classical regime

Next, we measure the evolution during a similar chirp but at a much smaller anharmonicity - $\beta/2\pi = 20$ MHz. Lowering the anharmonicity brings about more mixing between the levels for a given drive, and may therefore result in the simultaneous excitation of many levels. Figure 3.2.5c shows the measured time evolution under these conditions. Instead of sharp steps, we notice a broad excitation during the chirp, consisting of up to 6 levels. On top of that, we observe

large amplitude oscillations, as expected from autoresonant wavepacket dynamics [81]. The oscillations are also seen in phase space simulation[5] (see inset of Fig. 3.9c) where the phase of the localized distribution (crescent shape) oscillates during the chirp.

In Fig. 3.10 we compare our data with simulation. The simulation is calculated with no fit parameters and includes the effect of energy and phase damping. The energies at large anharmonicity (Fig. 3.10b) are estimated from spectroscopy data, while for small anharmonicity (Fig. 3.10d) they are extracted from the WKB calculation[6]. The experimental data and simulation agree qualitatively in both regimes. At large anharmonicity, the lengths of the "steps" are slightly different in the simulation due to the error in determining the bare transition frequencies (obtained from high power spectroscopy, where shifts and broadenings are significant). At small anharmonicity, we see a smearing of the oscillations at higher states. This is mainly due to the frequency dependent drive amplitude. In both measurements (and simulations), we used a frequency dependent drive which decays along the chirp $\sqrt{n(t)}$ (where $n(t)$ is the expected average state number as a function of time) to compensate for the increasing drive coupling at higher states which increases the mixing between the levels. This, however, does not affect the locking condition which is determined from the drive amplitude at the first transition [81, 86].

To check the stability of the generated wavepacket at small anharmonicity, we fix the amplitude and frequency of the drive at the end of the chirp to their final value (illustrated in Fig. 3.11a, in the case $\Omega_{\text{hold}} = \Omega$, where Ω_{hold} is the drive amplitude after the chirp). Figure 3.11b shows the resulting time evolution after the chirp. The phase-locked wavepacket is centered around $n \approx 18$ and is remarkably long-lived, despite the short decay time at these highly excited levels. We define the locking probability P_{locked} as the probability to be in the phase-locked state, taken for this measurement as the integrated probability

[5] A detailed description of our simulation methods is given in Appendix 5.

[6] For further details see 3.2.6

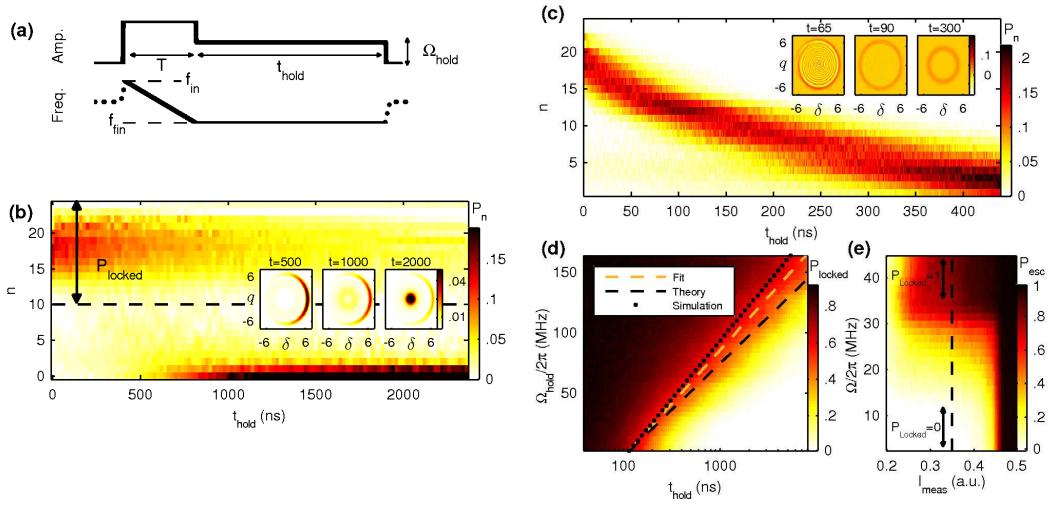

Figure 3.11: Decay of a wavepacket. (a) Time sequence of the decay measurement after the chirp. (b) Measured occupation probability (color-scale) as a function of level number and time after the chirp shown in Fig. 2c, with $\Omega_{hold}/2\pi = 190\,\mathrm{MHz}$ and (c) $\Omega_{hold} = 0$. Insets of (b) and (c) show the simulated Wigner plot at different times along the decay. (d) Measured locking probability (color-scale) as a function of time and amplitude of the drive after the chirp, with contours corresponding to $P_{locked}(t_{hold}, \Omega_{hold}) = 0.5$, obtained from data, theory and simulation. (e) Escape probability (color-scale) as a function of measurement amplitude I_{meas} and drive amplitude Ω after a chirp, with $\alpha/2\pi = 10\,\mathrm{MHz/ns}$ and $\beta/2\pi = 36\,\mathrm{MHz}$. To measure the locking probability, an intermediate I_{meas} is used (dashed line) at the end of the chirp.

for levels $n > 10$ [86]. The locking probability decays non-exponentially with a time constant $T_{locked} = 1.4\,\mu\mathrm{s}$, where T_{locked} is defined as the time it takes for the locking probability to decay to half of its initial value. The results of this experiment should be contrasted with the measurement shown in Fig. 3.11c, where $\Omega_{hold} = 0$. In this measurement, the energy expectation (proportional to the average level number) decays exponentially at roughly $T_1 \approx 300\,\mathrm{ns}$, consistent with the expected decay of a wavepacket in a nearly harmonic oscillator [88]. In phase space (insets of Fig. 3.11c) there is a quick (5 ns) delocalization into a pattern of circular fringes due to the non-negligible anharmonicity. The short lifetime-limited dephasing at $\langle n \rangle = 18$ smears out this pattern into a ring (30 ns) [89], shrinking at a constant rate $\Gamma_1 = 1/T_1$, as expected. When $\Omega_{hold} = \Omega$ (see insets of Fig. 3.11b), the locked population (crescent shape) remains localized, but slowly leaks out through the edge to the unlocked state, which freely decays as in Fig. 3.11c.

The results are explained within an effective barrier model [90, 91], where, the drive at the end of the chirp and the system's anharmonicity form an effective potential barrier for the population that is locked by the chirp. In this picture, the size of the potential barrier scales as the amplitude of the drive. We find from this theory that the resulting lifetime of the locked population is given by $T_{\text{locked}} \propto \exp(\eta\Omega_{\text{hold}}/2\pi)$ [90], where the parameter η depends on the system and drive frequencies (see below). To check this model experimentally, we measure the locking probability as a function of time after the chirp and of drive amplitude. In this measurement (see Fig. 3.11d) the chirp parameters are fixed, but the drive amplitude at the end of the chirp is varied[7]. We find that T_{locked} scales exponentially with Ω_{hold}, supporting the effective barrier picture. The holding lifetime increases by nearly two orders of magnitude to more than $10\,\mu$s. The factor η we extract from this data ($\eta \approx 26\,$ns), is in agreement with theoretical prediction ($\eta \approx 30\,$ns) and simulation ($\eta \approx 24\,$ns). Note that in this experiment, the chirp is only used to prepare the initial locked state.

The locking time $T_{\text{locked}} = W^{-1}$ (where W is the decay rate from the locked state of the nonlinear resonance) is calculated by Dykman et al. in the framework of quantum activation [90]. It is shown that in the case of weak damping and at low temperatures ($k_B T \ll h f_{01}$), the locking time is given by:

$$T_{\text{locked}} = c \exp(\eta\Omega/2\pi), \tag{3.14}$$

where, $\eta \approx 4/\sqrt{|f_{\text{fin}} - f_{01}|\,\beta/2\pi}$, and c is a constant on the order of T_1. This result is valid for intermediate drive amplitudes:

$$\frac{1}{2\pi T_1}\sqrt{\frac{4\,|f_{\text{fin}} - f_{01}|}{\beta/2\pi}} \ll \Omega/2\pi \ll |f_{\text{fin}} - f_{01}|\sqrt{\frac{4\,|f_{\text{fin}} - f_{01}|}{\beta/2\pi}}, \tag{3.15}$$

as is the case in our experiment, where these conditions translate to $4\,\text{MHz} \ll \Omega/2\pi \ll 3.7\,\text{GHz}$.

[7] This measurement is carried out with a different sample. The results are similar in both samples.

In this theory, the dynamics are considered to be classical while the noise is quantum, and is associated with zero-point fluctuations. Moreover, the expression for the locking time coincides with the classical formula for the escape time [92], when the classical temperature in [92] is replaced by an effective temperature, $T_{\text{eff}} = (hf_{01}/2k_{\text{B}})\coth(hf_{01}/2k_{\text{B}}T)$. A more intuitive, but equivalent theory for the locking time is given by Dykman et al. [91] where the escape time from an effective potential well associated with the phase-locked state is calculated. The potential barrier in this case scales as the drive amplitude.

We find good agreement between the simulation of this experiment at several anharmonicities and the scaling predicted by Eq. 3.14. The theoretical prediction of the factor η, calculated using the experimental parameters (see black dashed line in Fig. 3.11d) is within 15 % from that obtained in the simulation with the same parameters.

3.2.5 Quantum vs. classical response

The locking probability is directly measured using a calibrated measurement pulse. In Fig. 3.11e, as the drive amplitude is increased near the threshold ($\Omega_{th}/2\pi \approx$ 30 MHz), the highly excited (phase-locked) levels become more populated, as indicated by the increased escape probability at smaller measurement amplitudes. To measure the locking probability P_{locked}, we use a measurement amplitude that causes only the population in the upper levels to tunnel out (dashed line).

Although the state dynamics during the chirp is fundamentally different at large and small anharmonicities, it has common features in both regimes. In addition to the notable increase of the system's energy at relatively small drive amplitudes, both autoresonance and ladder climbing have a threshold in amplitude for phase-locking. As shown earlier, the threshold amplitude in autoresonance Ω_{th} scales as $\alpha^{3/4}$, while in the ladder climbing regime $\Omega_{th} \propto \alpha^{1/2}$. The change in scaling indicates a transition between the two regimes [84]. To map the transition, we measure the locking probability as a function of chirp rate, drive amplitude and

anharmonicity.

Following Marcus et al. [84] we plot the results (see Fig. 3.12a) in the dimensionless parameters space, $\Omega/\sqrt{\alpha}$ and $\beta/\sqrt{\alpha}$ [8]. The measured threshold, defined by $P_{\text{locked}}(\Omega/\sqrt{\alpha}, \beta/\sqrt{\alpha}) = 0.5$, changes scaling (the dependence of threshold amplitude on chirp rate) at thresholds where $\beta \approx \Omega$ (blue line). This condition is met when the broadening of the first transition (caused by the drive amplitude) is comparable to the frequency difference between neighboring transitions. This marks the transition between the classical and quantum regimes, where the energy levels are mixed or resolved [84, 86]. For comparison, the theoretical threshold lines of autoresonance and ladder climbing are shown on the same axes in red and black respectively. Our data converges to the theoretical scaling at the classical limit. At the quantum limit the threshold shows slow oscillations as a function of $\beta/\sqrt{\alpha}$, centered on the theoretical ladder-climbing threshold line with superimposed fast oscillations. The slow oscillations are reproduced by numerical simulation (see Fig. 3.12b) and are the result of multi-level Landau-Zener tunneling effects [86]. In the simulation, the amplitude of these oscillations decreases at larger $\beta/\sqrt{\alpha}$ values, converging to the theoretical ladder climbing threshold scaling [86]. We also find from simulation that the fast oscillations strongly depend on initial conditions (the distance from resonance), and are eliminated for sufficiently large chirp bandwidths.

3.2.6 Methods

To properly measure the locking probability P_{locked}, it is generally desirable to have the maximal possible chirp bandwidth $\Delta f = f_{\text{in}} - f_{\text{fin}}$ (f_{fin} being the final frequency of the chirp) in order to raise the energy expectation of the locked pop-

[8] The complete data set is obtained from measurements at four different anharmonicities ($\beta/2\pi = 170, 158, 35, 20$ MHz), in order to span a large range of $\beta/\sqrt{\alpha}$ while keeping the chirp time $T = 2\pi\left(f_{\text{in}} - f_{\text{fin}}\right)/\alpha$ shorter than $\sim T_1/2$, avoiding the effects of decay. We find that the threshold amplitude is affected by the decay for longer chirp times. The finite coherence time T_2 only weakly affects the threshold in our experiment (see 3.2.7 for further details on the effects of the coherence time).

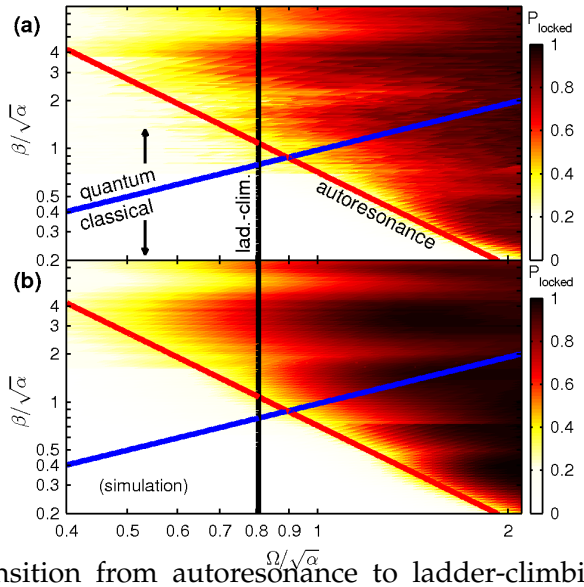

Figure 3.12: Transition from autoresonance to ladder-climbing. (a) Measured locking probability (color-scale) as a function of the dimensionless chirp parameters $\Omega/\sqrt{\alpha}$ and $\beta/\sqrt{\alpha}$. The red and black lines are the theoretical thresholds for autoresonance ($\Omega_{th}^{ar} = 0.82\alpha^{3/4}\beta^{-1/2}$) and ladder climbing ($\Omega_{th}^{lc} = 0.8\alpha^{1/2}$) [86]. The blue line ($\Omega = \beta$) marks the separation between the quantum and classical regimes [84]. (b) A simulation of the experiment shown in (a) with the same parameters, including the effects of decay and measurement at different β.

ulation higher. This leads to a better distinguishability between the locked and the unlocked population at the end of the chirp and correspondingly to an increased measurement fidelity of the locking probability, as illustrated in Fig 3.13. The AWG's bandwidth limitation results in an error of up to $\sim$10 % in P_{locked} at large anharmonicity (Fig. 3.13a), however it does not affect the threshold position in Fig. 4a. We use the maximal bandwidth (600 MHz) by varying the modulation frequency from 300 MHz to -300 MHz[9], and setting the oscillator frequency f_{LO} 200 MHz lower than the qubit frequency f_{01}. The additional 100 MHz of bandwidth beyond the qubit frequency is taken to reduce the sensitivity of the threshold to initial condition [93].

Data processing. To extract the state occupation probabilities P_{n}, we use the es-

[9]We use a smaller bandwidth only in some of the measurements shown in the parameter space diagram (Fig. 4a) in order to achieve larger $\beta/\sqrt{\alpha}$ values while keeping the chirps shorter than the decay time.

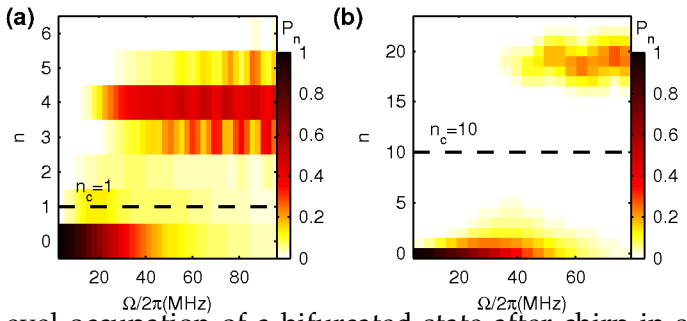

Figure 3.13: Level occupation of a bifurcated state after chirp in simulation, as a function of amplitude. Chirp parameters: (a) $\beta = 158\,\mathrm{MHz}$, $\alpha/2\pi = 6\,\mathrm{MHz/ns}$ and (b) $\beta = 20\,\mathrm{MHz}$, $\alpha/2\pi = 12\,\mathrm{MHz/ns}$. $\Delta f = 600\,\mathrm{MHz}$ and $f_{01} - f_{\mathrm{fin}} = 500\,\mathrm{MHz}$ in both. The locked and unlocked populations are discriminated by a level cutoff n_c (dashed lines), which is experimentally realized by a calibrated measurement pulse (see dashed line in Fig. 3e). At large anharmonicity relative to the chirp bandwidth (a), the locked and unlocked populations partially overlap, leading to a maximal error of $\sim$10 % in P_{locked} for the parameters used in the experiments.

cape probabilities vs. measurement amplitude data ("escape curve", $P_{\mathrm{esc}}(I_{\mathrm{meas}})$). We first measure the single-level escape curves by preparing the system in an $|n\rangle$ state, and then measuring the escape probability as a function of I_{meas} (see Fig. 3.14a). Once the single-level escape curves $P^{\mathrm{n}}_{\mathrm{esc}}(I_{\mathrm{meas}})$ are at hand, we decompose the measured escape curve of an arbitrary state into the single-level basis $P^{n}_{\mathrm{esc}}(I_{\mathrm{meas}})$ by optimizing the solution P_{n} to the set of J equations $P_{\mathrm{esc}}(I^{j}_{\mathrm{meas}}) = \sum_{n} P_n P^{n}_{\mathrm{esc}}(I^{j}_{\mathrm{meas}})$, where $j = 1, .., J$. Generating the $|n\rangle$ state becomes increasingly difficult at a larger n, due to the short lifetime of excited states. The procedure is even more problematic when the anharmonicity β is small and longer pulses are required to create the target state with reasonable fidelity. In practice, at the small anharmonicity regime, where $\beta/2\pi = 18\,\mathrm{MHz}$, it becomes impossible to prepare the system in an $|n\rangle$ state, even for $n > 1$. Instead, we use the first excited state escape curve, shifted by $\delta I_{\mathrm{meas}}(n) = I_{\mathrm{meas}}(0) - I_{\mathrm{meas}}(n)$ as an approximate escape curve. This approximation is supported by WKB calculation (see below). To determine the position of the escape curves we use the chirp data itself: for a given state, the measured escape curve contains information about the position of the single level escape curves. As seen in Fig. 3.14a, the position of these escape curves (defined as the point where the single-level escape curve increases to 0.5 of

its maximal value) is determined from the positions of the peaks in the derivative $\partial P_{\mathrm{esc}}(I_{\mathrm{meas}})/\partial I_{\mathrm{meas}}$. Due to the finite width of the single level escape curves, the peak corresponding to a certain level is visible only when the level occupation is sufficiently large. To find $I_{\mathrm{meas}}(n)$ for all the relevant levels, we sum the derivative over all the times along the chirp, as illustrated in Fig. 3.14b. The extracted $I_{\mathrm{meas}}(n)$ values are plotted in Fig. 3.14c (red circles).

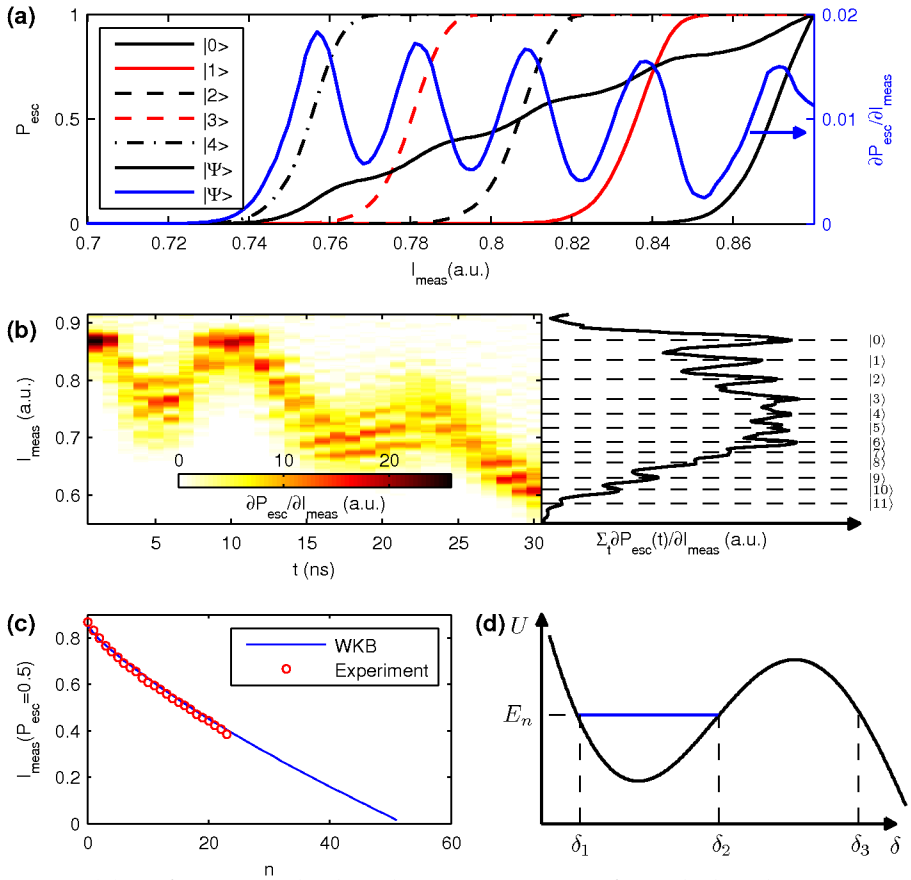

Figure 3.14: (a) Left axis: Calculated escape curves of single level states, and of the state $|\Psi\rangle = 1/\sqrt{5}\,(|0\rangle + |1\rangle + |2\rangle + |3\rangle + |4\rangle)$. Right axis: derivative of the escape curve of $|\Psi\rangle$. (b) Left panel: Derivative of the escape curve as a function measurement amplitude and time along the chirp shown in Fig. 2c. Right panel: Temporal sum of the data shown in the left panel, as a function of measurement amplitude. (c) Experimental and calculated positions of the escape curve. The WKB curve is calculated from the level dependent tunneling rates, based on the calculated energies using the best fitted circuit parameters. (d) Potential energy of the circuit used for WKB calculation, with classical turning points.

