A Purity Benchmarking Study of Superconducting Single-qubit Fluctuations

by

Auda Zhu

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Physics - Quantum Information

Waterloo, Ontario, Canada, 2023

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Quantum processes are susceptible to errors. Over the years, numerous noise models, such as two-level system noise and flux noise, have been proposed by physicists to describe the mechanisms behind the error sources affecting quantum processes. However, a comprehensive understanding of the quantum noise landscape, particularly on longer timescales, is still under active exploration. This thesis contributes to this ongoing effort, exploring long-term quantum noise through the lens of a superconducting Xmon transmon qubit.

In our study, we explore the long-term qubit noises by conducting continuous purity benchmarking experiments, utilizing a set of established metrics to gauge the quantum errors. These metrics, namely the average gate fidelity and unitarity, provides a more detailed characterization of quantum error compared to the commonly studied variables such as T_1 and frequency detuning, including characterization of the coherence property. These metrics are also the subject of intense discussions, particularly in the fields of quantum algorithms and quantum information processing hardware development. We measured the coherent and incoherent quantum error for very long time periods, up to 440 hours. Through these experiments. we gain valuable insights into the nature of the quantum noise and its impact on qubit coherence.

Following the experiments, we further attempt to reconcile our observations with well-established models, namely the two-level system and flux noise, through simultaneous measurements and comprehensive simulations. While we succeed in explaining certain aspects of the experimental results, our findings also highlight intriguing discrepancies between experimental observations and simulations, thus prompting further research.

Acknowledgements

I would like sincerely to thank my supervisor, Dr. Matteo Mariantoni for offering me this amazing and life changing opportunity proceeding in academia, and his tremendous help and support throughout last three years.

I would like to extend my sincere gratitude to my committee members, Dr. Jonathan Baugh and Dr. Adrian Lupascu, for their insightful feedback and generous support throughout this degree. Their understanding and patience have been an invaluable part of my learning and growth process.

I also owe a debt of gratitude to Dr. Jérémy H. Béjanin, with whom I had the privilege of working closely. His extensive knowledge, practical assistance, and willingness to share his expertise have been invaluable to my research. As a senior member in the group, his pedagogical approach to collaborations not only improved the quality of my work but also greatly enhanced all junior members' learning experience. Beyond our academic interactions, Dr. Jérémy H. Béjanin has also served as an exemplary role model in both my professional and personal life. I am immensely grateful for his guidance.

I must also express my appreciation for the extraordinary team members with whom I am privileged to work: Christopher Xu, Mohammad Ayyash, Noah Gorgichuk, and for Dr. Yosri Ayadi, with whom I had the honor to work previously. Their consistent dedication, unique talents, and unwavering support have made a significant impact on my academic journey.

On a personal note, I would like to express my deepest gratitude to those who provided me with unwavering support throughout my degree journey. My parents, Mengcun Zhu and Hongxin Zhang, and my friends, Boyan Liu and Han Huang, deserve a special mention for their enduring support and encouragement.

Throughout this journey, I have been fortunate to encounter countless touching stories and kind souls that have greatly enriched my experience and shaped my soul. It was impossible for me to reach this milestone without their support, and I regret that it is impossible to acknowledge them all here.

Finally, I reserve these special thanks for those extraordinary individuals who stood by me during a critical juncture in my life. Christopher and Alicia, Carter Tsai and Vicky Zhou, Yiran Wang and Mei Dong, Dr. Kesava KV Reddy, Dr. Waleed Abdulaziz Alsunbul, Dr. Yosri Ayadi, as well as Dr. Matteo Mariantoni. I am forever indebted to them for their kindness during that time of hardship.

Dedication

To my parents, Mengcun and Hongxin, for their unwavering support and love throughout my life journey. To all those who offered help and supported me through my *setsuna* of *shoumetsu*, a moment for life or extinction. This degree has reshaped me and my beliefs, and I carry its impact with me as I present this work.

Table of Contents

List of Figures

List of Tables

Personal Message

We are thinking of you, and we wish you a full and speedy recovery.
We are looking forward to have you back healthy.

Matteo, Yoon, Chris & Jérémy

The gift card I received in my despair, a reminder of the love and support from my team.

Chapter 1

Introduction

With recent breakthroughs in both experimental design and quantum algorithms[1, 2], superconducting qubits have emerged as a leading candidate for fault-tolerant quantum information processing devices. In order to achieve fault tolerance practically, it is crucial to accurately model and mitigate stochastic fluctuations that can degrade the performance of superconducting qubits. The physical mechanisms of the quantum errors and noises can be explored and studied via various metrics. Some broadly used metrics include relaxation time (T_1) and the dephasing time (T_2). In this thesis, we study the coherent and incoherent stabilities of single "Xmon" transmon qubit through long-time purity benchmarking experiments. The corresponding metrics obtained through the experiment are introduced in Chapter 2 and analyzed in Chapter 3.

This introductory chapter provides an brief overview of the critical components of our study. We will first delve into the essential instruments for the experiments - the superconducting resonator and the Superconducting Qubit - from the perspective of circuit quantum electrodynamics (QED), with a specific focus on frequency-tunable Xmon transmon qubits. This is followed by brief explanations of experimental procedures: how to practically readout and manipulate the state of the qubit. Finally, we justify the purpose of this study: exploring the physics behind the quantum noise. We also review a specific noise model pertinent to our investigation.
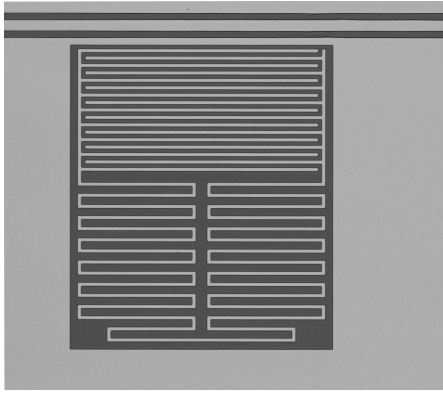
(a) Optical photo of a hanger-type lumped-element resonator
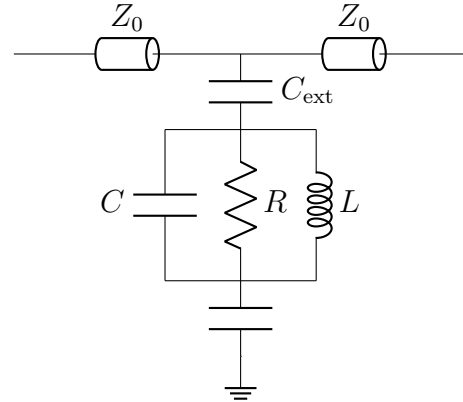
(b) Equivalent circuit diagram

Figure 1.1: Hanger-type lumped-element superconducting resonator

1.1 Circuit Quantum Electrodynamics

1.1.1 Superconducting Resonator

In the context of circuit QED, superconducting resonators are typically implemented as microwave resonators. They serve as the quantum mechanical counterparts to classical LC circuits and can be mathematically described as harmonic oscillators. Furthermore, superconducting resonators play a critical role as ancillary systems for qubit readout, providing a key mechanism for state measurement in superconducting qubits. Thus, these resonators act as an indispensable component within the circuit QED framework.

Figure 1.1a shows a photo of a planar lumped-element superconducting resonator[3]. The resonator consists of interdigital electrodes and a meandering line, both made of superconducting aluminum on top of a silicon wafer. The interdigital electrodes and meandering line effectively forms a parallel LC circuit, capacitively coupled to a coplanar waveguide (CPW) transmission line above. Due to the resistance of the silicon wafer, a resistor is added to the circuit in parallel to better model the corresponding energy dissipation. The classical circuit model of this lumped-element resonator is presented in Figure 1.1b.

Quantum Description

In order to derive the quantum description of the LC resonator, i.e. the quantum Hamiltonian of the system, we start from the classical circuit model. Since the typical energy relaxation time of a superconducting resonator is much longer than the oscillation period, we can assume the resistor is open and the loss from it is negligible, without loss of generality. This simplified system can be represented by a simplest LC circuit.

The current-voltage relationships across the capacitor C and the inductor L are:

$$v_L = -L\frac{di_L}{dt} \tag{1.1}$$

$$i_C = C\frac{dv_C}{dt} \tag{1.2}$$

Since the electric field within this closed system is conservative, we can directly write down the classical Lagrangian and Hamiltonian for the circuit as:

$$\mathcal{L} = \frac{1}{2}(Cv_C^2 - Li_L^2) \tag{1.3}$$

$$H = \frac{1}{2}(Cv_C^2 + Li_L^2) \tag{1.4}$$

Although the current i and voltage v are the most easily measured quantities in classical electrical engineering, it is more common to express this Hamiltonian in another pair of canonical conjugate variables: flux ϕ and charge q. The variation is given by

$$v_L = \frac{d\phi_L}{dt} \tag{1.5}$$

$$i_C = \frac{dq_C}{dt} \tag{1.6}$$

Combined with the Kirchhoff's circuit laws for the LC circuit, we obtain the Hamiltonian

$$H = \frac{1}{2L}\phi_L^2 + \frac{1}{2C}q_C^2 \tag{1.7}$$

Now it is time for the canonical quantization: we promote the variables ϕ_L and q_C to operators $\hat{\phi}$ and $\hat{q}$ in a Hilbert space and impose the canonical commutation relation

$$[\hat{\phi}, \hat{q}] = i\hbar \tag{1.8}$$

3

and now we have the quantum Hamiltonian for a superconducting resonator.

$$\hat{H} = \frac{1}{2L}\hat{\phi}^2 + \frac{1}{2C}\hat{q}^2 \tag{1.9}$$

Notice that Equation 1.9 possesses the exact mathematical structure as the Hamiltonian for the quantum harmonic oscillator. Drawing inspiration from the method developed by Paul Dirac[4], we proceed to change the operational basis by defining the following ladder operators:

$$a = \sqrt{\frac{C\omega}{2\hbar}}(\hat{\phi} + iL\omega\hat{q}) \tag{1.10}$$

$$a^\dagger = \sqrt{\frac{C\omega}{2\hbar}}(\hat{\phi} - iL\omega\hat{q}) \tag{1.11}$$

with $\omega = \frac{1}{\sqrt{LC}}$. Finally we obtain the diagonalized form of the quantum Hamiltonian:

$$\hat{H} = \hbar\omega(a^\dagger a + \frac{1}{2}) \tag{1.12}$$

1.1.2 Superconducting Qubit

An essential component of circuit QED, and the heroine of this thesis, is the superconducting qubit. Unlike resonators, these two-level systems (or multi-level systems in the case of a qu-"dit") possess distinct states that register information and evolve in alignment with the quantum nature of the world. This is typically achieved by introducing non-linearity in the energy spacing between each stationary state in a quantum resonator. There is a vast amount of different designs proposed in the field. We primarily focus on a type of qubit design called the transmon[5], which is derived from the charge qubit (Cooper-pair box), and its variant, the Xmon transmon qubit[6].

The non-linearity of a superconducting qubit typically arises from the Josephson element within the circuit. This element can either be a single junction or a pair arranged in a loop. This device is named after British physicist Brian Josephson. In its simplest form, a single Josephson junction is constructed by positioning a thin layer of dielectric material between two layers of superconducting material. The charge carriers, Cooper pairs in this case, go through the dielectric material by quantum tunneling to form the current. Josephson, in his Nobel-winning paper published in 1962[7], predicted the behavior of this specific circuit element using equations that are now well-known as the Josephson Equations:

$$i_J = I_0\sin(\varphi) \tag{1.13}$$

$$v_J = \frac{\hbar}{2e}\frac{d\varphi}{dt} \tag{1.14}$$

where I_0 is the critical current of the junction determined by the physical design[8], and e is the elementary charge constant. These equations establish relationships between the branch current i_J and the potential v_J to an order parameter: phase difference φ.

By integrating both sides of the second equation 1.14 in time, we obtain a relationship between the magnetic flux ϕ and the phase φ:

$$\phi_J = \frac{\hbar}{2e}\varphi = \frac{\Phi_0}{2\pi}\varphi \tag{1.15}$$

where $\Phi_0 = \frac{h}{2e}$ is the magnetic flux quantum. Substitute Equation 1.15 back to Equation 1.13 and we obtain a pivotal relation for the qubit design:

$$i_J = I_0\sin(\frac{2e}{\hbar}\phi) \tag{1.16}$$

This is where the non-linearity manifests. In a classical inductor, the current flowing through it exhibits a linear relationship with the magnetic flux, whereas in the case of a Josephson junction, the current is proportional to the non-linear sinusoidal function of the flux. This is why sometimes the Josephson junction is also called "non-linear inductor".

To harness this non-linearity, we substitute the inductor in the superconducting resonator with a Josephson junction (as shown in Figure 1.2a), transforming it into an anharmonic oscillator. The classical Hamiltonian of system reads:

$$H = -E_J\cos(\phi) + \frac{1}{2C}q_C^2 \tag{1.17}$$

where $E_j = \frac{\hbar I_0}{2e}$ is the Josephson energy. After the canonical quantization, we have:

$$H = -E_J\cos(\hat{\phi}) + \frac{1}{2C}\hat{q}^2 = -E_J\cos(\hat{\phi}) + 4E_c\hat{n}^2 \tag{1.18}$$

where $E_c = \frac{e^2}{2C}$ is the single-electron charging energy and $\hat{n}$ is the charge number operator rescaled from the charge operator $\hat{q}$ by $n = \frac{q}{2e}$.

The diagonalization of this Hamiltonian becomes significantly more complex compared to the case of resonators. Since the "potential" term in the Hamiltonian is no longer of quadratic form, the same basis induced by the harmonic ladder operator cannot diagonalize this Hamiltonian anymore. More detailed discussion of the diagonalizations, with specific assumptions made for different qubit designs are beyond the scope of this thesis. Here
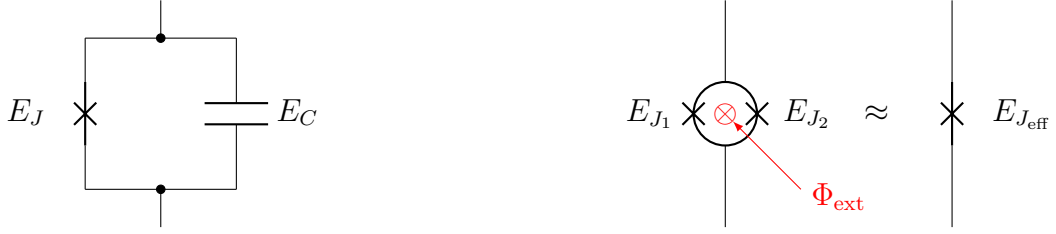
(a) Circuit of an anharmonic oscillator involving single Josephson junction

(b) Circuit diagram of a SQUID (two Josephson junction in parallel)

Figure 1.2: Circuit diagrams for an anharmonic oscillator and a SQUID

we only present the results for the transmon qubit[9, 5, 10]. The approximated two-level Hamiltonian reads

$$\hat{H} \approx -\frac{1}{2}\hbar\omega_{01}\hat{\sigma}_z \tag{1.19}$$

where $\hbar\omega_{01} = E_1 - E_0 \approx \sqrt{8E_J E_C} - E_C$ is the first qubit transition energy. In addition, the anharmonicity for a transmon qubit is approximately equal to $-E_C$.

Frequency-Tunable Xmon Transmon Qubits

The qubit transition frequency of a transmon qubit is determined by the Josephson E_J and charge energy E_C of corresponding circuit elements. In order to change the frequency of the qubit we need to make one of those parameters tunable. This is usually done by replacing the single Josephson junction in the circuit with two in a loop, forming a superconducting quantum interference device[11], or SQUID. The SQUID configuration allows for control of the effective Josephson energy E_J by applying an external magnetic field, which induces a flux through the loop. This flux modifies the E_J, and consequently biases the qubit transition frequency, enabling us to tune the qubit to desired frequencies.

Figure 1.2b presents a circuit diagram for one SQUID. J_1 and J_2 are two Josephson junctions with Josephson energies E_{J_1} and E_{J_2} respectively. Some external magnetic flux Φ_{ext} is applied through the loop. Given the flux quantization conditions in a superconducting loop, one can find the SQUID as one circuit element behave just like a single Josephson junction with an effective Josephson energy E_{eff}. The effective Hamiltonian can be expressed as:

$$H_{\text{SQUID}} = -E_{\text{eff}}\cos(\varphi) \tag{1.20}$$

$$E_{\text{eff}} = (E_{J_1} + E_{J_2})\sqrt{1 + \left(\frac{E_{J_2} - E_{J_1}}{E_{J_2} + E_{J_1}}\tan(\frac{\pi\Phi_{\text{ext}}}{\Phi_0})\right)^2}\cos(\frac{\pi\Phi_{\text{ext}}}{\Phi_0}) \tag{1.21}$$

where φ is the phase difference across the SQUID[12, 13].

Evolved from the transmon qubit, the Xmon transmon qubit uses a DC SQUID as its Josephson element, enabling it to be frequency-tunable. During operation, a DC current is applied around the SQUID, generating an external magnetic flux. This flux biases the frequency of the qubit. To control the qubit frequency, we adjust the amplitude of the applied current. As a result, the flux bias and hence the qubit frequency change accordingly.

In experiments, this current is generated by applying a DC voltage to the Z line of the qubit. Given that the impedances of the control lines can fluctuate over different samples and cool-downs, the absolute amplitude of the individually applied voltage becomes less meaningful. Nevertheless, the mapping between applied voltage and qubit frequency has been experimentally found quite robust and stable over time. Therefore, we do not explicitly differentiate between the concepts of qubit frequency and bias voltage for the same qubit sample under the same experimental setup. Furthermore, in this thesis, I use the terms "current bias", "voltage bias", and "frequency bias" interchangeably: all refer to the same concept of changing the qubit frequency to a specific value.

Gate Operation

The state of an Xmon transmon qubit can be manipulated using microwave pulses that are delivered via a capacitively coupled drive line. The Hamiltonian of a qubit driven by a classical field can be expressed as[14]:

$$\hat{H} = -\frac{\hbar}{2}\omega_Q\hat{\sigma}_z + i\hbar\varepsilon(t)\cos(\omega_d t + \phi_d)(\hat{\sigma}_+ - \hat{\sigma}_-) \tag{1.22}$$

Here, $\varepsilon(t)\cos(\omega_d t + \phi_d)$ represents the applied classical microwave pulses. $\omega_d t$ and ϕ_d are the frequency and phase of the drive pulse, respectively, while $\varepsilon(t)$ is the rescaled envelope of the pulse. The qubit is typically driven close to resonance, meaning the drive detuning $\Delta\omega = \omega_Q - \omega_d$ is small. After transforming to a frame rotating at the drive frequency ω_d and applying the rotating-wave approximation, we obtain the rotating frame driven Hamiltonian:

$$\hat{H}' = \frac{\hbar}{2}\Delta\omega\hat{\sigma}_z + \frac{\hbar}{2}\varepsilon(t)\Big(\cos(\phi_d)\hat{\sigma}_y - \sin(\phi_d)\hat{\sigma}_x\Big) \tag{1.23}$$

Solving the time-evolution dynamics for $\hat{H}'$ leads to the phenomenon known as Rabi oscillations, which represent the transition of the qubit between its ground and excited states. Importantly, by adjusting the relative phase ϕ_d, we can manipulate the phase difference of the qubit states. In the context of a single qubit, this is equivalent to rotating the qubit state within the Bloch sphere along an axis in the xy-plane. Any rotations within this plane can be achieved by meticulous tuning the envelope and relative phase of the pulse. This type of drive line is often referred to as the XY line.

All single-qubit operations can be represented by rotations within the Bloch sphere, corresponding to transformations in the special unitary group SU(2) or the special orthogonal group SO(3). These operations involve rotations about the x, y, or z axis of the Bloch sphere. By precisely controlling the amplitude and phase of the microwave pulses applied through the XY line, we can perform any arbitrary rotation around the x and y axes. Since the entire Lie algebra so(3) can be generated by any two non-parallel rotations[15], we can realize any desired single-qubit operation through combinations of these rotations in the xy-plane, fully spanning the SU(2) or SO(3). This level of control is crucial for implementing quantum algorithms in quantum computing.

1.2 Experimental Infrastructure

In this section, we briefly review the experimental techniques necessary for measuring the state of a qubit and executing gate operations.

