Magnetic flux noise in SQUIDs and qubits

By

Steven Anton

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor John Clarke, Chair
Professor Irfan Siddiqi
Professor Theodore Van Duzer

Fall 2013
Magnetic flux noise in SQUIDs and qubits

Copyright 2013
by
Steven Anton
Abstract

Magnetic flux noise in SQUIDs and qubits

by

Steven Anton

Doctor of Philosophy in Physics

University of California, Berkeley

Professor John Clarke, Chair

For over three decades, the presence of magnetic flux noise with a power spectral density scaling roughly as $S_\Phi(f) \propto 1/f^\alpha$, where $\alpha \lesssim 1$, has been known to limit the low-frequency performance of dc superconducting quantum interference devices (SQUIDs). In recent years, experiments indicate that this same noise persists to frequencies up to 1 GHz and is a dominant source of dephasing in flux-sensitive superconducting quantum bits (qubits). Thus, the reduction of flux noise presents a major hurdle towards the successful realization of scalable quantum computers that are based on flux-based qubits. In this thesis, we present experimental measurements, theoretical analyses, and numerical simulations that support a more detailed understanding of both the microscopic and macroscopic properties of flux noise.

Our experimental work begins with flux noise measurements of a large number of SQUIDs in the temperature range from 0.1 K to 4 K. We report on measurements of ten SQUIDs with systematically varied geometries and show that $\alpha$ increases as the temperature is lowered; in so doing, each spectrum pivots about a nearly constant frequency. The mean square flux noise, inferred by integrating the power spectra, grows rapidly with temperature and at a given temperature is approximately independent of the outer dimension of a given SQUID washer. We show that these results are incompatible with a model based on the random reversal of independent, spins that are located at the surface of the SQUID washer.

In the course of our flux noise measurements, we became aware of a spurious contribution to low-frequency critical current noise in Josephson junctions—normally attributed to charge trapping in the barrier—arising from temperature instabilities inherent in cryogenic systems. These temperature fluctuations modify the critical current via its temperature dependence. By computing cross-correlations between measured temperature and critical current noise in Al-AlOx-Al junctions, we show that, despite excellent temperature stability, temperature fluctuations induce observable critical current fluctuations. Particularly, because $1/f$ critical current noise has decreased with improved fabrication techniques in recent years, it is important to understand and eliminate this additional noise source.
Next, we introduce a numerical method of calculating the mean square flux noise $\langle \Phi^2 \rangle$ from independently fluctuating spins on the surface of thin-film loops of arbitrary geometry. By reciprocity, $\langle \Phi^2 \rangle$ is proportional to $\langle B(r)^2 \rangle$, where $B(r)$ is the magnetic field generated by a circulating current around the loop and $r$ varies over the loop surface. By discretizing the loop nonuniformly, we efficiently and accurately compute the current distribution and resulting magnetic field, which may vary rapidly across the loop. We use this method to compute $\langle \Phi^2 \rangle$ in a number of scenarios in which we systematically vary physical parameters of the loop. We compare our simulations to an earlier analytic result predicting that $\langle \Phi^2 \rangle \propto R/W$ in the limit where the loop radius $R$ is much greater than the linewidth $W$. We further show that the previously neglected contribution of edge spins to $\langle \Phi^2 \rangle$ is significant—even dominant—in narrow-linewidth loops.

To calculate theoretical dephasing rates in qubits, we consider flux noise with a spectral density $S_\Phi(f) = A^2/(f/1\text{ Hz})^\alpha$, where $A$ is of the order of $1 \mu\Phi_0\text{ Hz}^{-1/2}$ and $0.6 \leq \alpha \leq 1.2$; $\Phi_0$ is the flux quantum. For a qubit with an energy level splitting linearly coupled to the applied flux, our calculations of the dependence of the pure dephasing time $\tau_\phi$ of Ramsey and echo pulse sequences on $\alpha$ for fixed $A$ show that $\tau_\phi$ decreases rapidly as $\alpha$ is reduced. We find that $\tau_\phi$ is relatively insensitive to the noise bandwidth, $f_1 \leq f \leq f_2$, for all $\alpha$ provided the ultraviolet cutoff frequency $f_2 > 1/\tau_\phi$. We calculate the ratio $\tau_{\phi,E}/\tau_{\phi,R}$ of the echo ($E$) and Ramsey ($R$) sequences, and the dependence of the decay function on $\alpha$ and $f_2$. We investigate the case in which $S_\Phi(f_0)$ is fixed at the “pivot frequency” $f_0 \neq 1$ Hz while $\alpha$ is varied, and find that the choice of $f_0$ can greatly influence the sensitivity of $\tau_{\phi,E}$ and $\tau_{\phi,R}$ to the value of $\alpha$.

Finally, we conclude with a brief review of our principal results and conclusions. We also comment on promising avenues of future research.
To Mariana
# Contents

List of Figures v  
List of Tables vii  
Acknowledgments viii

## 1 Introduction 1

1.1 Overview .................................................. 1  
1.2 Superconductivity ........................................... 2  
  1.2.1 Flux quantization ..................................... 2  
  1.2.2 The Josephson effect ................................ 3  
1.3 Superconducting devices ..................................... 6  
  1.3.1 The dc SQUID ....................................... 6  
  1.3.2 Qubits ........................................ 9  
1.4 Noise .................................................. 10  
  1.4.1 Generic noise processes ................................ 10  
  1.4.2 Noise in SQUIDs and qubits .......................... 13  
1.5 Flux noise ............................................... 14  
  1.5.1 Previous measurements ................................ 14  
  1.5.2 Surface spin model .................................. 15  
  1.5.3 Proposed origin of spins .............................. 16  
  1.5.4 Flux noise in qubits ................................ 16

## 2 Experimental measurement system and procedures 17

2.1 Overview .................................................. 17  
2.2 Measurement overview ...................................... 18  
  2.2.1 Optimal circuit and device parameters ............... 19  
2.3 Implementation ........................................... 22  
2.4 Calibration and validation .................................. 25  
2.5 Measurement procedure .................................... 28
## Contents

3 **Scaling of flux noise with temperature** 33
   3.1 Introduction 33
   3.2 Methods 34
   3.3 Results 35

4 **Low-frequency critical current noise in Josephson junctions induced by temperature fluctuations** 44
   4.1 Introduction 44
   4.2 Methods 45
   4.3 Results 46