To check the validity of our estimate for the escape curves, we calculate them

numerically using the WKB approximation of the level dependent tunneling rates [94]:

$$\Gamma_n = f_n exp(-2iS_n/\hbar), \tag{3.16}$$

where $S_n = \int_{\delta_2}^{\delta_3} |p_n(\delta)|\, d\delta$ is the action, δ_{i} are the classical turning points defined in Fig. S2d, $p_n(\delta) = \sqrt{2m\,(E_n - U(\delta))}$ is the momentum, E_{n} and f_n are the energy and classical attempt frequency of/at the n^{th} level, $U(\delta)$ is the potential energy, $m = C\,(\Phi_0/2\pi)^2$ is the effective mass and $\Phi_0 = h/2e$ is a flux quantum. f_n is calculated using the classical oscillation time: $f_n = 1/\tau$, where $\tau = \oint dt = 2\int_{\delta_1}^{\delta_2} \frac{d\delta}{p(\delta)/m}$. The energies of the system are calculated by diagonalizing the system Hamiltonian (Eq. 1.14).

The circuit parameters are found by best-fitting the calculated lowest frequencies f_{01} and f_{12} to the measured ones and fixing the number of levels in the well to 50 (the number of levels in the well is obtained from extrapolating the experimental points in Fig. 3.14c to $I_{\mathrm{meas}}(P_{\mathrm{esc}} = 0.5) = 0$)[10]. The single-level escape curves are then given by, $P^{\mathrm{n}}_{\mathrm{esc}}(I_{\mathrm{meas}}) = 1 - \exp(-\Gamma_{\mathrm{n}}(I_{\mathrm{meas}})\Delta t)$, where Δt is the measurement pulse length. The calculated positions of the single-level escape curves are plotted in Fig. 3.14c (solid blue line).

We simulate the state dynamics of our N-level system under a frequency-chirped drive by propagating its density matrix ρ with the time evolution operator $U = \exp(iH_{\mathrm{N}}\Delta t)$ (see Eq. 3.12). The Rabi amplitude Ω is taken as a real constant during the chirp, and the detuning is a linearly decreasing function starting at +$2\pi\cdot$100 MHz and ending at -$2\pi\cdot$500 MHz, as done in the experiment. The cumulative anharmonicities $\epsilon_{0,\mathrm{n}}$ are calculated from the diagonalization of the system Hamiltonian (Eq. 1.14). The simulation neglects deviations of matrix elements

[10]The Hamiltonian in Eq. 1.14 is only an approximation of our system. In practice, the inductance may be flux dependent due to the small coupling between the circuit and the SQUID. For this reason, we find it difficult to obtain a good fit of the measured flux dependent energies to the calculated ones. We find that the parameters obtained from the fit have ∼10 % of uncertainty in the flux range used in the calculation.

due to the drive, beyond the harmonic oscillator approximation. Namely, for $m \neq n \pm 1$, we set $\langle n \left|\hat{\delta}\right| m\rangle = 0$, and for $m = n \pm 1$ we set $\langle n \left|\hat{\delta}\right| m\rangle \propto \sqrt{n+1}, \sqrt{n}$. We find for the first order matrix elements ($m = n \pm 1$) a maximal deviation of order $\sim 10^{-2}$ at the largest anharmonicity, and highest states. The second order matrix elements ($m = n \pm 2$) have a maximal value of order $\sim 10^{-1}$, relative to the first order term at the same m value. Higher order elements are smaller than $\sim 10^{-4}$. For $m = n$, the contribution to the energies for the range of drive amplitudes used in the experiment is small compared with the rotating frame energies. A separate simulation taking into account all the matrix elements, without the rotating wave approximation, yields similar results to within a $\sim 10^{-2}$ deviation (see Fig. 3.10b,d). Decoherence is taken into account using quantum operations [9] for amplitude and phase damping.

3.2.7 Discussion

Our experiment parameters allow us to observe the phase-locking threshold in the vicinity of the quantum-classical response boundary. Indeed, our data clearly shows a change in the amplitude-threshold scaling with chirp rate and anharmonicity. While in the classical regime, our data follows the autoresonance scaling, in the quantum regime our available $\beta/\sqrt{\alpha}$ range is insufficient for observing the expected ladder-climbing scaling. In practice, the maximal value we can obtain on the vertical axis of Fig. 3.12 is limited by both the maximal anharmonicity and the lifetime T_1; We find that for the anharmonicity we use in the quantum regime and chirp times on the order of T_1 or longer, the threshold shifts to larger amplitudes (compared to the ladder-climbing scaling). This is consistent with theoretical predictions [79, 95], indicating a non-negligible increase for $\varsigma \gtrsim 0.1$, where $\varsigma = T_1^{-1}\alpha^{-1/2}$. To measure the threshold deep in the quantum regime, we must optimize the ratio $\left(\beta/\sqrt{\alpha}\right)/\varsigma$. As the chirp rate cancels-out, we are left with the product βT_1. In addition, in order to properly distinguish between the phase-locked and the non-phase-locked population we must use an appropriate chirp

bandwidth, which is at least 2β. These factors limit the maximal value of $\beta/\sqrt{\alpha}$ to be below 10, where the threshold is oscillatory.

The coherence time T_2 only weakly affects the threshold in our experiment. This claim is supported by the fact that the measured threshold follows that of the decoherence free simulation despite the chirp time being longer than T_2. In this simulation, we see that at small anharmonicity, far off-diagonal elements of the density matrix (high-order coherence terms) of a high-amplitude phase-locked wavepacket are negligible and the phase space (Wigner) representation of the wavepacket is similar to the one calculated for a classical system [80, 86]. At large anharmonicity, we find that only the first order coherences of the state ($\rho_{i,i+1}$ terms of the density matrix) are non-zero and they are significantly populated for short times (compared to the relevant dephasing time), during transitions between neighboring levels.

3.3 Genetic optimization

Finite anharmonicity makes it possible to prepare states, composed of arbitrary superpositions of eigenstates within our system. However, Fourier broadening of the drive, together with power broadening of the transition energies causes nontrivial excitation at relatively small anharmonicities. Adding the short decay and coherence time, it becomes challenging to prepare a desired, yet pure state. We solve this difficulty, by optimizing the state produced with a feedback from the experiment. Our target state in the optimization algorithm is a superposition of eigenstates of our system. Although we can control only the level populations[11], these states can be used for obtaining accurate escape curves (see 3.3.3), Wigner images with l-fold symmetry (for states of the form $|\psi_l\rangle = \left(|0\rangle + e^{i\phi}|l\rangle\right)/\sqrt{2}$), measuring decoherence times and in multi-level gate optimizations 4.

[11] As we show in the next section, it is in fact possible measure all the phases as well using Wigner tomography, however this method was not yet implemented with an optimization algorithm

3.3.1 Algorithm formulation

Our optimization algorithm is based on guided evolution. Many other types of optimizations are possible, however a genetic-type is particularly powerful in multidimensional problems as we have. It is constructed in the following steps:
(a) We define a set of N_G pulse sequences (genomes). Each genome is a sequence of N_t complex numbers that represent the amplitude and phase of the drive's pulse envelope at each time step, with 1 ns time resolution. In addition, each genome is associated with a population overlap, defined as: $\chi = \sqrt{\overrightarrow{P_{\text{ideal}}}} \cdot \sqrt{\overrightarrow{P_{\text{meas}}}}$, where $\overrightarrow{P_{\text{ideal}}}$ and $\overrightarrow{P_{\text{meas}}}$ are the desired and measured population vectors . The genomes are initialized with random complex numbers, each having a maximal amplitude Ω_{max}, and $\chi = 0$.
(b) At each iteration of the algorithm (generation), we assume that the current set of genomes is sorted in order of decreasing χ. The upper N_a genomes are kept, and the next N_a genomes are replaced by a combination of the upper $N_a + 1$ genomes: $\overrightarrow{G_{N_a+k}} = \overrightarrow{\Upsilon}(G_k, G_{k+1}) + \overrightarrow{\epsilon}$, where $\overrightarrow{\Upsilon}(x, y)$ is a breeding function and $\overrightarrow{\epsilon}$ is a set of random complex numbers with a small maximal amplitude ϵ_{max} (noise). The bottom $N_G - 2N_a$ genomes are replaced by random complex numbers with an amplitude Ω_{max}.
(c) The bottom $N_G - N_a$ genomes of the next generation are then applied and the resulting probabilities $\overrightarrow{P^k_{\text{meas}}}$ for each genome are measured. After calculating χ of each, we sort our new set G in order of decreasing χ, and mark genome sequences that have (potentially) higher χ than the current optimal χ. For each genome having a potentially higher χ we repeat the measurement N_{rep} times, recalculate χ and reposition them in the set G. Step (b) is run again.
Breeding function. The breeding function $\overrightarrow{\Upsilon}(x, y)$ affects the efficiency of the algorithm. There are countless possible functions that can be used, and this point should be further explored in a future work. In this work, we frequently used a function that randomly selects an amplitude between the sets x and y at each time

step[12]. It imitates chromosomal cross-over which occurs in living organisms [96], and is found to work well for quantum state optimization problems, involving *coherent* dynamics. We also used a simpler breeding function that averages the input sets.

Fidelity measure. Our population overlap χ measures the distance between two states. There are many possibilities of defining the fidelity, and in general each gives a different bias for the algorithm towards finding a specific family of states. In our definition for example, χ is highly sensitive to the overall population that exists outside the subspace that is populated in the target state. However it is much less sensitive to the distribution among levels that are populated in the target state. Our definition is particularly useful for optimizing target states $|\psi_l\rangle$ for a Wigner measurement, because we wish to eliminate the population of states other than $|0\rangle$ and $|l\rangle$.

There are two important factors that limit the performance of any optimization algorithm in experimental systems: shot noise, and drifts in the physical parameters of the system.

Shot noise. Any finite number of repetitions results in some uncertainty in the measured probabilities. To determine n occupation probabilities in the final state we measure the tunneling probabilities after n measurement pulses having different amplitudes. In a typical optimization algorithm, we repeat the measurement of a single tunneling probability 900 times. Statistical analysis (see Appendix B for further details) gives a typical uncertainty (standard deviation) of $\sim$2 % in χ for sufficiently high values of χ ($\chi \gtrsim 80\,\%$) and therefore one cannot distinguish between two genomes with a χ difference which is smaller than $\sim$3 %. We therefore expect the algorithm to run significantly more slowly at χ's approaching unity, due to a false increase of χ. To increase the efficiency of the algorithm without increasing the number of repetitions of each measurement, we repeat the measurement only for genomes with potentially increased χ, as described in step (c).

Drifts. The response of the system to the application of a given genome can

[12] All optimizations done for Wigner tomography measurements used this breeding function

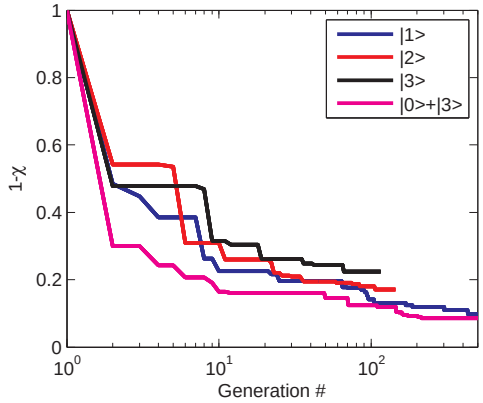

Figure 3.15: The error $1 - \chi$ of the best genome as a function of generation number for optimization of various optimized states.

change in time due to drifts in its physical parameters. We correct for drifts in the energies by performing a spectroscopic measurement of the qubit transition frequency every 10 minutes and adjusting the flux bias to restore the qubit frequency to its original value. In addition, we correct for drifts in the offset voltages applied to the IQ mixer that eliminate leakage of the microwave drive while it is turned off.

3.3.2 Optimization dynamics

The average trend of convergence to equilibrium with generation (iteration) number, depends on many parameters. Our algorithm parameters include, gene size (N_t), population size, amplitude of noise, breeding function and fidelity measure. On top of that, in any experimental realization the dynamics are affected by uncontrolled gene noise and uncertainties in the measured state. In this work, we do not attempt to study the effects of each parameter on the dynamics and efficiency of the algorithm, however from observing the data, we find several interesting characteristics.

Despite the large dimension of the optimization problem, the fidelity (population overlap or similar measures) increases sharply (on the order of 10 generations)

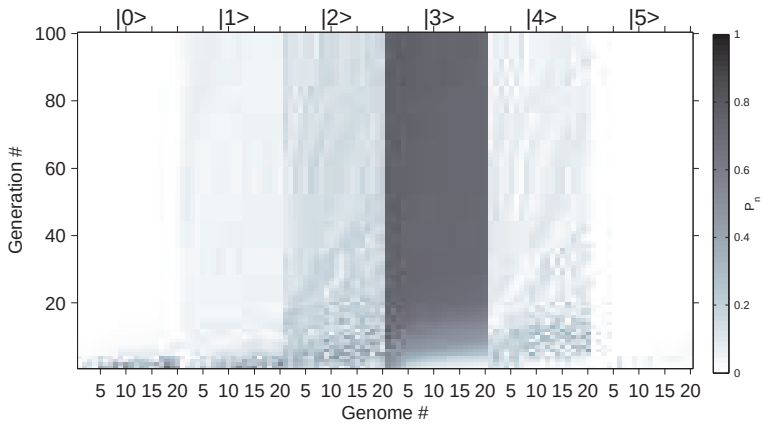

Figure 3.16: State populations evolution of the $|3\rangle$ state optimization. Level population P_n vs. iteration number and level number n for the 20 best genomes. The genome size is $N_t = 12$.

to close to a saturation value (Fig. 3.15). The error $1 - \chi$[13] then decreases logarithmically (on the average) with generation number and saturating to a value limited by the amount of noise we introduce into the algorithm. To achieve greater fidelity we decrease the noise manually. We find that the maximal achievable fidelity does not depend on initial conditions (in our case we choose random initial amplitudes).

In Figure 3.16 we show the complete data set of the $|3\rangle$ state optimization, including all the state occupation probabilities of the 20 best genomes. The best fidelity configuration propagates down to the lowest fidelity genome, making future generations more similar and the optimization more efficient. However, with increasing similarity, the algorithm converges towards a local fidelity maximum which is not necessarily the absolute maximum. In Fig. 3.17 similar trends are observed, however since the fidelity is not a monotonically increasing function of P_1 and P_3, these probabilities do not increase monotonically.

[13] In the case of optimization of single levels $|\psi\rangle = |l\rangle$ we use $\chi \equiv P_l$.

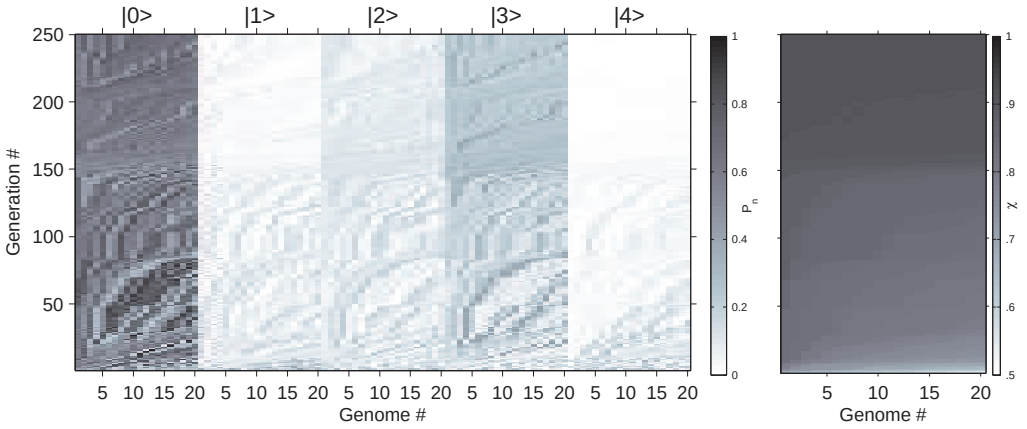

Figure 3.17: State populations evolution of the $|0\rangle + |3\rangle$ state optimization. Left: Level population P_n vs. iteration number and level number n for the 20 best genomes. The genome size is $N_t = 25$. Right: calculated fidelity.

3.3.3 Escape curves of optimized states

Using the optimization algorithm, one can achieve arbitrary superpositions given enough generations, without knowing the details of the system and system-drive interaction parameters. This property must be treated with caution, since the algorithm assumes a correct measurement of the population probabilities which is not always the case. Systematic errors in these measurements lead to errors in the optimized state. To independently confirm the correct state preparation we measure the escape curves of the optimized state. In Fig. 3.18 we plot the escape curves for several single-level optimizations, compared with the manually prepared states. All the states were prepared at small anharmonicity ($\beta/2\pi = 22\,\text{MHz}$) using a sequence length of 20 ns, 20 ns and 12 ns for the first, second and third excited states respectively. The escape curves of the optimized states we measure after the genetic algorithm are consistent with the measured level occupation of that state (reported at the end of the algorithm). Despite the short time-sequence, corresponding to a simultaneous excitation of several states due to Fourier broadening, the complex coherent dynamics lead to an efficient destructive interference of the amplitudes of all but one level.

To better understand these dynamics, we plot in Fig. 3.19 the optimized sequences

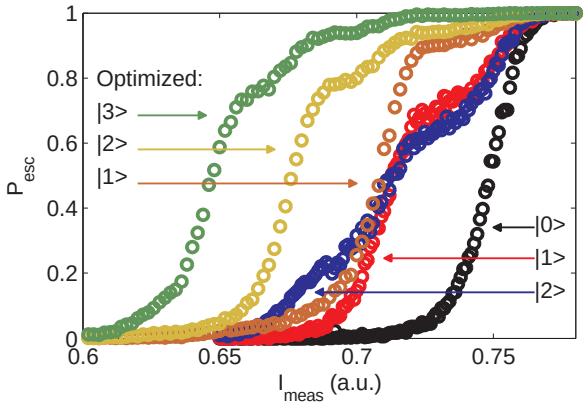

Figure 3.18: Optimized vs. non-optimized escape curves. Escape curves, measured immediately after state preparation in the case of non-optimized states (right) and genetically optimized states (left) at small anharmonicity ($\beta/2\pi = 22\,\mathrm{MHz}$). In manually prepared second excited state we use a long pulse, resonant with the two photon transition to minimize excitation of higher levels. This results in only $\sim 25\,\%$ of the population in this level. Preparation of higher excited states results in even a lower fidelity.