1.2.1 Homodyne Readout

For most types of superconducting qubits, including the Xmon transmon qubit, the readout of the qubit's energy state is performed through the strong interaction between the qubit and a resonator. Depending on the state of the qubit, the resonant frequency of the resonator will shift by different amount. This phenomenon is closely related to the AC Stark shift observed in quantum optics.

When the resonator is strongly coupled to a qubit that is dispersively detuned, the resonant frequency of the resonator will be shifted away from its classical frequency in a direction opposite to the qubit frequency. Denote the detuning of the qubit as $\Delta = \omega_Q - \omega_R$. The (angular) frequency shift when the qubit is in ground state $|g\rangle$ can be represented as χ in the following equation:

$$\chi = -\frac{g^2}{\Delta} \tag{1.24}$$

where g represents the coupling strength constant. On the other hand, when the qubit is in an excited state, this (angular) frequency shift, for a transmon, would be smaller and dependent on its charge energy[9]:

$$\Delta\omega \approx \frac{E_C + \Delta}{E_C - \Delta}\chi \qquad (1.25)$$

Thus, the measurement of the resonator's frequency shift is an effective method for destructively determining the qubit's state. This technique is commonly referred to as dispersive readout.

For dispersive readout, the state of the readout resonator is determined using homodyne detection. During this process, a pulse of a few microseconds in length at the frequency $(\omega_R + \chi)/2\pi$ is sent through the transmission line. This pulse is then highly attenuated before it reaches the resonator. After interaction with the resonator, the signal is amplified twice and transmitted back to the analog-to-digital converter (ADC) measurement device.

All the instruments involved in this process are synchronized from a 10 MHz external reference generated by a rubidium atomic clock. The timing of pulse generation and readout measurements is further fine-tuned by a triggering signal, ensuring precise alignment. As a result, the measured readout waveform is compared with the input signal to obtain the single frequency complex-valued transmission coefficient, denoted as S_{21}.

The behaviors of resonators are well understood. By analyzing the phase and amplitude of S_{21}, we can determine whether the resonator is at the measuring frequency, thus enabling us to infer the state of the qubit.

1.2.2   Qubit Control

In our experimental setup, qubit control is facilitated by driving pulses through a capacitively coupled XY line, as outlined in Section 1.1.2. Each experiment commences with a series of preliminary calibrations, including a Ramsey measurement, which fine-tunes the qubit frequency. The resultant frequency is then used to set the frequency of the sinusoidal carrier wave for the XY pulses, enabling close-resonant control of the qubit.

Once the frequency is established, we define the envelope of the pulses. The design of the pulse envelope is crucial to the qubit control. Since a qubit is not a truely two-level system, particularly in the case of transmon-like qubits which often exhibit weak anharmonicity on the order of 100 MHz, an improperly designed XY pulse can result in substantial leakage into the third energy level of the qubit. This leakage can lead to significant experimental errors.

To address this challenge, we employ cosine-shaped derivative removal by adiabatic gate (DRAG) pulses. This method allows us to execute qubit gate operations with a minimal error rate more efficiently. More detailed descriptions about the design and parameter calibrations are given in Reference [16, 17].

1.3 Quantum Noise

Quantum noise refers to the uncertainties and random fluctuations that occur in quantum systems. These fluctuations can be inherent from characteristics of quantum mechanics and also from various sources, including quantum decoherence, imperfect measurements, and environmental interactions. Understanding and mitigating quantum noise is therefore a fundamental challenge in the field of quantum information. For quantum computing, the ultimate goal is to mitigate these errors, aiming to keep the error threshold below a certain level at all times. This task requires a dynamic approach due to the time-dependent nature of quantum errors.

Beside practical considerations, studying and mitigating quantum noises also provide a valuable context for advancing our understanding of quantum mechanics and environmental interactions.

For instance, on-chip superconducting devices are known to suffer largely from interactions with defects in amorphous dielectric materials, which can be modelled as two-level systems (TLS)[18, 19]. Experimental studies have shown that the TLSs introduce stochastic fluctuations in qubit frequency as well as energy relaxation and dephasing channels[20, 21]. These experimental results can be well-explained by the General Tunneling Models[22]. Such stochastic channels inevitably affect the gate fidelity for quantum algorithms. Moreover, studying these quantum errors not only helps us understand and mitigate noise in quantum systems but also provides a unique lens through which we can explore fundamental physical phenomena.

In this section we briefly review a popular model characterizing the physical impacts of TLSs to the system and justify how these impacts could contribute to the quantum errors.

1.3.1 Two-Level System

The dynamics of TLS noises can generally be described by the *generalized tunneling model* (GTM)[22], which states that superconducting quantum devices, including qubits and ancilla resonators, interact semi-resonantly with an ensemble of two-level systems described

earlier. Each TLS is characterized by its own transition frequency E_i and coupling strength g_i. Moreover, the TLSs interact with the thermal bath, leading to energy relaxations $\Gamma_{1,i}^{\text{TLS}}$ and stochastic fluctuations in the TLS frequency E_i. Overall, the Hamiltonian of a driven qubit coupled to TLSs in the rotating frame of qubit drive frequency reads

$$\hat{H}_{\text{int}} = \Delta_q \sigma_+^{\text{q}} \sigma_-^{\text{q}} + \sum_i \frac{\Delta_{\text{TLS}}}{2} \sigma_z^{\text{TLS}} + \sum_i g_i \left(\sigma_+^{\text{q}} \otimes \sigma_-^{\text{TLS}} + h.c. \right) + \left(\frac{\Omega}{2} \sigma_x^{\text{q}} - \frac{\Omega^*}{2} \sigma_y^{\text{q}} \right) \quad (1.26)$$

where $\Delta_{\text{q}} = 2\pi(f_{\text{q}} - f_{\text{drive}}), \Delta_{\text{TLS}} = 2\pi(E_i - f_{\text{drive}})$, f_{q} is the qubit transition frequencies, E_i is the transition frequnecy of the i-th TLS in the ensemble, and f_{drive} is the drive frequency. Both the energy relaxation rate $\Gamma_{1,i}^{\text{TLS}}$ and the frequency of the TLS fluctuate stochastically due to the coupling to the thermal bath.

Energy Relaxation

The energy relaxation from TLS coupling is known to be one of the major mechanisms contributing to qubit relaxation. Given that the energy relaxation time T_1 for TLSs is typically much shorter than that of the qubit, and considering the experimental limitations in precisely estimating TLS parameters, it is practical to simplify this process to an Markovian amplitude damping channel. The corresponding Lindblad operator for the master equation is:

$$L_1^{\text{TLS}} = \sqrt{\Gamma_1^{\text{TLS}}} \mathbb{I}^{\text{q}} \otimes \sigma_-^{\text{TLS}} \quad (1.27)$$

Meanwhile, the inherent energy relaxation of the qubit itself can be represented as a combination of phase and amplitude damping channels:

$$L_1^{\text{q}} = \sqrt{\Gamma_1^{\text{q}}} \sigma_-^{\text{q}} \otimes \mathbb{I}^{\text{TLS}}, \qquad L_2^{\text{q}} = \sqrt{2\Gamma_2^{\text{q}}} \sigma_+^{\text{q}} \sigma_-^{\text{q}} \otimes \mathbb{I}^{\text{TLS}} \quad (1.28)$$

The rapid energy relaxation of TLSs accelerates qubit decoherence, thereby leading to quantum errors. The stochastic nature of TLSs further introduces fluctuations to these quantum errors.

Frequency Shift and Leakage

The qubit and the TLS interact via the electric dipole interaction. When a qubit is coupled to a TLS, this interaction results in a shift in the eigen-frequencies of the system's

Hamiltonian, leading to a detuning of the qubit frequency. Since both the DRAG pulses involved in qubit operation and the qubit readout depend on precise qubit frequency, any such detuning shift could result in a phase shift, imprecise axis rotations, and leakage, thereby introducing errors into the quantum process.

In practice, it is possible to mitigate this type coherent errors by system calibrations[23]. However, recent studies suggest that the TLS's frequency and energy relaxation rate can vary significantly over time due to the coupling with the stochastic thermal bath[20, 22, 19, 24]. This uncertainty necessitates frequent re-calibrations in order to maintain accurate control of the qubit. While this approach may be feasible for short experiments with a limited number of qubits, it is neither practical nor scalable for larger-scale experiments.

Chapter 2

Method

In quantum information processing, the performance of algorithms and protocols is fundamentally limited by the presence of noise and errors within the system. Therefore, understanding and estimating these noise channels are crucial steps toward the design and implementation of robust quantum computing systems. The goal of this chapter is to formulate an experimental approach for estimating both coherent and incoherent errors associated with a single superconducting qubit.

In Section 2.1, we delve into some widely recognized mathematical measures specifically tailored for quantitative characterization of quantum errors. These include the error rate or average gate infidelity, and the unitarity measure. These metrics serve as a benchmark for the degree of deviation from the desired quantum state.

Following that, we explore two robust and scalable protocols for estimating coherent and incoherent errors - randomized benchmarking and purity benchmarking. These protocols operate by experimentally measuring the metrics introduced in the earlier section.

In the final part of this chapter, Section 2.3, we propose two experimental procedures aimed at continuously monitoring the fluctuations in the qubit errors. The rationale behind these procedures is to enable a more in-depth understanding of the error dynamics and how they affect the quantum system's performance over a long-time scale.

2.1   Metrics

Conceptually, each quantum experiment involves three operational procedures: preparation, dynamic transformation, and measurement, Each of which is carried out by an

abstract quantum instrument which applies a specific transformation to an input density operator. We can define an *implementation* of a quantum transformation [25] as a map $\Phi : \mathbb{G} \longrightarrow \mathcal{L}\left(\mathcal{L}\left(\mathcal{H}_D\right)\right)$, where $\mathbb{G}$ is an arbitrary unitary group and $\mathcal{L}\left(\mathcal{L}\left(\mathcal{H}_D\right)\right)$ is the space of quantum channels. The implementation of g describes the actual channel implemented on the quantum state with a request of gate operation g. The purpose of this definition is to establish a connection from the group elements (usually the group of unitary transformation) to the corresponding quantum channels that the state undergoes, allowing us to characterize the behavior of the quantum instrument. For an ideal instrument, the implementation Φ_0 of an element $g \in \mathbb{G}$ can be expressed as

$$\Big(\Phi_0\left(g\right)\Big)(\rho) = U_g \rho U_g^\dagger \tag{2.1}$$

where U_g is an unitary representation of the group element g. This equation implies that this implementation applies the unitary transformation given by the element exactly as desired. We refer to such perfect implementation as a *reference implementation* and denote it with a subscript "$_0$". It is worth noting that the map $\Phi_0(g)$ defined above is a homomorphism and preserves the associativity from group $\mathbb{G}$. Hence it is well-defined under gate composition. In addition, the unitary representation of gates used in quantum computation is often faithful, making the reference map isomorphic.

Real-world implementations are subject to a variety of imperfections, such as imprecision in the instrument control and environmental noise. These imperfections can lead to error channels in the actual implementations, deviating the transformation from their ideal reference. Hence the nature of quantum errors can be seen as the deviation of a practical implementation from its reference.

The study of the error process could be done in a fully deterministic way by reconstructing the entire implementation matrix. One experimental protocol for achieving this is quantum process tomography (QPT)[26]. However, it is well-known that QPT is not scalable in terms of system size, as the complexities of both query and side computations grow exponentially with the system. Additionally, QPT requires the state preparation and measurement (SPAM) errors to be much lower than the error process being characterized, which is often hard to achieve in real-world scenarios[27, 28]. These limitations make QPT less practical and efficient for characterizing quantum errors.

Besides the full characterizations, it is also possible to learn about quantum errors through partial characterization. To characterize quantum errors in real-world implementations, it is necessary to develop measures that can quantify the extent to which an implementation deviates from its ideal reference. These measures can provide statistical insights into the nature of quantum errors, even if they do not provide a complete picture

14

of the error behavior. Such a measure could be naturally induced by a metric distance between the implementation map and its reference.

Over time, multiple metrics have been developed to serve different research interests. For example, two of the most popular definitions are known as diamond distance[29, 30] $\epsilon_\diamond$, which is popular in the theoretical studies of quantum algorithm and fault-tolerant threshold theorems, and the average gate fidelity[31] $\mathcal{F}_{\mathrm{avg}}$ (or $\overline{\mathcal{F}}$), which is more experimental accessible, especially since the development of randomized benchmarking. The average gate fidelity can be used to describe the errors averaged uniformly over the space of states and the operations, while the diamond distance can be thought as the gate fidelities in the worst case scenarios.

Interestingly, average gate fidelities provide bounds for the diamond distance on both sides via the Wallman-Flammia bounds[32]:

$$(1 - \mathcal{F}_{\mathrm{avg}})\frac{D+1}{D} \leq \frac{1}{2}\epsilon_\diamond \leq \sqrt{1 - \mathcal{F}_{\mathrm{avg}}}\sqrt{D(D+1)}$$

This relation implies that examining either of the two metrics can also provide valuable insights into the other. In practical terms, the average gate fidelity is more easily accessible, especially with the development of randomized benchmarking. Therefore, this thesis primarily focuses on the average gate fidelity as a means of characterizing quantum errors.

As indicated by the name, the average gate fidelity is an average of the gate fidelity $\mathcal{F}_g$ over all pure input states, weighted by the Fubini-Study measure. For an implementation of gate g, the average gate fidelity can be defined as[33]:

$$\mathcal{F}_{\mathrm{avg}}\left(\Phi\left(g\right), \Phi_0\left(g\right)\right) := \int_{\mathcal{H}_D} d\psi \ F_g\left(\Phi\left(g\right), \Phi_0\left(g\right), |\psi\rangle\langle\psi|\right)$$

$$= \int_{\mathcal{H}_D} d\psi \ \langle\psi|U_g^\dagger \left(\Phi\left(g\right) \circ |\psi\rangle\langle\psi|\right) U_g|\psi\rangle \tag{2.2}$$

where $d\psi$ is the Fubini-Study measure over the pure states. It is convenient to write the *pure error process* of the implementation $\Phi(g)$ as

$$\mathcal{E}_g := \Phi_0^\dagger\left(g\right) \circ \Phi\left(g\right) = U_g^\dagger\left(\Phi(g) \circ (\cdot)\right) U_g \tag{2.3}$$

which is a composition of an imperfect implementation and the inverse of the ideal transformation. Essentially, this process characterizes the pure error process relative to the imperfect implementation of identity. It should be noted that by making the assumption that $\mathcal{E}_g$ represents the pure error process, we assume implicitly that the error occurs prior

to the unitary process. However, this assumption is generally satisfied under the Markovian limit.

We write $\mathcal{E}_g$ in Kraus decomposition:

$$\mathcal{E}_g(\rho) = \sum_k A_k^{(g)} \rho A_k^{(g)\,\dagger} \tag{2.4}$$

The average gate fidelity can be expressed as[34, 35, 36]:

$$\mathcal{F}_{\text{avg}}(\mathcal{E}_g) = \int_{\mathcal{H}_D} d\psi\, \langle\psi|\mathcal{E}_g(|\psi\rangle\langle\psi|)|\psi\rangle = \frac{\sum_k |\text{Tr}[A_k^{(g)}]|^2 + D}{D^2 + D} \tag{2.5}$$

Finally, we can define a quantity to characterize the average quantum errors associated with the pure error process, the *error rate* ϵ:

$$\epsilon(\mathcal{E}_g) = 1 - \mathcal{F}_{\text{avg}}(\mathcal{E}_g) \tag{2.6}$$

The error rate is sometimes called infidelity of the error process since it is a subtraction of the gate fidelity. This error rate is a useful quantity for experimentally characterizing the overall errors in practice, because it is easily accessible via randomized benchmarking.

2.1.1 Coherence and Unitarity

While the error rate ϵ is a commonly used metric for quantifying the overall quality of a quantum gate operation, it does not provide insight into the specific types of errors that contribute to the overall error rate. To address this, a useful approach is to divide the error rate based on the coherence of the error process, which characterizes how well the process preserves the coherence of the quantum state.

Since any coherent quantum operation can be represented by a unitary transformation, a good measure for the coherence is the unitarity of the process. The unitarity of a quantum process should remain invariant under unitary transformations and reach its maximal value if and only if the process itself is fully unitary.

The significance of this division is that the coherent error does not destroy or decrease the coherence of the system. In fact, it is usually caused by systematic control errors in the gate implementations[37, 38]. Hence it in principle can be corrected via better calibrations or other error correction algorithms. In addition, the coherent error rate could provide a metric of how well the instruments are experimentally calibrated. The incoherent error on

the other hand, is usually caused by decoherence channels such as energy relaxation and dephasing, which are harder to improve in reality.

A natural way to define the unitarity is by taking the average change in purity of a pure state when a channel is applied, since the purity of the state persists when the channel is unitary and reduces when non unitary. A more robust and popular definition is given by [39], where the unitarity is defined as the average purity of output states subtracted with identity components:

$$u(\mathcal{E}_g) = \frac{D}{D-1} \int d\psi \, \mathrm{Tr}\Big(\mathcal{E}_g'(\psi)^\dagger \mathcal{E}_g'(\psi)\Big) \tag{2.7}$$

where $\mathcal{E}_g'(A) = \mathcal{E}_g(A) - \frac{\mathrm{Tr}\mathcal{E}_g(A)}{\sqrt{D}}\mathbb{I}$ for all trace-less Krause operators A, and the integral is over all pure states. $u(\mathcal{E})$ reaches to the maximum 1 if and only if the process is unitary. The unitarity also provides a upper bound of errors in terms of average gate fidelity that can be corrected using unitary correction operations:

$$\epsilon_{\mathrm{coh}}(\mathcal{E}_g) \le \epsilon(\mathcal{E}_g) - \frac{D-1}{D}(1 - \sqrt{u(\mathcal{E})}) \tag{2.8}$$

Based on this bound, we can partition the total error rate ϵ by defining the incoherent error rate as:

$$\epsilon_{\mathrm{inc}}(\mathcal{E}_g) = \frac{D-1}{D}(1 - \sqrt{u(\mathcal{E}_g)}) \tag{2.9}$$

and the corresponding coherent error rate

$$\epsilon_{\mathrm{coh}} = \epsilon - \epsilon_{\mathrm{inc}} \tag{2.10}$$

The quantities ϵ_{coh} and ϵ_{inc} can be experimentally estimated using a modified randomized benchmarking protocol, often referred to as purity benchmarking [39, 40].

2.2 Randomized Benchmarking

An experimental protocol was proposed in 2005 by J. Emerson et al. [41] to study quantum errors through partial characterization. It was later named as randomized benchmarking (RB) by E. Knill & R. Laflamme (2007) [42]. This protocol, which involves using random gate sequences and the Haar average, offers a statistical approach to estimating quantum errors. Later, E. Magesan et al. made further improvements and developed the first well-grounded and practically scalable version of RB protocol in 2012 [43]. RB has since become

a popular method for estimating the average error rate of a quantum gate or circuit, and has been used to optimize the performance of quantum computing systems.

Let $\mathcal{S}^{m+1} = \{g_1, g_2, ..., g_m, g_{\text{inv}}\}$ be a sequence consisting of m gates drawn randomly from the unitary group $\mathbb{G}$, and an inverse gate $g_{\text{inv}} = (g_m \circ ... \circ g_1)^{-1}$ determined classically. We evaluate the implementation of the entire sequence:

$$\Phi(\mathcal{S}^{m+1}) = \bigcirc_{i=1}^{m+1} \Phi(g_i) = \bigcirc_{i=1}^{m+1} (\Phi_0(g_i) \circ \mathcal{E}_{g_i}) \tag{2.11}$$

where $\mathcal{E}_{g_i}$ is the individual pure error process for each gate. The $\bigcirc_{i=1}^{m+1}(\cdot)$ symbol here denotes the sequential composition of the operation from $i = 1$ to $i = m + 1$. Again, by writing down the second equation in Equation 2.11, we made an assumption that the error process has negligible memory effects on the time scale of gate operation, which is also known as a Markovian process. Since the last gate is the inverse of the composition of all previous m gates, the reference $\Phi_0(\mathcal{S})$ is an identity map and the $\Phi(\mathcal{S})$ can be treated as a pure error process. The key result from RB is that the Haar average of such implementations over the unitary group $\mathbb{G}$ is equivalent to a linear combination of depolarizing channels[31, 43]:

$$\frac{1}{|\mathbb{G}|} \int_{\mathbb{G}} d\mu(g_1, ..., g_m) \, \Phi(g_{\text{inv}}) \circ \Phi(g_m) \circ ... \circ \Phi(g_1) = \sum_{\sigma} A_{\sigma}(\lambda_{\sigma}, m) \tag{2.12}$$

where $\mu(g)$ is the Haar measure over the unitary group $\mathbb{G}$, A represent a depolarizing channel and λ is the corresponding decay rate, and the sum $\sum_{\sigma}$ is over all irreducible sub-representations of original representation U_g.