5 **Mean square flux noise in SQUIDs and qubits: numerical calculations** 52
   5.1 Introduction 52
   5.2 Methods
      5.2.1 Previous methods 53
      5.2.2 Numerical technique 55
   5.3 Results and discussion
      5.3.1 Role of loop radius 60
      5.3.2 Role of loop linewidth for fixed hole size 63
      5.3.3 Role of loop aspect ratio 64
      5.3.4 Role of loop linewidth for fixed aspect ratio 64
      5.3.5 Role of penetration depth 64
      5.3.6 Role of edge spins 66
      5.3.7 Role of film thickness 67
   5.4 Concluding remarks 68

6 **Dephasing in qubits due to flux noise** 70
   6.1 Introduction 70
   6.2 Model 70
   6.3 Results
      6.3.1 Dephasing times versus $\alpha$ 72
      6.3.2 Dephasing times versus cutoff frequencies 73
      6.3.3 The ratio $\tau_{\phi,E}/\tau_{\phi,R}$ 73
      6.3.4 Dependence of decay function on $\alpha$ and ultraviolet cutoff frequency 77
      6.3.5 $S_\phi(f)$ pivoting about $f_0 \neq 1$ Hz as $\alpha$ is varied 79
   6.4 Concluding remarks 80

7 **Concluding remarks** 81
   7.1 Summary of results 81
   7.2 Future directions 83

Bibliography 84
A Calculating a stitched spectrum from time series 91
B Fitting spectra 95
List of Figures

1.1 Josephson junction $I_c$-$\Phi$ characteristics ................................................. 6
1.2 dc SQUID schematic and $I_c$-$\Phi$ characteristics ........................................ 7
1.3 The three-junction flux qubit ............................................................................. 10
1.4 White and $1/f$ noise ........................................................................................ 11
1.5 The role of random telegraph signals in $1/f$ noise ............................................ 12
1.6 $T_1$ and $T_2$ processes shown on the Bloch sphere ........................................ 14
2.1 Schematic of measurement system ................................................................. 18
2.2 Bias configuration and operating points of measured SQUID ...................... 20
2.3 Schematic of measurement system ................................................................. 23
2.4 Photo of sample box ......................................................................................... 24
2.5 Circuit schematics of current supply boxes .................................................... 24
2.6 LabVIEW time capture acquisition program .................................................. 26
2.7 Current noise spectra for various biasing of measured SQUID ...................... 30
2.8 Equivalent flux noise spectrum showing scaling with $\partial I_{\text{loop}}/\partial \Phi_m$ ........ 31
3.1 Configuration of measurement system ............................................................. 35
3.2 Example power spectra $S_{\Phi}(f)$ .................................................................. 36
3.3 Temperature dependence of $1/f$ flux noise for 10 SQUIDs ......................... 37
3.4 Fitted power spectra $S_{\Phi}(f)$ for SQUIDs I ................................................ 38
3.5 Fitted power spectra $S_{\Phi}(f)$ for SQUIDs II ................................................ 39
3.6 Mean-square flux noise: inferred and predicted ............................................ 41
3.7 Scaling of noise power with $R$ ........................................................................ 42
4.1 Temperature dependence of critical current .................................................... 46
4.2 Temperature and critical current noise for three settings of temperature controller 47
4.3 Cross-correlations of the three temperature and critical current time series .... 48
4.4 Cross-spectra and spectral densities ............................................................. 49
4.5 Critical current noise caused by charge-trapping compared to that due to temperature variations ................................................................. 50
5.1 Current density in square loop .......................................................................... 55
5.2 Current density across linewidth and corresponding magnetic field ............ 57
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3</td>
<td>Numerically computed and analytic $\langle \Phi^2 \rangle$ versus $R_{\text{ins}}$ and $R/W$ for constant $W$.</td>
<td>61</td>
</tr>
<tr>
<td>5.4</td>
<td>Comparison of numerical and analytic current density and magnetic field.</td>
<td>62</td>
</tr>
<tr>
<td>5.5</td>
<td>$\langle \Phi^2 \rangle$ versus $W$ for two constant hole dimensions.</td>
<td>64</td>
</tr>
<tr>
<td>5.6</td>
<td>$\langle \Phi^2 \rangle$ versus $R/W$ for two constant values of $R$.</td>
<td>65</td>
</tr>
<tr>
<td>5.7</td>
<td>$\langle \Phi^2 \rangle$ versus $W$ for two constant values of $R/W$.</td>
<td>65</td>
</tr>
<tr>
<td>5.8</td>
<td>$\langle \Phi^2 \rangle$ versus $\lambda$ for two fixed geometries.</td>
<td>66</td>
</tr>
<tr>
<td>5.9</td>
<td>Relative contributions to $\langle \Phi^2 \rangle$ of inside and outside edge spins.</td>
<td>67</td>
</tr>
<tr>
<td>5.10</td>
<td>$\langle \Phi^2 \rangle$ versus film thickness $b$ for two fixed geometries.</td>
<td>68</td>
</tr>
<tr>
<td>6.1</td>
<td>Computed values of $\tau_{\phi,R}$ and $\tau_{\phi,E}$ vs $\alpha$.</td>
<td>72</td>
</tr>
<tr>
<td>6.2</td>
<td>Normalized Ramsey dephasing times for $0.6 \leq \alpha \leq 1.2$ in steps of 0.1.</td>
<td>74</td>
</tr>
<tr>
<td>6.3</td>
<td>Computed values of $\tau_{\phi,E}(f_2)/\tau_{\phi,E}(f_2 \to \infty)$ vs $f_2$.</td>
<td>75</td>
</tr>
<tr>
<td>6.4</td>
<td>Ratio $\tau_{\phi,E}/\tau_{\phi,R}$ vs $\alpha$.</td>
<td>76</td>
</tr>
<tr>
<td>6.5</td>
<td>Computed decay function $g(t)$ versus $t/\tau_{\phi}$.</td>
<td>78</td>
</tr>
<tr>
<td>6.6</td>
<td>Computed dephasing times $\tau_{\phi}$ vs $\alpha$.</td>
<td>79</td>
</tr>
<tr>
<td>A.1</td>
<td>Spectral density of white and $1/f$ noise.</td>
<td>93</td>
</tr>
<tr>
<td>A.2</td>
<td>Histograms of white noise.</td>
<td>94</td>
</tr>
</tbody>
</table>
List of Tables

3.1 Dimensions and inductances of SQUIDs I and II. . . . . . . . . . . . . . . . . . . . . 34
Acknowledgments

As a graduate student at UC Berkeley, I have had the distinct pleasure of interacting with a large number of extremely talented colleagues, many of whom became close friends as well. In my early years, I worked on a joint project between Robert Dynes, Irfan Siddiqi, and John Clarke. During that time, I was under the close guidance of Shane Cybart, who introduced me to a wealth of experimental concepts and techniques that proved invaluable in my later work. Both Shane and Stephen Wu, a classmate who also worked on the project, became and have remained very good friends both inside and outside the lab.