(used in the measurement shown in Fig. 3.3.3) as a function frequency, relative to the relevant transition frequencies. Although the exact form of each curve is complex, the gross structure can be intuitively understood. In the red curve (optimization of state $|1\rangle$), the strongest peak is slightly shifted above the first transition, so that it is suppressed on the second and third transitions, but sufficiently strong on the first to excite only the first level. The additional peaks resemble the ringing pattern of a sinc function, implying sharp ends of the pulse in the time domain. The width of the peaks is determined by the temporal width of the pulse and the total width of the Fourier envelope is determined by the bandwidth of the system. In addition, it is clear that the ringing pattern is asymmetric relative to the main peak and decreases non-monotonically, indicating an additional frequency component. This component suppresses the second order peak of the sinc pattern which otherwise would overlap with the third transition, causing undesired excitations. The second and third level's optimized sequences show a similar structure. However, they differ in the position of their main peak and its width. The relative amplitudes of the yellow curve at the first and second transitions, is close to the $\sqrt{2}$ ratio, expected from a spin 1 rotation. Treating the first three levels as an

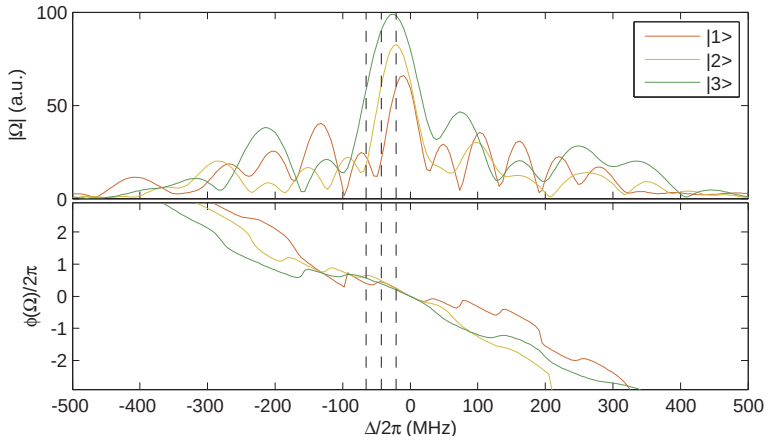

Figure 3.19: Fourier transforms of optimized sequences of the $|1\rangle$ (red), $|2\rangle$ (yellow) and $|3\rangle$ (green) states. In calculating the Fourier transform we use the DAC values of the optimized sets, convolved with the impulse response function of the microwave pulse shaping system. The impulse response function (using a 1 ns DAC pulse) is measured with a sampling oscilloscope after the microwave amplifier (see Sec. 2.2), and is found to be approximately a 1.6 ns FWHM Gaussian. All curves are plotted as a function of detuning from the LO frequency. The dashed vertical lines mark the first (0-1) second (1-2) and third (2-3) transition frequencies.

effective spin 1 system, the Hamiltonian can be brought to a form which is similar to the $SU(2)$ generator of rotation about x, if the drive contains two frequencies, resonant with the first and second transitions, with the appropriate ratio [97]. For such a case, the time evolution of the amplitude of the second excited state resembles that of a spin system in response to a resonant drive. Similarly, in the green curve, an amplitude ratio between the first and second and between the second and third transition frequencies is close to the expected from a rotation of a spin 3/2. Note that due to the small anharmonicity in our optimization, it is incorrect to use the effective coupling Hamiltonian used in Ref. [97] and therefore we expect deviations. For example, the phases of all frequency components should be the same in such a scheme, however we find different phases in Fig. 3.19, indicating there exists a time delay between them.

3.3.4 Conclusions

We demonstrate a working genetic optimization algorithm on a quantum system, involving a multidimensional parameter space ($d > 20$), where the dynamics are coherent. The algorithm is tested on the generation of states composed of one and two levels, and achieves exceptionally high fidelities relative to the ones achieved by manual calibrations. This method allows for an accurate measurement of single level escape curves (thereby improving multi-level measurements), improved measurements of coherence and decay times and for future gate optimizations.

3.4 State tomography in the Josephson phase circuit

In any quantum computing realization, it is essential to be able to readout the full quantum state of the system in the computational subspace. In addition to reading-out the result of an algorithm, state tomography is used for optimizing gates and characterizing an unknown process that occurs in the system. This technique has already been demonstrated in superconducting circuits for single, two and three qubits, however to date only one demonstration of state tomography of a d-level quantum system (qudit) has been demonstrated. A qudit-based quantum computing scheme is equivalent to the traditional qubit scheme, but uses less physical resources for a given computational subspace. Superconducting quantum information processing devices typically require more than two electrical leads per element (a quantum circuit unit) for control and readout that must be fed down to the millikelvin stage of a refrigeration system. In addition, as the number of elements grow, device design becomes more challenging due to undesired cross-talks between the elements. The qudit approach offers a promising simplification. For example, using a two-elements device, one can replace a four-qubit device using four levels in each qudit.

In this work we demonstrate arbitrary state preparation and quantum state tomography of a quantum three-level state (qutrit). We use this method to fully characterize decoherence and energy relaxation within this subspace. Although this method can in principle be generalized to higher-dimensions, tomography pulse calibrations become more complex and prone to errors due to decoherence. To solve this difficulty we use a different approach. We measure the state in phase space using Wigner tomography, and then transform it into a density matrix under the harmonic approximation. This method allows us to measure up to 5 states with relatively small errors, and can be potentially extended to large anharmonicities as well. We test it on superposition of eigenstates and on phase-locked wavepackets.

3.4.1 Standard state tomography

The density matrix of an arbitrary qudit state is obtained by applying a set of orthogonal unitary rotations, followed by a measurement in some chosen basis. In our system, we use pulses that are resonant with the qudit transition frequencies as unitary rotations, and then measure the level occupation probabilities. For a d-level state (a $d \times d$ density matrix), we have $d^2 - 1$ independent parameters that must be measured. These include $d - 1$ diagonal elements (d occupation probabilities summed up to unity), and $d(d-1)/2$ complex numbers (number of independent off-diagonal elements of an hermitian matrix). In general, any density matrix can be decomposed into a set of orthogonal hermitian matrices. It is convenient to choose this basis to be the generators of the special unitary group of degree n, SU(n). In this basis, the density matrix can be written as $\rho = \frac{1}{d}(I + \left\langle \vec{\mathbf{A}} \right\rangle \cdot \vec{\mathbf{A}})$, where $\mathbf{A}_i$ are the generators of the group and $\langle \mathbf{A_i} \rangle$ are their expectation values [98]. Many other bases are possible, however as we show below, the SU(n) generators are particularly convenient to work with.

3.4.2 Two and three-level standard state tomography

For a qubit system ($d = 2$), the generators are the Pauli matrices. To measure their expectation values, we use $\pi/2$ rotations around the x and y axes. For example, to measure $\langle \sigma_x \rangle$ we apply a $\pi/2$ rotation around the y axis. This is equivalent to applying the unitary transformation $\mathrm{U} = \exp(-i\pi\sigma_y/4)$ on the density matrix $\tilde{\rho} = \mathrm{U}\rho\mathrm{U}^\dagger$, which yields the following excited state probability: $P_e = \tilde{\rho}_{11} = \frac{1}{2}(1 + \rho_{01} + \rho_{10}) = \frac{1}{2}(1 + \langle \sigma_x \rangle)$. Similarly, applying a $\pi/2$ rotation around the x axis yields $P_e = \frac{1}{2}(1 - \langle \sigma_x \rangle)$. $\langle \sigma_z \rangle$ is directly measured in the energy basis (no rotations are required) and is given by $\langle \sigma_z \rangle = 1 - 2P_e$.

Tomography pulses are calibrated using a Rabi oscillations experiment. For example, to perform a $\pi/2$ rotation around the x axis, we apply a resonant pulse at zero phase. We then measure P_e as a function of microwave pulse length.

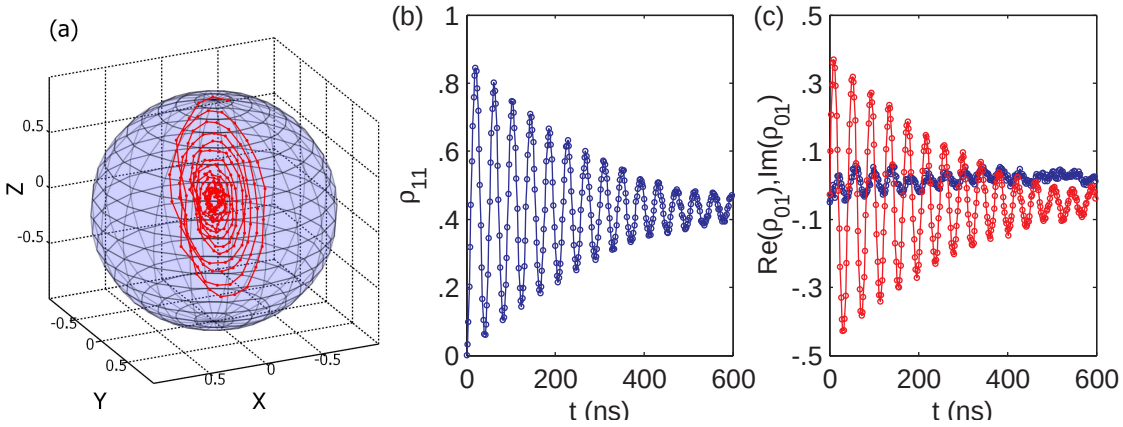

Figure 3.20: Qubit state tomography during Rabi oscillations. (a) trajectory of the state on the Bloch sphere. (b) Excited state probability vs. time. (c) Real (blue) and imaginary (red) parts of the off-diagonal element of the density matrix.

Such a measurement yields coherent oscillations with a period determined by the microwave drive amplitude (see Fig. 2.10). Since these oscillations decay in approximately 200 ns due to decoherence, we typically use pulse lengths of approximately 10 ns[14]. Instead of fixing the amplitude and finding the $\pi/2$ pulse length, we usually fix the time and vary the amplitude. This gives a higher resolution of rotation angles at short oscillation periods, allowing for a more accurate calibration. To perform a $\pi/2$ rotation we choose the drive amplitude to be a quarter of a full oscillation period, assuming the area theorem for a two-level system $\theta = \int \Omega(t)dt$, where θ is the acquired polar angle on the Bloch sphere[15]. To perform rotation about the y axis, we set the phase of the pulse to be $\pi/2$. In Fig. 3.20 we perform a state tomography measurement during a Rabi oscillations experiment. The Rabi amplitude in this experiment is $\Omega/2\pi = 24$ MHz and the detuning $\Delta/2\pi = 2$ MHz. The angle of the Rabi vector $\arctan(\Delta/\Omega)$ agrees with the resulting angle of precession ($\sim 5^\circ$) .

[14]Using shorter pulses will usually cause errors due to excitations of higher levels.

[15]The oscillations in a amplitude-Rabi experiment change in amplitude and period at large drive amplitude due to power shifts of the qubit resonance caused by the third level. At still larger amplitudes (comparable or larger than the anharmonicity), power broadening of the transitions causes non-negligible excitations of the third level and the dynamics become more complicated. At intermediate amplitudes, only power shifts are important and we iteratively correct the excitation frequency with a spectroscopy measurement at the updated drive amplitude.

State tomography on the qutrit can be performed using similar rotations. Here, we use the SU(3) generators, also known as the Gell-Mann matrices:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Note that the λ matrices can be constructed using the Pauli matrices, written in different subspaces. For example, $\lambda_1, \lambda_2, \lambda_3$ are constructed from the Pauli matrices on the $0-1$ subspace. λ_4, λ_5 are constructed from σ_x and σ_y in the $0-2$ basis, and so on. We can therefore find the expectation value of each of these matrices by $\pi/2$ qubit rotations around the x and y axes of the $0-1$, $1-2$ and $0-2$ qubits followed by measurement of level probabilities P_0, P_1, P_2 (for λ_3, λ_8 no rotations are required):

$$(\pi/2)_y^{01} \rightarrow P_1' = (1 - P_2 + \langle\lambda_1\rangle)/2 \quad (\pi/2)_x^{01} \rightarrow P_1' = (1 - P_2 - \langle\lambda_2\rangle)/2$$
$$(\pi/2)_y^{02} \rightarrow P_2' = (1 - P_1 - \langle\lambda_3\rangle)/2 \quad (\pi/2)_x^{02} \rightarrow P_2' = (1 - P_1 - \langle\lambda_4\rangle)/2$$

$$(\pi/2)_y^{12} \rightarrow P_2' = (1 - P_0 + \langle\lambda_5\rangle)/2 \quad (\pi/2)_x^{12} \rightarrow P_2' = (1 - P_0 - \langle\lambda_6\rangle)/2$$
$$\langle\lambda_3\rangle = P_0 - P_1 \qquad \langle\lambda_8\rangle = \tfrac{1}{\sqrt{3}}(1 - 3P_3)$$

To understand how these rotations are applied in our system, consider the qutrit Hamiltonian in the rotating frame of the drive, with the rotating wave approximation applied:

$$H_{\text{qutrit}} = \hbar \begin{pmatrix} 0 & \Omega/2 & 0 \\ \Omega^*/2 & -\Delta & \sqrt{2}\Omega/2 \\ 0 & \sqrt{2}\Omega^*/2 & \beta - 2\Delta \end{pmatrix}$$

To perform a rotation in the $0-1$ subspace, we set $\Delta = 0$. If $\beta \gg \Omega$, the second transition is far detuned from the drive and the population in the third level remains constant. The upper level causes only a shift in the resonance which is

found using spectroscopy. A $\pi/2$ pulse in the $0-1$ subspace is then calibrated using an amplitude-Rabi measurement, similar to the qubit tomography calibration.

To perform a rotation in the $1-2$ subspace, we set $\Delta = \beta$. If $\beta \gg \Omega$, the first transition is far detuned and the population in the first level remains constant. Note that the diagonal elements in this subspace are shifted by $-\beta$ due to the frequency difference between the $0-1$ and $1-2$ transitions. Since our goal is to have stationary phases in the measured density matrix during free evolution, we account for the phase accumulation in the third level by adding a phase $\beta\tau$ to the drive for $1-2$ rotations, where τ is the time elapsed from the beginning of the experiment. For a $\pi/2$ pulse of the same length, we use a drive amplitude which is approximately $\sqrt{2}$ smaller due to the increased coupling.

Rotations in the $0-2$ subspace are more difficult to perform, since there is no direct transition between the first and the third level. Numerical simulations show that the Rabi amplitude for a direct transition $\langle 2|\delta|0\rangle$ in our devices, working at the largest anharmonicity, is smaller by almost 2 orders of magnitude relative to that of the first transition. An additional option is to use the two-photon transition between these levels, which occurs at $\Delta = \beta/2$. In this case, transitions occur via a virtual level, however they are still much weaker than the direct transition. Using stronger pulses will cause power broadening of the first transition and therefore undesired rotation in the $0-1$ subspace and therefore one must use longer pulses instead. From numerical simulations, we find that it is necessary to use $\sim 100\,\text{ns}$ long pulses to achieve $\pi/2$ rotations in the $0-2$ subspace with reasonable fidelity. These are longer than the coherence time of the qubit, and as we show in Fig. (3.22) much longer than the decay time of ρ_{02}, and therefore we neglect this option.

We notice that the expectation value of λ_4 and λ_5 can be measured in the energy basis after a two-step rotation. To measure $\langle\lambda_4\rangle$, we apply a $\pi/2$ rotation about the y axis in the $0-1$ subspace, followed by a π rotation about the y axis in the

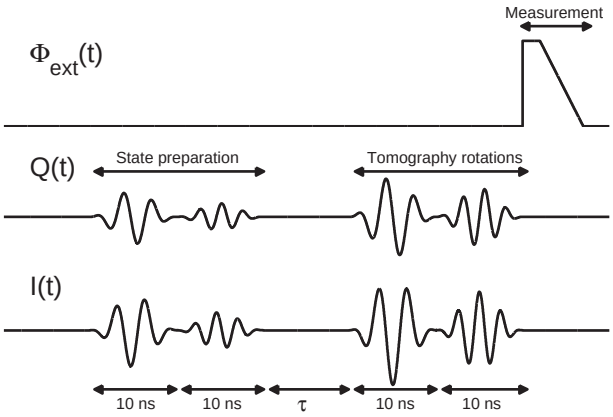

Figure 3.21: Qutrit tomography experimental sequence. The lower two waveforms describe the signal in the I and Q channels of the IQ mixer. The upper channel is the flux bias channel. These sequences are preceded by a standard reset of the qubit in the ground state of the computational well. The LO frequency in the experiment was $f_{\mathrm{LO}} = 12.690\,\mathrm{GHz}$ while the qubit transition frequencies (at low power), were $f_{01} = 12.757\,\mathrm{GHz}$, $f_{12} = 12.545\,\mathrm{GHz}$. The high-power pulse amplitudes used for $\pi/2$ rotations were: $f_{01}^{\pi/2} = 12.764\,\mathrm{GHz}$, $f_{12}^{\pi/2} = 12.501\,\mathrm{GHz}$ and for π rotations: $f_{01}^{\pi} = 12.792\,\mathrm{GHz}$, $f_{12}^{\pi} = 12.472\,\mathrm{GHz}$.

$1-2$ subspace. We then find $P_1' = (1 - P_1 - \langle\lambda_4\rangle)/2$. To measure $\langle\lambda_5\rangle$, we apply a $\pi/2$ rotation about the x axis in the $0-1$ subspace, followed by a π rotation about the y axis in the $1-2$ subspace. We then find $P_1' = (1 - P_1 + \langle\lambda_5\rangle)/2$. The second pulse is applied with an additional phase shift $\beta\tau_1$ to correct for the phase accumulation of the third level during the first tomography pulse.

In order to perform accurate rotations, our tomography pulses must be short compared to the shortest decoherence time in the measured subspace. However, for sufficiently short pulses, spectral broadening of the pulse and inevitable power broadening of the transitions[16] become comparable to the anharmonicity, and our two-level approximations are no longer valid. Although this trade-off cannot be circumvented due to the fundamental frequency-time uncertainty of pulses, with proper pulse-shaping one can achieve better control over the spectral shape and thereby improve the rotations. For example, square pulses have the advantage of being well defined in time, allowing for easy concatenation of pulses[17]. How-

[16] To keep the rotation angle constant, one must compensate for shorter pulse durations with a larger amplitude

[17] To reduce unnecessary delays between pulses (during which the state can decohere), we con-

ever, in the frequency domain they are described by a sinc function, which decays slowly with frequency, causing undesired transitions to other levels. In our experiments we use inverted-cosine shaping: $A(t) = (A_0/2)\left(1-\cos\left(\frac{2\pi t}{T}\right)\right)$ for $0 \leq t \leq T$. This function has the advantage of being well defined in both time *and* frequency, with $\Delta T_{\mathrm{FWHM}} = T/2$ and $\Delta\nu_{\mathrm{FWHM}} = 2/T$. The main disadvantage of this shape, compared to a square shape is that it requires twice the peak amplitude for a given rotation, causing a larger shift and transition broadening at its center. Using a 4-level simulation with these pulse shapes we find that the errors in the rotation matrices ($U = \prod_j \exp\left(-iH_{\mathrm{qutrit}}(t_j)dt/\hbar\right)$) are smaller than 3 %.
Using two-step rotations (first in the $0-1$ subspace and then in the $1-2$ subspace), we are able to prepare arbitrary 3-level states. To measure the decay of all the diagonal elements accurately, we prepare in each experiment an equal superposition state with arbitrarily chosen phases. We prepared the states: $\left|\psi_0^{01}\right\rangle = \frac{|0\rangle+e^{i2\pi/3}|1\rangle}{\sqrt{2}}$, $\left|\psi_0^{12}\right\rangle = \frac{|1\rangle+|2\rangle}{\sqrt{2}}$,$\left|\psi_0^{02}\right\rangle = \frac{|0\rangle-|2\rangle}{\sqrt{2}}$ (figures. 3.22a-c). We extract state fidelities using the standard definition $\mathcal{F} = Tr\sqrt{\sqrt{\rho_i}\rho_m\sqrt{\rho_i}}$, where ρ_i and ρ_m are the ideal and measured density matrices. We find $F^{01} = 97\,\%$,$F^{12} = 89\,\%$ and $F^{02} = 88\,\%$. We attribute the low state preparation fidelity for states involving the third level, to the cumulative error in a two-step rotation pulse and the longer sequence time (20 ns) during which decoherence occurs. In figures. 3.22d-f we plot the decay dynamics of the diagonal elements. We find that the dynamics are consistent with exponential decay from each level. The extracted decay times are summarized in Tab. 3.2. From Fig. 3.22d we extract a decay time for the second level $T_1^{11} = 170 \pm 5$ ns. From figures 3.22e,f we extract a similar decay time for the second level $T_1^{22} = 86 \pm 5$ ns. This is consistent with a decay of a nearly harmonic oscillator, where $\Gamma_{n,n-1} \propto \left|\langle n\left|\delta\right|n-1\rangle\right|^2 \approx n$ [88]. In figures. 3.22g-i we plot the dynamics in the off-diagonal elements. For simplicity, we show only the absolute values, which are insensitive to small detunings that may cause an oscillatory decay. While the off-diagonal element $|\rho_{01}|$ starts at nearly 0.5 in the

catenate them. Concatenation of pulses with a slowly decaying envelope causes significant errors and therefore it is crucial to have pulses that are well defined in time.