E. Magesan et al. brought a significant improvement to the RB protocol by formalizing the use of Clifford gates, demonstrating that the unitary group $\mathbb{G}$ could be replaced by groups made of Clifford gates, while still maintaining the Equation 2.12 under some minor assumptions. It's important to note that the use of Clifford gates was initiated by E. Knill et al[42], with the work of E. Magesan et al. serving to theoretically justify and ground this practice.

This improvement is significant because it makes the RB protocol more practical and scalable by avoiding the need for Haar-averaging over the unitary group, which is difficult and not scalable in practice. In particular, for a single qubit and under the assumption of gate-independent errors, the error process can be simplified to a single depolarizing channel using the Clifford group-based protocol.

For a better demonstration of RB protocol, let's consider a single qubit prepared in a pure state $\rho_0 = |\alpha\rangle\langle\alpha|$, and a set of POVM $\{E_i\}_{i=0,1,\pm,\pm i}$ induced by the Pauli eigenstates.

The probability of obtaining the measurement outcome E_i after applying a random gate sequence $\mathcal{S}^{m+1}$ can be expressed as:

$$P(m, E_i) = \text{Tr}\Big(E_i\big(\Phi(g_{\text{inv}}) \circ \Phi(g_m) \circ ... \circ \Phi(g_1) \circ \rho_0\big)\Big) \tag{2.13}$$

Next, we average $P_{|0\rangle}$ over different random gate sequences drawn from the Clifford group $\mathcal{C}$:

$$\overline{P}(m, E_i) = \frac{1}{|\mathcal{C}|} \int_{\mathcal{C}} d\mu(g_1, ..., g_m) \, E_i\big(\Phi(g_{\text{inv}}) \circ \Phi(g_m) \circ ... \circ \Phi(g_1) \circ \rho_0\big) \tag{2.14}$$

$$= \Big(\sum_{\sigma} A_{\sigma}(\lambda_{\sigma}, m)\Big) \tag{2.15}$$

where μ is the Haar measure on a Clifford group. The theory of RB shows that the probability of the outcome E_i follows a linear combination of exponential decays. For a single qubit, there are only two irreducible subrepresentations for the Pauli representations of Clifford gates. Furthermore, by taking the assumption that the quantum errors are much smaller than the gate fidelity and are independent of gates, the error process can be approximated to a single depolarizing channel. In this case, the averaged probability $\overline{P}(m, E_i)$ can be fitted with one simple exponential decay:

$$\overline{P}(m, E_i) = Ap^m + B \tag{2.16}$$

where A, B, p are fitting parameters, and m is the number of random gates in the gate sequence . The average gate fidelity $\mathcal{F}_{\text{avg}}$ and error rate ϵ are then

$$\mathcal{F}_{\text{avg}} = \frac{1}{2}p + \frac{1}{2} \tag{2.17}$$

$$\epsilon = 1 - \mathcal{F}_{\text{avg}} = \frac{1}{2} - \frac{1}{2}p \tag{2.18}$$

The error rate ϵ provides an average estimate of the total quantum errors per gate for the implementation. As it is calculated by fitting over various gate sequence lengths m, it is typically considered to be free of SPAM errors. To distinguish it with other error rates appearing later, we denote this total error rate with a subscript "$_{\text{total}}$".

2.2.1 Experimental Protocol

The goal of RB protocol is to experimentally track changes in the survival probability of a quantum state while applying increasingly long random gate sequences. To do this,

the protocol involves an initial state preparation ρ_0, applying a sequence of m randomly chosen gates and an inverse gate at the end of the sequence, and measuring the survival probability of the state. By repeating this process, the average survival probability should converge to an exponential decay as given by Equation 2.16. To ensure that the convergence is unbiased, the gate sequence should be generated uniformly at random from a Clifford group and cumulative gate set should span the entire group.

The preparation of the initial state for a transmon qubit is typically achieved through energy relaxation process in which the qubit is allowed to naturally lose energy and relax to the energy ground state $\rho_0 = |g\rangle\langle g|$. Because the thermal energy at the qubit operation temperature is much lower than the qubit transition energy, the qubit can be prepared in the ground state with high fidelity. In this case, it is natural to choose the energy ground state $|g\rangle$ and the first excited state $|e\rangle$ in the energy basis as computational basis. To measure the survival probability of the state, we use a set of POVMs $\{E_i\}_{i=\pm z, \pm x, \pm y}$ induced by the Pauli eigenstates. In particular, the survival probability is then given by the measurement probability of E_{-z}.

The RB protocol proceed as:

1. Determine an increasing sequence of distinct integers $\mathcal{M} = \{m_i\}_{i=1}^n$, with each m_i which will determine the length of the different gate sequences used in the protocol.

2. For each m_i in the sequence $\mathcal{M}$, generate N different randomized gate sequences $\mathcal{S}_{j=1,...,N}^{m_i+1} = [g_1, ..., g_{m_i}, g_{m_i+1}]$, where $g_1, ..., g_{m_i}$ are gates uniformly drawn at random from the 1-qubit Clifford group. The final gate g_{m_i+1} is computed classically so that the composition of this entire sequence is effectively equal to the identity, ensuring that the gate sequence is reversible.

3. For each gate sequence $\mathcal{S}_j^{m_i+1}$, experimentally apply the sequence and measure the qubit state survival probability:

$$P(\mathcal{S}_j^{m_i+1}, \Pi_{-z}) = \mathrm{Tr}(\Pi_{-z}\mathcal{S}_j^{m_i+1}(\rho_0)) \tag{2.19}$$

Then, average the survival probability P over N different sequences $\mathcal{S}_j^{m_i+1}$ with same m_i:

$$\overline{P}(m_i + 1) = \sum_{j=1}^{N} \frac{P(\mathcal{S}_j^{m_i+1}, \Pi_{-z})}{N} \tag{2.20}$$

4. Fit the average survival probability $\overline{P}$ to m_i using the Equation 2.16

The sequence $\mathcal{M}$ is determined based on the physical device being tested. Each m_i should be chosen so that $\overline{P}(m_i + 1)$ can show a clear exponential decay with minimal uncertainties. In this way, the fittings later would be more accurate. For example, a qubit with higher quality might require longer sequences, i.e., larger m_i to display a clear decay than the qubit with lower quality. The number of random sequences for each sequence length, k, needs to be sufficiently large, so that the random gates drawn for each m_i span the Clifford group. This would make the convergence less biased.

2.2.2   Purity Benchmarking

The coherence of the errors can be estimated by the unitarity u of the process, which is based on the averaged state purity change. The PB protocol is designed to experimentally estimate the unitarity of a gate operation.

As a modified RB protocol, the goal of purity benchmarking (PB) is to track changes in the purity of the state while applying increasingly long random gate sequences, instead of the survival probability. For a single transmon qubit, we choose the usual Pauli representation and the same POVM as in RB. The state of the qubit can be expressed by a vector $\vec{\alpha}$ within the Bloch sphere:

$$\rho = \frac{1}{2}(\mathbb{I} + \vec{\alpha} \cdot \vec{\sigma})$$

where $\vec{\sigma}$ is the unit Pauli vectors. The purity of the state can then be expressed by the squares of the 2-norm of α:

$$\mathrm{Purity}(\rho) = \mathrm{Tr}(\rho^2) = \frac{1}{2}(1 + |\alpha|^2)$$

Note that when we define the unitarity, an identity is subtracted from the purity, leading to a quantity $\mathcal{P}$ as:

$$\mathcal{P}(\rho) = |\alpha|^2 = \langle \sigma_x \rangle^2 + \langle \sigma_y \rangle^2 + \langle \sigma_z \rangle^2 \tag{2.21}$$

As justified by [39], the average of $\mathcal{P}$ is directly related to the unitarity by a exponential fit:

$$\overline{\mathcal{P}}(m) = A' u(\mathcal{E}_g)^m + B' \tag{2.22}$$

where A', B', u are fitting parameters, and m is the number of gates in the sequence. Fitted u is the averaged unitarity per gate. The incoherent and coherent errors for a single qubit are then respectively:

$$\epsilon_{\mathrm{inc}}(\mathcal{E}_g) = \frac{1}{2}(1 - \sqrt{u(\mathcal{E}_g)}) \tag{2.23}$$

21

$$\epsilon_{\text{coh}}(\mathcal{E}_g) = \epsilon_{\text{total}}(\mathcal{E}_g) - \epsilon_{\text{inc}}(\mathcal{E}_g) \tag{2.24}$$

To experimentally measure the quantity $\mathcal{P}$, we only need to modify the protocol for the RB[40]. Specifically, in RB protocol the measured survival probability is the measurement output of POVM element E_{-z} , since the initial state is just the (almost) pure "$-z$"state

$$\langle E_{-z} \rangle = \langle \frac{\mathbb{I} + \sigma_z}{2} \rangle = \frac{1}{2} + \frac{1}{2}\langle \sigma_z \rangle \tag{2.25}$$

For the quantity $\mathcal{P}$, we need more measurements for $\langle \sigma_x \rangle, \langle \sigma_y \rangle$. This could simply be done by repeating Step 3 of the RB another 2 times for every $\mathcal{S}_j^{m_i+1}$, with each time replace the observable σ_z with σ_x, σ_y. The quantity $\mathcal{P}$ for a specific gate sequence can be calculated with three measurements on observables $\langle \sigma_x \rangle, \langle \sigma_y \rangle, \langle \sigma_z \rangle$. We average $\mathcal{P}$ over N different random sequences with the same length. The averaged $\overline{\mathcal{P}}$ can then be fitted with Equation (2.22) to obtain the unitarity and hence, coherent and incoherent errors.

2.3 Experimental Protocol for Time Fluctuations of Quantum Errors

Quantum errors fluctuate both coherently and incoherently over time. To study the time fluctuations of coherent and incoherent errors, we propose a experiment protocol based on the original PB. This protocol allows us to monitor the fluctuations of both coherent and incoherent errors over an long period of time for multiple qubit transition frequencies using a single Xmon transmon qubit. The beginning part of the protocol follows standard RB. First, we ran a few standard RB and PB experiments to determine the proper sequence length set $\mathcal{M}$ for a physical qubit and the other devices. For example, Figure 2.1 shows the results of RB and PB experiments for the target qubit and instruments. Based on the decay curve, we choose $\mathcal{M} = \{2, 6, 13, 25, 50, 100, 200\}$. These sampling points in gate sequence lengths can adequately describe the exponential decays in average gate fidelity and unitarity. The choice of $\mathcal{M}$ is not unique, as the choice of points in the set affects both the accuracy and efficiency of the measurements. In general, a larger set or longer gate sequence lengths lead to better fittings but longer measurement times.

Next, instead of directly generating and measuring N random sequences for each length m_i, we measure a single random sequence for each length and iterate the process one after another. Each iteration is considered an execution of a PB sequence, and uses randomly generated gate sequences that are independent from those used in other PB sequence. The random Clifford gates are generated by algorithms described in [9, 44].
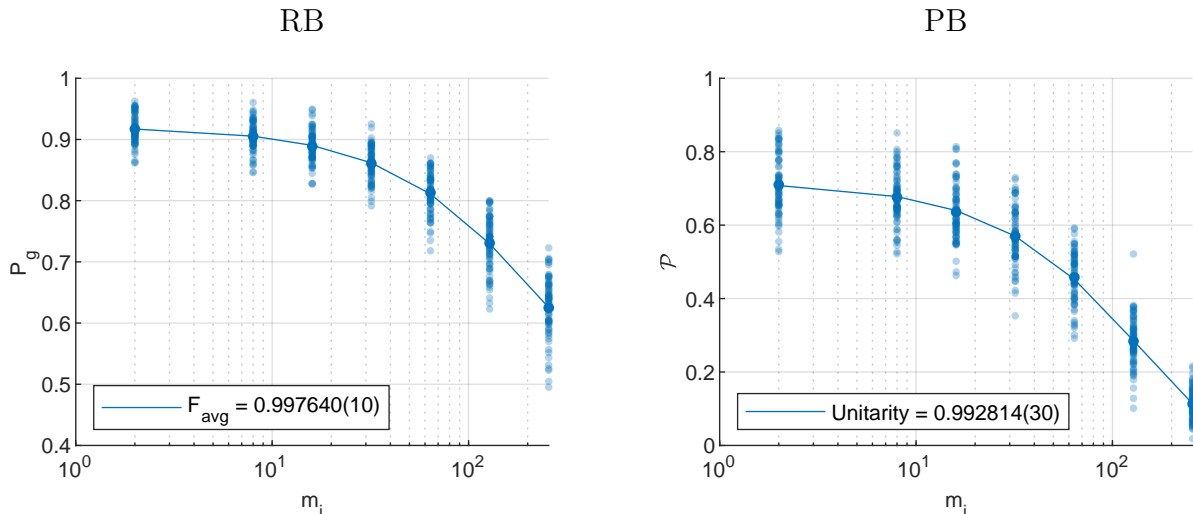
Figure 2.1: Single-qubit RB and PB: For each length m_i, 40 random sequences S^{m_i+1} are averaged. 7 sequence lengths have been chosen to reflect the decay in both survival probability and the quantity $\mathcal{P}$.

Since the Xmon transmon qubit transition frequency is tunable via flux bias, we can estimate the errors at different frequencies by executing one PB sequence per flux bias point. The overall execution of the PB sequences at different designated flux bias points is called a cycle. The cycles can be repeated for as long as desired. At the end of each cycle, a short time buffer is added to ensure that the duration of each cycle is consistent.

This protocol enables us to perform PB continuously over time, as a conventional PB measurement typically takes several minutes to complete. Diagrams of this proposed protocol is shown in Figure 2.2,2.3.

After completing the measurements, the data collected from the cycles is partitioned into segments of length N, with each segment representing a complete PB measurement. The measurement data for each segment is then processed using the usual PB protocol, which involves fitting the exponential decay of the average gate fidelity and unitarity to obtain the coherent and incoherent errors.

This proposed protocol also offers greater flexibility in data post-processing. Since we only measure a single random sequence for each length m_i during each cycle, we can easily adjust the number of sequences N collected for each length m_i to optimize the trade-off between measurement time and statistical uncertainty. Additionally, the use of cycle segmentation enables for the possibility of overlapping segments, which can provide more
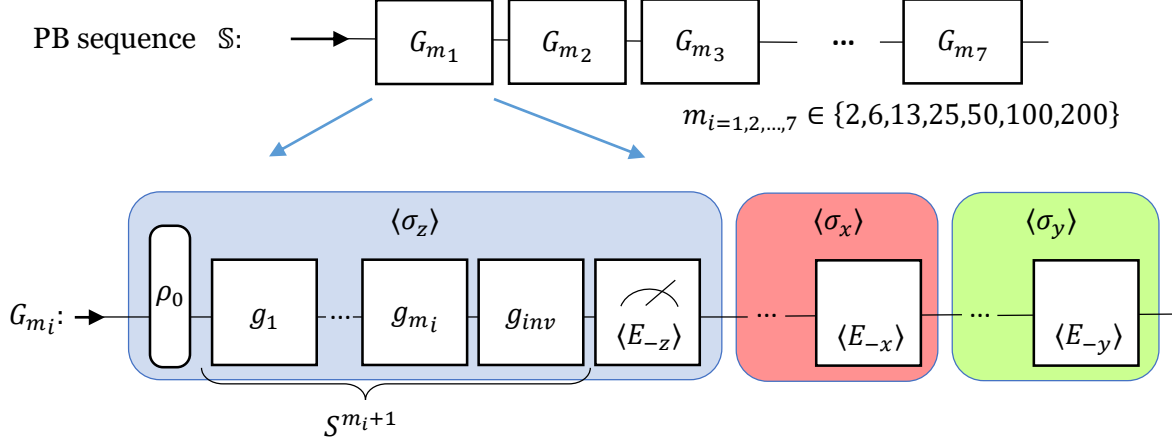
Figure 2.2: Diagram for PB Sequence: each PB sequence is composed of $|\mathcal{M}|$ sub-sequences G, with each sub-sequence G matched with a specific $m_i \in \mathcal{M}$. For each G_{m_i}, one random sequence $\mathcal{S}^{m_i+1}$ is generated. Then the random sequence is implemented 3 times. Each time we change the measurement basis and measure one of the observables $\langle E_{-z} \rangle, \langle E_{-x} \rangle, \langle E_{-y} \rangle$.

data points for analysis and increase the accuracy of the extracted error rates for lower frequency.

2.3.1 Elimination of Offset Term

The advancements in engineering and environmental noise suppression methods have greatly reduced the magnitude of quantum errors and fluctuations, which has made it increasingly challenging to study the temporal dynamics of these errors. Due to the random nature of the PB protocol, the estimation uncertainty of the measurement can often be so high that the effect of the noise channels is obscured by the fitting uncertainty, making the measured fluctuations unreliable. In order to improve this deficiency, various improvements have been proposed in the literature for RB and its variants[45, 46, 47]. Based on these improvements, the proposed protocol can be further optimized by a common way of eliminating the offsets in the fitting models, yielding more accurate error estimations.

Ideally, after applying a sufficiently long gate sequence the qubit state decoheres and converges to a maximally mixed state where $\langle \sigma_z \rangle = 0$. With this process the overall fidelity should be 0, which implies the offset term B in the Equation 2.16 being 0. However this
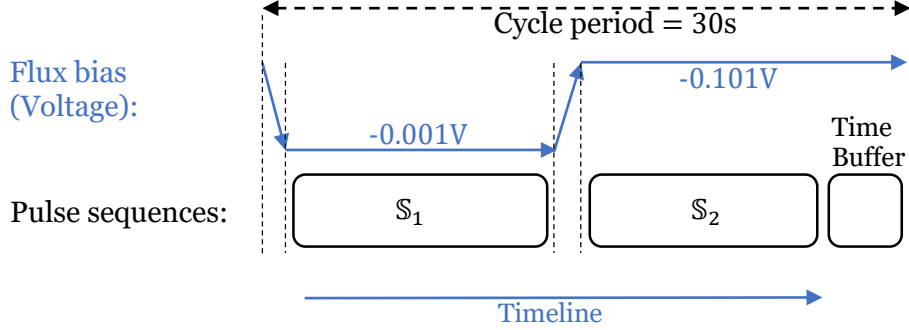
Figure 2.3: Cycle diagram: Each cycle consists of multiple runs of PB sequences, with each run executed at a different flux bias. One example is shown in the diagram. At first the flux bias voltage is set to $-0.001V$, which shifts the qubit transition frequency to a specific value. A PB sequence is then executed with the voltage held constant. The voltage then ramps to $-0.101V$, which shifts the qubit frequency to another value, and another PB sequence is executed. This process can be repeated for multiple bias voltages. At the end of each cycle, there is a time buffer to ensure the cycle period is exactly 30 seconds. The correlation between the flux bias voltage and qubit frequency is fixed. By iterating cycles on the qubit, we can monitor errors at different qubit frequencies over time.

isn't always true in practice. In fact, it is almost certain that there is a non-zero bias in the state's converging limit, which results in non-trivial offset term B in the exponential decay fitting. For a non-linear fitting algorithm, such as least square fitting, the third fitting parameter B in the fitting model can dramatically increase the fitting uncertainty. Hence, the goal of this optimization is to eliminate this fixed offset without the need to estimate its value.