After my transition from this first project to a new project headed by John, I became an official Clarkian, with all the rights, privileges, and honors thereunto appertaining. Of these, I will most fondly remember the daily lunch group, the core of which consisted of Sarah Busch, Michael Hatridge, Jedediah Johnson, and honorary members Shane and Stephen. Nathan Kelso, who brought a packed lunch far too often, is fondly remembered notwithstanding. Shortly after I joined the group, Emile Hoskinson arrived and he, Jed, and I would take regular early-morning and lunch-time bike rides, sharing cycling news and trash talk, as appropriate. Many of my fondest early memories revolve around this motley troupe.

Seemingly as soon as meeting these new friends, Nathan—followed by Michael, Sarah, and finally Jed—graduated and moved on, taking with them the customary dynamics of our group. In the year before Sarah moved across the country, we grew particularly close and I had come to rely on her professional expertise as well as her friendship and more mature perspective of graduate school. Similarly, I grew especially close to Jed during his final year as both of us grappled with planning our respective futures. His example, advice, and positive attitude were of no small comfort while I struggled during my own final years.

The departure of old members, however, was marked by an influx of new members. Andrew Fefferman, Sean O’Kelley, and Jeffrey Birenbaum joined the group to work with me on the same project. Andrew, our postdoc, was instrumental in reviving the old dilution refrigerator. Despite our overlapping work, I will often not explicitly mention the contributions of Jeff and Sean. This omission, rather than implying a lack of assistance from them, should instead be interpreted as their being key elements in every aspect of this work. I can also declare with certainty that having Jeff in the lab made the long hours and late nights considerably more bearable, despite our disagreements regarding the appropriate amount of flavor food should have.

Visiting members to the Clarke group upheld the long-standing tradition of being bright,
motivated, collegial—and international!—scientists and greatly enriched my experience. I enjoyed working with Ida Sogmaes (from Norway), Fred Øisjöen (Sweden), Koos Zevenoven (Finland), Paul SanGiorgio (USA, via Switzerland), Kai Buckenmaier (Germany), Anders Pedersen (Denmark), and Hui Dong (China). One regular visitor of whom I became particularly fond was Jørn Hansen, a professor in Denmark. To me, Jørn represented the ideal modern scientist, setting a standard to which I will always aspire to reach.

My list of collaborators whom I wish to acknowledge is long, for which I am grateful. At UIUC, Allison Dove, Gus Olson, and Zach Yosovits in Jim Eckstein’s group were instrumental, fabricating high-quality materials that formed the basis of many SQUIDs that we measured. Also at UIUC, Chris Nugroho and Vlad Orlyanchik in Dale Van Harlingen’s group did an incredible amount of fabrication work. I personally thank Chris for his tireless efforts, endless dedication, and professional acumen, none of which can be overstated in magnitude or significance. At NIST, Boulder, Gene Hilton, Sherry Cho, and Kent Irwin gratefully provided us with numerous SQUIDs, which they made custom for our application. In particular, I thank Gene for his last-minute deposition of SiNx on several sets of devices and I thank Sherry for being the most outgoing and engaging scientist I know. At MIT Lincoln Labs, Vlad Bolkhovsky, Danielle Braje, George Fitch, Matthew Neeley, and Will Oliver were responsible for fabricating numerous qubit-sized SQUIDs used in our measurements. Moreover, Will’s interest in our experiments was always a source of encouragement for me. The SQUIDs graciously provided to us by NIST and MIT were the basis of a fundamental shift in our understanding of flux noise.

I’m also happy to acknowledge the advice and expertise of Fred Wellstood, the early pioneer in the study of flux noise. At KIT, I received significant support in the theoretical understanding of flux noise from Clemens Müller, Sasha Shnirman, Gerd Schönh, Juan Atalaya, and Pablo Schad. For their interest and contagious enthusiasm, I thank John Martinis, Clare Yu, and Mike Weissman. I also sincerely thank Coenrad Fourie for his dedication and expertise in helping to crack a three-year problem that we had very nearly abandoned. With his help, we published one of our most well-received papers, landing us on the cover of Superconducting Science and Technology. He has been a joy to work with.

The colleague I thank most enthusiastically is my adviser, John Clarke. In the years since I joined his group, we have spent countless hours working together on a wide range of scientific problems, from engineering a cryostat to debugging a measurement to interpreting data to writing—and defending!—our results. I am truly fortunate to have benefited from his more than four decades working at the frontiers of science and advising the world’s brightest new scientists. I aspire to honor this relationship by taking the skills I have learned under his tutelage to my own frontiers.

Based on the fact that I am graduating as a sane human being, it can correctly be inferred that I had extensive support outside of my world of research. I first wish to acknowledge my friends on Field Theory: Thierry Botter, Jon Blazek, Jonas Kjäll, Ben MacBride, Ben Westbrook, Nic Ross, Jens Bardarson, Brad Cenko, and Daniel Flanigan, just to name a few. It’s amazing what one hour of soccer a week can do to one’s mental health. I thank my dear parents Elizabeth, David, and Johanna who, from my earliest memories, provided me...
examples of scientific excellence and the support to pursue my own dreams—as long as they were scientific in nature and preferably included a PhD! Kahalley, I thank for keeping me young. I am eternally grateful to my grandparents Hector and Lois for their encouragement and for inviting me into their home during my time in Berkeley. To my grandmother Gloria, I will always cherish our many dinners, giving us the opportunity to get to know each other not possible on different coasts, and your unfailingly positive attitude and pride of my accomplishments. To my new family Nélida, Reinaldo, Rosana, and Claudia, I thank you for embracing me as a son and brother and for sharing with me my greatest blessing.