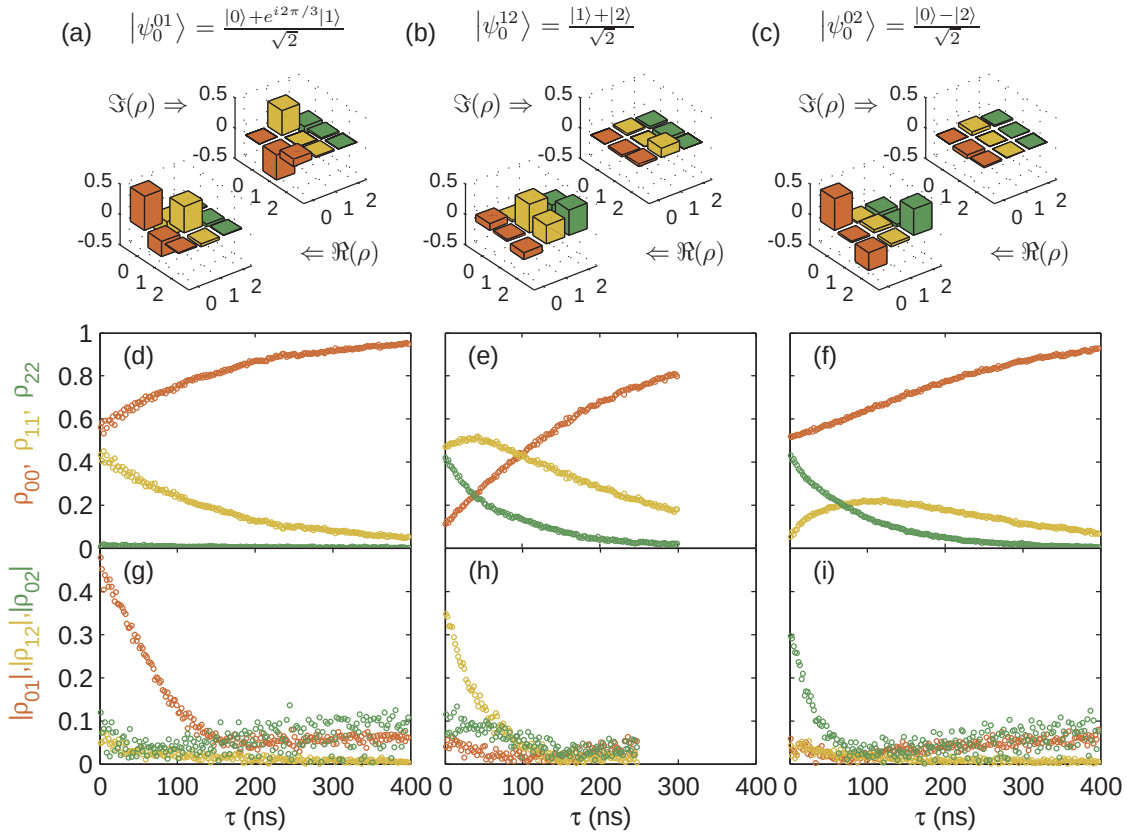

Figure 3.22: Qutrit tomography. (a)-(c) Real and imaginary parts of the density matrix immediately after state preparation for the equal superposition states: $|\psi_0^{01}\rangle$, $|\psi_0^{12}\rangle$,$|\psi_0^{02}\rangle$. (d)-(e) dynamics during free evolution of the diagonal elements for the initial states (a)-(c). (g)-(i) dynamics during free evolution of the off-diagonal elements for the initial states (a)-(c)

superposition state $|\psi_0^{01}\rangle$, the initial coherences in the other superposition states start off at a much lower value, due to decoherence during state preparation. The states' purity $\wp = Tr(\rho^2)$ are consequently much smaller for the superpositions involving the third level: $\wp^{01} = 99\,\%$, $\wp^{12} = 68\,\%$ and $\wp^{02} = 64\,\%$. The low purity we find in the second and third states are easily understood from the decay dynamics. In Fig. 3.22g we find that the ρ_{01} decays like a Gaussian, with a time constant $T^{01} = 75 \pm 5\,\text{ns}$. From figures 3.22e,f we find similar decay shapes, but at shorter time scales. The extracted decay times are $T^{12} = 65 \pm 5\,\text{ns}$ in Fig. 3.22e and $T^{02} = 26 \pm 5\,\text{ns}$ in Fig. 3.22f. Note that the noise in the off-diagonal elements is much larger than in the diagonal ones. In addition, ρ_{02} is noisier than other el-

n	0	1	2
0	-	75 ± 5	26 ± 5
1	75 ± 5	170 ± 5	65 ± 5
2	26 ± 5	65 ± 5	86 ± 5

Table 3.2: Qutrit decay times. The table shows decay times in nanoseconds of the density matrix elements.

ements. This is consistent with the fact that diagonal elements are measured with no rotation pulses, and ρ_{02} is measured with 2 rotation pulses. Note also that ρ_{02} and ρ_{01} increase slowly at long times due to errors in the two-step rotation, however, since the time scale over which they change is much longer than the decay time, we expect only a small correction for the extracted time scale. This effect is taken into account in the errors. The sharp decrease in coherence time of the high-order element is somewhat less severe than that found in a harmonic oscillator, subject to uncorrelated noise [99, 89].

3.4.3 Discussion

We demonstrate arbitrary state preparation and measurement in the lowest three levels in the well using pulses that are resonant with the first and second transitions (at f_{01} and f_{12}). Using an inverted cosine pulse shaping, we have reduced the gate errors to only a few percents. Further reduction of the errors (down to 10^{-4}) was later demonstrated using derivative reduction by adiabatic gating (DRAG) [75]. This method is applicable to any qudit rotations, where leakage and phase errors due to higher levels is minimized.

We find that the second-order coherence (ρ_{02}) decays much faster than the first order, however not as much as in the harmonic oscillator subjected to uncorrelated noise. Diagonal element decay rates increase, as expected, in proportion to level number. Due to the short coherence times, we get non-negligible errors in our pulses. Although using DRAG correction with shorter pulses, it is possible in principle to keep error level below 1 percent, ultimately, this method is not scal-

able to higher dimensions. For example, 4-level tomography will have to use 3 steps of rotations which give larger errors, take more time to apply, and are subjected to much higher decoherence rates in the $0-4$ subspace. We believe that to perform accurate tomography of up to 5 levels with current decoherence rates, it is essential to use single-step gates. Such gates may be accomplished using a generalized spin approach [97], however there is currently no solution to this hard problem. Other approaches including gate optimization algorithms (using genetic optimization for example) are also possible, however they require state tomography capabilities to run. As described in the next section, we solve this difficulty using phase-space tomography.

3.4.4 Wigner tomography

In systems with continuous degrees of freedom, e.g. the relative position and momentum of atoms in a molecule, wavepackets are often formed and therefore phase space distributions are better suited to characterize the state and its dynamics [100]. In particular, the Wigner distribution offers direct information about expectation values and purity, and provides a convenient framework to test the quantum-classical correspondence [101, 102]. In addition, since the Wigner distribution holds complete information about the state, it can be directly transformed into a density matrix and therefore its measurement is useful for quantum state tomography [9] as well. Numerous experiments measured the Wigner distribution of harmonic systems [103, 104, 105]. However, while anharmonic systems exhibit a wider variety of phenomena, a full quantum state reconstruction has so far been limited to atomic and molecular systems [106, 107, 108, 109].

Measuring the Wigner distribution in an anharmonic system poses several challenges. First, the phase of each level ϕ_n in the rotating frame advances in an increasing rate with n, causing wavepackets to disperse in phase space during the tomography pulse. For example, in a cubic potential this rate is given by $\dot{\phi}_n \approx \beta n(n-1)/2$, where $\beta = 2\pi\left(f_{10} - f_{21}\right)$ is the anharmonicity and $n = 0, 1, 2, \ldots$. Sec-

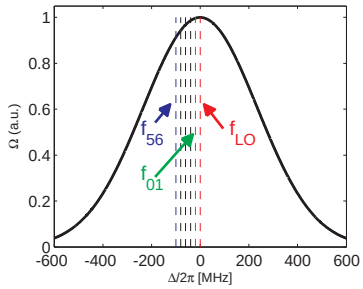

Figure 3.23: Tomography pulse envelope in the frequency domain. Dashed lines are the lowest five transitions $f_{n,n+1}$ and the LO frequency position in the experiment. The pulse envelope (solid line) is the normalized Fourier transform of a 1.6 ns FWHM Gaussian function. The effective drive amplitude of the lowest 6 transitions is the same, to within a 10 % variation. This is consistent with our data and simulations (see 3.4.4.1), showing small systematic errors within the corresponding subspace. In principle, more transitions can fit within this region by positioning the LO lower in frequency.

ond, in our system it is impossible to measure the probability distribution in the phase space quadratures directly, and therefore the prominent method for phase-space tomography is measurement of the parity after a coherent displacement [104]. In order to achieve an approximate displacement operation, one has to apply a pulse which is simultaneously resonant with all the transitions within the measured subspace. We find that both restrictions can be practically met by both reducing the anharmonicity and applying sufficiently short tomography pulses. Ideally, the Wigner distribution is proportional to the expectation value of the parity operator after a coherent displacement [110]. Due to the finite anharmonicity in our system, we use short, Gaussian-shaped resonant pulses as an approximate coherent displacement, while working at a weak anharmonicity. We control the phase and size of the complex displacement α by setting the phase and amplitude of the microwave pulse, using the relation $\alpha = -(1/2)\int \Omega(t)dt$ for a harmonic system, where $\Omega(t)$ is the time-dependent, Gaussian-shaped Rabi amplitude. For our experimental parameters[18], the pulse is simultaneously resonant with many transitions between consecutive levels, ~ 6 of which are subject to an amplitude variation of less than 10 % from the peak amplitude (see Fig. 3.23).

[18] We use a 1.6 ns FWHM pulse and an anharmonicity $\beta/2\pi = 20$ MHz.

Populations measurement. To obtain level populations of the lowest N levels we measure the escape probabilities for $N-1$ different measurements pulses and transform them into occupation probabilities using the single level escape curves (see Sec. 3.2.6). However in this experiment, we use a novel measurement pulse shape, to reduce distortions of the escape curves of higher levels. This measurement pulse is shown in Fig. 3.24. Typically, one must use a pulse with a slowly decaying end, in order to reduce the effect of population re-trapping in the original well after the pulse ends. The effect is similar to standard re-trapping observed in DC-SQUIDS [111], caused by low energy relaxation rates. Despite the fact that the pulse amplitude is large enough to quench the well, the wavepacket has not had the time to decay sufficiently in the large well to avoid re-trapping. This results in ground state escape curves saturating at a value smaller than unity. The effect of re-trapping becomes more prominent as the anharmonicity is reduced due to the corresponding deepening of the potential well, and therefore requires increasing both τ_2 and τ_3 to eliminate re-trapping. At the working bias point used in our experiment, we must use $\tau_2 = 25\,\text{ns}$ and $\tau_3 = 15\,\text{ns}$. These parameters are non-negligible relative to the decay time and cause artificial increase of the extracted population at lower levels, which in turn distorts the Wigner image. To avoid this effect we add a small step in amplitude at the beginning of the pulse ($\tau_1 = 1\,\text{ns}$) to preselect the population we want to escape, and then reduce the amplitude by an amount δA (a fixed $\sim 10\%$ of the pulse channel range). Even the small difference δA is sufficient to protect undesired population at lower levels from tunneling out during the long waiting time in which the escaped population decays in the external well.

Although the new pulse shape has removed the visible distortions in the measured Wigner images, it is not an optimal one. While the Wigner measurements are sensitive to the shape of the pulse, we find it impractical to study the effects of pulse shape via Wigner tomography, since each change of measurement pulse involves re-calibration of the single-level escape curves, a genetic optimization, and

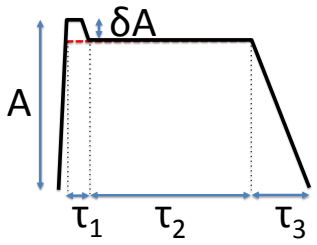

Figure 3.24: Measurement pulse used in the experiment. All the pulse parameters except for the total amplitude A are fixed.

a long ($\sim$ 1 hour) measurement of the Wigner function. Instead, we propose measuring the escape curves as a function of pulse parameters after a tomography pulse is applied on the ground state. By analyzing the derivative of the escape curves we can track changes in the population of each state (without knowing the single level escape curves) and infer what is the optimal measurement pulse. This point is left for a future work.

Density matrix fit. We extract the density matrix from the populations of the displaced states [104]. We use 200 homogeneously distributed random displacements within a $|\alpha| < 2$ circle to fit the density matrix, while restricting the density matrix to a 6×6 subspace. It should be noted that both the measured Wigner distribution and the extracted density matrix represent the state after a rotation that occurs during the tomography pulse. To get more accurate phases, one can apply an inverse propagator on the density matrix $U = \exp(-iH_0\Delta t/\hbar)$, where H_0 is the drive-free Hamiltonian and Δt is the effective pulse length for the rotation.

For a given anharmonicity β, a resonant pulse can be well approximated by a harmonic displacement in the limit $\beta T\,|\alpha|\,m^2/4 \ll 1$, where $|\alpha|$ is the size of the applied displacement, m is the maximal occupied level after a displacement and T is the pulse duration (see 3.4.4.1 for more details). This condition limits the maximal displacement to be well below the size of the distribution in phase space ($|\alpha_{max}| \ll 1$), for initial states occupying up to 5 of the lowest levels. However, we find in simulation that while the phases of the displaced state are very sensitive to the above condition, the probabilities are not, and therefore our approximation remains accurate for much larger displacements. We find that for our experiment

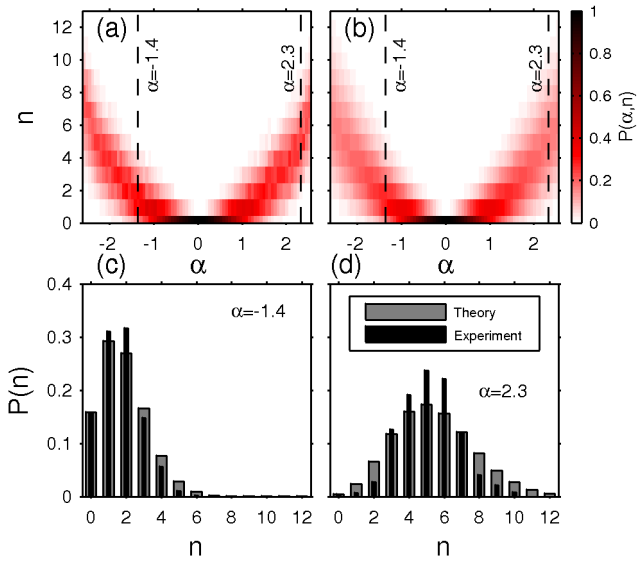

Figure 3.25: Tomographic pulse. Occupation probabilities for levels $n \leq 12$, as a function of $|\alpha|$ in the experiment (a) and theory (b). (c) and (d) show a histogram along the dashed lines shown in (a) and (b).

parameters, up to 5 levels can be measured accurately. The expectation value of the parity operator is calculated from the measured occupation probabilities and is given by, $W(\alpha) = (2/\pi) \sum_n (-1)^n P_n$ [110]. To test the effect of this pulse on our system, we initiate our system in the ground state and measure the occupation probabilities immediately after a short microwave pulse of total area α. Figure 3.25 shows the results of this measurement as a function of state number n and α (Fig. 3.25a), compared with the expected probabilities in a harmonic oscillator (Fig. 3.25b), $P(\alpha, n) = (1/n!) \exp(-|\alpha|^2) |\alpha|^{2n}$. As expected, the probability distribution in n is narrower compared to the harmonic system for higher amplitudes due to our finite bandwidth. To compare the data and theory quantitatively, we plot a histogram of the distribution (Fig. 3.25c,d) for $\alpha = -1.3$ and $\alpha = 2.2$. At $\alpha = -1.3$ our data fits well with a harmonic displacement. At the largest displacement values ($\alpha = 2.2$) the deviation from harmonicity becomes more apparent as expected.

To benchmark our method, we apply it on a set of superpositions of eigenstates,

l	$P(\|0\rangle)$	$P(\|1\rangle)$	$P(\|2\rangle)$	$P(\|3\rangle)$	$P(\|4\rangle)$	$P(\|5\rangle)$	τ (ns)	χ (%)
1	0.52	0.47	0.01	0	0	0	15	99.8
2	0.69	0.05	0.24	0.02	0	0	30	93.4
3	0.58	0.03	0.11	0.27	0.01	0	25	90.5
4	0.70	0.01	0.03	0.07	0.17	0.02	40	88

Table 3.3: Results of the optimization algorithm.

$|\psi_l\rangle = \left(|0\rangle + e^{i\phi}|l\rangle\right)/\sqrt{2}$[19]. These states have a simple structure in phase space and cannot disperse since they contain one measurable phase that results in only a free rotation. As we operate our system in the small anharmonicity regime, preparing such states with high fidelity requires longer pulse sequences. Due to our short decay and coherence times ($T_1 = 130\,\text{ns}$, $T_2 > 150\,\text{ns}$) we apply optimized control sequences to reduce decoherence effects to a minimum (see Sec. 3.3 for more details). The results of the optimization algorithm are summarized in Table 3.3.

As expected, χ decreases for a larger n, mostly due to our relatively short preparation pulse duration τ. To increase the value of χ for these states we increase τ, at the expense of the states' purity and optimization run-time. Our optimization algorithm does not take into account phases, however for sufficiently high values of χ, the Wigner image is dominated by only one phase, as can be seen in Fig. 3. In Fig 3a we plot the results of the tomography measurement on the ground state, and states $|\psi_l\rangle$ described in Table 1. These are compared with the expected images (Fig. 3b), calculated using the measured occupation probabilities, obtained independently by the optimization algorithm[20]. In Fig. 3c we plot the density matrices of the measured states, extracted using the harmonic oscillator wavefunctions. The diagonal elements in the extracted density matrices agree well with the measured occupation probabilities up to level $n = 4$. However, simulation shows that off-diagonal elements deviate increasingly for $n > 2$ (see

[19] ϕ is an unknown phase that can be determined via Wigner tomography. Since our optimization algorithm is only sensitive to state populations and not to phases, we cannot know it in advance.

[20] Due to the non-negligible occupation probabilities we get at larger l states, we add phases to the Wigner distribution to account for additional features in the image

3.4.4.1 for a detailed analysis). We find that the origin of these errors is primarily the significant decay and decoherence that occur at highly excited levels during the tomography pulse. The errors in the measured Wigner distribution can be further reduced with currently available higher coherence samples.

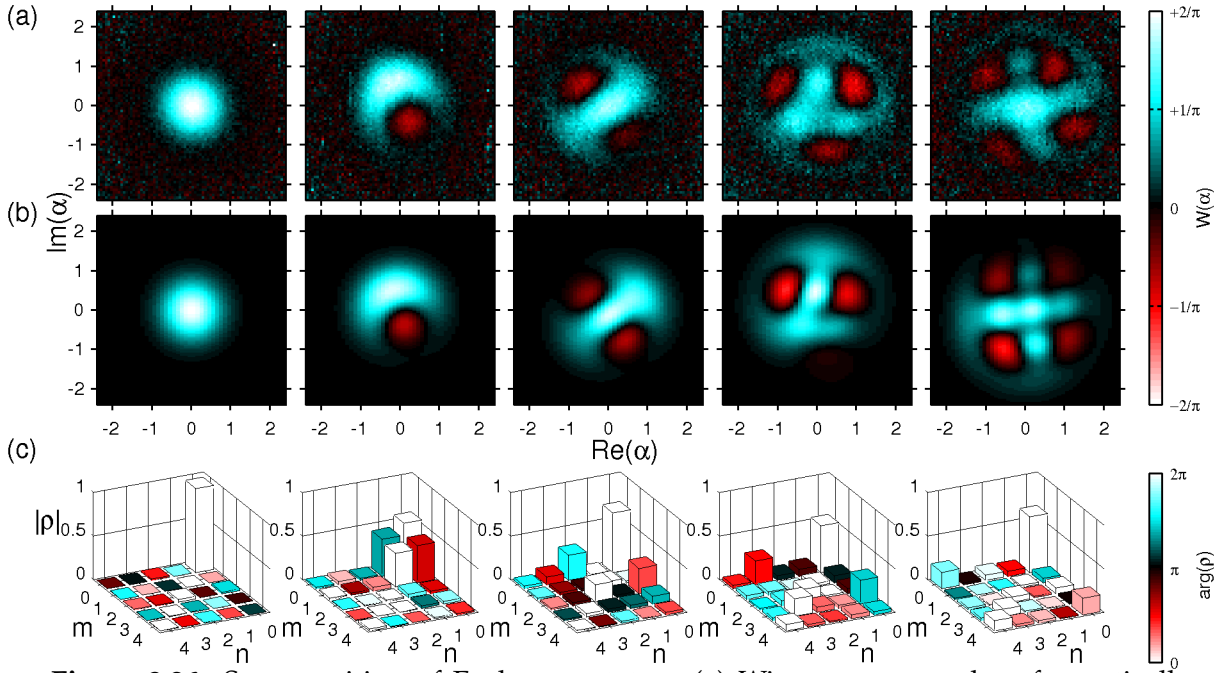

Figure 3.26: Superposition of Fock-type states. (a) Wigner tomography of genetically optimized superposition of Fock-type states and (b) calculated Wigner distributions of these states, using level occupations from table 1 and phases of the extracted density matrices. (c) Extracted density matrices from the measurements shown in (a).