Suppose the qubit is at the state of converging limit after applying a sufficient long gate sequence, the measurement outcome $\langle E_{-z} \rangle$ when the qubit is supposed to recover to $|-z\rangle$ should be the same as to recover to an orthogonal state $|+z\rangle$:

$$\langle E_{-z} \rangle_{-z} \approx \langle E_{-z} \rangle_{+z} = 1 - \langle E_{+z} \rangle_{+z} = B$$

for some offset value B. By the model given in Equation 2.16, we have

$$Ap^m + B = \langle E_{-z} \rangle_{-z} \approx 1 - \langle E_{+z} \rangle_{+z} = A'p^m + (1 - B) \tag{2.26}$$

for some parameters A, A'. We define a quantity $\bar{p}$ such as:

$$\bar{p} = \frac{\langle E_{-z} \rangle_{-z} + \langle E_{+z} \rangle_{+z}}{2} = \frac{A - A'}{2}p^m + \frac{1}{2} = A''p^m + \frac{1}{2} \tag{2.27}$$

25

Now by estimating the new quantity $\bar{p}$, we are able to eliminate the offset term from the exponential decay by fixing it to $\frac{1}{2}$. The average gate fidelity and total error can be calculated with the fitting parameter p following Equation 2.17,2.18 To experimentally measure the $\langle E_{+z} \rangle_{+z}$, we would just need to repeat the corresponding gate sequence $\mathcal{S}$ again, and composite an extra X rotation at the end with the inverse gate before the usual measurement.

A similar improvement for PB was proposed in Ref [47], where by introducing another quantity $\hat{b}$, the unitarity can be fitted by a simple exponential decay model without the offset term.

$$\hat{b}(m_i + 1) = \frac{1}{N} \sum_{j=1,\ldots,N} \sum_{i=x,y,z} (\langle \sigma_i \rangle^2_{\mathcal{S}_j^{m_i+1}} - \frac{1}{N} \sum_{j=1,\ldots,N} \langle \sigma_i \rangle^2_{\mathcal{S}_j^{m_i+1}}) \tag{2.28}$$

Here, "σ_i"-s represent the usual Pauli matrices. The expressions $\langle \sigma_i \rangle_{\mathcal{S}_j^{m_i+1}}$ denote the expectation value of the observable σ_i, after applying the gate sequence $\mathcal{S}_j^{m_i+1}$ to the qubit. The unitarity can be determined by fitting the $\hat{b}$ quantity to the simple decay model expressed as:

$$\hat{b}(m_i + 1) = A\,u^{m_i} \tag{2.29}$$

The expectation values of Pauli matrices can be experimentally determined by evaluating the POVM elements E_{-z}, E_{-x}, E_{-y}. Finally, we eliminate the offset terms from all fitting models used in the protocol.

In summary, by introducing new quantities $\bar{p}$ and $\hat{b}$, we can effectively eliminate the offset terms present in the decay models. To accommodate this change, the protocol of sub-sequence G in PB sequence (see Figure 2.2) needs to be modified to include a fourth component. In the fourth component, a X rotation is composed with the inverse gate at the end of the sequence to measure E_{+z}. In essence, this modification equates to measuring $E_{-z}, E_{+z}, E_{-x}, E_{-y}$ in a sub-sequence G.

Chapter 3

Result

To investigate the time fluctuations of coherent and incoherent errors in superconducting qubits, we conducted a series of PB experiments over extended time periods using a single superconducting Xmon transmon qubit. Throughout each experiment, we iterated PB cycles to acquire the time series data. Subsequently, we processed this data by averaging the time series within a moving window and fitting it to the models developed in Chapter 2. The error metrics were then calculated from these fitting parameters.

The measurements were done at two different qubit transition frequencies for all the experiments to account for frequency-dependent fluctuations. The selection of these frequencies was decided by both the visibility of the qubit readout and the spectral background of the qubit.

This chapter begins with one preliminary experiment, which is generally unsatisfactory in terms of the fitting uncertainty. This led to the introduction of a modified model that includes the elimination of offsets (Section 2.3.1). The results of the experiments performed by the modified protocol are then presented and discussed in Section 3.2.

3.1 Preambles

Our initial long-time experiment was conducted using the "standard" PB protocol, wherein the exponential decay fitting models (Equation 2.16 and 2.22) includes the offset terms B, B'. The exact protocol is described in Figure 2.2, 2.3. The results of coherent errors are presented as time series in Figure 3.1, with error bars denoting the uncertainties from fitting.
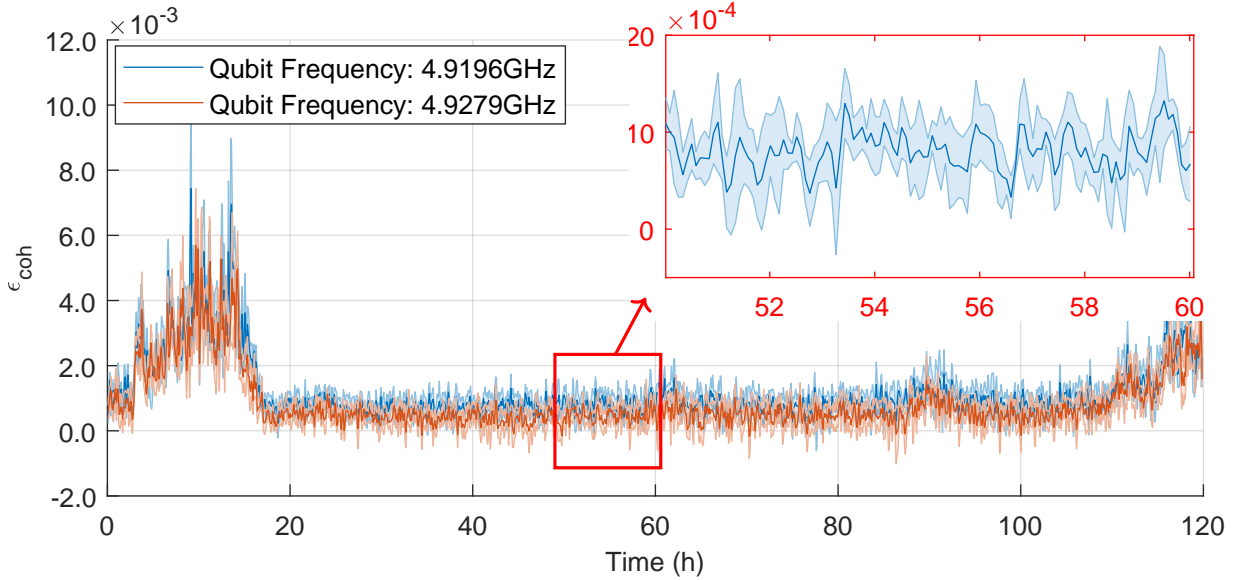
Figure 3.1: Time series of coherent error ϵ_{coh} acquired by "standard" PB protocol.

The time series for both frequencies already show intriguing patterns, including several notable telegraphic shifts. Initially, the coherent error is large and fluctuates significantly. Just before the 20-hour mark, a telegraphic shift occurs, causing the coherent error to drop and stabilize. Some minor shifts are observed at approximately 60 and 90 hours. Toward the end of the experiment, the coherent error begins to increase, and another telegraphic shift, similar to the initial one, is imminent. This substantial shift pattern appears to be recurring. It would be interesting to further explore the physics behind these shifts. In addition, the time series patterns for both frequencies highly coincides, indicating the sources of the coherent error is mostly frequency-independent.

Before examining the data in more depth, we noticed that the error bars for the time series of the coherent error are relatively large in comparison to the amplitude of the actual fluctuations. This discrepancy could render the fluctuations almost indistinguishable from the fitting uncertainties. The situation worsens for the coherent error due to uncertainty propagation, which can be expressed as:

$$s_e(\epsilon_{\text{coh}}) = \sqrt{(s_e(\epsilon))^2 + (s_e(\epsilon_{\text{inc}}))^2} \tag{3.1}$$

This equation is derived from the relation in Equation 2.24. $s_e(\cdot)$ represents the fitting uncertainty in terms of standard error.

28

To evaluate the clarity, or the "visibility", of the true fluctuations, we approximate the concept of the "signal-to-noise" ratio (SNR) and define the following quantity "SNR" for the fitted time series as:

$$\text{"}SNR\text{"} := \frac{\sigma(\varepsilon)}{\langle s_e(\varepsilon) \rangle} \tag{3.2}$$

where $\sigma(\varepsilon)$ is the standard deviation of ε, and $\langle s_e(\varepsilon) \rangle$ is the averaged standard error of the least square fitting for ε. The "SNR" for the coherent error time series in Figure 3.1 is approximately 1.1, indicating the fitting uncertainties are almost equal to the actual noise fluctuation. At this stage, the true qubit noises are primarily masked by these fitting uncertainties. Additional experiments also yielded similar "SNR"; which raises concerns about the reliability of these measurements.

By averaging more cycles, we may potentially reduce the fitting uncertainties and thereby increase the "SNR": However, this approach is not universally beneficial. As we increase the number of cycles averaged for each data point, the sampling frequency of the time series decreases, leading to a bigger loss of high-frequency information in the data. Furthermore, the "SNR" may remain unchanged or even worsen due to distortions caused by strong, fast-oscillating noises.
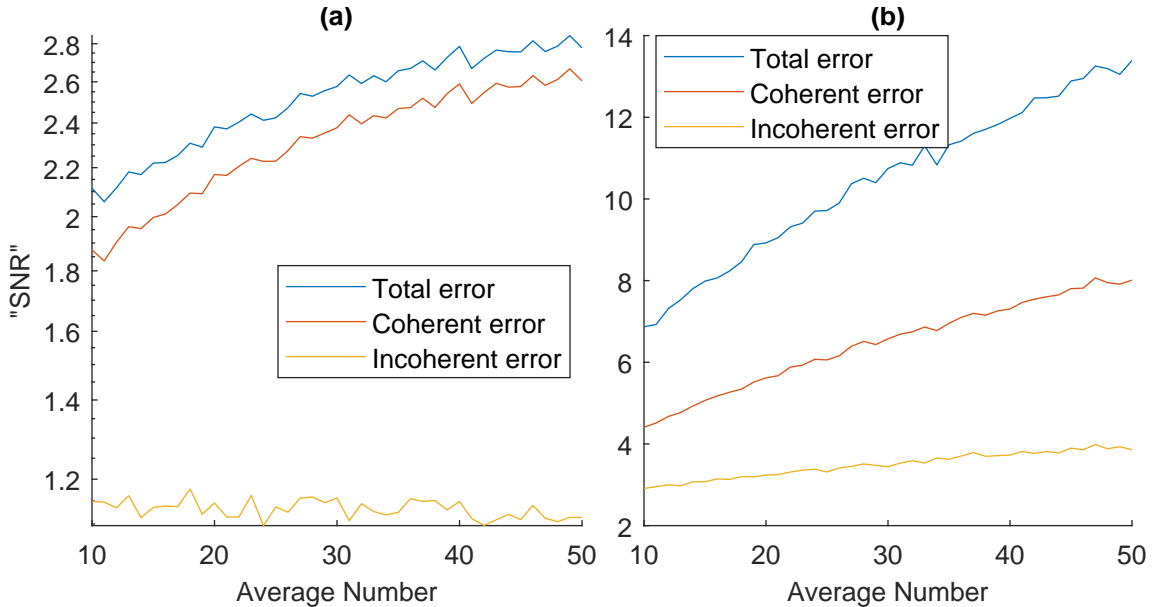

Figure 3.2: Impact of averaging more cycles on the "SNR".

29

Figure 3.2(a) plots the "SNR" with respect to the number of cycles averaged in the previous experiment. We see the "SNR" slightly increased for total and coherent error, and barely moved for incoherent error. Despite the slight improvements in total and coherent error by averaging more cycles, the uncertainties still remain significant, especially for the incoherent errors in this experiment. This indicates a need for further improvements to the experimental protocol to adequately reduce these measurement uncertainties.

To reduce the fitting uncertainties, we tested the modified protocol described in Section 2.3.1. Figure 3.2(b) shows the improved "SNR" achieved with this new protocol. In comparison to the previous protocol, the new models significantly increase the "SNR": Notably, the lowest "SNR" nearly triples, compensating for only a 67% increase in the sampling period (from 30s to 50s). In order to achieve a balance between minimizing fitting uncertainties and maintaining reasonable sampling frequencies, we have adopted the modified protocol. Furthermore, to increase the number of data points in the time series while ensuring the random gate sequences are unbiased, we use a window size of 20 cycles with a 50% overlap for the remaining analysis.

3.2 Fluctuations in Single-qubit Errors

Using the protocol outlined in Section 2.3.1, we performed two experiments spanning over 440 hours in total. Each experiment began with a precise calibration of the read out and gate control parameters. These parameters were then maintained consistently throughout the experiment. For the multi-frequency measurements, both experiments used the same set of flux bias settings. To some extent, the second experiment can be seen as a continuation of the first one after a system calibration. However, it's important to note that the two datasets cannot be merged into a single experiment due to the time interruption and system re-calibration.

Before performing the PB experiments, we characterized the qubit device by measuring T_1 and T_2 over a range of qubit frequencies to study the qubit decoherence profile. First, we measured the qubit frequencies in the presence of flux bias by applying a DC bias voltage to the SQUID, which shifted the qubit frequency. We scanned multiple bias voltages and measured the frequency shift using Ramsey measurements. Figure 3.3(a) shows the qubit frequencies at different flux bias voltages. The measured frequency responses fit well with the theory presented in Section 1.1.2. Finally, we measured the qubit T_1 and T_2 at each bias voltage, as shown in Figure 3.3(b).

In Figure 3.3(b), the spectral lines of both T_1 and T_2 exhibit a clear "dip" at approximately 4.598GHz (0.11V flux bias), indicating the qubit decoheres significantly faster
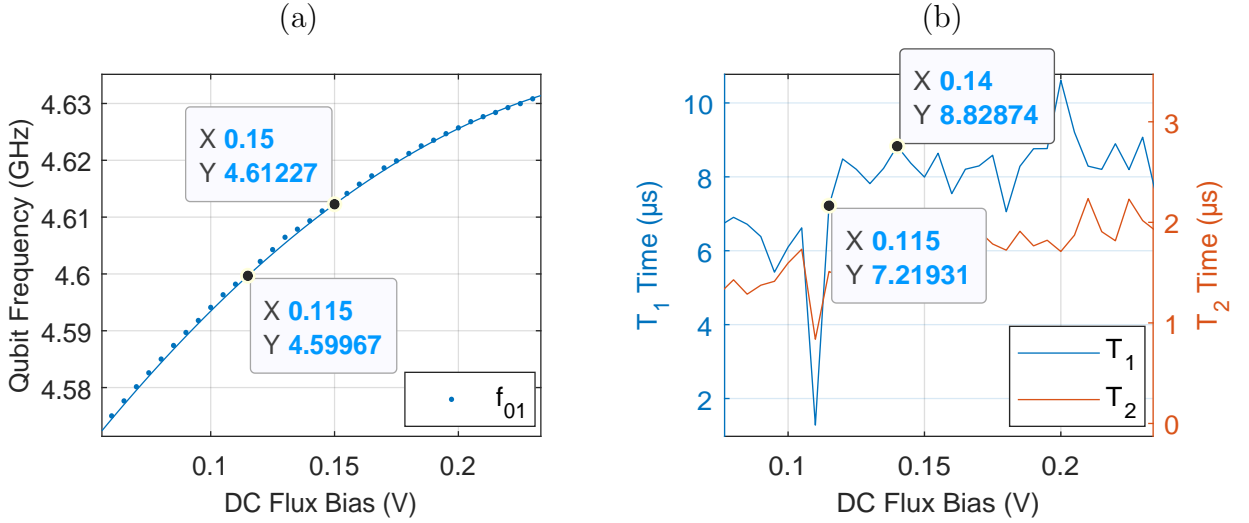
30

Figure 3.3: Qubit calibrations: (a) Qubit frequencies at different flux bias voltages. Dots represent the measured qubit frequency at each flux bias voltage, and the line is the fit according to the Equation 1.19 and 1.21. The tips mark the bias voltages used in the PB experiments. (b) T_1 and T_2 at different bias voltages.

around this frequency. This phenomenon can be attributed to a frequency-dependent coupling to the environment, which is often due to interactions with the TLS. Later swap spectroscopy experiments have supported this assumption.

To investigate the frequency dependency of the qubit noise, particularly in relation to the impact of TLS coupling, we selected 0.115V and 0.14V as two bias voltages for the PB experiments. The first bias corresponds to a qubit frequency that falls directly next to the bandwidth of the coupling (i.e., close-resonant coupling), while the second bias is placed in a more moderate region, further detuned from the TLS.

3.2.1 Time Series

The time series of both experiments are shown in Figure 3.4 and Figure 3.5. The first experiment was prematurely interrupted due to an unforeseen equipment malfunction, resulting in only about 202 hours of data collection. Fortunately, this incident did not compromise the validity of the data gathered before and after the malfunction.