My deepest and most heartfelt thanks go to the love of my life, Mariana. The completion of this thesis in reality marks a shared success and ends a shared struggle. Her unwavering devotion to me and my aspirations was a constant and unlimited source of support and comfort during hard times. Our time together and travels across the globe never failed to invigorate and inspire me. And thankfully, now she can no longer tease me for not being a Doctor!
Chapter 1

Introduction

1.1 Overview

In all observable phenomena there is an inevitable contribution from random fluctuations. Often, these fluctuations originate from outside the system and arise from our inability to make perfect measurements. However, nearly all physical systems at some level exhibit intrinsic fluctuations as well. Some of these intrinsic fluctuations originate from fundamental laws of physics, such as the random thermal motion of gas particles, whereas others can be reduced or eliminated, for example through better materials or device manufacturing. Due to the deleterious effects of both internal and external fluctuations, researchers often go to great lengths to understand and reduce them as much as possible.

In this thesis we focus our attention on the random fluctuations in magnetic flux intrinsic to two devices: the dc superconducting quantum interference device (SQUID) and superconducting quantum bit (qubit). The SQUID has been used as an ultra-sensitive detector of magnetic field since the 1960’s while the qubit, a much newer device, forms the basis for a potential form of computing that exploits the laws of quantum mechanics to achieve vastly superior computational efficiency. As we shall see, flux noise degrades the performance of both devices, and so it is our goal to shed light on this particular noise process with the ultimate goal of significantly reducing it.

In the remainder of this chapter, we introduce previous work and key concepts necessary to understand the work of this thesis. In Chapter 2 we begin with our experimental work and introduce our measurement system and flux noise characterization procedures. In Chapter 3 we review our measurements of the temperature dependence of flux noise and attempt to reconcile our results to existing theory. A subtle contribution to SQUID critical noise—critical current noise—of which we became aware in the course of our measurements, is described in Chapter 4. In Chapter 5, we transition to our theoretical work, where we first develop a computational model to predict the noise in both qubit and SQUID loops based on our current understanding of flux noise. Next, in Chapter 6 we present an analysis that quantifies the theoretical effects of flux noise on qubits. Finally, we conclude with some remarks regarding our current understanding of flux noise and consider promising future
1.2. SUPERCONDUCTIVITY

1.2 Superconductivity

Since the first observation of the phenomenon in 1911 by H. Kamerlingh Onnes in Leiden, superconductivity has formed the basis of an extremely rich field of both theoretical and experimental research. Onnes observed that certain metals, when cooled below a material-specific characteristic temperature $T_c$ on the order of a few kelvin (roughly $-270 \, ^\circ C$), exhibited zero electrical resistance. The origin of this astonishing phenomenon eluded researchers for decades until the breakthrough theory in 1957 by Bardeen, Cooper, and Schrieffer (BCS) [1]. Only the year before, in 1956, Cooper had shown that electrons in a metal with an arbitrarily small net attraction can form a bound pair [2]. The destruction of this bound state requires an energy $2\Delta_s$, which is zero at $T_c$ and maximum at $T = 0$. BCS showed that this energy gap leads to a condensed ground state ensemble-average wave function within the superconductor. In the years since, BCS theory has proven to be remarkably successful in describing a number of aspects of superconductivity in bulk materials, such as the temperature dependence of the gap energy and various thermodynamic properties; the Meissner effect, whereby magnetic flux is expelled from bulk superconductors cooled through $T_c$; and the critical magnetic field, above which superconductivity can no longer exist in a material.

1.2.1 Flux quantization

Of the diverse and complicated properties of superconductivity, we focus on one in particular, namely that the electrons in a superconductor condense into a single ensemble-average wave function, which is generally written as

$$\psi = |\psi(r)| e^{i\theta(r)}, \quad (1.1)$$

where $\theta$ is the phase of the electron pairs at $r$. Furthermore, $\psi$ is usually normalized such that $\psi^*\psi$ equals the number density of Cooper pairs $n_s(r)$, which is normally taken to be spatially constant.

Next, we consider the momentum of a Cooper pair.\footnote{This derivation closely follows that of Van Duzer and Turner [3].} Classically, it is

$$\mathbf{p} = 2m\mathbf{v} - 2e\mathbf{A}, \quad (1.2)$$

where $\mathbf{v}$ is the velocity of the pair, and $m$ and $e$ are the electron mass and charge, respectively. The total momentum for $n_s$ pairs condensed into the same state and, therefore, with the same momentum, is

$$n_s\mathbf{p} = n_s(2m\mathbf{v} - 2e\mathbf{A}). \quad (1.3)$$
Substituting the quantum mechanical momentum operator \(-i\hbar \nabla\), we compute the expectation value of \(\mathbf{p}\) as

\[
n_s \mathbf{p} = \langle \psi | -i\hbar \nabla | \psi \rangle = -i\hbar (i\nabla \theta) \langle \psi | \psi \rangle = n_s \hbar \nabla \theta
\]  

so that

\[
\mathbf{p} = \hbar \nabla \theta.
\]  

Noting that the pair-current density can be expressed as \(\mathbf{J} = -2n_s e \mathbf{v}\), we see that

\[
\hbar \nabla \theta = -\mathbf{J} \frac{m}{n_s e} - 2e \mathbf{A}.
\]  

Because \(\psi\) represents a physically observable quantity, it must be single-valued at any given position. In other words, the integrated phase gradient around any closed path must be zero or some integer \(n\) multiple of \(2\pi\):

\[
\hbar \oint \nabla \theta \cdot d\mathbf{l} = -\oint \left( \frac{m}{n_s e} + 2e \mathbf{A} \right) \cdot d\mathbf{l} = 2\pi \hbar n = \hbar n.
\]  

If we consider the example of a superconducting ring and take our contour path to be deep inside the superconductor, where the current density vanishes, then we find

\[
\hbar \oint \nabla \theta \cdot d\mathbf{l} = -2e \oint \mathbf{A} \cdot d\mathbf{l} = -2e \int_S (\nabla \times \mathbf{A}) \cdot d\mathbf{S} = -2e \oint_S \mathbf{B} \cdot d\mathbf{S} = -2e \Phi_S = \hbar n
\]  

by Stokes’ theorem, where \(\mathbf{B}\) is the magnetic field in the loop and \(\Phi_S\) is the total magnetic flux through the surface \(S\). Clearly, \(\Phi_S\) can take only discrete values that are integer multiples of the flux quantum

\[
\Phi_0 = \frac{\hbar}{2e} \approx 2.068 \times 10^{-15} \text{ Wb}.
\]  

As we will see, this periodicity \(\Phi_0\) in flux threading a superconducting loop will prove to be an essential part to the operation of SQUIDs and qubits. Moreover, the fact that \(\Phi_0\) is so small is central to explaining the high sensitivities of both devices.