3.4.4.1 Systematic errors

To quantify the errors in the measured Wigner function and density matrix caused by finite anharmonicity and decoherence, we perform numerical simulations . In our simulations, we propagate an anharmonic system, initialized with a pure state $\rho = |\psi\rangle\langle\psi|$, with resonant pulses. Each pulse is assumed to be resonant with the first transition and has a Gaussian envelope with the same width as in the experiment. We use the rotating wave approximation for constructing the Hamiltonian, and neglect corrections to the drive coupling beyond the harmonic approximation, namely $\langle n|\,\delta\,|m\rangle = 0$ for $n \neq m \pm 1$ and $\langle n|\,\delta\,|m\rangle \propto \sqrt{n}, \sqrt{n+1}$

for $m = n \pm 1$. This results in the following time dependent unitary propagator: $\mathrm{U}(t) = \exp\left(i\frac{\delta t}{2}\left[\left(\Omega(t)a^\dagger + \Omega(t)^* a\right) + \beta a^\dagger a(a^\dagger a - 1)\right]\right)$, where δt is the time step in the simulation, $\Omega(t)$ is the time dependent drive envelope amplitude and $\beta = 2\pi\left(f_{01} - f_{12}\right)$ is the anharmonicity. Decoherence and energy relaxation are taken into account using quantum operations, assuming only two phenomenological parameters: the energy relaxation time of the qubit T_1 and its pure dephasing time T_2.

We expect to have negligible systematic errors due to finite anharmonicity when the second term in the exponent of the propagator becomes negligible. By separating the terms in the exponent to first order using the Zassenhaus formula, and assuming a constant drive amplitude Ω of total duration T, we can approximate the propagator to $U \approx D(\alpha)\Gamma(T)\exp(-i\frac{T\beta}{4}[\left(\alpha a^\dagger - \alpha^* a\right), a^\dagger a(a^\dagger a - 1)])$ where $D(\alpha)$ is the displacement operator, $\alpha \equiv i\Omega T$, and $\Gamma(T) = \exp(i\frac{T\beta}{2}a^\dagger a(a^\dagger a - 1))$ is an overall dispersion operator. The second term can be considered negligible in the limit, $|\alpha|\, T\beta m^2/4 \ll 1$, where m is the maximal occupied state. For our experimental parameters ($\beta/2\pi = 20\,\mathrm{MHz}$, $T = 1.6\,\mathrm{ns}$), the second term can be neglected only for $|\alpha| \ll 1$, however, the error in the Wigner distribution, obtained from the state populations after the pulse are negligible even for $\alpha \approx 2$, as we show in our analysis.

We plot the errors in the Wigner distribution and the extracted density matrices, as a function of the anharmonicity β, initial state and decoherence parameters. The density matrix is extracted, as in the experimental data, using only the populations of the lowest 6 levels.

To quantify the error in the Wigner distribution, we calculate the cross-correlation (at zero offsets) between the ideal Wigner distribution of the initial state, and the one obtained from the expectation value of the parity operator after a set of displacements. The cross correlation is defined as

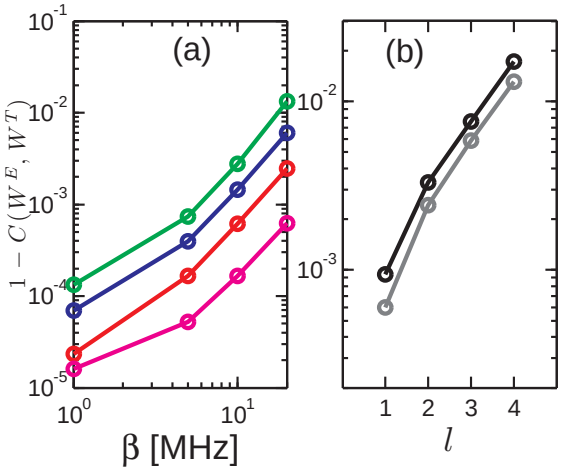

Figure 3.27: Systematic errors in the Wigner distribution, for initial states $|\psi_l\rangle = (|0\rangle + |l\rangle)/\sqrt{2}$, where $l = 1,2,3,4$. The cross-correlation deviation as a function of the anharmonicity (a) and initial state. In (a), decoherence is not included, and the magenta, red, blue, green lines correspond to the initial state $l = 1,2,3,4$ respectively. In (b) decoherence is included and the anharmonicity is fixed to $\beta = 20\,\mathrm{MHz}$. The black and gray lines correspond to decoherence parameters $T_1 = 150\,\mathrm{ns}$, $T_2 = 200\,\mathrm{ns}$ and $T_1 = 600\,\mathrm{ns}$, $T_2 = 600\,\mathrm{ns}$ respectively.

$$C(f(x,y),g(x,y)) = \sum_{x',y'} \frac{(f(x',y') - \langle f\rangle)(g(x',y') - \langle g\rangle)}{\sqrt{\sum_{x'',y''}(f(x'',y'') - \langle f\rangle)^2 \sum_{x,y}(g(x'',y'') - \langle g\rangle)^2}},$$

where $\langle f\rangle, \langle g\rangle$ are the average values of f, g. The results are plotted in Fig. 3.27. As seen in the figure, the error increases sharply with both anharmonicity and maximal populated level l.

To quantify the error in the extracted density matrices, we use the standard fidelity definition: $\langle \circ \mathcal{F}\rangle = 1 - Tr\sqrt{\sqrt{\rho^F}\rho^I\sqrt{\rho^F}}$, where ρ^I is the initial density matrix and ρ^F is the density matrix obtained from a fit to the populations of the displaced states. We choose the initial density matrix ρ^I to be $|\psi_l\rangle\langle\psi_l|$, where $|\psi_l\rangle = (|0\rangle + |l\rangle)/\sqrt{2}$. In addition, we calculate the error in the non-zero off-diagonal elements of the density matrix using the following definitions: $\delta\rho_{0l} = \left||\rho^F_{0l}| - |\rho^I_{0l}|\right|$ is the error in the amplitude of the matrix element, and $\delta\phi(\rho_{0l}) = \left|\phi\left(\rho^F_{0l}\right) - \phi\left(\rho^I_{0l}\right)\right|/2\pi$ is the normalized error in the phase of the matrix element. As seen in the Fig. 3.28, all the error measures are negligible (smaller than ~ 0.05)

when decoherence is neglected. However, when included, both the fidelity and the amplitude of the off-diagonal elements have non-negligible errors. For currently available samples having $T_1 > 600\,\text{ns}$, and correspondingly longer T_2, the errors due to decoherence can be substantially reduced; In this case, the errors become smaller than 0.1 in all the measures. As expected, the phase error is quite insensitive to decoherence.

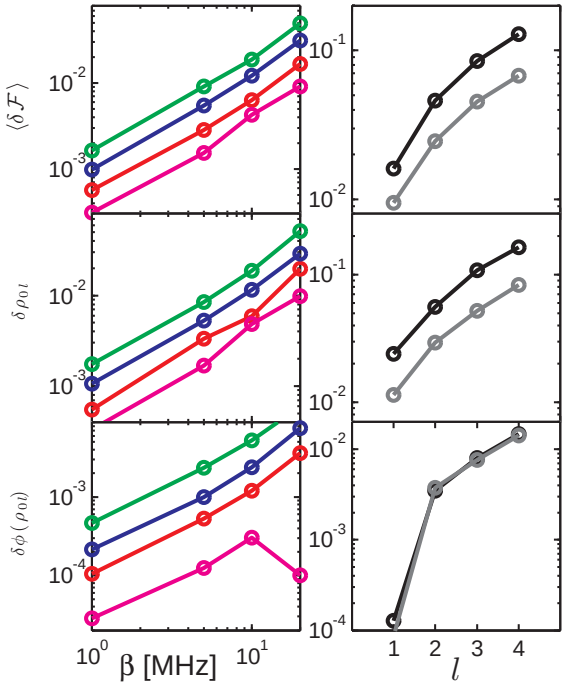

Figure 3.28: Systematic errors in the extracted density matrix, for initial states $|\psi_l\rangle = (|0\rangle + |l\rangle)/\sqrt{2}$, where $l = 1, 2, 3, 4$. The left column of plots shows the fidelity error, and errors in the off-diagonal matrix element ρ_{0l} as a function of anharmonicity, in the absence of decoherence. The magenta, red, blue and green lines correspond to the initial state $l = 1, 2, 3, 4$ respectively. The same error measures are plotted in the right column, as a function of initial state l, with decoherence included, at $\beta = 20\,\text{MHz}$. The black and gray lines correspond to decoherence parameters $T_1 = 150\,\text{ns}$, $T_2 = 200\,\text{ns}$ and $T_1 = 600\,\text{ns}$, $T_2 = 600\,\text{ns}$ respectively.

3.4.4.2 Shot-noise comparison between Wigner and Standard state tomography

As pointed out earlier, phase space is a convenient basis to experimentally acquire global information about the state (e.g. phase distribution, average energy

etc.) *fast*, on the expense of accurate knowledge of the state in the eigenstate basis. Therefore, when extracting the density matrix in the latter basis from a Wigner measurement, one requires an excess number of measurements compared to standard state tomography (SST) to achieve similar uncertainties. It turns out that this is only true for a certain class of states. For states that are dispersed in phase space, such as states composed of a coherent superposition of a small number of eigenstates, SST requires significantly less tomography pulses than Wigner tomography (WT) to achieve a comparable error in the density matrix but the *same* number of pulses for states that are localized in phase space. The overhead in the number of measurements is an important parameter from an experimental standpoint, and therefore we perform numerical simulations to calculate it. In the following we describe our simulation methods for both cases.

WT. We start with a pure initial state $|\psi\rangle$ and calculate the density matrix after a coherent displacement, using the coherent displacement operator. We keep the diagonal elements of the final density matrix, for a set of N_W random displacements α that are uniformly distributed in the complex plane. For each experiment (a particular displacement) we calculate an ensemble of M possible outcomes for the measurements of the diagonal elements, assuming r repetitions in the experiment, and a binomial distribution for the escape probabilities, from which the diagonal elements $P(n)$ are calculated. For each outcome, we extract the density matrix and calculate its fidelity $\mathcal{F} = Tr\sqrt{\sqrt{\rho^{tom}}\rho^{ideal}\sqrt{\rho^{tom}}}$. We then calculate the average fidelity $\langle\mathcal{F}\rangle$ of the ensemble to find the experimental error $\langle\delta\mathcal{F}\rangle = 1 - \langle\mathcal{F}\rangle$ due to shot noise.

SST. We start with a pure initial state $|\psi\rangle$. We construct a set of $N_{SST} = d^2$ orthogonal, unitary operations U_j to span a d-level subspace. The set of operations is chosen to be the generators of $SU(n)$ for convenience. From the diagonal elements of the rotated density matrices, we are able to extract the expectation values of each generator $\langle U_j\rangle = Tr(\rho U_j)$, and therefore reconstruct the original density matrix: $\rho = \sum_j \langle U_j\rangle \mathbf{U}_j$ [98]. As before, for each diagonal we calculate an ensemble

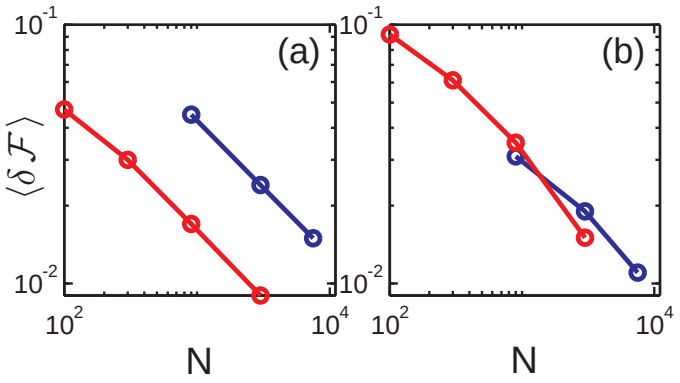

Figure 3.29: Error in the fidelity due to shot noise in standard state tomography and WT. $\langle \circ \mathcal{F} \rangle$ vs. N in WT (blue) and SST (red) of the initial states $|\psi\rangle = (|0\rangle + |4\rangle)/\sqrt{2}$ (a) and $|\psi\rangle = \left(1/\sqrt{5}\right) \sum_{k=0}^{4} |k\rangle$ (b). Both follow the $1/\sqrt{N}$ trend.

of M possible measurement outcomes due to shot noise. From each of the measurements in the ensemble, we extract the expectation values of the operators U_j and calculate the corresponding density matrix. The average fidelity error of the ensemble of density matrices is then evaluated.

Figure 3.29 shows the results of both simulations. We plot $\langle \delta \mathcal{F} \rangle$ as a function of $N = R$, where R is the number of repetitions of the experiment per tomography pulse. In the WT simulation we define N relative to the SST case: $N = R\left(\frac{N_W}{N_{SST}}\right)$, where N_W is the number of tomography pulses in the Wigner simulation, and N_{SST} is the number of tomography pulses in SST. While N_{SST} is fixed, we vary N in the WT simulation by using many displacement pulses while fixing R, and in the SST simulation we vary N by changing the statistics R. In all Wigner simulations we fix R to be 900 (as in our experiment) and in the SST simulations we vary R from 100 to 3000.

We see that for an initial state, composed of a coherent superposition of only 2 states (Fig. 3.29a), SST outperforms WT by a factor of 8, in terms of the fidelity $\mathcal{F}$. In contrast, for states composed of a coherent superposition of 5 eigenstates, the amount of information that is extracted per pulse in SST and WT is similar. This is because our chosen state is partially localized in phase space, and therefore requires less displacement pulses to extract its properties.

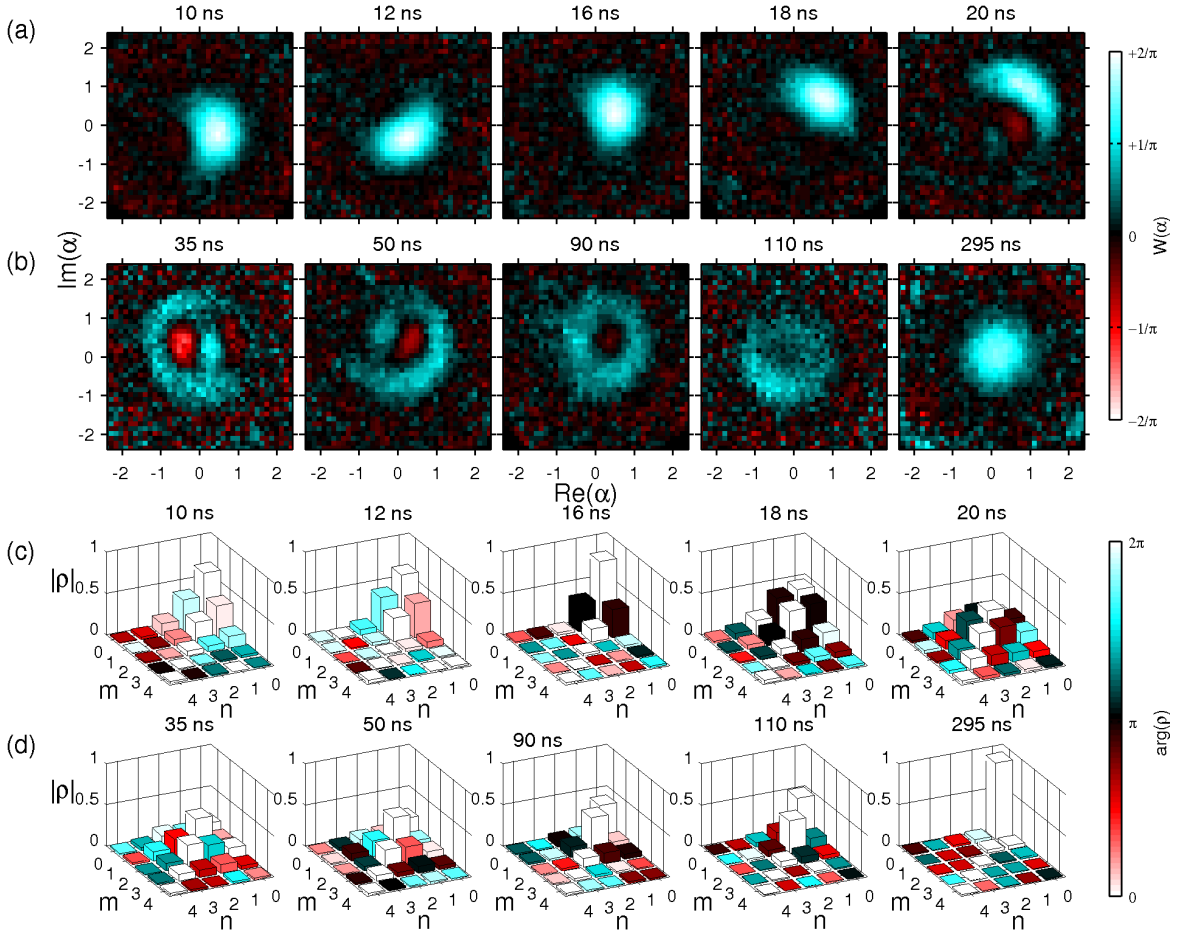

Figure 3.30: Dynamics of a phase-locked wave-packet. (a) Wigner tomography during a chirp ($T = 20\,\text{ns}$, $\Omega/2\pi = 66\,\text{MHz}$, $f_{in} - f_{01} = 320\,\text{MHz}$, $f_{fin} - f_{01} = -50\,\text{MHz}$), and (b) during drive-free evolution. (c) and (d) show the extracted density matrices for (a) and (b) respectively.

3.4.5 Wavepacket dynamics in phase-space

Wigner tomography of anharmonic oscillators is particularly useful to detect phase-locked states [81] since these are characterized by non-dispersive dynamics in phase space. Phase-locking to the drive may occur when the system is driven with a frequency chirp, such that the system's oscillation frequency follows that of the drive. To measure this effect, we initiate the system in the ground state and apply a negative frequency chirp, with a drive amplitude above the phase-locking threshold and a final frequency centered close to the transition frequency f_{23}. The chirp's temporal length and bandwidth $|f_{\text{fin}} - f_{\text{in}}|$ are chosen to be short (20 ns)

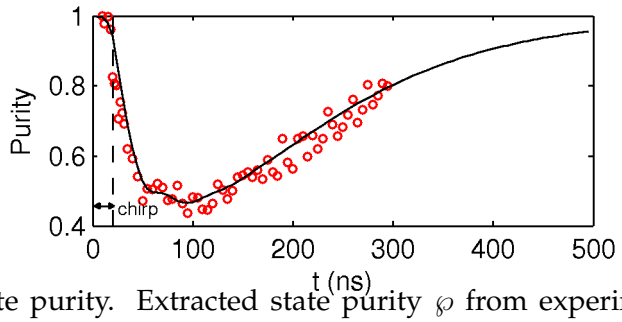

Figure 3.31: State purity. Extracted state purity $\wp$ from experiments (red circles) and simulation (solid line) during chirp and decay. The simulation includes the effect of the decay and dephasing and agrees with the measured decay time.

and large (600 MHz) respectively, in order have a broad excitation of states and for the excitation to be adiabatic [112]. Figure 4 shows the result at selected times along the chirp (Fig. 4a), and after the drive is turned off (Fig. 4b). The axes in the images are rotated at each time frame to fit the rotating frame of the drive. During the chirp, we see a displacement of the ground state distribution that gradually acquires a constant phase as the drive crosses the linear resonance ($f = f_{01}$). This happens, as expected around $t = 16$ ns and the shape of the wavepacket becomes crescent-like. After the drive is turned off ($t = 21$ ns), phase-locking is lost and the wavepacket disperses due to the finite anharmonicity. At $t = 35$ ns, the state has completely dispersed, however it still contains significant coherences, as indicated by the negative values in the Wigner plot and the large off-diagonal elements in the extracted density matrix (Fig. 4d). The state then dephases into a ring shape which shrinks in radius towards the ground state. To track the decoherence dynamics in this experiment, we extract the state's purity evolution directly from the Wigner distribution: $\wp = \pi \int d^2\alpha \left|W(\overrightarrow{\alpha})\right|^2$ [100]. The result is shown in Fig. 5 (red circles), and compared with a simulation (solid line). As expected, the purity remains high during the chirp, and then decays as a result of decoherence. The purity reaches a minimum and then ascends towards unity with an exponential rate, consistent with the measured decay time ($T_1 = 120$ ns) and in accordance with simulation.