Due to the qubit fluctuations and the calibration performed before the second experiment, the qubit drive frequencies changed slightly for the same two flux biases. Moreover,
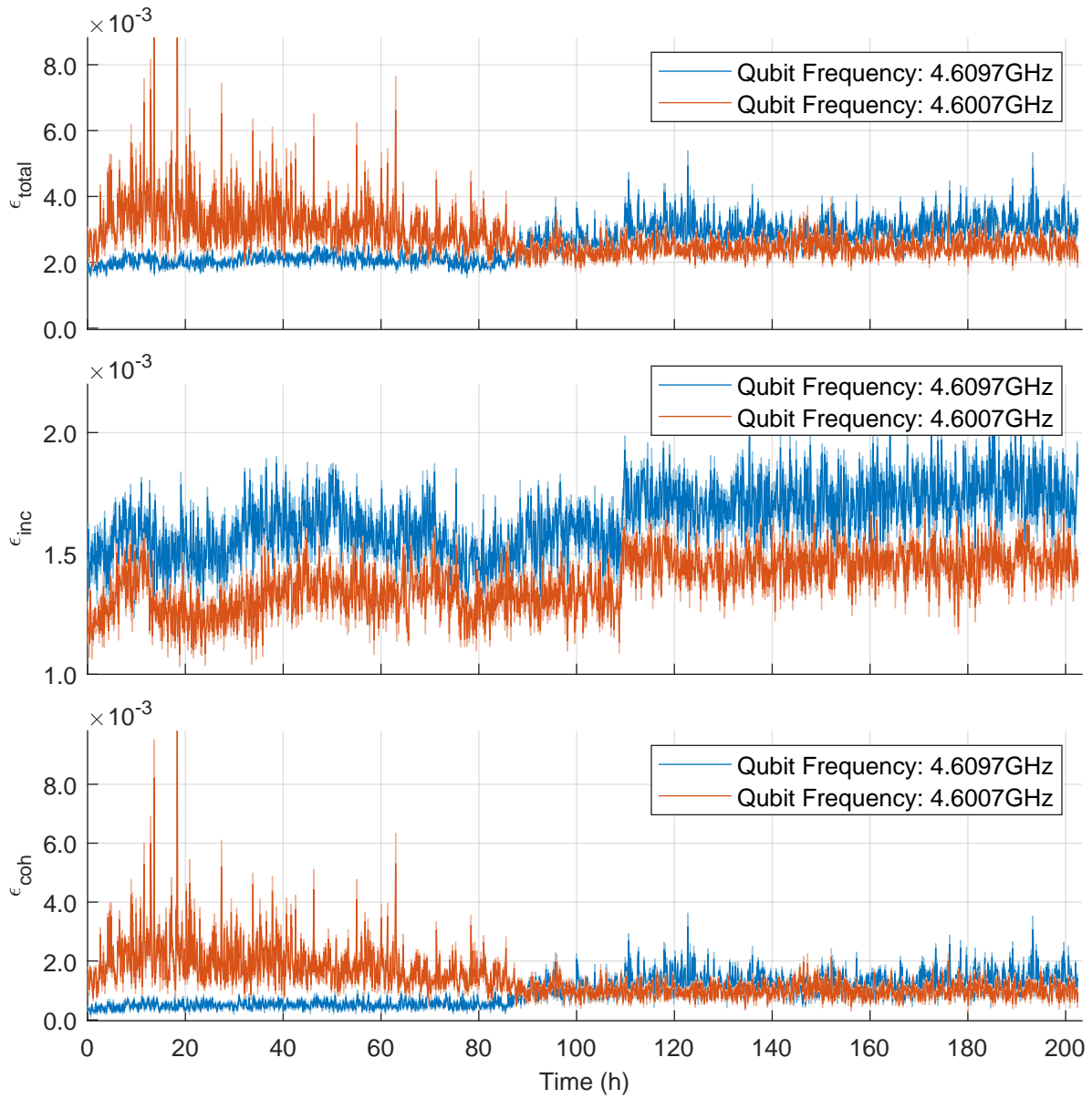
Figure 3.4: Time series for the first experiment. The red lines correspond to bias voltage at 0.115V. The blue lines correspond to 0.14V.
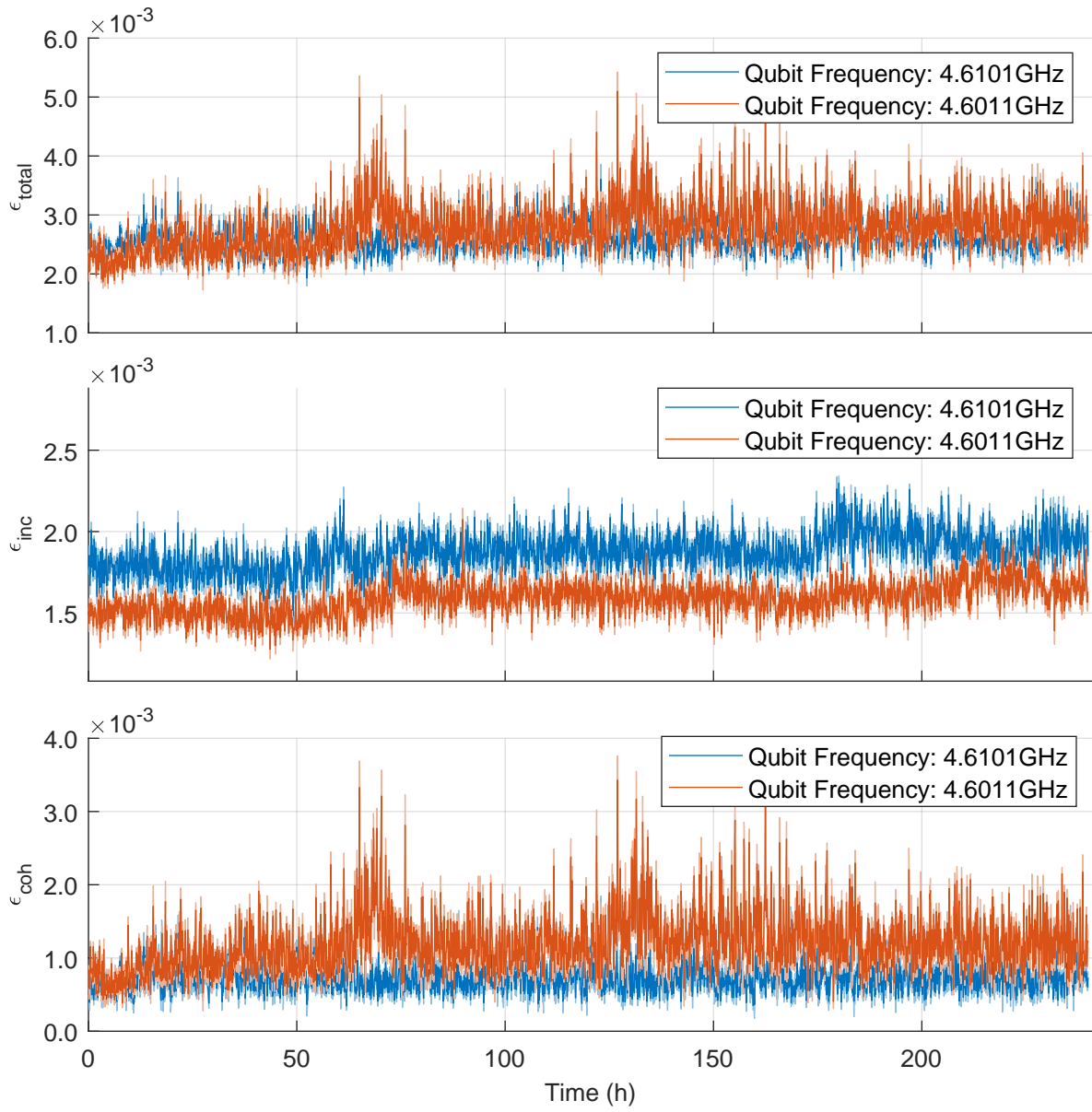
Figure 3.5: Time series for the second experiment. The red lines correspond to bias voltage at 0.115V. The blue lines correspond to 0.14V.

the calibration modified the parameters for both driving and readout pulses. Therefore, the two time series cannot be directly merged to generate a longer time series. Nevertheless, analyzing both time series is still insightful, as they are both subject to similar noise channels.

Interestingly, we find that the fluctuating patterns of coherent error closely resemble those of total error in both experiments. This suggests that the noises in total error might be primarily driven by coherent noise channels. In other words, the coherent error, typically resulting from systematic control errors like imperfect gate rotations or leakage[37, 38, 48], appears to be dominant factors in qubit fidelity fluctuations.

Although incoherent error still contributes considerably to the total errors, its fluctuation patterns differ from coherent and total errors, and the correlations between incoherent and coherent errors are very low. At first glance, incoherent error fluctuates relatively mild and exhibit less magnitude spread. However, there are also some arguably telegraphic shifts. For example, there are jumps in incoherent error at approximately 110 hours for both frequencies in the first experiment. Interestingly, both the coherent and total error seem to remain unaffected by the jumps. Conversely, in the first experiment (Figure 3.4), there is a clear telegraphic shift at approximately 90 hours in both total error and coherent error, but incoherent errors seem to remain unaffected. Further exploration is needed to understand the correlations between fluctuations in incoherent error, coherent noise, and qubit fidelity.

In our investigation of frequency-dependent qubit noise, we selected the frequencies represented by the red lines in Figure 3.4 and 3.5 to be closer to the TLS in the frequency domain. This selection was based on the expectation that enhanced TLS coupling for these frequencies would result in more errors and greater fluctuations. TLS is known to be a source of incoherent error. It's worth noting that the TLS coupling can also shift the effective qubit frequency, leading to imprecise DRAG pulses and leakage which contributes to coherent error.

However, the outcomes of both experiments reveal more complex patterns. During the initial part of the first experiment, the red lines in total and coherent error demonstrated notably higher amplitude and greater instability compared to the blue ones, aligning with the intuition that stronger TLS coupling induces more errors. However, the incoherent errors behaved contrarily to this expectation, with the blue line amplitudes consistently surpassing those of the red.

Interestingly, the noise pattern in the incoherent error seemed independent of the two frequencies at which we measured the qubit, which contrasts with its coherent counterpart. Further combining the fact that the telegraphic shifts in both errors are not well

synchronized, these features provide a key insight into the origins of both coherent and incoherent errors. In the second experiment, these patterns persisted. The red lines in total and coherent still show much more fluctuations compared to the blue lines, while the contradiction about the incoherent errors also persists, showing some consistency of both time series.

In order to check for any systematic drifting that may have occurred during the experiments, we re-measured T_1 and T_2 of the qubit for the range of qubit frequency again after the second experiment. The results are shown in Figure 3.6. The plot reveals a shift of approximately 16MHz in both the "dips"of T_1 and T_2, indicating a stochastic frequency shift in the TLS. After this frequency shift, both bias points used in the experiments were positioned significantly away from the "dip", and both had similar T_1 and T_2 values. Unfortunately, the experiments lack direct evidence indicating the precise time and impact of this shift. Further improvements in the experimental protocol could help capture these pivotal moments in future studies. For instance, a T_1 scan can be incorporated into each PB cycle to monitor the movement of the TLS frequency. Furthermore, a Ramsey measurement could also be integrated into each PB cycle to track the flux noises. Additional investigations on the TLS noises and qubit frequency fluctuations will be addressed later in Chapter 4.

3.2.2 Time Drift

Long-time noise measurements requires a stable clock reference for sampling. Given that our experimental protocol is structured such that all measurements are conducted in a series, a single deviation in the clock reference could impact all subsequent measurements, leading to a global time drift. Furthermore, many statistical analysis methods depend on equally spaced samples, including time domain variances and frequency spectrum analyses. Therefore, maintaining a stable clock reference is critical for the forthcoming analysis.

To maintain a consistent sampling rate, we time each execution of the PB cycle and integrated a time buffer at the end of each cycle to offset any time variations. The timer takes reference from the wall-clock time of the computer, and the computer clock synchronizes with a network time protocol (NTP) server[49] every four hours. These measures helped us maintain the period time coefficient of variation below 0.1% over a span of more than 240 hours.

However, several unexpected drifts in the clock reference still occurred during the first experiments. Figure 3.7 presents the time period recorded for each cycle in both experiments. While the period of each cycle remained stable for the second experiment, the
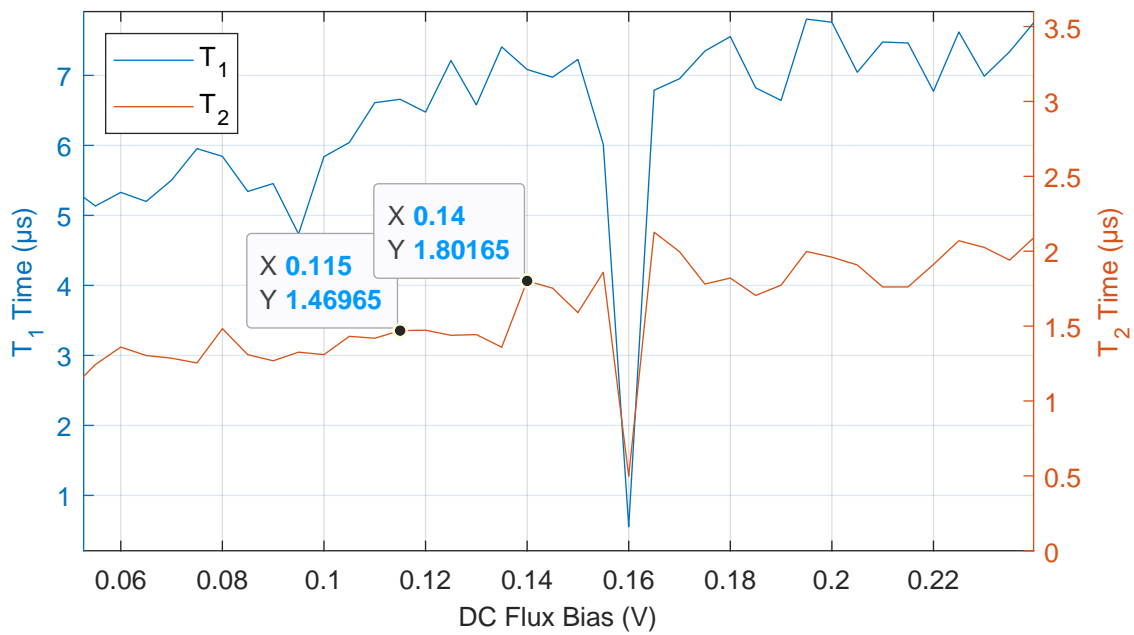
Figure 3.6: Qubit calibration after the second PB experiment: T_1 and T_2 at different bias voltages. The tips mark the bias voltages used in the PB experiments.
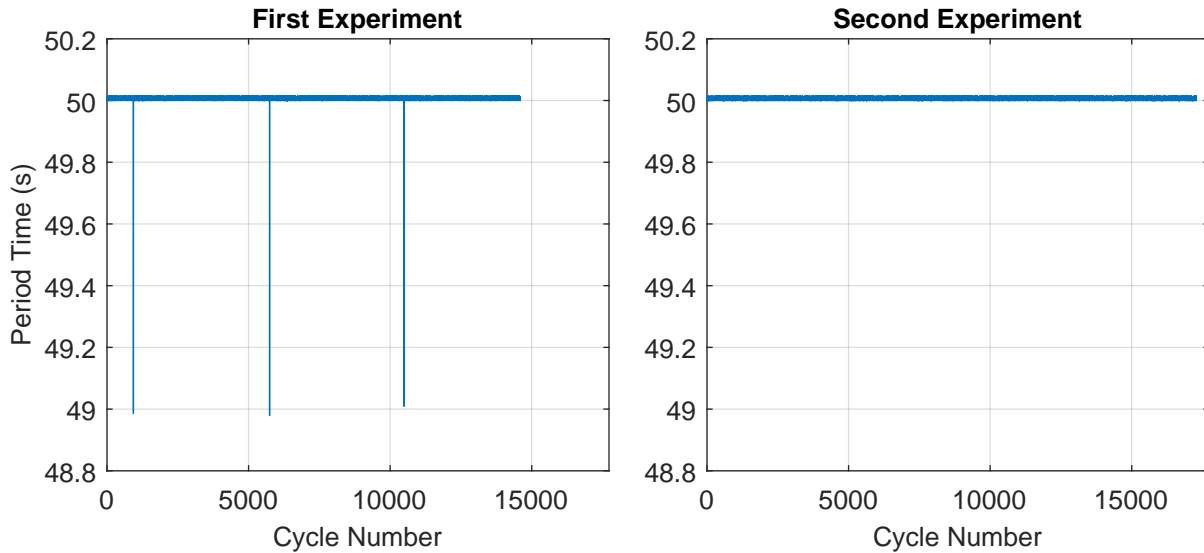
Figure 3.7: Period time for each cycle in both experiments. While the most of cycles have a consistent 50s sampling period, there are three outliers with 1 second shorter in the first experiment.

first experiment exhibited three outliers, each ending about one second early. Each outlier caused the clock reference to drift by roughly one second. Unfortunately, we don't have sufficient information and indicators to identify the cause of these outliers.

Prior to proceeding with statistical analysis, we conducted simulations with different power-law noise processes to examine the influence of such time drifts. Figure 3.8 shows the Allan deviations for simulated white noise and random walk noise, both with and without the exact time drifts as shown in Figure 3.7. The fact that, in both power-law models, the time drifts did not significantly impact the Allan deviations reinforces our confidence in the subsequent analysis. Nevertheless, improving the stability of the clock reference remains a desirable goal for future experiments.

3.2.3 Statistical Analysis

Following the visual validation of the time series, we proceeded to analyze the frequency components of the noise. Our focus is mainly on the low-frequency noises inherent in both coherent and incoherent errors. For both experiments, we set the cycle time to 50 seconds. After applying the moving-window averaging, the final base sampling period
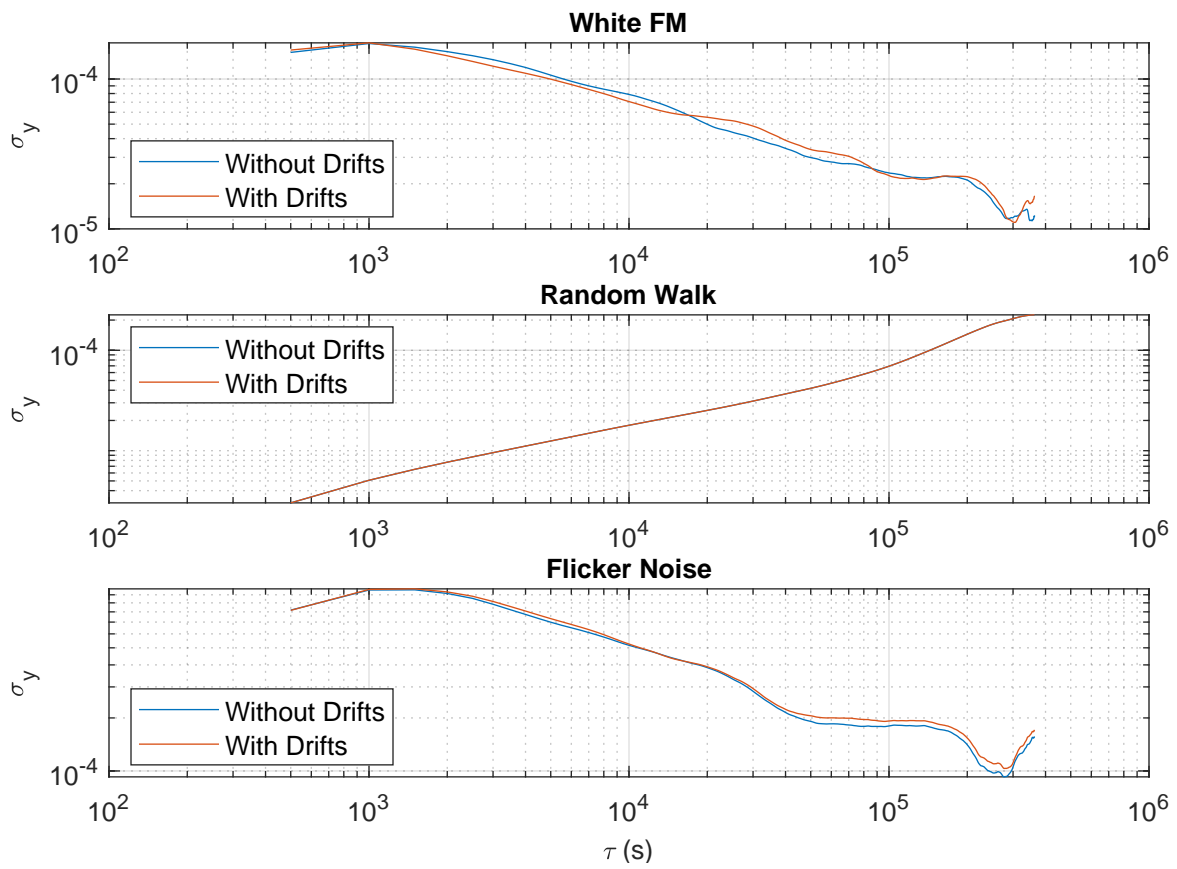
Figure 3.8: Simulations for time drifts.

for frequency domain analysis becomes $50s * 20$ cycles $* 50\%$ overlapping$= 500s$, which corresponds to a sampling frequency as low as 2 milliHz. The 50% percent overlapping is introduced to double the number of available sample points for later analysis. Although this approach induces statistical dependency between neighboring points, thus compromising the reliability of high-frequency analysis, it does not significantly impact our primary focus: the study of low-frequency noise. In fact, we visually confirmed that, with a 50 percent overlap, fluctuations with frequencies lower than $5 * 10^{-4}$Hz converged satisfactorily.

Starting with the time series, the most direct approach to studying noise in the frequency domain is through power spectral density (PSD). However, it is well-known that PSD is less ideal for analyzing low-frequency noises. One reason is that the PSD spectrum has a relatively low resolution in the low-frequency range due to the linear frequency division in discrete Fourier transform (DFT). Achieving a reasonable resolution for low-frequency noise would often require impractically long time series. Although other research suggests potential improvements using the logarithmic Fourier transform (LFT) [50], the implementation of this method would require extensive discussions on statistical modifications in the PSD algorithm, which is beyond the scope of this thesis.

Another challenge of using the PSD arises from the extraordinarily low Nyquist frequency of the time series in this experiment. For instance, the Nyquist frequency of the time series in Fig 3.4 is $1m$Hz, equivalent to a period of 1000 seconds. In the meanwhile, the potential noises in this experiment could span a wide frequency range. This could result in significant spectral leakage as well as aliasing in the spectrum[51, 52]. An example of a PSD for a time series can be found in Figure A.1. The spectrum is less insightful and provides minimal useful information.

Instead of the PSD, we turn to methods from time-domain analysis, computing the overlapping Allan deviation (ADEV) for all time series presented in Section 3.2.1. The (overlapping) Allan deviation is a powerful and common tool for identifying noise processes, especially those occurring over extended timescales at low frequencies. A brief introduction to the Allan deviation is presented in Appendix A.

Similar to the power spectral density, certain power-law noise processes exhibit linear patterns in the log-space plot of the Allan deviation. Each type of these noise models produces a distinct line with a characteristic slope. As a result, we can categorize the dominant noise model at different frequencies based on the slope of the curve. This provides an intuitive and useful tool for examining noise trends. Table 3.1 lists some common power-law noise processes and their associated ADEV curve patterns and slopes in log-space[20, 53].