### 1.2.2 The Josephson effect

A second effect originating in the existence of a macroscopic wave function within a superconductor was predicted in 1962 by B. D. Josephson, who showed that two superconductors separated by a thin insulating layer can support a supercurrent \(I_s\) via the quantum mechanical tunneling of Cooper pairs across the barrier [4, 5]. He showed the magnitude of \(I_s\) to be related to the phase difference \(\delta\) between the two superconducting electrodes up to a maximum of \(I_0\), the critical current of the junction,

\[
I_s = I_0 \sin(\delta).
\]  

\[ (1.10) \]
The following year, P. W. Anderson and J. M. Rowell experimentally verified what is now referred to as the dc Josephson effect [6]. Josephson also predicted that the presence of a voltage difference $V$ across the barrier would lead to a time-dependent phase,

$$\frac{\partial \delta}{\partial t} = \frac{2e}{\hbar} V \approx 2\pi \times 483.6 \text{ MHz/} \mu \text{V}, \quad (1.11)$$

which is known as the ac Josephson effect. While originally derived for tunnel barriers, subsequent analysis and experiments have shown that both effects are generally applicable to any weakly-coupled superconductors. For non-tunnel barriers, however, the dependence of $I_s$ on $\delta$ is often non-sinusoidal.

From just these two relations, we can form a quite detailed understanding of the operation of Josephson junctions. For a dc current bias $I$ below the critical current ($I \leq I_0$), we readily see that a static phase difference develops across the barrier. As soon as the junction is current biased above the critical current or is voltage biased, the dynamics rapidly become quite complicated due to the nonlinear current-phase relationship [Eq. (1.10)]. To calculate the time evolution of the junction, it is common to use the resistively and capacitively shunted junction (RCSJ) model. Here, we recognize that in addition to the supercurrent supported by the junction, additional channels of current flow exist through the inevitable resistance $R$ and capacitance $C$ that exist across the junction. The resistance can be internal to the junction or be an external shunt, and the capacitance results from the geometry of two closely spaced superconducting electrodes. The total current $I$ is the sum of the currents in these three channels:

$$I = I_0 \sin(\delta) + V/R + CdV/dt, \quad (1.12)$$

or, using Eq. (1.11) to write Eq. (1.12) in terms of phase,

$$I = I_0 \sin(\delta) + \frac{\hbar}{2e} \frac{\partial \delta}{\partial t} + C \frac{\hbar}{2e} \frac{\partial^2 \delta}{\partial t^2}. \quad (1.13)$$

We may notice that Eq. (1.13) is analogous to the differential equation describing the motion of a particle of mass $m$ moving in some viscous fluid with friction coefficient $\xi$ in some spatially varying potential $w(x)$ [7]:

$$m \ddot{x} + \xi \dot{x} = -\frac{\partial w(x)}{\partial x}. \quad (1.14)$$

In fact, if we recast Eq. (1.13) as

$$C \left( \frac{\hbar}{2e} \right)^2 \frac{\partial^2 \delta}{\partial t^2} + \frac{1}{R} \left( \frac{\hbar}{2e} \right)^2 \frac{\partial \delta}{\partial t} = -\frac{\hbar}{2e} \frac{\partial}{\partial t} \left[ -I_0 \cos(\delta) - I\delta \right] \quad (1.15)$$
then the analogy becomes obvious:

\[ x \rightarrow \delta \quad (1.16a) \]

\[ m \rightarrow C \left( \frac{\hbar}{2e} \right)^2 \quad (1.16b) \]

\[ \xi \rightarrow \frac{1}{R} \left( \frac{\hbar}{2e} \right)^2 \quad (1.16c) \]

\[ w(x) \rightarrow E_J[-\cos(\delta) - (I/I_0)\delta]. \quad (1.16d) \]

Here, the Josephson coupling energy \( E_J \equiv (h/2e)I_0 \) defines the characteristic energy scale of the potential, which we recognize as a tilted washboard potential. The depth of each well is \( E_J \) and the degree of tilt is proportional to \( I \). At zero current bias, the particle is trapped in a single well, oscillating at a frequency \( \omega_p \equiv (E_J/m)^{1/2} = (2eI_0/hC)^{1/2} \). Because the time-averaged phase is zero, no static voltage is observed. As the bias current increases past the critical current \( (I > I_0) \), the particle starts to roll down the washboard and a voltage develops across the junction. If the bias current is brought back to below the critical current, one of two distinct forms of behavior will occur: if the damping is large, the particle will progressively slow until it becomes trapped in a single well; if the damping is small, the particle will continue to roll down the washboard and a voltage will still be observed, despite the fact that \( I < I_0 \). This latter hysteretic effect is eliminated by increasing the viscous force, which, by Eq. (1.16c), is equivalent to lowering the junction resistance \( R \). Typically, this is accomplished by fabricating a resistive metal shunt across the junction. Stewart and McCumber introduced the parameter

\[ \beta_c \equiv \frac{2\pi I_0 R^2 C}{\Phi_0} \quad (1.17) \]

and independently found the criterion for nonhysteretic junctions to be \( \beta_c < 1 [8, 9] \). As hysteretic junctions have quite complicated noise properties, in this thesis we will consider only nonhysteretic junctions.

In the common case where the junction is highly overdamped \( (\beta_c \ll 1) \), we can readily solve the voltage dependence on bias current. Dropping the capacitance term, Eq. (1.13) becomes

\[ \frac{\hbar}{2e} \frac{\partial \delta}{\partial t} = I_0 R \left( \frac{I}{I_0} - \sin \delta \right). \quad (1.18) \]

We can solve this nonlinear first order differential equation by separating variables and integrating. In particular, we use the fact that \( \delta \) is periodic at the angular frequency \( 2\pi/\tau \), where \( \tau \) is the period of a single oscillation. For \( I > I_0 \), the right-hand side of Eq. (1.18) is always positive and \( \delta(t) \) is constantly increasing. Therefore, \( \delta(t + \tau) = \delta(t) + 2\pi \). Using this
1.3. SUPERCONDUCTING DEVICES

Figure 1.1: Normalized current–voltage characteristics of Josephson junction (solid blue) and resistor (dashed red). At large current biases \((I \gg I_0)\), the two curves converge.