3.4.6 Discussion

The harmonic response approach works well at small anharmonicities using our maximal available bandwidth. However, it cannot be applied in the large anharmonicity regime, where most of the qudit experiments are applied. This problem can be circumvented by quickly changing the bias at the end of the experiment to the small anharmonicity regime, where the tomography pulse is applied. This can be done in less than 2 ns, where decoherence effects are negligible. The accumulated phases can be accounted for using [113]:

$$\phi_{n,n+1} = 2\pi \int^{T_b} \left[f_{n,n+1}(\Phi_{\text{exp}}) - f_{n,n+1}(\Phi_{\text{meas}}) \right] dt,$$

(where $f_{n,n+1}(\Phi)$ is the bias dependent transition frequency between levels n and $n+1$) or calibrated out via Ramsey measurements[21].

As shown earlier, the maximal overhead in a Wigner measurement is only 8 fold relative to standard state tomography. We can therefore integrate a Wigner measurement into the optimization algorithm to optimize arbitrary states (including phases). Having perfected arbitrary state preparation and measurement it is then possible to optimize arbitrary gates via process tomography [9].

[21] A technical difficulty arises when trying to implement this approach, since it requires that in each experiment the phase of the tomography pulse be locked to the phase of the pulses used for state preparation. This can be done in a number of ways: (a) A large bandwidth, high sampling rate AWG capable of fully creating arbitrary microwave pulses up to a 9 GHz carrier frequency, without using a mixer. Such an instrument is commercially available but very expensive. (b) Cascading several mixers, each modulated with a standard AWG (300 MHz bandwidth). This will probably cause phase distortions that will have to be corrected, due to the dispersion of IQ mixers. (c) Use two independent oscillators (with two separate control AWGs), combined together. To lock their phases, one can either program the AWG to start the DAC sequence when the phases are locked (using a timer) or create a trigger for the DAC using the combined signals of the oscillators.

CHAPTER 4

Summary and Outlook

The Josephson phase-circuit is a remarkable device. It allows an operator to create complex quantum states and measure them, and is flexible enough support these capabilities at qualitatively different regimes. Using these features we mapped-out the response of this system to a resonant excitation and found a clear transition between classical autoresonance and quantum ladder climbing. The phase-circuit's architecture has the unique characteristics, required to observe this transition in a single device and with relatively modest resources. To demonstrate control and measurement of arbitrary states in this system, we used quantum state tomography on the lowest three levels and found their decoherence parameters. While the second excited state decays twice as fast as the first excited state, as expected, we find that the coherence between the ground state and the second excited state decays slower than expected in a nearly harmonic system, subject to uncorrelated noise. However, the short coherence time of our device limits the accuracy of our measurement. To perform more accurate measurements on these system, and generalize this measurement to higher qudit dimensions, shorter control pulses must be devised. We circumvented this difficulty at the small anharmonicity regime by using Wigner tomography with short pulses. Using genetic optimizations, we successfully generated non-classical states at short times, despite the small anharmonicity, and directly observed phase-locking in phase space during a chirped excitation.

4.1 Future work

Multi-level states hold promising prospects for QIP. This work is merely one of the first steps towards fulfilling this vision [97]. In the following, we suggest several directions related to this work, which we believe will make an important contribution.

4.1.1 Measuring multi-level decoherence rates

An essential characteristic of a qubit device is its T_1 and T_2 values. This determines the number of gates that can be performed with high-fidelity, or the maximal fidelity with which a certain gate can be performed. The same holds for qudits. Qudits are characterized by $d-1$ relaxation times T_1^n, and $d(d-1)/2$ decoherence times T_2^{mn}. In addition, as stated earlier, knowledge of T_2^{mn}, can help classify the type of noise in the system. Although it is straight-forward to measure these (simply by preparing states containing a significant population in the relevant density matrix element and performing standard state tomography at a variable delay), we find this task to be difficult to apply accurately for $d > 3$ with current phase-qubit devices. One promising method that can be readily used is arbitrary state preparation by genetic optimization, demonstrated in Sec. 3.3. To measure relaxation times, one can simply prepare a state with sufficient population in level n, restricting the population at higher levels to be small, and measure this level's population at variable delays. To obtain decoherence times, one may prepare the state at level n, and perform a generalized Ramsey sequence: $(\pi/2)^{nm} \rightarrow \tau \rightarrow (\pi/2)^{mn}$, followed by a measurement of P_n. A $(\pi/2)^{nm}$ gate should be optimized by a *gates* optimization algorithm which we currently lack, however the Ramsey sequence does not require its pulses to be identical to the $(\pi/2)^{nm}$ gate. We only need to make sure that the first pulse brings the state to an equal superposition of m and n, and that the second pulse brings that superposition (up to an arbitrary phase) to level n. Even when other states do not rotate

according to the $(\pi/2)^{nm}$ gate, as long as equal superpositions of n and m rotate back to level n, the Ramsey signal will be the same. Both pulses are then optimized using the standard genetic algorithm[1]. The envelope in this measurement does not decay to 0.5, due to the decay of probability out of the $m-n$ subspace, however this can be easily corrected using the measured T_m and T_n.

4.1.2 Optimizing multi-level gates

To optimize gates, one should normally perform quantum process tomography (QPT). QPT involves standard state tomography on a complete set of basis states that the process is applied upon. Since state tomography of a qudit scales as d^2, QPT scales as d^3, and so does the number of parameters, characterizing the process. However, to optimize gates we are usually interested in a single merit that describes how far is our gate from the desired one. A simple solution is to use Random Gate Benchmarking [114]. In this method a set of random gates, composed from the d^2 orthogonal unitary gates ($\pi/2$ rotations about all the qubit subspaces in the x and y directions) are applied to calculate gate fidelity [2]. Despite the simplified algorithm for optimizing such gates, there is a one major challenge involved: calibrating the d^2 $\pi/2$ rotations. As concluded in Sec. 3.4, this cannot be done easily, even for three levels, due to the short coherence time. To perform such gates, one would have to find clever methods of performing these pulses in a short time (~ 5 ns in total). This could possibly be done using optimal control theory. Once these pulses are at hand, arbitrary single qudit gates can be demonstrated. One interesting case is the Quantum Fourier Transform (QFT), on a 4-level state [3]. This is the simplest possible QFT, equivalent to that applied with two qubits.

Only after single qudit gates are demonstrated, one can utilize them for multi-qudit quantum computing. Similar to the qubit case, universal quantum comput-

[1] To optimize the second pulse, we use the sequence for the optimized first pulse at the beginning of the sequence, which is kept constant.

[2] A preliminary demonstration of this method for a qudit using simulations is available in a private writeup by Uri Vool.

[3] See private writeup by Uri Vool

ing in qudits is possible with the combination of arbitrary single qudit gates and one two-qudit gate [115]. Two-qudit gates are more difficult to implement. In a phase circuit qutrit, for example, there are two relevant frequencies. Both frequencies will contribute to the entangling gate when the qudits are brought into resonance, and therefore the interaction will yield a more complex state evolution. Alternatively, one could bring the lower transition of one qutrit into resonance with the upper transition of the second, producing a more controlled interaction. In a 4-level qudit, however, it is not possible to do such a thing between all pairs of transitions without passing through another. One possible solution would be to use a tunable resonator as a filter for certain transition frequencies.

4.1.3 Cat states

Using chirps, it is possible to generate a superposition of classically separable states ("Schrödinger cat states"), and perform Wigner tomography on these states. This can be done with a series of chirps, each with an amplitude near the threshold. If for example, the phase-locking amplitude of the first chirp is equal to 0.5, it will create a superposition of the ground state, and a classically separable wavepacket. The second chirp, starting from the same frequency as the first, but with a larger amplitude can shift the ground state population into a different wavepacket. To avoid dispersion of the first wavepacket during the creation of the second, one should lock it with a constant drive component at the final frequency of the first chirp (a holding sequence). The result should be a coherent superposition of two classically distinct wavepackets, or "cat-states". If the two populations are sufficiently separated in frequency, the two drive components will not cause any mixing between them because they will be off-resonance. We expect one drive to cause energy shifts on the levels controlled by the other drive, and therefore one will have to account for that in order to accurately control the final states.

Bibliography

[1] A. Einstein, B. Podolsky, and N. Rosen. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review Online Archive (Prola)*, 47(10):777–780, May 1935. pages

[2] David Bohm. A suggested interpretation of the quantum theory in terms of "hidden" variables. i. *Phys. Rev.*, 85:166–179, Jan 1952. pages

[3] John S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1964. pages

[4] Alain Aspect, Jean Dalibard, and Gérard Roger. Experimental test of bell's inequalities using time- varying analyzers. *Phys. Rev. Lett.*, 49:1804–1807, Dec 1982. pages

[5] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, Mar 1993. pages

[6] Richard Feynman and Peter W. Shor. Simulating physics with computers. *SIAM Journal on Computing*, 26:1484–1509, 1982. pages

[7] Peter W. Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In *35th Annual Symposium on Foundations of Computer Science, 20-22 November 1994, Santa Fe, New Mexico, USA*, pages 124–134. IEEE, 1994. pages

[8] R.L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21:120–126, 1978. pages

[9] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information (Cambridge Series on Information and the Natural Sciences)*. Cambridge University Press, 2004. pages

[10] C. H. Bennett and G. Brassard. Quantum Cryptography: Public Key Distribution and Coin Tossing. In *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing*, pages 175–179, New York, 1984. IEEE Press. pages

[11] Artur K. Ekert. Quantum cryptography based on bell's theorem. *Phys. Rev. Lett.*, 67:661–663, Aug 1991. pages

[12] David P. Divincenzo. The Physical Implementation of Quantum Computation. *Fortschr. Phys.*, 48(9-11):771–783, 2000. pages

[13] L. Allen and J. H. Eberly. *Optical Resonance and Two-Level Atoms*. Dover Publications, December 1987. pages

[14] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. Microscopic theory of superconductivity. *Phys. Rev.*, 106:162–164, Apr 1957. pages

[15] R. C. Jaklevic, J. Lambe, A. H. Silver, and J. E. Mercereau. Quantum Interference Effects in Josephson Tunneling. *Physical Review Letters*, 12:159–160, February 1964. pages

[16] B. D. Josephson. Possible new effects in superconductive tunnelling. *Physics Letters*, 1(7):251–253, July 1962. pages

[17] Vinay Ambegaokar and Alexis Baratoff. Tunneling between superconductors. *Phys. Rev. Lett.*, 10:486–489, Jun 1963. pages

[18] B. S. Deaver and W. M. Fairbank. Experimental Evidence for Quantized Flux in Superconducting Cyclinders. *Physical Review Letters*, 7:43–46, July 1961. pages

[19] A. J. Leggett. Macroscopic quantum systems and the quantum theory of measurement. *Progress of Theoretical Physics Supplement*, 69:80–100, 1980. pages

[20] John M. Martinis, Michel H. Devoret, and John Clarke. Energy-level quantization in the zero-voltage state of a current-biased josephson junction. *Phys. Rev. Lett.*, 55(15):1543–1546, Oct 1985. pages

[21] Michel H. Devoret, John M. Martinis, Daniel Esteve, and John Clarke. Resonant activation from the zero-voltage state of a current-biased josephson junction. *Phys. Rev. Lett.*, 53:1260–1263, Sep 1984. pages

[22] JOHN CLARKE, ANDREW N. CLELAND, MICHEL H. DEVORET, DANIEL ESTEVE, and JOHN M. MARTINIS. Quantum mechanics of a macroscopic variable: The phase difference of a josephson junction. *Science*, 239(4843):992–997, 1988. pages

[23] A. D. O'Connell, M. Hofheinz, M. Ansmann, Radoslaw C. Bialczak, M. Lenander, Erik Lucero, M. Neeley, D. Sank, H. Wang, M. Weides, J. Wenner, John M. Martinis, and A. N. Cleland. Quantum ground state and single-phonon control of a mechanical resonator. *Nature*, 464(7289):697–703, April 2010. pages

[24] Markus Ansmann, H. Wang, Radoslaw C. Bialczak, Max Hofheinz, Erik Lucero, M. Neeley, A. D. O'Connell, D. Sank, M. Weides, J. Wenner, A. N. Cleland, and John M. Martinis. Violation of Bell's inequality in Josephson phase qubits. *Nature*, 461(7263):504–506, September 2009. pages

[25] 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(7315):574–578, September 2010. pages

[26] Matthias Steffen, M. Ansmann, Radoslaw C. Bialczak, N. Katz, Erik Lucero, R. McDermott, Matthew Neeley, E. M. Weig, A. N. Cleland, and John M. Martinis. Measurement of the Entanglement of Two Superconducting Qubits via State Tomography. *Science*, 313(5792):1423–1425, September 2006. pages

[27] Michel H. Devoret, John M. Martinis, and John Clarke. Measurements of macroscopic quantum tunneling out of the zero-voltage state of a current-biased josephson junction. *Phys. Rev. Lett.*, 55:1908–1911, Oct 1985. pages

[28] John M. Martinis and Kevin Osborne. Superconducting qubits and the physics of josephson junctions, 2004. pages

[29] Jens Koch, Terri M. Yu, Jay Gambetta, A. A. Houck, D. I. Schuster, J. Majer, Alexandre 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, Oct 2007. pages

[30] Vladimir E. Manucharyan, Jens Koch, Leonid I. Glazman, and Michel H. Devoret. Fluxonium: Single cooper-pair circuit free of charge offsets. *Science*, 326(5949):113–116, 2009. pages

[31] V Bouchiat, D Vion, P Joyez, D Esteve, and M H Devoret. Quantum coherence with a single cooper pair. *Physica Scripta*, 1998(T76):165, 1998. pages

[32] J. E. Mooij, T. P. Orlando, L. Levitov, Lin Tian, Caspar H. van der Wal, and Seth Lloyd. Josephson persistent-current qubit. *Science*, 285(5430):1036–1039, 1999. pages

[33] John M. Martinis, S. Nam, J. Aumentado, and C. Urbina. Rabi oscillations in a large josephson-junction qubit. *Phys. Rev. Lett.*, 89(11):117901, Aug 2002. pages

[34] John M. Martinis, K. B. Cooper, R. McDermott, Matthias Steffen, Markus Ansmann, K. D. Osborn, K. Cicak, Seongshik Oh, D. P. Pappas, R. W. Simmonds, and Clare C. Yu. Decoherence in josephson qubits from dielectric loss. *Phys. Rev. Lett.*, 95(21):210503, Nov 2005. pages

[35] H. Wang, M. Hofheinz, J. Wenner, M. Ansmann, R. C. Bialczak, M. Lenander, Erik Lucero, M. Neeley, A. D. O'Connell, D. Sank, M. Weides, A. N. Cleland, and John M. Martinis. Improving the coherence time of superconducting coplanar resonators. *Applied Physics Letters*, 95(23):233508, 2009. pages

[36] Jeffrey S Kline, Michael R Vissers, Fabio C S da Silva, David S Wisbey, Martin Weides, Terence J Weir, Benjamin Turek, Danielle A Braje, William D Oliver, Yoni Shalibo, Nadav Katz, Blake R Johnson, Thomas A Ohki, and David P Pappas. Sub-micrometer epitaxial josephson junctions for quantum circuits. *Superconductor Science and Technology*, 25(2):025005, 2012. pages

[37] Martin P. Weides, Jeffrey S. Kline, Michael R. Vissers, Martin O. Sandberg, David S. Wisbey, Blake R. Johnson, Thomas A. Ohki, and David P. Pappas. Coherence in a transmon qubit with epitaxial tunnel junctions. *Applied Physics Letters*, 99(26):262502, 2011. pages

[38] G.C. Tettamanzi, C.I. Pakes, A. Potenza, C.H. Marrows, S. Prawer, and D.N. Jamieson. Superconducting microbridge junctions fabricated using focused ion beams. In *Nanoscience and Nanotechnology, 2006. ICONN '06. International Conference on*, july 2006. pages

[39] M. D. Shaw, R. M. Lutchyn, P. Delsing, and P. M. Echternach. Kinetics of nonequilibrium quasiparticle tunneling in superconducting charge qubits. *Phys. Rev. B*, 78:024503, Jul 2008. pages

[40] John M. Martinis, M. Ansmann, and J. Aumentado. Energy decay in josephson qubits from non-equilibrium quasiparticles. 2009. pages

[41] Matthew Neeley, M. Ansmann, Radoslaw C. Bialczak, M. Hofheinz, N. Katz, Erik Lucero, A. O'Connell, H. Wang, A. N. Cleland, and John M. Martinis. Transformed dissipation in superconducting quantum circuits. *Physical Review B (Condensed Matter and Materials Physics)*, 77(18):180508, 2008. pages

[42] Radoslaw C. Bialczak, R. McDermott, M. Ansmann, M. Hofheinz, N. Katz, Erik Lucero, Matthew Neeley, A. D. O'Connell, H. Wang, A. N. Cleland, and John M. Martinis. $1/f$ flux noise in josephson phase qubits. *Phys. Rev. Lett.*, 99:187006, Nov 2007. pages

[43] S. Sendelbach, D. Hover, A. Kittel, M. Muck, John M. Martinis, and R. McDermott. Magnetism in squids at millikelvin temperatures. *Physical Review Letters*, 100(22):227006, 2008. pages

[44] Aaron D. O'Connell, M. Ansmann, R. C. Bialczak, M. Hofheinz, N. Katz, Erik 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. *Applied Physics Letters*, 92(11):112903, 2008. pages

[45] M. Tinkham. Effect of fluxoid quantization on transitions of superconducting films. *Phys. Rev.*, 129:2413–2422, Mar 1963. pages

[46] J. M. Schmidt, A. N. Cleland, and John Clarke. Resonant tunneling in small current-biased josephson junctions. *Phys. Rev. B*, 43:229–238, Jan 1991. pages

[47] N. Katz, M. Ansmann, Radoslaw C. Bialczak, Erik Lucero, R. McDermott, Matthew Neeley, Matthias Steffen, E. M. Weig, A. N. Cleland, John M. Martinis, and A. N. Korotkov. Coherent state evolution in a superconducting qubit from partial-collapse measurement. *Science*, 312(5779):1498–1500, 2006. pages

[48] Erik Lucero, M. Hofheinz, M. Ansmann, Radoslaw C. Bialczak, N. Katz, Matthew Neeley, A. D. O'Connell, H. Wang, A. N. Cleland, and John M.