Figures 3.9 and 3.10 display the overlapping Allan deviations for all the error time series

Noise Type	Spectral Form	ADEV Form	ADEV Slope in Log-Space
White FM	$h_0 f^0$	$\left(\frac{h_0}{2}\right)^{\frac{1}{2}} \tau^{-\frac{1}{2}}$	-0.5
Flicker FM	$h_{-1} f^{-1}$	$(2ln(2)h_{-1})^{\frac{1}{2}} \tau^0$	0
Random Walk FM	$h_{-2} f^{-2}$	$\frac{(2\pi)^2}{6} h_{-2}^{\frac{1}{2}} \tau^{\frac{1}{2}}$	0.5
Linear Drift	...	...	1

Table 3.1: Power-law noise processes and their associated ADEV patterns in log-space.

from Figure 3.4 and 3.5, respectively. The plots have been truncated at time intervals greater than 10^4 seconds ($\tau > 10^4 s$) to avoid convergence issues at higher frequencies. Additional reference lines with distinct slopes, corresponding to relevant power-law noise processes, have been added to each plot, aiding in the analysis of the dominant noise trend.

The ADEV patterns exhibited in the first experiment align well with power-law processes, characteristically forming "well"-like patterns. At higher frequencies, noise is dominated by a 1/f process. As frequency decreases (or as τ increases), the curves for the coherent and total errors exhibit an interesting evolution. Initially, they ascend and gradually adopt patterns resembling a frequency linear drift process. Then, at the lowest frequency (the right-most part of the plots), the noise curves flatten, revealing a divergence from qubit frequencies.

Interestingly, for the second flux bias setting, the noise curves flatten both earlier and more rapidly, displaying clear patterns of a random walk at the end. In contrast, the first bias setting's noise curves show only modest trends towards flattening. This discrepancy could stem from the heating effect, causing long-time frequency linear drifts. Specifically, the voltage setting for the flux bias 1 is approximately 22% higher than that for flux bias 2, translating into a roughly 48% higher input power to the qubit. This extra power ultimately dissipates as heat, potentially amplifying the linear drifts. Simultaneously, the incoherent patterns for both flux biases are similar, a conclusion also supported by time series analysis.

In the second experiment, the noise patterns display less complexity. The majority of ADEVs can be fitted with flicker noises, except for the coherent noises for flux bias 1 which are predominantly white noises. Interestingly, the coherent noise for flux bias 1 and the incoherent noise for bias 2 manifest small, hill-like bumps. These patterns are noteworthy since they don't align with power-law processes. Such bumps could signal a Lorentzian noise pattern, a phenomenon often associated with the interaction with TLSs[20, 54].
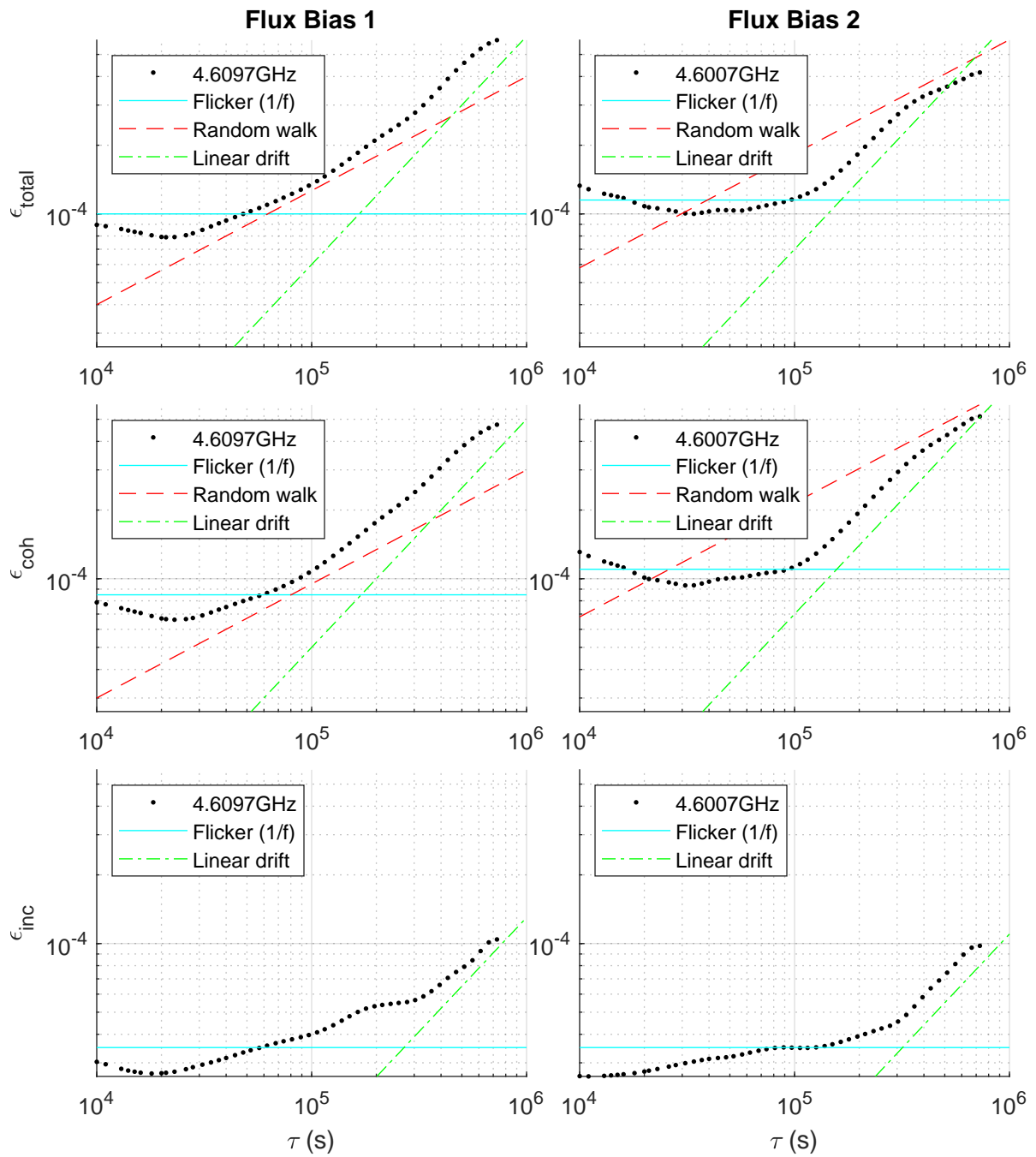
Figure 3.9: Overlapping Allan deviations for the first experiment.
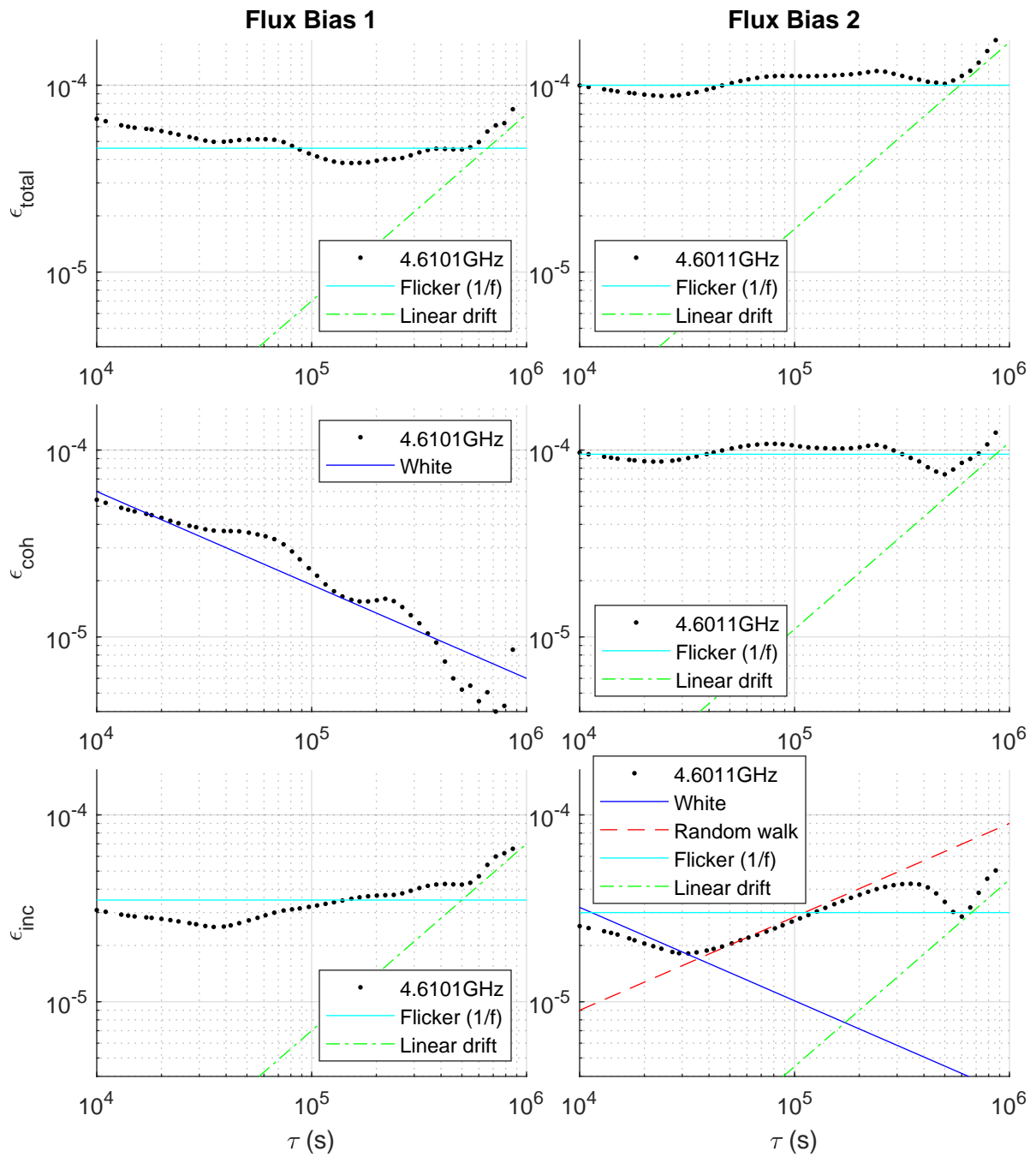
Figure 3.10: Overlapping Allan deviations for the second experiment.

Overall, these experiments offer valuable insights into the behaviors of coherent and incoherent qubit noise under different experimental conditions. The presence of a stochastic fluctuating TLS adds intriguing dynamics to the system. For a more comprehensive understanding of the qubit noise process involved in these experiments, it would be beneficial to characterize the qubit noise using multiple metrics simultaneously, such as qubit energy relaxation time T_1 and frequency shift. We will explore such studies in the subsequent chapter, offering a deeper dive into these intriguing phenomena.

Chapter 4

Discussion

4.1 Simultaneous Qubit Relaxation Measurements

In order to enhance the characterization of this quantum system, we sought to monitor additional metrics, particularly those closely related to TLS fluctuations. Thus, we incorporated method for monitoring qubit relaxation time T_1 into our PB protocol, drawing on well-established research such as References [21, 20].

To implement these modifications,we first select 11 evenly spaced qubit frequency values. These include the two frequency values at which the PB cycles are executed. We convert these frequency values into flux bias voltages based on the results of qubit calibration experiments. These flux bias voltages are then fixed for the duration of the experiment.

We make an assumption that the amount of external flux generated by a given bias voltage remains consistent and stable throughout the experiment. This assumption is reasonable given that the qubit and the external flux lines are heavily isolated, and the bias voltages are supplied by a dedicated, isolated voltage source. This ensures a stable and consistent external flux throughout the experiment.

For each cycle, we sweep the bias voltage across the selected 11 values. At each of these bias voltages, the T_1 time of the qubit is measured. Once the T_1 measurements for all 11 bias values are completed, the PB measurements proceed.

As mentioned in Chapter 3, a telegraphic shift in TLS frequency may be reflected as a shift in the "dip"of the T_1 spectrum. Therefore, this new modified protocol provides additional (semi-)simultaneous measurements that can be used to characterize the TLS coupling and flux noise, along with the PB measurements.
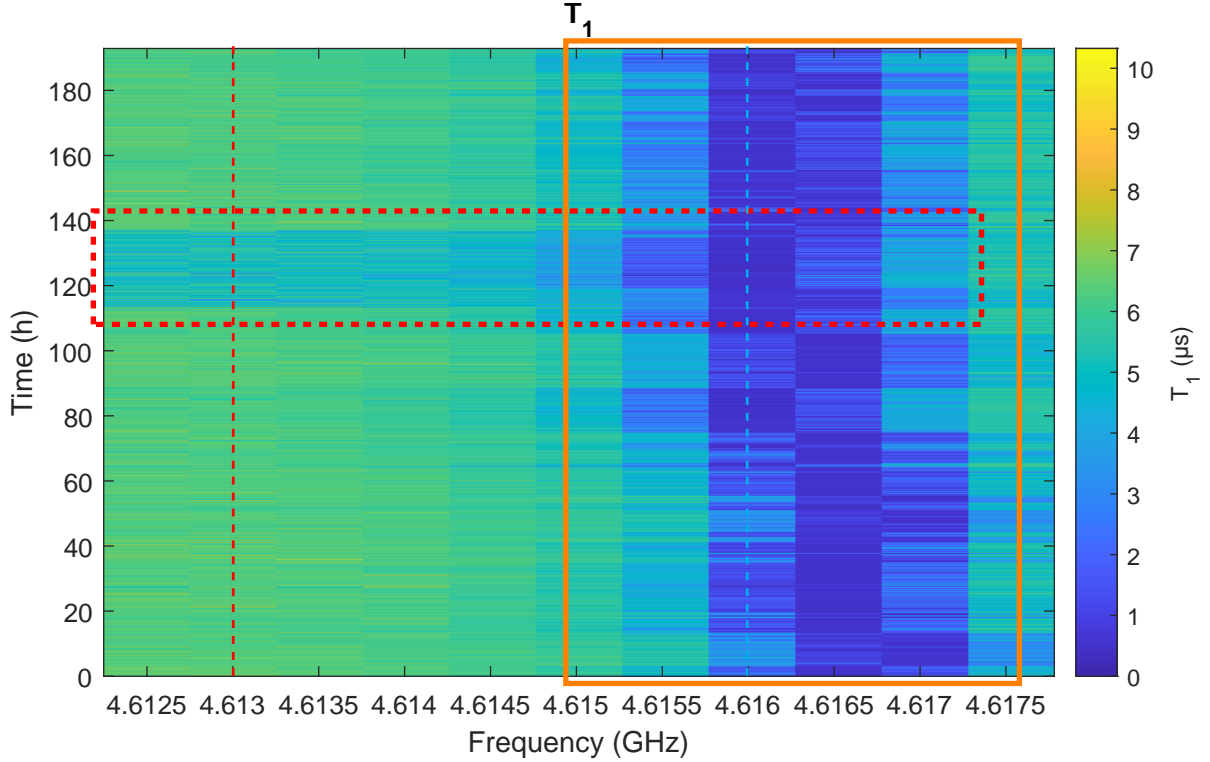
Figure 4.1: Color map for T_1 fluctuations. The x-axis represents the qubit center frequencies. The y-axis indicates the experiment time. The colors represent the T_1 time of the qubit. Two dashed vertical lines mark the frequencies where PB measurements were conducted. Spectral-diffusion patterns in the experiment are highlighted with boxes.

Following the enhancement of our protocol, we executed another 192-hour time fluctuation experiment. The results are shown in Figure 4.1, 4.2 and 4.3. Enhanced characterization comes at the cost of an decreased repetition rate. By incorporating (semi-)simultaneous T_1 measurements, we see the repetition period of PB cycles increase nearly four-fold, resulting in a cycle period of 180s. Given this extensive sampling period, performing statistical analyses similar to those in Section 3.2.3 may no longer be meaningful. Nevertheless, the time series still offer valuable insights into the fluctuations of the qubit.

Figure 4.1 presents a color map depicting a spectrotemporal representation of T_1 versus qubit frequency and time. A visual examination reveals two distinctive stochastic spectral-diffusion patterns. From approximately 110 hours to 140 hours, we observe a slow, wide-band telegraphic pattern, which is highlighted by a dashed red box. Another prominent

pattern, highlighted by the solid orange box, resembles a blend of band-limited diffusive and fast telegraphic patterns. These patterns can be qualitatively explained by the presence of the TLS, as suggested by a more comprehensive study conducted by our group[21].

PB measurements were undertaken at two specific bias voltages, namely the 2nd and 8th out of the total 11 selected points. The corresponding qubit frequencies are approximately indicated in the T_1 color map by two dashed vertical lines. It's crucial to note, however, that these lines correspond to the selected bias voltage points and may not represent the exact qubit frequencies. This distinction arises due to potential fluctuations in the qubit frequencies, suggesting that the frequencies employed for the T_1 and PB measurements may not be perfectly align. As a result, we use the bias voltage as our reference point instead of the qubit frequency. The time series of the coherent and incoherent errors, as well as the T1 fluctuation at corresponding bias voltages, for both frequencies are depicted in Figure 4.2 and 4.3.

The first PB frequency situates within the area of combined band-limited diffusive and fast telegraphic spectral patterns, placing it closer to the potential TLS frequency (indicated by the "dip"). As a result, both error types for this first PB frequency exhibit stronger fluctuations compared to those at the second frequency. Additionally, the overall level of the coherent error for the first PB frequency consistently exceeds that for the second frequency throughout the experiment. Interestingly, the overall levels of the incoherent error for both frequencies remain virtually the same.

Another noteworthy observation is that, after approximately 70 hours, the center frequency of the "dip" in the T_1 spectrum begins to slowly shift towards both PB frequencies. Simultaneously, we observe gradual increases in the coherent error at both PB frequencies. Given that the overall level and the intensity of fluctuations of the coherent error for the first PB frequency are higher, these observed behaviors of the coherent errors could be qualitatively explained by the coupling with the TLS.

Finally, we turn our attention to the most notable event observed in this experiment: the slow, wide-band telegraphic event. From approximately 110 hours to 135 hours, all time series plots exhibit a synchronized telegraphic shift. During this window, both the T_1 time at each PB frequency and all error measures alter in a telegraphic manner, reflecting a distinctive event within the system. This event is captured not only in the metrics for both coherent and incoherent errors but also in the T_1 relaxation time. Also notably, this is the only instance where a telegraphic shift occurs in the incoherent errors for both frequencies.