In fact, we find that

\[
\int_0^\tau dt = \int_0^{2\pi} d\delta \frac{\hbar}{2e I_0 R} \left( \frac{I}{I_0} - \sin \delta \right)^{-1} \tag{1.19}
\]

\[
\tau = \frac{\hbar}{2e I_0 R} \left[ \left( \frac{I}{I_0} \right)^2 - 1 \right]^{-1/2}. \tag{1.20}
\]

Relating the voltage to the time average phase via Eq. (1.11), we see that

\[
V(I \geq I_0) = R \left( I^2 - I_0^2 \right)^{1/2}, \tag{1.21}
\]

as shown in Fig. 1.1. Note that for \(I \gg I_0\), the bulk of the current is carried through the shunt resistor \(R\) and Eq. (1.21) reduces to \(V \approx IR\).

1.3 Superconducting devices

1.3.1 The dc SQUID

We have seen in Eq. (1.8) that the phase gradient integrated along a closed contour within a superconducting ring is related to the total flux \(\Phi_{\text{TOT}}\) threading the ring,

\[
\oint \nabla \theta \cdot dl = -\frac{2e}{\hbar} \int_S \mathbf{B} \cdot d\mathbf{S} = -2\pi \frac{\Phi_{\text{TOT}}}{\Phi_0} = 2\pi n, \tag{1.22}
\]
which leads to flux quantization. If, however, the loop is interrupted by one or more Josephson junctions, then the contour integration must include the phase differences across the junctions in addition to the accumulated phase generated by the presence of magnetic flux. Such a configuration is commonly known as a superconducting quantum interference device (SQUID), a name whose origin becomes clear when we consider a ring with two junctions.

The accumulated phase around the clockwise contour $C$, shown in Fig. 1.2(a), includes the phase differences$^2$ across the junctions $\delta_1$ and $\delta_2$:

$$\oint_C \nabla \theta \cdot \, d\mathbf{l} = \delta_2 - \delta_1 - 2\pi \frac{\Phi_{\text{TOT}}}{\Phi_0} = 2\pi n, \quad (1.23)$$

or

$$2\pi \frac{\Phi_{\text{TOT}}}{\Phi_0} \pmod{2\pi} = \delta_2 - \delta_1. \quad (1.24)$$

For simplicity, we begin by ignoring the self-inductance of the SQUID loop. In this case, the total flux is equal to the applied flux ($\Phi_{\text{TOT}} = \Phi_a$). If the junctions are identical ($I_{0,1} = I_{0,2} = I_0$), from Eq. (1.10) we see that the maximum supercurrent $2I_0$ is supported for $\delta_1 = \delta_2 = \pi/2$, which satisfies Eq. (1.24) only if $\Phi_a$ is an integer multiple of $\Phi_0$. For

---

$^2$In general, the phase differences across the junctions are measured in the same direction as the positive current bias. Because the directional contour passes one junction in the positive direction and the other negative direction, one of the phases necessarily picks up a negative sign. This choice is arbitrary and does not affect any of the physics.
bias currents above the SQUID critical current $I_c = 2I_0$, a voltage will develop across the SQUID. If instead $\Phi_a = (n + 1/2)\Phi_0$, then $\delta_2 - \delta_1 = \pi$ and

$$\frac{I_{s,1} + I_{s,2}}{I_0} = \sin(\delta_1) + \sin(\delta_2) = \sin(\delta_1) + \sin(\delta_1 + \pi) = \sin(\delta_1) - \sin(\delta_1) = 0.$$  \hfill (1.25)

In this situation, no net supercurrent can flow across the device and $I_c = 0$. In this way, the critical current oscillates between its maximum and minimum values with a periodicity in flux of $\Phi_0$, as shown in Fig. 1.2(b). This effect is analogous to the two-slit experiment with incident plane waves, where we equate the phases $\delta_1$ and $\delta_2$ to the path lengths of the waves, $\Phi_a$ to the perpendicular distance along the wall, and $I_c$ to the beam intensity; $\Phi_0$ then becomes the fringe spacing.

At flux biases other than $\Phi_a = n\Phi_0$, the phases differences across the junctions are necessarily not equal ($\delta_1 \neq \delta_2$) in order to satisfy Eq. (1.24). Correspondingly, the currents passing through the junctions are not equal, which means that some current $J$ circulates the loop as shown in Fig. 1.2(a). This circulating current introduces a flux $JL$ into the loop via the loop’s self-inductance $L$. Accounting for the loop inductance, $\Phi_{\text{TOT}}$ is the sum of the applied flux $\Phi_a$ and the induced flux $JL$: $\Phi_{\text{TOT}} = \Phi_a + JL$. Introducing a unitless self-inductance parameter $\beta_L \equiv 2I_0L/\Phi_0$, Eq. (1.24) becomes (dropping the modulo $2\pi$)

$$2\pi \frac{\Phi_{\text{TOT}}}{\Phi_0} = 2\pi \frac{\Phi_a}{\Phi_0} + \pi \beta_L \frac{J}{I_0} = \delta_2 - \delta_1.$$  \hfill (1.26)

For $\beta_L \ll 1$, the effects of the loop inductance can be neglected ($\Phi_{\text{TOT}} \approx \Phi_a$). In this limit, we see that the normalized modulation depth is independent of $\beta_L$ or, in other words, the flux sensitivity at a particular flux bias is proportional to the critical current: $\partial I_c/\partial \Phi_a \propto I_c$ [Fig. 1.2(b)]. As $\beta_L$ approaches unity, however, the flux contribution due to the circulating current becomes significant and the $I_c-\Phi_a$ characteristics are altered. In fact, if $\beta_L > 1$, multiple values of $J$ exist that satisfy Eq. (1.26). This situation can introduce nonlinearities and complex dynamics, which we will not consider in this thesis; all of our SQUIDs will have $\beta_L < 1$ and often $\beta_L \ll 1$.

This “interference” and corresponding periodicity in flux exists for all superconducting rings interrupted by Josephson junctions. We have just considered the case of a ring with two junctions, which is commonly known as the dc SQUID because it is often biased using static (or dc) currents. In contrast, a ring with a single junction cannot be biased as such because the static current would simply pass through the superconducting arm with no junction; this device must be operated at high frequencies and is correspondingly known as an rf SQUID. The great sensitivity of both devices can be seen to originate in the periodicity in $\Phi_0$, which is a very small flux. The earth’s magnetic field through a SQUID loop of radius 0.5 mm, for example, corresponds to roughly 15,000 $\Phi_0$. Put another way, in this device a flux quantum is equivalent to a magnetic field of just 2.6 nT. As we shall see, optimized SQUIDs can resolve changes in flux on the order of $\mu \Phi_0$, or magnetic fields of order fT. Through suitable designs, this exquisite sensitivity and resolution can be used to read other physical quantities. For example, an inductively coupled input coil fabricated on top of a SQUID washer can readily be used to measure currents of order pA.
1.3.2 Qubits

Much as the name implies, quantum bits (qubits) are the quantum analog to classical bits. Whereas classical bits can take on only one of two values, a qubit state $|\Psi\rangle$ can be an arbitrary linear superposition of two quantum states, say, $|0\rangle$ and $|1\rangle$,

$$|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle,$$

where $\alpha$ and $\beta$ are properly normalized complex probability amplitudes. Analogous to the binary logic gates in classical computing, quantum logic gates operate on the qubit states. However, with no analog to classical computing, qubits can be prepared in entangled states. These two properties—superposition and entanglement—empower quantum circuits to implement algorithms that are vastly more efficient at certain tasks than their classical counterparts. The oft-cited Shor’s algorithm, for instance, is significantly faster at integer factorization than classical algorithms [10].