Martinis. High-fidelity gates in a single josephson qubit. *Phys. Rev. Lett.*, 100(24):247001, Jun 2008. pages

[49] Matthias Steffen, John M. Martinis, and Isaac L. Chuang. Accurate control of josephson phase qubits. *Phys. Rev. B*, 68:224518, Dec 2003. pages

[50] John M. Martinis, S. Nam, J. Aumentado, K. M. Lang, and C. Urbina. Decoherence of a superconducting qubit due to bias noise. *Phys. Rev. B*, 67:094510, Mar 2003. pages

[51] Y. Nakamura, Yu. A. Pashkin, T. Yamamoto, and J. S. Tsai. Charge echo in a cooper-pair box. *Phys. Rev. Lett.*, 88:047901, Jan 2002. pages

[52] W A Phillips. Two-level states in glasses. *Reports on Progress in Physics*, Volume 50, Number 12:1657, 1987. pages

[53] Alexander Shnirman, Gerd Schön, Ivar Martin, and Yuriy Makhlin. Low- and high-frequency noise from coherent two-level systems. *Phys. Rev. Lett.*, 94(12):127002, Apr 2005. pages

[54] Li-Chung Ku and Clare C. Yu. Decoherence of a josephson qubit due to coupling to two-level systems. *Phys. Rev. B*, 72(2):024526, Jul 2005. pages

[55] M J Kirton, M J Uren, S Collins, M Schulz, A Karmann, and K Scheffer. Individual defects at the si:sio 2 interface. *Semiconductor Science and Technology*, 4(12):1116, 1989. pages

[56] N.S. Saks. A technique for suppressing dark current generated by interface states in buried channel ccd imagers. *Electron Device Letters, IEEE*, 1(7):131 – 133, jul 1980. pages

[57] P. Bushev, C. Müller, J. Lisenfeld, J. H. Cole, A. Lukashenko, A. Shnirman, and A. V. Ustinov. Hybrid quantum system surveyed using multi-photon spectroscopy. *ArXiv e-prints*, May 2010. pages

[58] A. Lupaccu, P. Bertet, E. F. C. Driessen, C. J. P. M. Harmans, and J. E. Mooij. One- and two-photon spectroscopy of a flux qubit coupled to a microscopic defect. *Phys. Rev. B*, 80(17):172506, Nov 2009. pages

[59] Y. Nakamura, H. Terai, K. Inomata, T. Yamamoto, W. Qiu, and Z. Wang. Superconducting qubits consisting of epitaxially grown nbn/aln/nbn josephson junctions. *Applied Physics Letters*, 99(21):212502, 2011. pages

[60] Matteo Mariantoni, H. Wang, T. Yamamoto, M. Neeley, Radoslaw C. Bialczak, Y. Chen, M. Lenander, Erik Lucero, A. D. OConnell, D. Sank, M. Weides, J. Wenner, Y. Yin, J. Zhao, A. N. Korotkov, A. N. Cleland, and John M. Martinis. Implementing the quantum von neumann architecture with superconducting circuits. *Science*, 334(6052):61–65, 2011. pages

[61] Matthew Neeley, M. Ansmann, Radoslaw C. Bialczak, M. Hofheinz, N. Katz, Erik Lucero, A. O'Connell, H. Wang, A. N. Cleland, and John M. Martinis. Process tomography of quantum memory in a josephson-phase qubit coupled to a two-level state. *Nat Phys*, 4(7):523–526, Jul 2008. pages

[62] I. Martin, L. Bulaevskii, and A. Shnirman. Tunneling spectroscopy of two-level systems inside a josephson junction. *Phys. Rev. Lett.*, 95:127002, Sep 2005. pages

[63] R. W. Simmonds, K. M. Lang, D. A. Hite, S. Nam, D. P. Pappas, and John M. Martinis. Decoherence in josephson phase qubits from junction resonators. *Phys. Rev. Lett.*, 93(7):077003, Aug 2004. pages

[64] SangKook Choi, Dung-Hai Lee, Steven G. Louie, and John Clarke. Localization of metal-induced gap states at the metal-insulator interface: Origin of flux noise in squids and superconducting qubits. *Phys. Rev. Lett.*, 103:197001, Nov 2009. pages

[65] T. C. DuBois, M. C. Per, S. P. Russo, and J. H. Cole. Delocalised oxygen as the origin of two-level defects in Josephson junctions. *ArXiv e-prints*, June 2012. pages

[66] Lara Faoro, Joakim Bergli, Boris L. Altshuler, and Y. M. Galperin. Models of environment and T_1 relaxation in josephson charge qubits. *Phys. Rev. Lett.*, 95:046805, Jul 2005. pages

[67] Lara Faoro and Lev B. Ioffe. Microscopic origin of critical current fluctuations in large, small, and ultra-small area josephson junctions. *Phys. Rev. B*, 75:132505, Apr 2007. pages

[68] M Schechter and P C E Stamp. What are the interactions in quantum glasses? *Journal of Physics: Condensed Matter*, 20(24):244136, 2008. pages

[69] T. A. Palomaki, S. K. Dutta, R. M. Lewis, A. J. Przybysz, Hanhee Paik, B. K. Cooper, H. Kwon, J. R. Anderson, C. J. Lobb, F. C. Wellstood, and E. Tiesinga. Multilevel spectroscopy of two-level systems coupled to a dc squid phase qubit. *Phys. Rev. B*, 81(14):144503, Apr 2010. pages

[70] James P. Sethna. Decay rates of tunneling centers coupled to phonons: An instanton approach. *Phys. Rev. B*, 25(8):5050–5063, Apr 1982. pages

[71] David R. Lide. *CRC Handbook of Chemistry and Physics, 88th Edition*. CRC, June 2007. pages

[72] J. Lisenfeld, C. Müller, J. H. Cole, P. Bushev, A. Lukashenko, A. Shnirman, and A. V. Ustinov. Measuring the temperature dependence of individual two-level systems by direct coherent control. *Phys. Rev. Lett.*, 105:230504, Dec 2010. pages

[73] Jürgen Lisenfeld, Clemens Müller, Jared H. Cole, Pavel Bushev, Alexander Lukashenko, Alexander Shnirman, and Alexey V. Ustinov. Rabi spectroscopy of a qubit-fluctuator system. *Phys. Rev. B*, 81:100511, Mar 2010. pages

[74] Grigorij J. Grabovskij, Torben Peichl, Jurgen Lisenfeld, Georg Weiss, and Alexey V. Ustinov. Strain tuning of individual atomic tunneling systems detected by a superconducting qubit. *Science*, 338(6104):232–234, 2012. pages

[75] Erik Lucero, Julian Kelly, Radoslaw C. Bialczak, Mike Lenander, Matteo Mariantoni, Matthew Neeley, A. D. O'Connell, Daniel Sank, H. Wang, Martin Weides, James Wenner, Tsuyoshi Yamamoto, A. N. Cleland, and John M. Martinis. Reduced phase error through optimized control of a superconducting qubit. *Phys. Rev. A*, 82(4):042339, Oct 2010. pages

[76] I. Siddiqi, R. Vijay, F. Pierre, C. M. Wilson, M. Metcalfe, C. Rigetti, L. Frunzio, and M. H. Devoret. Rf-driven josephson bifurcation amplifier for quantum measurement. *Phys. Rev. Lett.*, 93:207002, Nov 2004. pages

[77] Carlton M. Caves. Quantum limits on noise in linear amplifiers. *Phys. Rev. D*, 26:1817–1839, Oct 1982. pages

[78] M. Metcalfe, E. Boaknin, V. Manucharyan, R. Vijay, I. Siddiqi, C. Rigetti, L. Frunzio, R. J. Schoelkopf, and M. H. Devoret. Measuring the decoherence of a quantronium qubit with the cavity bifurcation amplifier. *Phys. Rev. B*, 76:174516, Nov 2007. pages

[79] O. Naaman, J. Aumentado, L. Friedland, J. S. Wurtele, and I. Siddiqi. Phase-locking transition in a chirped superconducting josephson resonator. *Phys. Rev. Lett.*, 101(11):117005, Sep 2008. pages

[80] K. W. Murch, R. Vijay, I. Barth, O. Naaman, J. Aumentado, L. Friedland, and I. Siddiqui. Quantum fluctuations in the chirped pendulum. *Nat Phys*, 7:105–108, 2011. pages

[81] J. Fajans and L. Friedland. Autoresonant (nonstationary) excitation of pendulums, plutinos, plasmas, and other nonlinear oscillators. *American Journal of Physics*, 69(10):1096–1102, 2001. pages

[82] R. Malhotra. The origin of Pluto's peculiar orbit. *Nature*, 365(7324):819–821, 1993. pages

[83] G. B. Andresen, M. D. Ashkezari, M. Baquero-Ruiz, W. Bertsche, P. D. Bowe, E. Butler, C. L. Cesar, S. Chapman, M. Charlton, A. Deller, S. Eriksson, J. Fajans, T. Friesen, M. C. Fujiwara, D. R. Gill, A. Gutierrez, J. S. Hangst, W. N. Hardy, M. E. Hayden, A. J. Humphries, R. Hydomako, M. J. Jenkins, S. Jonsell, L. V. Jorgensen, L. Kurchaninov, N. Madsen, S. Menary, P. Nolan, K. Olchanski, A. Olin, A. Povilus, P. Pusa, F. Robicheaux, E. Sarid, Nasr, D. M. Silveira, C. So, J. W. Storey, R. I. Thompson, D. P. van der Werf, J. S. Wurtele, and Y. Yamazaki. Trapped antihydrogen. *Nature*, 468(7324):673–676, December 2010. pages

[84] G. Marcus, L. Friedland, and A. Zigler. From quantum ladder climbing to classical autoresonance. *Phys. Rev. A*, 69(1):013407, Jan 2004. pages

[85] E. M. Lifshitz and L. D. Landau. *Course of Theoretical Physics, Volume I: Mechanics*. Butterworth-Heinemann, 3 edition, January 1976. pages

[86] I. Barth, L. Friedland, O. Gat, and A. G. Shagalov. Quantum versus classical phase-locking transition in a frequency-chirped nonlinear oscillator. *Phys. Rev. A*, 84:013837, Jul 2011. pages

[87] M. Born and V. Fock. Zeitschrift für physik a hadrons and nuclei. 51(3-4):165–180, 1928. pages

[88] H. Wang, M. Hofheinz, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, J. Wenner, A. N. Cleland, and John M. Martinis. Measurement of the decay of fock states in a superconducting quantum circuit. *Phys. Rev. Lett.*, 101(24):240401, Dec 2008. pages

[89] H. Wang, M. Hofheinz, M. Ansmann, R. C. Bialczak, Erik Lucero, M. Neeley, A. D. O'Connell, D. Sank, M. Weides, J. Wenner, A. N. Cleland, and John M.

Martinis. Decoherence dynamics of complex photon states in a superconducting circuit. *Phys. Rev. Lett.*, 103(20):200404, Nov 2009. pages

[90] M. I. Dykman and V. N. Smelyanskiy. Quantum theory of transitions between stable states of a nonlinear oscillator interacting with a medium in a resonant field. *JETP*, 67:1769–1776, 1988. pages

[91] M. I. Dykman, I. B. Schwartz, and M. Shapiro. Scaling in activated escape of underdamped systems. *Phys. Rev. E*, 72(2):021102, Aug 2005. pages

[92] M. I. Dykman and M. A. Krivoglaz. Theory of fluctuational transitions between stable states of a nonlinear oscillator. *Soviet Journal of Experimental and Theoretical Physics*, 50:30–+, July 1979. pages

[93] L Friedland. Efficient capture of nonlinear oscillations into resonance. *Journal of Physics A: Mathematical and Theoretical*, 41(41):415101, 2008. pages

[94] Joachim Ankerhold. *Quantum Tunneling in Complex Systems: The Semiclassical Approach (Springer Tracts in Modern Physics)*. Springer, 2007. pages

[95] J. Fajans, E. Gilson, and L. Friedland. The effect of damping on autoresonant (nonstationary) excitation. *Physics of Plasmas*, 8(2):423–427, 2001. pages

[96] Harriet B. Creighton and Barbara McClintock. A correlation of cytological and genetical crossing-over in zea mays. *Proceedings of the National Academy of Sciences*, 17(8):492–497, 1931. pages

[97] Matthew Neeley, Markus Ansmann, Radoslaw C. Bialczak, Max Hofheinz, Erik Lucero, Aaron D. O'Connell, Daniel Sank, Haohua Wang, James Wenner, Andrew N. Cleland, Michael R. Geller, and John M. Martinis. Emulation of a Quantum Spin with a Superconducting Phase Qudit. *Science*, 325(5941):722–725, August 2009. pages

[98] R. T. Thew, K. Nemoto, A. G. White, and W. J. Munro. Qudit quantum-state tomography. *Phys. Rev. A*, 66:012303, Jul 2002. pages

[99] D. F. Walls and G. J. Milburn. Effect of dissipation on quantum coherence. *Phys. Rev. A*, 31:2403–2408, Apr 1985. pages

[100] Wolfgang P. Schleich. *Quantum Optics in Phase Space*. Wiley-VCH, 1 edition, February 2001. pages

[101] W. H. Zurek. Decoherence and the transition from quantum to classical – REVISITED. *ArXiv Quantum Physics e-prints*, June 2003. pages

[102] Itamar Katz, Ron Lifshitz, Alex Retzker, and Raphael Straub. Classical to quantum transition of a driven nonlinear nanomechanical resonator. *New J. Phys.*, 10(12):125023+, December 2008. pages

[103] P. Bertet, A. Auffeves, P. Maioli, S. Osnaghi, T. Meunier, M. Brune, J. M. Raimond, and S. Haroche. Direct measurement of the wigner function of a one-photon fock state in a cavity. *Phys. Rev. Lett.*, 89:200402, Oct 2002. pages

[104] Max Hofheinz, H. Wang, M. Ansmann, Radoslaw C. Bialczak, Erik Lucero, M. Neeley, A. D. O'Connell, D. Sank, J. Wenner, John M. Martinis, and A. N. Cleland. Synthesizing arbitrary quantum states in a superconducting resonator. *Nature*, 459(7246):546–549, May 2009. pages

[105] C. Eichler, D. Bozyigit, C. Lang, L. Steffen, J. Fink, and A. Wallraff. Experimental state tomography of itinerant single microwave photons. *Phys. Rev. Lett.*, 106:220503, Jun 2011. pages

[106] T. J. Dunn, I. A. Walmsley, and S. Mukamel. Experimental determination of the quantum-mechanical state of a molecular vibrational mode using fluorescence tomography. *Phys. Rev. Lett.*, 74:884–887, Feb 1995. pages

[107] T. C. Weinacht, J. Ahn, and P. H. Bucksbaum. Measurement of the amplitude and phase of a sculpted rydberg wave packet. *Phys. Rev. Lett.*, 80:5508–5511, Jun 1998. pages

[108] Esben Skovsen, Henrik Stapelfeldt, Søren Juhl, and Klaus Mølmer. Quantum state tomography of dissociating molecules. *Phys. Rev. Lett.*, 91:090406, Aug 2003. pages

[109] Hirokazu Hasegawa and Yasuhiro Ohshima. Quantum state reconstruction of a rotational wave packet created by a nonresonant intense femtosecond laser field. *Phys. Rev. Lett.*, 101:053002, Aug 2008. pages

[110] Antoine Royer. Wigner function as the expectation value of a parity operator. *Phys. Rev. A*, 15:449–450, Feb 1977. pages

[111] Michael Tinkham. *Introduction to Superconductivity: Second Edition (Dover Books on Physics)*. Dover Publications, June 2004. pages

[112] Yoni Shalibo, Ya'ara Rofe, Ido Barth, Lazar Friedland, Radoslaw Bialczack, John M. Martinis, and Nadav Katz. Quantum and classical chirps in an anharmonic oscillator. *Phys. Rev. Lett.*, 108:037701, Jan 2012. pages

[113] N. Katz, M. Ansmann, Radoslaw C. Bialczak, Erik Lucero, R. McDermott, Matthew Neeley, Matthias Steffen, E. M. Weig, A. N. Cleland, John M. Martinis, and A. N. Korotkov. Coherent state evolution in a superconducting qubit from partial-collapse measurement. *Science*, 312(5779):1498–1500, 2006. pages

[114] Easwar Magesan, Jay M. Gambetta, B. R. Johnson, Colm A. Ryan, Jerry M. Chow, Seth T. Merkel, Marcus P. da Silva, George A. Keefe, Mary B. Rothwell, Thomas A. Ohki, Mark B. Ketchen, and M. Steffen. Efficient measurement of quantum gate error by interleaved randomized benchmarking. *Phys. Rev. Lett.*, 109:080505, Aug 2012. pages

[115] G. K. Brennen, S. S. Bullock, and D. P. O'Leary. Efficient circuits for exact-universal computationwith qudits. *Quantum Info. Comput.*, 6(4):436–454, July 2006. pages

CHAPTER 5

Appendices

A Simulation methods

Throughout this work, we use numerical simulations to predict the coherent response of our system to control pulses. We do it by solving Eq. 1.1 for the density matrix. In the absence of decoherence, the dynamics are completely determined by the system's Hamiltonian:

$$\hat{H} = -\frac{2e^2}{C}\frac{d^2}{d\hat{\delta}^2} - \frac{\left(I_c \cos\hat{\delta} + I_d(t)\hat{\delta}\right)\Phi_0}{2\pi} + \frac{1}{2L}\left(\Phi_{\text{ext}} - \frac{\hat{\delta}\Phi_0}{2\pi}\right)^2, \tag{5.1}$$

where $I_d(t) = I(t)\cos\omega t + Q(t)\sin\omega t$ is the microwave drive. By diagonalizing the bare system Hamiltonian we find the system's eigenstates. When the phase is localized in only one potential well (as is the case in our system after reset) we can safely ignore all the untrapped states. In this case it is convenient to work in the eigenstate basis to describe the state dynamics. The Hamiltonian matrix of the system in this basis, when the coupling terms are separated, is[1]:

$$H_{\text{N}} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & E_1 & 0 & & 0 \\ 0 & 0 & E_2 & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & E_{\text{N}-1} \end{pmatrix} + \frac{I_d(t)\Phi_0}{2\pi} \begin{pmatrix} \delta_{00} & \delta_{01} & \delta_{02} & \cdots & \delta_{0,\text{N}-1} \\ \delta_{10} & \delta_{11} & \delta_{12} & & \delta_{1,\text{N}-1} \\ \delta_{20} & \delta_{21} & \delta_{22} & & \delta_{2,\text{N}-1} \\ \vdots & & & \ddots & \vdots \\ \delta_{\text{N}-1,0} & \delta_{\text{N}-1,0} & \delta_{\text{N}-1,0} & \cdots & \delta_{\text{N}-1,N-1} \end{pmatrix}, \tag{5.2}$$

[1] We assume a zero ground state energy for convenience.

where $|n\rangle$ denotes the state-vector of the n^{th} energy level inside the well, $\delta_{m,n} = \langle m \left|\hat{\delta}\right| n\rangle$ and $\mathrm{N}-1$ is the highest excited level. From the solved eigenstates we find that the off-diagonal matrix elements are close to that of an harmonic oscillator. For example, at large anharmonicity ($\beta/\omega_{01} \approx 0.02$), where qubit experiments are performed, we find for the lowest 4 levels [2]:

$$\hat{\delta}/\left\langle 0\left|\hat{\delta}\right|1\right\rangle = \begin{pmatrix} 0.035 & 1 & 0.052 & 0.002 \\ 1 & 0.25 & 1.425 & 0.096 \\ 0.052 & 1.425 & 0.5 & 1.760 \\ 0.002 & 0.096 & 1.760 & 0.765 \end{pmatrix}.$$

The first off-diagonal ($m = n \pm 1$) is within $\sim 1\,\%$ close to the harmonic oscillator values ($\sqrt{n}, \sqrt{n+1}$), but increases for higher n due to the growing asymmetry of the potential at higher energies. The second off-diagonal ($m = n \pm 2$) is more than an order of magnitude smaller, and the third off-diagonal ($m = n \pm 3$) is almost 3 orders of magnitude smaller, and increases for both at larger n (not shown for the third off-diagonal).

At small anharmonicity ($\beta/\omega_{01} \approx 0.002$), where Wigner tomography and autoresonance experiments are performed, we find:

$$\hat{\delta}/\left\langle 0\left|\hat{\delta}\right|1\right\rangle = \begin{pmatrix} 0.014 & 1 & 0.014 & 0.002 \\ 1 & 0.076 & 1.416 & 0.025 \\ 0.014 & 1.425 & 0.138 & 1.736 \\ 0.002 & 0.025 & 1.736 & 0.202 \end{pmatrix}.$$

Here we find similar trends, but much closer to the harmonic oscillator values. It seems that the diagonal elements in both cases have non-negligible values, however these should be compared to the already large diagonal elements - the bare energies of the system. Since we are ultimately interested in the time evolution in the rotating frame (discussed below), it is more appropriate to compare these to the rotating frame energies. We will come back to this comparison after discussing the rotating wave approximation.

[2] Since most of the probability is displaced from $\delta = 0$ by the position $\delta_{\min}$ of our potential minimum, we use $\hat{\delta} - \delta_{\min}$ instead of $\hat{\delta}$. This has no measurable effect since it only adds an energy shift to the whole Hamiltonian.