In conclusion, the time series analysis conducted in this additional experiment offers partial support for the argument that the interaction with TLS qualitatively contributes to quantum errors. The impact on coherent errors is clear from earlier analysis. The closer the
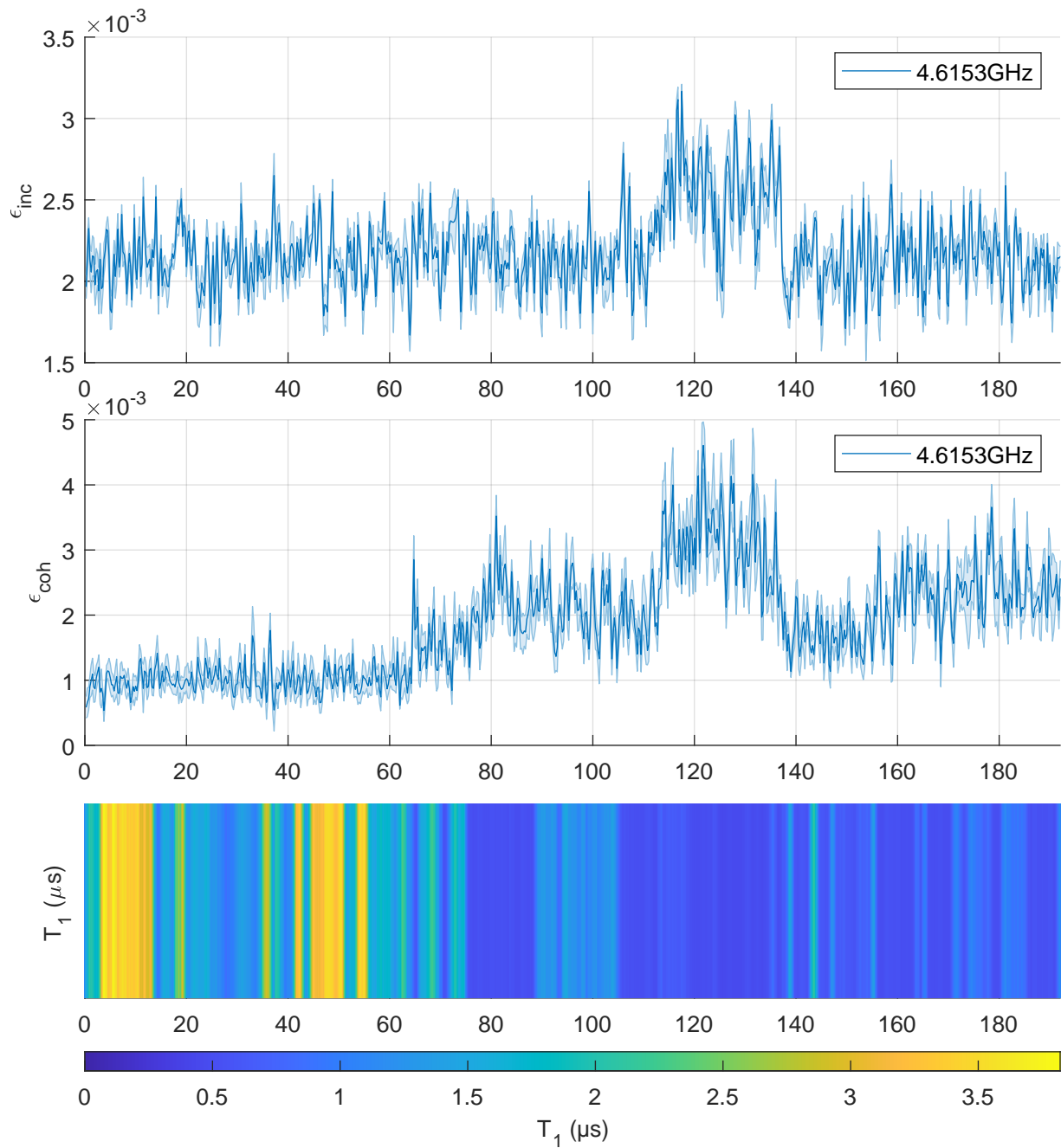
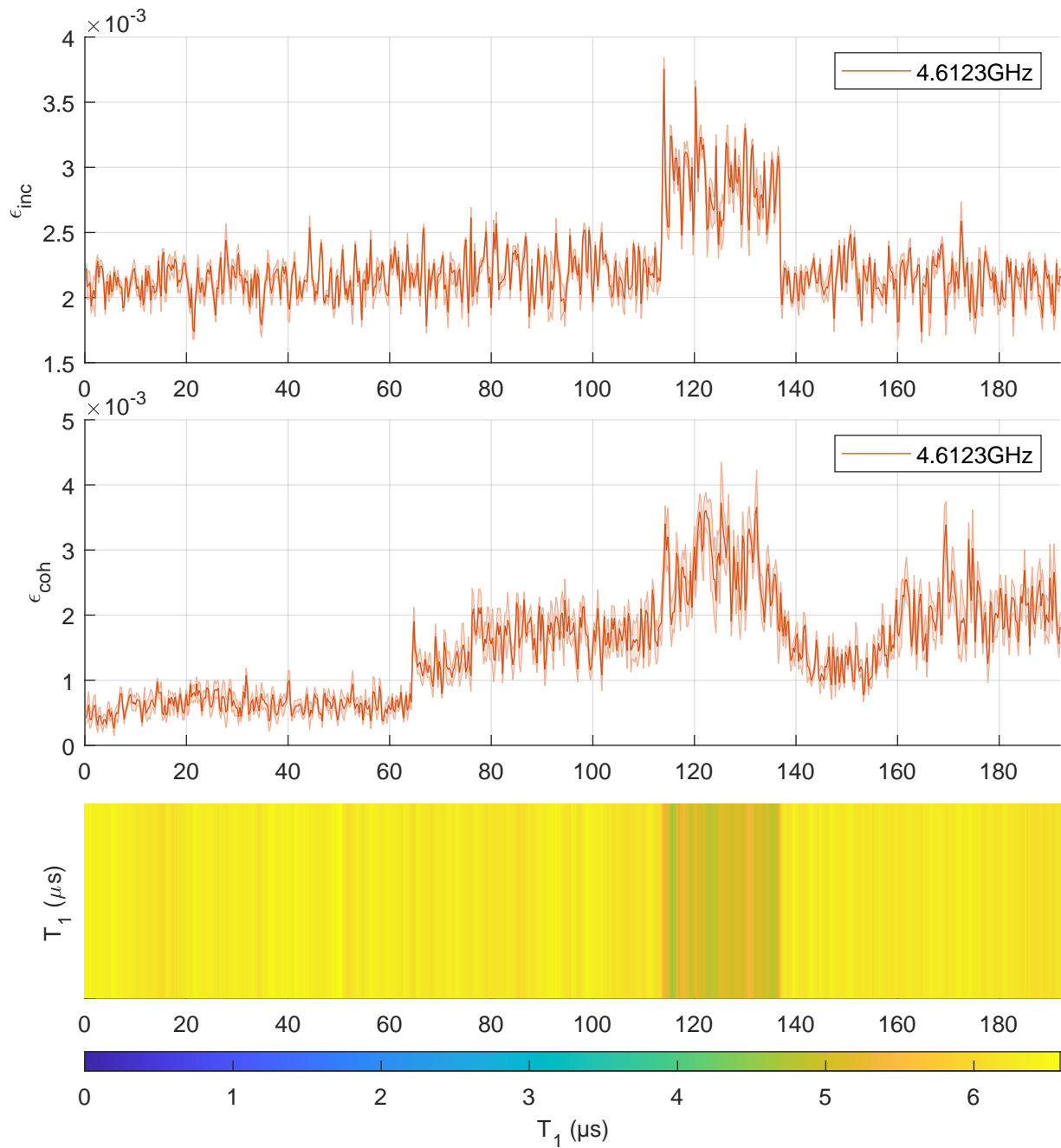Figure 4.2: Time series of errors and T_1 time for the first PB frequency

Figure 4.3: Time series of errors and T_1 time for the second PB frequency

qubit frequency is to the TLS, the larger and more intense the fluctuations of the coherent error become. This phenomenon might be explained by an increase in qubit leakage as changes in TLS coupling alter the system dynamics. Techniques developed to counteract leakage, such as DRAG pulses, may become less effective in the presence of stronger TLS coupling. However, this hypothesis requires further investigation and validation.

On the other hand, the dynamics of incoherent error fluctuations observed in this experiment remain more complex and demand further exploration. While the incoherent error for the PB frequency closer to the TLS exhibits slightly more intense fluctuations, its average amplitude remains largely the same as the other frequency. This observation is somewhat counter-intuitive, as one would typically expect the rate of decoherence and the incoherent error to be larger and faster at the qubit frequency closer to the TLS. The reasons behind these observed dynamics necessitate further investigation and a more robust set of evidence before a conclusive interpretation.

An intriguing avenue for further exploration arises from the observation of qubit frequency fluctuations. As evidenced in the figures presented in this section, the frequencies employed for PB and T_1 measurements at the same bias voltage exhibit considerable deviation (approximately 0.7MHz). This disparity is not unexpected, as the frequency of an Xmon transmon qubit is known to be sensitive to flux noise. Consequently, our initial assumption might not be proper in this experiment. Despite maintaining a fixed bias voltage throughout the experiment, the qubit frequency may still fluctuate. Given that all frequency parameters in the experiment remained constant for the duration of the study, such qubit frequency fluctuations could introduce additional complexities in the analysis of errors.

4.2   Simulations on TLS Noises and Flux Noises

To further substantiate the hypothesis that the fluctuation patterns observed in error metrics in the previous sections could be attributed to both TLS and flux noises, it is possible to further incorporate the measurements for qubit frequencies into the PB protocol. The simplest way to accurately measure the qubit frequency is through Ramsey measurements. However, our trial experiments confirmed that adding even one Ramsey measurement for each of the 11 selected bias voltages would extend the cycle period time to an excessive 340 seconds. Considering a moving window of length 20 cycle and 50% overlap, this leads to a sampling period of about 57 minutes, which is less practical for PB measurements. Therefore, we decided to conduct our exploration through comprehensive simulations of the PB experiment.

The model in our simulations is composed of a qubit and a TLS. To account for potential leakage scenarios, we included three energy levels within the qubit space. The qubit and the TLS are coupled via an electric dipole interaction. After applying a rotating-wave approximation, the total Hamiltonian takes the form of the Jaynes-Cummings model, where the system consists of a three-level anharmonic resonator (representing the actual qubit) and a "qubit" (representing the TLS). The energy splitting values and the coupling strength were determined based on estimates from the physical device.

The qubit control is implemented by adding the two driving terms from Equation 1.23. The drive pulse envelope $\varepsilon(t)$ used in the simulation follows the design of cosine-shaped DRAG pulses, identical to the pulses implemented in the physical experiments. The parameters for these DRAG pulses are calibrated based on the system Hamiltonian prior to conducting the PB simulations, thereby minimizing leakage. Similar calibration procedures were also employed in the physical experiments.

Given that the assumptions of the PB protocol mandate the error processes to be gate-independent and in the Markovian limit, the simulation is conducted by numerically solving the master equation with the system Hamiltonian. To emulate the system's decoherence, an amplitude damping channel is applied to the TLS subspace, and a combined amplitude and phase damping channel is applied for the qubit subspace. The Lindbladian operators are given from Equations 1.27 and 1.28, respectively. The decay parameters Γ_1^{TLS}, Γ_1^q, and Γ_2^q are drawn from physical estimations.

The simulation emulates the entire process of a standard PB experiment. Initially, both the qubit and TLS are prepared in the pure ground state. Following this, the qubit is driven by a pulse sequence derived from a random gate sequence $\mathcal{S}$. The time-dependent evolution of the system, under the influence of the qubit drive pulses and decoherence channels, is then simulated by numerically solving the master equation. Upon completion of the pulse sequence, the probability of the qubit being excited ($P_{|e\rangle}$ or $\langle E_{+z}\rangle$) is assessed by directly evaluating the density matrix elements. This method ensures that the execution of the PB sequence is free of SPAM errors, thus reducing the complexity in computation and post-analysis. We simulate 42 different random gate sequences for the same sequence length m_i (i.e., $N = 42$) and select $\mathcal{M} = \{2, 6, 13, 25, 50, 100, 200\}$. Following the procedures outlined in Chapter 2, the average gate fidelities and the quantity $\hat{b}$ are computed and fitted to derive the coherent and incoherent errors.

To investigate the contributions of TLS and qubit frequency shifts to the errors, as well as to validate that the TLS and flux noise channels could plausibly account for the error fluctuations observed in Section 4.1, we simulated both coherent and incoherent errors across a range of TLS frequency detunings Δ_t and qubit frequency detunings δ. The results
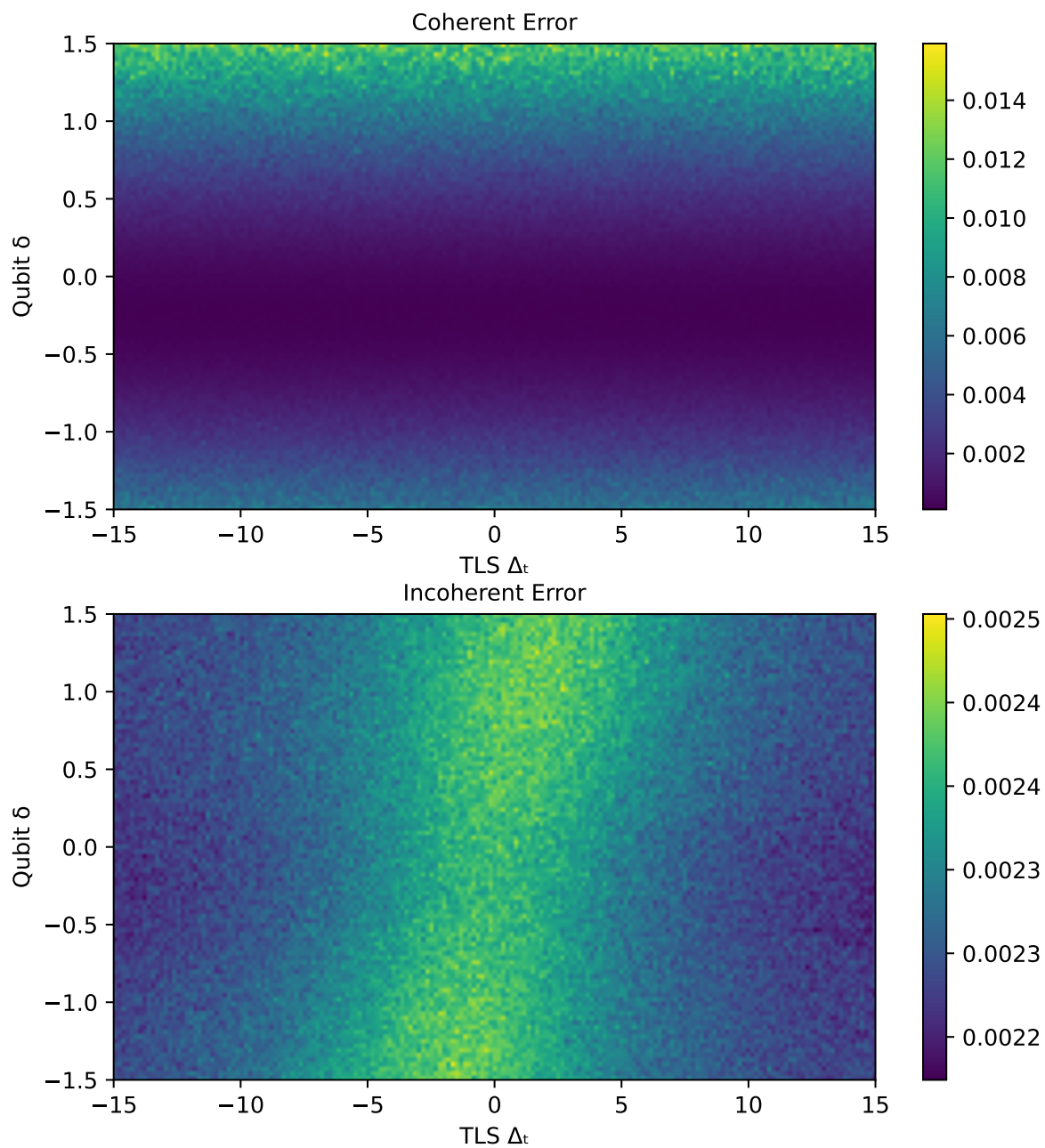
Figure 4.4: Color maps of simulated coherent and incoherent errors. The x-axis represents the TLS detuning (Δ_t) and the y-axis represents the qubit detuning (δ), with both detuning frequencies shown in MHz. The color intensity indicates the amplitude of the respective error metrics.

are illustrated as color maps in Figure 4.4. The units for the x and y axes are both in MHz. The source code of this simulation is available upon request.

From these simulations, it becomes evident that the coherent error is highly sensitive to the detuning of the qubit frequency. Conversely, the detuning of the TLS has a less pronounced effect on the coherent error. This observation supports the hypothesis that the coherent noise for both frequencies could be primarily attributed to fluctuations in the qubit frequency, which are potentially linked to flux noise. However, the role of the TLS appears to be minimal in the generation of coherent noise.

The simulations of the incoherent error, on the other hand, reveal a richer set of features. The amplitude of the incoherent error seems to vary slowly with the detuning of the qubit frequency. In the range of qubit frequency variation implied by the coherent errors, the amplitude of incoherent error remains relatively stable, suggesting that qubit frequency fluctuations may not be a significant factor in inducing changes in incoherent error. However, the fluctuations from the TLS detuning do show notable variations, indicating a potential correlation between incoherent noise and TLS noise.

It is also noteworthy that the overall variation level of incoherent error in this simulation is relatively small. This could potentially explain the lack of telegraphic events in incoherent noise compared to coherent noise. Moreover, this observation might hint at the existence of additional noise channels, which were not explored in this simulation, contributing to the incoherent noise. Such channels could be responsible for the wide-band telegraphic event observed in Figure 4.2 and 4.3.

Chapter 5

Conclusion

In this thesis, we explore qubit dynamics and their interactions with the environment, using a set of novel metrics. These metrics, distinct from the commonly studied variables such as T_1 and frequency, are garnering increased attention in the quantum community, particularly in the field of quantum algorithms and hardware development.

The metrics being explored are average gate fidelity and unitarity of the implementation of the quantum gates. The average gate fidelity offers estimations of the total averaged quantum errors for quantum gates, while the unitarity offers an assessment of the coherent portion of these total quantum errors. This analysis is significant because, in principle, coherent errors could be mitigated through systematic optimizations, including refining control protocols and calibrating experimental instruments. On the other hand, the incoherent component of the error should ultimately be constrained by quantum decoherence of the quantum device, which is harder to improve.

In Chapter 3, we conducted two long-term experiments and subsequently performed a detailed analysis of the results. The experiments were conducted utilizing modified PB protocols introduced in Chapter 2. These investigations provide an initial understanding of the behavior of the quantum noises derived from those novel metrics. Additionally, the continuous measurements of these metrics over extended periods allowed us to characterize the coherence of noise at very low frequencies, providing insights into how long-term qubit fluctuations correlates to the coherent and incoherent errors.

Subsequently, we aimed to link the observed fluctuations in quantum errors with well-established models, namely, the TLS and flux noises. This endeavor involved the implementation of simultaneous measurements and comprehensive simulations, as discussed in

Chapter 4. Our goal was to connect these novel observations with existing knowledge, thereby contributing to a more complete understanding of quantum dynamics.

While our experiments and simulations generally align, they also present intriguing discrepancies. To delve deeper into these differences, one might consider incorporating additional metrics into the PB cycle for better characterization. These could include Ramsey measurements to simultaneously estimate the qubit frequency fluctuations. Conducting more experiments over extended periods would also be beneficial in capturing telegraphic events.

A significant challenge for adding more measurements for other metrics is the extended sampling period of a cycle. The execution time for a single cycle in the experiments presented in Chapter 3 is 50 seconds, and tens of cycles are required to fit just one error value, which is already a long duration.This issue is worsen when incorporating other measurements. Introducing relaxation time and qubit frequency shift measurements into the PB cycle extends the error sampling period to a daunting 3400 seconds. Although the results from relaxation time and frequency measurements are both intriguing and consistent, the measurements of the error metrics become less meaningful due to the long cycle period. Hence, adjustments are needed prior to incorporating additional measurements into the PB cycle.

To reduce the cycle period time, we experimentally benchmarked the time efficiency of the PB protocol and found that initializing the instruments and uploading the random sequences for each cycle take approximately 10 seconds. This duration is significant relative to the cycle period. Better integration of the instruments, both at the hardware and software levels, could potentially optimize the cycle period to some degree. However, the T_1 and Ramsey measurements wouldn't benefit significantly from this optimization. Protocol-wise, since the T_1 and Ramsey measurements usually take much longer, we could execute and measure more than one PB sequence per frequency during each cycle. Although this would extend the cycle period, we could reduce the length of the moving window accordingly, ultimately reducing the sampling period for the error points. Further studies are required to evaluate these potential improvements.

References

[1] J. Gambetta, Quantum-centric supercomputing: The next wave of computing, `https://research.ibm.com/blog/next-wave-quantum-centric-supercomputing` (2022), accessed: 2023.

[2] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, *et al.*, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505–510 (2019).

[3] J. Béjanin, Y. Ayadi, X. Xu, C. Zhu, H. Mohebbi, and M. Mariantoni, Fluctuation spectroscopy of two-level systems in superconducting resonators, Phys. Rev. Appl. **18**, 034009 (2022).

[4] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Clarendon Press, 1930).

[5] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Charge-insensitive qubit design derived from the cooper pair box, Phys. Rev. A **76**, 042319 (2007).

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

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

[8] V. Ambegaokar and A. Baratoff, Tunneling between superconductors, Phys. Rev. Lett. **10**, 486–489 (1963).

[9] J. H. Béjanin, *Advances in Superconducting Circuit Quantum Electrodynamics*, PhD Thesis, University of Waterloo (2022).

[10] J. A. Schreier, A. A. Houck, J. Koch, D. I. Schuster, B. R. Johnson, J. M. Chow, J. M. Gambetta, J. Majer, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Suppressing charge noise decoherence in superconducting charge qubits, Phys. Rev. B 77, 180502 (2008).