The near-dissipationless environment and existence of a macroscopic quantum wavefunction in superconducting structures suggest that these circuits are natural candidates for implementing qubits. Importantly, the existence of dissipative thermal quasiparticle excitations in superconductors drops exponentially with temperature and are often negligible at dilution refrigerator temperatures. These conditions were recognized as precursors to macroscopic quantum tunneling [11], which was expected to be the dominant flux transition mechanism of SQUIDs below approximately 0.1 K [12]. The macroscopic quantum nature of the phase variable $\delta$ was demonstrated convincingly through a series of experiments on Josephson junctions in the 1980’s [13–16]. These measurements showed that a phase particle trapped in a well in the washboard potential exhibited discrete energy levels. Furthermore, the particle could tunnel through the potential barrier separating an adjacent well at a lower potential.

Since the initial demonstrations of macroscopic quantum coherence, the superconducting qubit field has matured tremendously. A large number of qubit implementations now exist, which can be roughly classified as charge, flux, and phase qubits. The interested reader is referred to Refs. 17 and 18 for detailed reviews of the various flavors of qubits and underlying theory. In this thesis we are primarily concerned with flux-based qubits.

In particular, we consider the three-junction qubit, shown schematically in Fig. 1.3(a). In this qubit, the two quantum states correspond to the magnetization of the qubit: magnetic flux $\Phi$ pointing up $|\uparrow\rangle$ or down $|\downarrow\rangle$. Alternatively, we can think of the states as a supercurrent $I_q$ that is either counterclockwise $|\circlearrowleft\rangle$ or clockwise $|\circlearrowright\rangle$. At an applied flux bias $\Phi_a = \Phi_0/2$, the circulating states are degenerate in energy. In this case, the energy eigenstates, which are separated by energy $\Delta$, are the symmetric (ground state) and anti-symmetric (first excited state) combinations of counter-rotating circulating currents. As the flux bias is moved away from $\Phi_0/2$, the energy splitting $\nu = (\Delta^2 + \varepsilon^2)^{1/2}$ diverges hyperbolically with $\varepsilon \equiv 2I_q(\Phi_a - \Phi_0/2)$ as shown in Fig. 1.3(b). A key advantage of the flux qubit is the tunability of $\nu$ via flux bias. In addition, the high anharmonic energy levels of the qubit allows the controlled transition between specific energy states. Principal challenges...
1.4 Noise

1.4.1 Generic noise processes

In general, we classify any process that varies randomly in time to be “noisy.” Frequently, these noise processes are unwanted and can obscure signals, which may carry information one desires to extract. Through suitable design, it is often possible to reduce the deleterious impact of these random fluctuations by either reducing their magnitude or amplifying the signal. In a large number of situations, however, the noise may originate from a fundamental physical process that is not possible to alter. The discrete charge of electrons, for instance, gives rise to shot noise, which is significant when measuring small electric currents [19]. Johnson-Nyquist noise originates from the thermal agitation of charge carriers within a resistive element [20, 21], a phenomenon that was subsequently shown to be a specific case of the general fluctuation-dissipation theorem [22, 23].

While noise processes are observed in the time domain, they are often most easily analyzed in the frequency domain. The conversion from the time to frequency domain of a discretely sampled signal is most commonly achieved by computing the Fast Fourier Transform (FFT) of the signal (see Appendix A), which gives us the approximate spectral density of the signal versus frequency $S(f)$, the spectrum. By viewing the noise spectrum, we can quickly distinguish between two important and extremely common types of noise. “White” noise has a frequency-independent spectral density ($S(f) \propto \text{const}$), whereas “1/f” noise varies...
1.4. NOISE

Figure 1.4: White and $1/f$ noise. (a) Computer-generated white and (b) $1/f$ noise in the time domain. (c) Spectral density of time series shown in (a). (d) Spectral density of time series shown in (b).
1.4. NOISE

Figure 1.5: The role of random telegraph signals in 1/f noise. (a) Computer-generated RTS in the time domain. (b) Spectral density of RTS shown in (a). (c) The spectral densities of nine RTS signals (blue lines) and their sum (red line) are shown. Here, $1/\tau_i \in \{10^{-4}, 10^{-3}, \ldots, 10^4\}$ so that the sum varies as $1/f$ (dashed black line).

roughly inversely with frequency ($S(f) \propto 1/f$). In the time domain, white-noise processes are uncorrelated between successive samples, but 1/f-noise processes display long time-scale correlations. These properties are shown schematically in Fig. 1.4.

In this thesis we shall be primarily interested in 1/f-noise processes. Given the ubiquity of 1/f noise, it may come as no surprise that it can often be explained in terms of a simple model [24]. Consider first an individual two-level system, such as a particle in a double-well potential. If the energy levels are approximately equal and the barrier is low enough, the particle may hop between wells via thermal activation or quantum tunneling. These hops, which occur randomly with an average time $\tau$ between hops, give rise to a random telegraph signal (RTS) [Fig. 1.5(a)]. The spectral density of such an RTS (assuming the two levels are 0 and 1) is given by

$$S_{RTS}(f, \tau) = \frac{1}{4\pi} \frac{2/\tau}{(2\pi f)^2 + (2/\tau)^2}$$  \hspace{1cm} (1.28)

and is often referred to simply as a “Lorentzian” [Fig. 1.5(b)]. For frequencies much less than the characteristic flipping rate ($f \ll 1/\tau$), the spectral density is constant ($S_{RTS} = \tau/8\pi$), whereas for frequencies much greater than the characteristic flipping rate ($f \gg 1/\tau$) the spectral density scales as $1/f^2$.