The rotating wave approximation (RWA)

The solution to Liouville's equation for the Hamiltonian H_{N} contains fast-oscillating terms, at frequencies corresponding to the difference between energy levels. However, in most experiments relevant to our system, the short time-scale dynamics is trivial and the interesting dynamics occurs at a much longer time-scale. When the fast oscillating terms are averaged out (RWA), the dynamics are more simple to trace and the computational effort in solving the equations is reduced significantly. Here we provide a formal derivation of the RWA for our system, which is a generalization of the standard two-level system RWA. For simplicity we work with a state-vector, however, the result is applicable for a density matrix as well. The time evolution of the state vector $|\psi\rangle$ can be written as: $|\psi(t)\rangle = \sum_n a_n(t)e^{-i\omega_n t}$, where $\omega_n = E_n/\hbar$ and $a_n(t)$ vary much slower than the exponent. It is therefore appropriate to define a drive-oriented rotating-wave transformation[3]:

$$V^{\dagger} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & e^{i\omega t} & 0 & 0 & 0 \\ 0 & 0 & e^{i2\omega t} & 0 & 0 \\ \vdots & 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^{i\mathrm{N}\omega t} \end{pmatrix}. \tag{5.3}$$

In the rotating frame, the time evolution is given by $|\tilde{\psi}(t)\rangle = \sum_n a_n(t)e^{-i\beta_n}$, where $\beta_n = \omega_n - n\omega t$ are the cumulative anharmonicities. The above transformation makes sense for our system because $\beta_n \ll \omega_n$. To find the rotating frame Hamiltonian, $\tilde{H}$, we write the Schrödinger equation for $|\tilde{\psi}\rangle$:

$$i\hbar\frac{\partial}{\partial t}|\tilde{\psi}\rangle = \tilde{H}\,|\tilde{\psi}\rangle\,. \tag{5.4}$$

Using the transformation V we obtain,

[3]It is possible to use a system oriented transformation, with the phases constructed with the bare energies (also called the Heisenberg picture), but this leads to a time dependent Hamiltonian after the RWA (time dependence comes about from system-drive detuning), while in our case, only the drive envelope $A(t)$ depends on time in the final Hamiltonian.

$$i\hbar\frac{\partial}{\partial t}\left(V^{\dagger}\left|\psi\right\rangle\right)=\tilde{H}V^{\dagger}\left|\psi\right\rangle .$$

Expanding the left hand side, and moving the second term to the write, we get:

$$i\hbar V\frac{\partial}{\partial t}\left|\psi\right\rangle=-i\hbar\frac{\partial V}{\partial t}\left|\psi\right\rangle+\tilde{H}V\left|\psi\right\rangle . \tag{5.5}$$

By multiplying to the left by V, we get a relation between the original and the rotating frame Hamiltonians:

$$H=V\tilde{H}V^{\dagger}-i\hbar V\frac{\partial V^{\dagger}}{\partial t},$$

where we used the unitarity property $V^{\dagger}V=1$. The relation can be easily inverted to yield:

$$\tilde{H}=V^{\dagger}HV-i\hbar V^{\dagger}\frac{\partial V}{\partial t}. \tag{5.6}$$

Using this transformation on the lowest four levels (for simplicity) of the Hamiltonian H_{N}, we get:

$$\tilde{H}_4=V^{\dagger}\begin{pmatrix}\hbar\tilde{\Omega}\tilde{\delta}_{00} & \hbar\tilde{\Omega} & 0 & 0\\ \hbar\tilde{\Omega} & E_1+\hbar\tilde{\Omega}\tilde{\delta}_{11} & \sqrt{2}\hbar\tilde{\Omega} & 0\\ 0 & \sqrt{2}\hbar\tilde{\Omega} & E_2+\hbar\tilde{\Omega}\tilde{\delta}_{22} & \sqrt{3}\hbar\tilde{\Omega}\\ 0 & 0 & \sqrt{3}\hbar\tilde{\Omega} & E_3+\hbar\tilde{\Omega}\tilde{\delta}_{33}\end{pmatrix}V-\hbar\begin{pmatrix}0 & 0 & 0 & 0\\ 0 & \omega & 0 & 0\\ 0 & 0 & 2\omega & 0\\ 0 & 0 & 0 & 3\omega\end{pmatrix},$$

where we have defined $\tilde{\delta}_{m,n}=\delta_{m,n}/\delta_{01}$ and $\tilde{\Omega}=\frac{I_d(t)\Phi_0}{2\pi\hbar}\delta_{01}$. For off-diagonal elements we have used the harmonic approximation, in accordance with the above examples. Expanding the left term and combining with the second term, we get:

$$\tilde{H}_4=\hbar\begin{pmatrix}\tilde{\Omega}\tilde{\delta}_{00} & \tilde{\Omega}e^{-i\omega t} & 0 & 0\\ \tilde{\Omega}e^{i\omega t} & -\Delta+\tilde{\Omega}\tilde{\delta}_{11} & \sqrt{2}\tilde{\Omega}e^{-i\omega t} & 0\\ 0 & \sqrt{2}\tilde{\Omega}e^{i\omega t} & \beta_2-2\Delta+\tilde{\Omega}\tilde{\delta}_{22} & \sqrt{3}\tilde{\Omega}e^{-i\omega t}\\ 0 & 0 & \sqrt{3}\tilde{\Omega}e^{i\omega t} & \beta_3-3\Delta+\tilde{\Omega}\tilde{\delta}_{33}\end{pmatrix},$$

where we have defined the frequency detuning $\Delta = \omega - \omega_{01}$. By explicitly writing the drive's time dependence and defining the Rabi amplitude $\Omega(t) = \frac{A(t)\Phi\delta_{01}}{2\pi\hbar}e^{i\phi(t)}$, we get $\tilde{\Omega}e^{-i\omega t} = \Omega(t)\frac{1+e^{-2i(\omega t+\phi)}}{2}$, and $\tilde{\Omega}e^{i\omega t} = \Omega^*(t)\frac{e^{2i(\omega t+\phi)}+1}{2}$ for the off-diagonal terms, where $A(t) = \sqrt{I(t)^2 + Q(t)^2}$ and $\phi = \arctan\left(Q(t)/I(t)\right)$. The fast oscillating term is non-resonant and has negligible effect on the dynamics when $|\Omega|, \Delta \ll \omega$[4]. The neglection of the rapidly oscillating term is called the rotating wave approximation (RWA)[5]. With the rotating wave approximation applied, we get:

$$\tilde{H}_4^{\text{RWA}} = \hbar \begin{pmatrix} \tilde{\Omega}\tilde{\delta}_{00} & \Omega/2 & 0 & 0 \\ \Omega^*/2 & -\Delta + \tilde{\Omega}\tilde{\delta}_{11} & \sqrt{2}\Omega/2 & 0 \\ 0 & \sqrt{2}\Omega^*/2 & \beta_2 - 2\Delta + \tilde{\Omega}\tilde{\delta}_{22} & \sqrt{3}\Omega/2 \\ 0 & 0 & \sqrt{3}\Omega^*/2 & \beta_3 - 3\Delta + \tilde{\Omega}\tilde{\delta}_{33} \end{pmatrix}. \quad (5.7)$$

The first diagonal $\tilde{\Omega}\tilde{\delta}_{00}$ is a common energy offset and therefore can be ignored. The remaining terms on the diagonal containing $\tilde{\Omega}$, oscillate at high frequency. This is equivalent to the detuning changing sign very fast compared to the timescale of changes in the state. Simulations show that these components, too, have a zero net effect at small amplitudes. Their contribution becomes non-negligible at higher amplitudes, where the RWA already breaks down.

By extrapolation, we conclude that the Hamiltonian of our system in the rotating frame and under the RWA for the lowest N levels can be written:

[4]For $|\Omega| \sim \omega$, state amplitudes vary with time at a frequency comparable to the rate of change of Ω and therefore one cannot assume that its effect averages out. Similarly, for $\Delta \sim \omega$, the state's phases vary with time at a frequency comparable to the rate of change of Ω, making the approximation invalid.

[5]The neglection of the matrix elements $m = n \pm 2$ (and above) is seen to be even more justified when working in the rotating frame, since these have fast oscillations at ω and 3ω (no DC component) that average out during the slow state evolution.

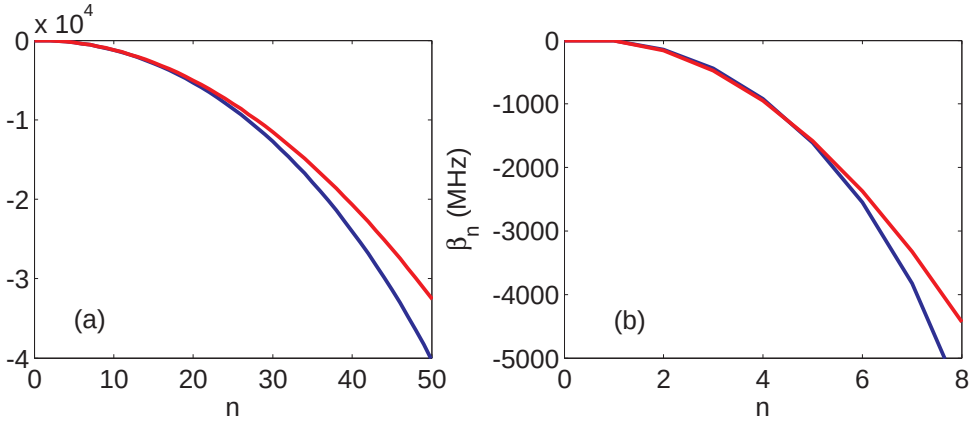

Figure 5.1: Cumulative anharmonicity vs. level number. (a) actual β_n (blue) at small anharmonicity ($\beta/\omega_{01} \approx 0.002$) and its quadratic approximation (blue). (b) actual β_n (blue) at small anharmonicity ($\beta/\omega_{01} \approx 0.02$) and its quadratic approximation (blue).

$$\tilde{H}_{\mathrm{N}}^{\mathrm{RWA}} = \hbar \begin{pmatrix} 0 & \Omega/2 & 0 & \cdots & 0 \\ \Omega^*/2 & -\Delta & \sqrt{2}\Omega/2 & 0 & \vdots \\ 0 & \sqrt{2}\Omega^*/2 & \beta_2 - 2\Delta & \sqrt{3}\Omega/2 & 0 \\ \vdots & 0 & \sqrt{3}\Omega^*/2 & \ddots & \sqrt{\mathrm{N}-1}\Omega/2 \\ 0 & \cdots & 0 & \sqrt{\mathrm{N}-1}\Omega^*/2 & \beta_{\mathrm{N}-1} - (\mathrm{N}-1)\,\Delta \end{pmatrix}, \tag{5.8}$$

where the circuit bias Φ_{ext} is taken into account through the anharmonicities β_n. In most of the simulations we make a further approximation, where we take the cumulative anharmonicity to be $\beta_n = \beta\frac{n(n-1)}{2}$. As shown in Fig. 5.1, the approximation is accurate to within a few percents at the bottom of the well (~ 20 of the lowest levels at small anharmonicity and ~ 5 of the lowest levels at large anharmonicity).

Decoherence To describe decohering states in the simulation, we work with a density matrix.

The solution to Eq. 1.1 in the rotating frame can be obtained using a propagator: $\rho(t) = \mathrm{U}^\dagger(t,t_0)\rho(t_0)\mathrm{U}(t,t_0)$, where $\mathrm{U}(t,t_0) \approx \mathcal{T}\left[\prod_i \exp(-i\tilde{H}_{\mathrm{N}}^{\mathrm{RWA}}(t_i)\Delta t/\hbar)\right]$

in the limit $\Delta t \ll t_d$[6] and $\mathcal{T}$ denotes time-ordering. Decoherence is incorporated to this simulation using quantum operations[9]. In this formalism, the state of the system after interacting with the environment is obtained by a partial trace over the environment:

$$\mathcal{E}(\rho) = \mathrm{Tr}_{\mathrm{env}}\left[\mathrm{U}\left(\rho \otimes \rho_{\mathrm{env}}\right)\mathrm{U}^{\dagger}\right],$$

where U is the time-evolution operator for the whole system (principal system and environment). This can be written using the *operator sum representation* as:

$$\mathcal{E}(\rho) = \sum_k E_k \rho E_k^{\dagger}, \tag{5.9}$$

where $E_k \equiv \langle e_k | \mathrm{U} | e_0 \rangle$ are the Kraus operators, which act on the principal system space, $|e_k\rangle$ is an orthonormal basis for the environment and $|e_0\rangle$ is the initial state of the environment.

To simulate energy relaxation, we model the system and environment as coupled harmonic oscillators. The interaction Hamiltonian in this case is: $H_{\mathrm{AD}} = \chi\left(a^{\dagger}b + b^{\dagger}a\right)$, where $\chi = 1/T_1$ is the interaction strength, $a, a^{\dagger}$ are the system's operators and $b, b^{\dagger}$ are the bath's operator. In this case, the Kraus operators are given by [9]:

$$E_k = \sum_n \sqrt{\binom{n}{k}} \sqrt{(1-\gamma)^{n-k}\gamma^k}\, |k\rangle\langle n|, \tag{5.10}$$

where $\gamma = 1 - e^{-\chi \Delta t}$. These operators, applied in the subspace of the N-level Hamiltonian at each time step of the simulation, result in an exponential decay from level n, with an energy relaxation rate equal to n/T_1 and an exponential decay of the coherence $\rho_{m,n}$ at a rate $\frac{m+n}{2T_2'}$[7], as expected from an uncorrelated bath. To simulate pure dephasing, we use a similar model, however with a different coupling Hamiltonian: $H_{\mathrm{PD}} = \epsilon a^{\dagger}a\left(b + b^{\dagger}\right)$ [9], where $\epsilon = \sqrt{2/T_2 \Delta t}$. In this

[6] t_d is the characteristic timescale for changes of the state in the rotating frame.

[7] T_2' is the qubit decoherence time due to energy relaxation

case, the Kraus operators are given by[8]:

$$E_k = \sum_n \frac{\left((-i\epsilon\Delta t)\, a^\dagger a\right)^{2n-k}}{(n-k)!\sqrt{k!n!}} e^{\left((\epsilon\Delta t)a^\dagger a\right)^2/2}. \tag{5.11}$$

Application of these operators at each time-step leads to an exponential decay of the coherence $\rho_{m,n}$ at a rate $\frac{(m+n)^2}{T_2}$, as expected from an uncorrelated bath [99]. The above models are consistent with recent measurements of superconducting resonators [89], however similar measurements for our system are not available yet[9]. In each time step in the simulation we can apply both processes as necessary.

B Measurement error analysis

In Sec. 3.3, we described how to implement a genetic algorithm on the phase-circuit to optimize states. During run-time, we dynamically changed the number of repetitions of each tunneling measurement, in order to reach the required precision in the population overlap estimation χ. Since the fidelity depends on all the occupation probabilities P_n, its precision is determined by the precision of P_n, and therefore we are interested in finding these. In the phase circuit, the calculated occupation probability of each level depends on measurements of tunneling probabilities, taken at several pulse amplitudes.

Since each tunneling measurement yields two possible results (tunneled\not tunneled), each tunneling probability distributes binomially. From the binomial distribution, we know that the variance in the *average* of an escape probability $\langle P^k_{\rm esc}\rangle$, taken with N_r repetitions is given by $\sigma_k^2 \equiv {\rm Var}\langle P^k_{\rm esc}\rangle = \langle P^k_{\rm esc}\rangle \left(1 - \langle P^k_{\rm esc}\rangle\right)/N_r$. The occupation probabilities may have different distributions. Since tunneling events X_k for different measurement amplitudes k are *independent variables*, the variance of P_n is given by ${\rm Var}P_n = \sum_k A^2_{nk}\sigma^2_k$, where the matrix A is defined such that

[8] Derivation is available in a private writeup by Roy Resh and Yoni Shalibo.

[9] Our qutrit measurements indicate a slightly weaker scaling of the decay of coherence with level number. Our simulations probably overestimate decoherence rates.

$P_n = \sum_k A_{nk} P_{\text{esc}}^k$.

To find the variance in $\chi = \chi(\vec{P_n})$, we expand it to first order in the error about the average. For each experiment we get a different set of occupation probabilities Y_n such that $\langle Y_n \rangle = P_n$. In terms of the error $x_n = Y_n - \langle Y_n \rangle$ the measurement outcome can be approximated by,

$$\chi(\vec{x}) \approx \chi(0) + \sum_j \frac{\partial \chi}{\partial x_j} x_j. \tag{5.12}$$

The variance of χ is then given by $\sigma_\chi^2 = \langle \chi^2 \rangle - \langle \chi \rangle^2 = \sum_{j,k} \frac{\partial \chi}{\partial x_j} \frac{\partial \chi}{\partial x_k} \left(\langle x_j x_k \rangle - \langle x_j \rangle \langle x_k \rangle \right) =$ $\text{Tr}\left(\mathbf{J}\text{cov}\mathbf{x}\right)$, where $\mathbf{J}_{jk} = \frac{\partial \chi}{\partial x_j} \frac{\partial \chi}{\partial x_k}$ is the Jacobian matrix and $(\text{cov}\mathbf{x})_{jk} = \langle x_j x_k \rangle - \langle x_j \rangle \langle x_k \rangle$ is the covariance matrix of $\mathbf{x}$. For our population overlap definition, we get $\mathbf{J}_{jk} = \frac{1}{4}\sqrt{\frac{P_j P_k}{x_j x_k}}$. The covariance matrix $\mathbf{x}$ is calculated from the covariance matrix of the escape probabilities $\sigma_{jk} = \delta_{jk} \text{Var}\langle P_{\text{esc}}^k \rangle$, and is given by $A \sigma A^T$.

In a stochastic simulation we find that using a non-negative least-square fitting algorithm to find the occupation probabilities from measurements yields much smaller variance values for the population overlap. Therefore, in our genetic algorithm experiments we find the variance by multiple repetitions of the set of measurements (composed of N_r repetitions).

C Pulse correction

As explained in Sec. 2.2, imperfection in the microwave pulse shaping chain cause ringing in output pulse envelope and mixing between the I and Q channel. The mixing cannot be modeled as simple rotation in IQ space due to the ringing component of the cross-talk, and therefore any correction must use a combination of both the I and the Q channels. In Fig. 5.2 we plot an example of the I and Q components of the output signal for a 1 ns pulse in the I (blue) or Q (red) channels, measured with a sampling oscilloscope. These are calculated from the high-frequency signal by performing a numerical demodulation. We see a

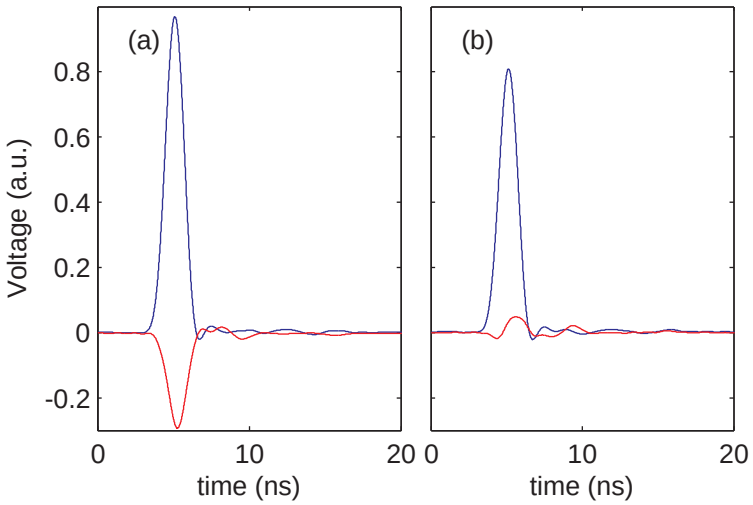

Figure 5.2: Ringing of the I and Q components of the signal. Envelope of the measured signal for a 1 ns input pulse applied on the I port (a) and the same input applied on the Q port. The I component of the demodulated signal is shown in blue and the Q component is shown in red.

non-negligible signal in the Q component, while a pulse is applied only in the I channel and vice versa for a signal in the Q channel. Moreover, we find that the ringing cannot be eliminated in one of the extracted components by phase shifting the demodulation (IQ rotation).

In the case where only a single channel contains a ringing output, we can eliminate ringing within bandwidth limits by reprogramming the DAC for a different pulse, which, in the presence of ringing will result in the desired output. Knowing the impulse response of the microwave chain, one can calculate its output for any given DAC output using convolution:

$$g(t) = \int f(\tau)h(t-\tau)d\tau, \tag{5.13}$$

where $f(t)$ is the DAC output and $h(t)$ is the impulse response function of the microwave chain. Using the convolution theorem, one gets $\tilde{g}(\nu) = \tilde{f}(\nu)\tilde{h}(\nu)$, where the tilde sign denotes a Fourier transform. The relation can then be inverted to obtain the *required* input given the *desired* output $f(t) = \mathcal{F}^{-1}\left[\tilde{g}(\nu)/\tilde{h}(\nu)\right]$. The last relation is generally known as deconvolution.

We can now generalize Eq. 5.13 for two channels, where the output of one is allowed to depend on the input of the other:

$$g_I(t) = \int I(\tau) h_{II}(t-\tau) d\tau + \int Q(\tau) h_{IQ}(t-\tau) d\tau, \tag{5.14}$$

and

$$g_Q(t) = \int I(\tau) h_{QI}(t-\tau) d\tau + \int Q(\tau) h_{QQ}(t-\tau) d\tau, \tag{5.15}$$

where h_{ij} is the impulse response of channel j in the i component of the output. By convolving both equations, we get:

$$\begin{pmatrix} \tilde{h}_{II} & \tilde{h}_{IQ} \\ \tilde{h}_{QI} & \tilde{h}_{QQ} \end{pmatrix} \begin{pmatrix} \tilde{I} \\ \tilde{Q} \end{pmatrix} = \begin{pmatrix} \tilde{g}_I \\ \tilde{g}_Q \end{pmatrix}.$$

By inverting the relation, we get:

$$\begin{pmatrix} \tilde{I} \\ \tilde{Q} \end{pmatrix} = \frac{1}{\Delta h} \begin{pmatrix} \tilde{h}_{QQ} & -\tilde{h}_{IQ} \\ -\tilde{h}_{QI} & \tilde{h}_{II} \end{pmatrix} \begin{pmatrix} \tilde{g}_I \\ \tilde{g}_Q \end{pmatrix}, \tag{5.16}$$

where $\Delta h = \tilde{h}_{II}\tilde{h}_{QQ} - \tilde{h}_{IQ}\tilde{h}_{QI}$. We then get the required input for a desired output by an inverse Fourier transform of Eq. 5.16. This method is only limited by the DACs' bandwidth ($\sim$ 300 MHz in our case). With this bandwidth limit we are able to eliminate the ringing to $\sim$ 1 % of the peak amplitude. Since the response changes as a function of frequency, we apply a frequency dependent correction, calibrated with 10 MHz steps. In our experiments, where the LO frequency can have any value, we choose the correction corresponding to the closest calibrated frequency.