[11] R. C. Jaklevic, J. Lambe, A. H. Silver, and J. E. Mercereau, Quantum interference effects in josephson tunneling, Phys. Rev. Lett. 12, 159–160 (1964).

[12] D. I. Schuster, *Circuit quantum electrodynamics*, Ph.D. thesis, Yale University (2007).

[13] J. B. Ketterson, *The Physics of Solids* (Oxford University Press, 2016).

[14] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, Circuit quantum electrodynamics, Rev. Mod. Phys. 93, 025005 (2021).

[15] J. Stillwell, *Naive Lie Theory*, Undergraduate Texts in Mathematics (Springer New York, 2008).

[16] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, Simple pulses for elimination of leakage in weakly nonlinear qubits, Phys. Rev. Lett. 103, 110501 (2009).

[17] Z. Chen, J. Kelly, C. Quintana, R. Barends, B. Campbell, Y. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Lucero, E. Jeffrey, A. Megrant, J. Mutus, M. Neeley, C. Neill, P. J. J. O'Malley, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Korotkov, and J. M. Martinis, Measuring and suppressing quantum state leakage in a superconducting qubit, Phys. Rev. Lett. 116, 020501 (2016).

[18] W. A. Phillips, Two-level states in glasses, Reports on Progress in Physics 50, 1657 (1987).

[19] C. Müller, J. H. Cole, and J. Lisenfeld, Towards understanding two-level-systems in amorphous solids: insights from quantum circuits, Reports on Progress in Physics 82, 124501 (2019).

[20] J. J. Burnett, A. Bengtsson, M. Scigliuzzo, D. Niepce, M. Kudra, P. Delsing, and J. Bylander, Decoherence benchmarking of superconducting qubits, npj Quantum Information 5, 54 (2019).

[21] J. H. Béjanin, C. T. Earnest, A. S. Sharafeldin, and M. Mariantoni, Interacting defects generate stochastic fluctuations in superconducting qubits, Phys. Rev. B 104, 094106 (2021).

[22] L. Faoro and L. B. Ioffe, Interacting tunneling model for two-level systems in amorphous materials and its predictions for their dephasing and noise in superconducting microresonators, Phys. Rev. B **91**, 014201 (2015).

[23] N. F. Ramsey, A molecular beam resonance method with separated oscillating fields, Phys. Rev. **78**, 695–699 (1950).

[24] J. Burnett, L. Faoro, I. Wisby, V. Gurtovoi, A. Chernykh, G. Mikhailov, V. Tulin, R. Shaikhaidarov, V. Antonov, P. Meeson, *et al.*, Evidence for interacting two-level systems from the 1/f noise of a superconducting resonator, Nature communications **5**, 4119 (2014).

[25] J. Helsen, I. Roth, E. Onorati, A. Werner, and J. Eisert, General framework for randomized benchmarking, PRX Quantum **3**, 020357 (2022).

[26] I. L. Chuang and M. A. Nielsen, Prescription for experimental determination of the dynamics of a quantum black box, Journal of Modern Optics **44**, 2455–2467 (1997).

[27] J. M. Chow, J. M. Gambetta, L. Tornberg, J. Koch, L. S. Bishop, A. A. Houck, B. R. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Randomized benchmarking and process tomography for gate errors in a solid-state qubit, Phys. Rev. Lett. **102**, 090502 (2009).

[28] R. C. Bialczak, M. Ansmann, M. Hofheinz, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, H. Wang, J. Wenner, M. Steffen, *et al.*, Quantum process tomography of a universal entangling gate implemented with josephson phase qubits, Nature Physics **6**, 409–413 (2010).

[29] J. Watrous, Semidefinite programs for completely bounded norms, Theory of Computing **5**, 217–238 (209).

[30] N. Johnston, D. W. Kribs, and V. I. Paulsen, Computing stabilized norms for quantum operations via the theory of completely bounded maps (2007), arXiv:0711.3636 [quant-ph] .

[31] Y. Sanders, *Characterizing Errors in Quantum Information Processors*, PhD Thesis, University of Waterloo (2016).

[32] J. J. Wallman and S. T. Flammia, Randomized benchmarking with confidence, New J. Phys. **16**, 103032 (2014), publisher: IOP Publishing.

[33] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).

[34] M. Horodecki, P. Horodecki, and R. Horodecki, General teleportation channel, singlet fraction, and quasidistillation, Physical Review A **60**, 1888 (1999).

[35] M. D. Bowdrey, D. K. Oi, A. J. Short, K. Banaszek, and J. A. Jones, Fidelity of single qubit maps, Physics Letters A **294**, 258–260 (2002).

[36] M. A. Nielsen, A simple formula for the average gate fidelity of a quantum dynamical operation, Physics Letters A **303**, 249–252 (2002).

[37] J. J. Morton, A. M. Tyryshkin, A. Ardavan, S. C. Benjamin, K. Porfyrakis, S. A. Lyon, and G. A. D. Briggs, Bang–bang control of fullerene qubits using ultrafast phase gates, Nature Physics **2**, 40–43 (2006).

[38] S. Kimmel, G. H. Low, and T. J. Yoder, Robust calibration of a universal single-qubit gate set via robust phase estimation, Physical Review A **92**, 062315 (2015).

[39] J. Wallman, C. Granade, R. Harper, and S. T. Flammia, Estimating the coherence of noise, New J. Phys. **17**, 113020 (2015).

[40] G. Feng, J. J. Wallman, B. Buonacorsi, F. H. Cho, D. K. Park, T. Xin, D. Lu, J. Baugh, and R. Laflamme, Estimating the coherence of noise in quantum control of a solid-state qubit, Phys. Rev. Lett. **117**, 260501 (2016), publisher: American Physical Society.

[41] J. Emerson, R. Alicki, and K. Życzkowski, Scalable noise estimation with random unitary operators, J. Opt. B: Quantum Semiclass. Opt. **7**, S347 (2005).

[42] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Randomized benchmarking of quantum gates, Phys. Rev. A **77**, 012307 (2008).

[43] E. Magesan, J. M. Gambetta, and J. Emerson, Scalable and robust randomized benchmarking of quantum processes, Phys. Rev. Lett. **106**, 180504 (2011), publisher: American Physical Society.

[44] R. Koenig and J. A. Smolin, How to efficiently select an arbitrary Clifford group element, Journal of Mathematical Physics **55**, 10.1063/1.4903507 (2014), 122202.

[45] J. T. Muhonen, A. Laucht, S. Simmons, J. P. Dehollain, R. Kalra, F. E. Hudson, S. Freer, K. M. Itoh, D. N. Jamieson, J. C. McCallum, *et al.*, Quantifying the quantum gate fidelity of single-atom spin qubits in silicon by randomized benchmarking, Journal of Physics: Condensed Matter **27**, 154205 (2015).

[46] M. A. Fogarty, M. Veldhorst, R. Harper, C. H. Yang, S. D. Bartlett, S. T. Flammia, and A. S. Dzurak, Nonexponential fidelity decay in randomized benchmarking with low-frequency noise, Phys. Rev. A **92**, 022326 (2015).

[47] R. Harper, I. Hincks, C. Ferrie, S. T. Flammia, and J. J. Wallman, Statistical analysis of randomized benchmarking, Phys. Rev. A **99**, 052350 (2019), publisher: American Physical Society.

[48] J. J. Wallman, M. Barnhill, and J. Emerson, Robust characterization of leakage errors, New Journal of Physics **18**, 043021 (2016).

[49] Microsoft time server, `time.windows.com` (n.d.), accessed in 2023.

[50] G. V. Haines and A. G. Jones, Logarithmic fourier transformation, Geophysical Journal **92**, 171–178 (1988).

[51] H. Nyquist, Certain factors affecting telegraph speed, The Bell System Technical Journal **3**, 324–346 (1924).

[52] C. Shannon, Communication in the presence of noise, Proceedings of the IRE **37**, 10–21 (1949).

[53] T. Lindström, J. Burnett, M. Oxborrow, and A. Y. Tzalenchuk, Pound-locking for characterization of superconducting microresonators, Review of Scientific Instruments **82**, 104706 (2011).

[54] P. V. Klimov, J. Kelly, Z. Chen, M. Neeley, A. Megrant, B. Burkett, R. Barends, K. Arya, B. Chiaro, Y. Chen, A. Dunsworth, A. Fowler, B. Foxen, C. Gidney, M. Giustina, R. Graff, T. Huang, E. Jeffrey, E. Lucero, J. Y. Mutus, O. Naaman, C. Neill, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, S. Boixo, R. Babbush, V. N. Smelyanskiy, H. Neven, and J. M. Martinis, Fluctuations of energy-relaxation times in superconducting qubits, Phys. Rev. Lett. **121**, 090502 (2018).

[55] C. H. Yang, K. W. Chan, R. Harper, W. Huang, T. Evans, J. C. C. Hwang, B. Hensen, A. Laucht, T. Tanttu, F. E. Hudson, S. T. Flammia, K. M. Itoh, A. Morello, S. D.

Bartlett, and A. S. Dzurak, Silicon qubit fidelities approaching incoherent noise limits via pulse engineering, Nat Electron **2**, 151–158 (2019), number: 4 Publisher: Nature Publishing Group.

[56] R. Kueng, D. M. Long, A. C. Doherty, and S. T. Flammia, Comparing experiments to the fault-tolerance threshold, Phys. Rev. Lett. **117**, 170502 (2016), publisher: American Physical Society.

[57] Y. R. Sanders, J. J. Wallman, and B. C. Sanders, Bounding quantum gate error rate based on reported average fidelity, New J. Phys. **18**, 012002 (2015), publisher: IOP Publishing.

[58] J. J. Wallman, Bounding experimental quantum error rates relative to fault-tolerant thresholds (2016), arXiv:1511.00727 [quant-ph].

[59] C. Müller, J. Lisenfeld, A. Shnirman, and S. Poletto, Interacting two-level defects as sources of fluctuating high-frequency noise in superconducting circuits, Physical Review B **92**, 035442 (2015).

[60] D. Gottesman, Theory of fault-tolerant quantum computation, Phys. Rev. A **57**, 127–137 (1998).

[61] S. Aaronson and D. Gottesman, Improved simulation of stabilizer circuits, Phys. Rev. A **70**, 052328 (2004).

[62] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, and J. M. Gambetta, Efficient z gates for quantum computing, Phys. Rev. A **96**, 022330 (2017).

[63] J. Béjanin, C. Earnest, and M. Mariantoni, The quantum socket and demuxyz-based gates with superconducting qubits, arXiv preprint arXiv:2211.00143 (2022).

[64] D. Allan, Statistics of atomic frequency standards, Proceedings of the IEEE **54**, 221–230 (1966).

[65] W. Riley and D. Howe, Handbook of frequency stability analysis (2008).

APPENDICES

Appendix A

Time-domain Analysis: Overlapping Allan Deviation

The Allan variance (AVAR) was initially developed in 1966 by David W. Allan to analyze frequency stability in precision clocks, oscillators, and other frequency sources[64]. Since its inception, AVAR has been widely adopted in the fields of telecommunications, navigation, and metrology. To enhance the statistical confidence of stability estimates, an overlapping version of AVAR was later introduced. This overlapping version, and its deviation, has since become the most common measures of time-domain frequency stability. The terms "AVAR" and "ADEV" are now commonly used to refer to these overlapping versions, which have largely superseded the original Allan variance and deviation[65].

For an averaging time τ, the overlapping Allan deviation σ is defined by the following equation:

$$\sigma^2(\tau) = \frac{1}{2m^2(M-2m+1)} \sum_{j=1}^{M-2m+1} \left\{ \sum_{i=j}^{j+m-1} (\bar{y}_{i+m} - \bar{y}_i) \right\}^2 \tag{A.1}$$

In this equation, M is the total number of samples, τ_0 is the base measurement period, and $\bar{y}$ is the time-averaged fractional frequency. The averaging time τ is a multiple of the base measurement period ($\tau = m\tau_0$), where m is an integer representing the number of base periods.

Similar to the analysis using PSDs, we compute the ADEV over different averaging time τ. By examining the ADEV values corresponding to different τ, one can reveal various characteristics of time-domain stability, including insights into the type of noise and the frequency dependency.
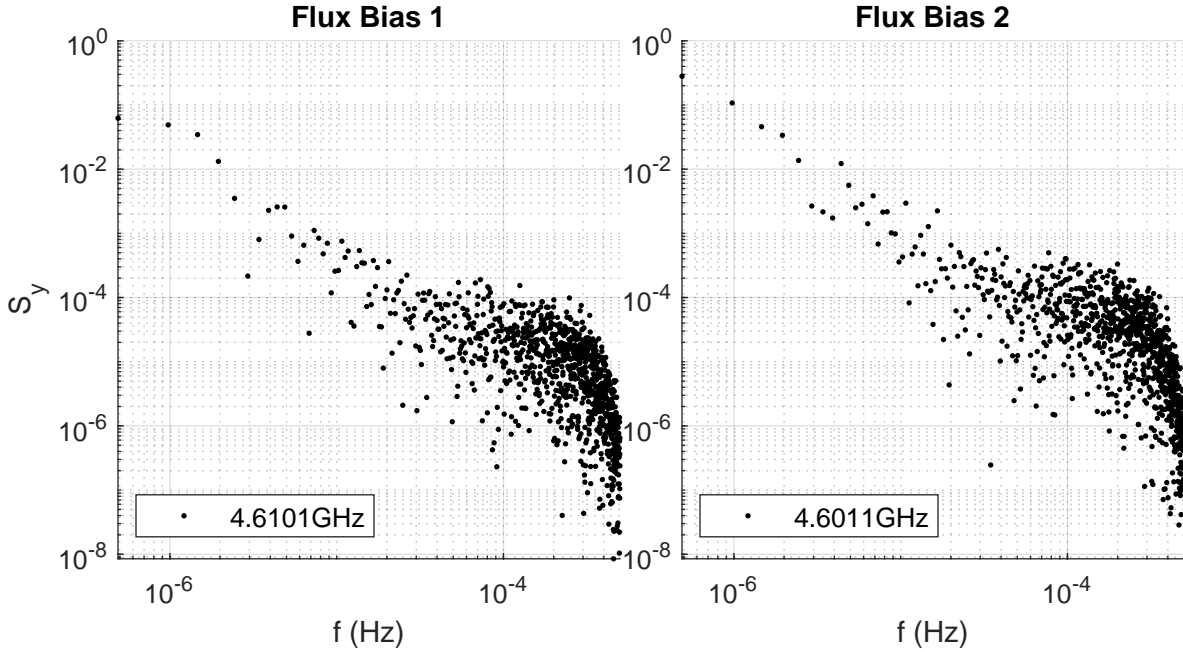
Figure A.1: Examples of PSD plots for time series of coherent error in the second experiment (Figure 3.10).

In Particularly, when plotted in logarithmic space, the ADEV exhibits linear patterns for power-law noises, with each type of noise having distinct slopes. Table 3.1 presents some power-law noises and their characteristic slopes in the log-space $\sigma - \tau$ plot.

Power-law noise processes typically produce a "well"-shaped curve in the Allan deviation plot. The "walls" of this well often exhibit slopes of -1/2 and 1/2, which are typically indicative of white noise and random walk noise processes, respectively. As more types of noise are incorporated into the noise model, the range of potential slope values expands. Despite these additional noise types, the 'well'-shaped pattern in the Allan deviation plot persists.

Next, we provide PSD plots in Figure A.1 for time series of coherent error in the second experiment (Figure 3.10). When compared with the corresponding ADEV plots (second row in Figure 3.10), the PSDs exhibit very limited resolution on the lower frequency side, and overall, they are significantly more noisy and challenging to analyze. More detailed explanations of why ADEV is preferred for low-frequency analysis can be found in Reference. [20, 64].

Appendix B

Single-qubit Clifford Gates

The protocol for randomized benchmarking requires random gate sequences to be drawn uniformly from a quantum operation group. This effectively corresponds to taking the Haar-average over this group. Ideally, for the most comprehensive estimation, this group would have been chosen as the entire Unitary group. However, this approach is neither scalable nor experimentally approachable. As a result, subsequent theories and experimental implementations have shifted their focus to a special subset of the Unitary group: the Clifford set. This appendix chapter provides a brief review of the single-qubit Clifford set, and outlines our experimental implementations of the Clifford gates in our purity benchmarking experiments.

In quantum information theory, n-qubit Clifford gates are defined as the normalizers of the n-qubit Pauli group. In other words, they are operations that transform n-qubit Pauli states into other n-qubit Pauli states[60]. These n-qubit Clifford gates form a subgroup of the Unitary group. Although the Clifford group itself does not constitute a quantum universal gate set, it plays a fundamental role in several key areas. These include, but are not limited to, stabilizer codes, error corrections, entanglement studies, and as an essential component in the construction of quantum universal gate sets. These areas represent significant portions of research and application in the field of quantum information. Therefore, characterizing qubit errors by uniformly averaging random gates from the Clifford group, rather than the entire Unitary group, still provides a sufficiently comprehensive approach for most applications.

From an experimental standpoint, there are two advantages using the Clifford group for randomized benchmarking. First, the Clifford group is a finite discrete set, which simplifies the Haar integral in Equation 2.14 to a finite summation. Although the cardinality

of the Clifford group still increases exponentially with the number of qubits, thereby not being fully scalable, it does provide a practical approach for performing the Haar-average. Second, Clifford gates can be efficiently computed and simulated using classical computers (Gottesman-Knill theorem[61]). This is particularly advantageous in randomized benchmarking , as it allows us to efficiently calculate the final inverse gate (g_{inv}) and generate the random sequence ($\mathcal{S}^{m_i}$).

Clifford groups can be generated by compositions of Pauli matrices. Table B.1 lists the 24 elements in single-qubit Clifford group and the respective decompositions we used to implement the gate operations. Note that the representation and the decomposition of each Clifford gates is not unique. The notation $R_{(x,y,z)}(\theta)$ represents a Clifford gate as a Block-sphere rotation by an angle θ around the axis $\vec{u} = (x, y, z)$. The decompositions are expressed in terms of primitive gates. To avoid the usage of flux pulses, we further decompose the z rotations into combinations of x,y rotations[9, 62, 63].

Clifford gate	Decomposition	Clifford gate	Decomposition
I	I	$R_{(1,0,0)}(\pi/2)$	$X/2$
$R_{(1,0,0)}(\pi)$	X	$R_{(-1,0,0)}(\pi/2)$	$-X/2$
$R_{(0,1,0)}(\pi)$	X, Z	$R_{(0,1,0)}(\pi/2)$	$X/2, Z/2, -X/2$
$R_{(0,0,1)}(\pi)$	Z	$R_{(0,-1,0)}(\pi/2)$	$X/2, -Z/2, -X/2$
$R_{(1,1,1)}(2\pi/3)$	$X/2, Z/2$	$R_{(0,0,1)}(\pi/2)$	$Z/2$
$R_{(1,1,-1)}(2\pi/3)$	$-Z/2, X/2$	$R_{(0,0,-1)}(\pi/2)$	$-Z/2$
$R_{(1,-1,1)}(2\pi/3)$	$Z/2, X/2$	$R_{(1,0,1)}(\pi)$	$X/2, Z/2, X/2$
$R_{(1,-1,-1)}(2\pi/3)$	$X/2, -Z/2$	$R_{(1,0,-1)}(\pi)$	$X/2, -Z/2, X/2$
$R_{(-1,1,1)}(2\pi/3)$	$Z/2, -X/2$	$R_{(0,1,-1)}(\pi)$	$Z, X/2$
$R_{(-1,1,-1)}(2\pi/3)$	$-X/2, -Z/2$	$R_{(0,1,1)}(\pi)$	$Z, -X/2$
$R_{(-1,-1,1)}(2\pi/3)$	$-X/2, Z/2$	$R_{(1,1,0)}(\pi)$	$X, Z/2$
$R_{(-1,-1,-1)}(2\pi/3)$	$-Z/2, -X/2$	$R_{(-1,1,0)}(\pi)$	$X, -Z/2$

Table B.1: Single-qubit Clifford gates and their decompositions. Adapted from References [9, 63]. The order of the decomposition sequence is from left to right. For instance, "-Z/2,X/2" presents applying a rotation of $\pi/2$ around the z-axis first, followed by a rotation of $\pi/2$ around the x-axis.