If a system consists of many such uncorrelated fluctuations, the total noise is naturally the sum of the noise of all individual fluctuators. Furthermore, if the distribution of flipping rates $\tau_i$ varies as $1/\tau$—that is, an equal number of fluctuators per decade—then the resulting spectral density scales as $1/f$ [Fig. 1.5(c)]. A thermally activated process with a constant distribution of barrier heights $E_i$ is a common origin of such a distribution in $\tau_i$, since the flipping rate $1/\tau_i \propto \exp(-E_i/k_B T)$. Note that a varying density of barrier heights will give rise to noise with a spectral density that varies as $1/f^\alpha$, where $\alpha \neq 1$ [25]. In this way,
we see an intimate connection between a $1/f$-like noise spectrum and underlying individual fluctuators; if one hopes to reduce a particular noise source, it is essential first to understand the nature of the fluctuators.

### 1.4.2 Noise in SQUIDs and qubits

A systematic study of noise in superconducting devices was published in 1976 by Clarke and Hawkins, who studied noise in tunnel junctions [26]. They characterized two distinct types of noise: at low frequencies, the voltage noise spectrum of current-biased junctions in the voltage state varied roughly as $1/f$ but was dominated by a white noise at high enough frequencies. They identified the white noise as originating from Johnson noise in the resistive shunt, added to prevent hysteresis on the $I-V$ characteristic; the $1/f$ noise resulted from fluctuations in the critical current of the junctions. Subsequent theory and measurements have convincingly shown critical current fluctuations to originate from the random trapping and release of electrons in the junction barrier, which locally decrease the conductivity and cause the critical current to fluctuate [27–35].

Concurrently, Clarke et al. studied noise in dc SQUIDs [36]. Naturally, the SQUID exhibited a similar voltage noise to that of Josephson junctions. In addition, they observed a magnetic flux noise component that exceeded their predictions by two orders of magnitude. Subsequent measurements by Koch et al. resolved this discrepancy by identifying a third noise source [37]. They found evidence that an apparent flux $1/f$ noise is generated by the SQUIDs themselves with a magnitude that greatly exceeded the contribution from critical current noise within the junctions. In the following years, Wellstood et al. characterized the geometry, material, and temperature dependence of the flux noise [38–40].

Flux qubits exhibit both critical current and flux $1/f$ noise. However, because the junctions in qubits are not resistively shunted, there is no white noise component. While critical current noise is capable of causing decoherence in qubits [41], it is not currently believed to limit coherence times [35]. Flux noise, however, is known to be a dominant source of dephasing in flux-sensitive devices [18, 42–50].

### Decoherence in qubits

For a qubit to be effective, its quantum state must be sensitive in a predictable way to experimentally controlled degrees of freedom. In addition, it must be largely insensitive to all other degrees of freedom. One of the principal advantages of the superconducting qubit—namely, the facility with which its state is controlled externally—also presents one of its largest obstacles. Due to its strong coupling to numerous degrees of freedom, the superconducting qubit is extremely susceptible to decoherence through a number of mechanisms, both internal and external to the qubit. Here, it is typical to borrow the formalism originally developed for nuclear magnetic resonance (NMR) spectroscopy [51], where we map the NMR spin polarization (up or down) and angle to the static applied field to the qubit state (ground or excited) and phase on the Bloch sphere [52].
Two distinct mechanisms of decoherence are characterized by times $T_1$ and $T_2$. The random relaxation of the qubit from its excited state to its ground state occurs in time $T_1$. This process necessarily involves an exchange of energy between the qubit and its environment. Importantly, noise at the frequency equal to the energy gap of the qubit states can stimulate both the relaxation and excitation of the qubit. The time over which the phase of the qubit is randomized is $T_2$, which is related to $T_1$ by

$$\frac{1}{T_2} = \frac{1}{2T_1} + \frac{1}{\tau_\phi}, \quad (1.29)$$

where $1/\tau_\phi$ is the pure dephasing rate. The pure dephasing rate is often dominated by $1/f$-noise processes, which, in flux-sensitive devices, is primarily flux $1/f$ noise. In contrast to $T_1$ processes, the dephasing rate is susceptible to noise over a broad bandwidth and does not involve energy exchange. The two fundamental decoherence processes ($T_1$ and $\tau_\phi$) are illustrated schematically on the Bloch sphere in Fig. 1.6.

### 1.5 Flux noise

#### 1.5.1 Previous measurements

Flux $1/f$ noise was first identified as an independent noise source in dc SQUIDs by Koch et al. [37]. They found that the flux noise was approximately independent of SQUID
1.5. FLUX NOISE

size for effective loop areas spanning six orders of magnitude, a fact implying that the noise is generated locally to the SQUID. Detailed flux noise measurements performed by Wellstood et al. on dozens of SQUIDs fabricated on Si, quartz, and sapphire substrates with loops made from Nb, Pb, and PbIn revealed that flux noise was relatively insensitive to materials and temperature, as well as to loop area [38-40]. They characterized spectra by fitting them to the functional form $S_\Phi(f) = A^2/(f/1\,\text{Hz})^\alpha$, where they found $S_\Phi(1\,\text{Hz}) = A^2 \approx 7\pm3\,\mu\Phi_0/\text{Hz}^{1/2}$ and $0.5 \lesssim \alpha \lesssim 1$ at a temperature of about 0.1 K. Many subsequent measurements have confirmed this frequency scaling and noise magnitude. [38, 50, 53, 54].

1.5.2 Surface spin model

A microscopic model put forth by Koch et al. explained flux noise as arising from the random reversal of electron dipoles that were uniformly distributed on the surface of the substrate [55]. As each dipole changed orientation, the amount of flux in the inductively coupled SQUID loop correspondingly changed, generating an RTS. For a broad distribution of flipping rates, these fluctuations could give rise to the $1/f$-like spectral density observed experimentally. By numerically computing the average dipole coupling strength and by making certain assumptions regarding the spectral density, they inferred a surface spin density of $\sigma = 5 \times 10^{17}/\text{m}^2$ or about one spin per $2\,\text{nm}^2$.

Such a high density of spins was initially thought unlikely. However, two independent measurements of paramagnetism have since corroborated this value. First, Sendelbach et al. measured the temperature-dependent magnetism in a dc SQUID cooled in a perpendicular magnetic field [54]. They found that the flux offset in the SQUID scaled with temperature $T$ as $1/T$ and that the scaling factor was proportional to the field intensity in which the SQUID was cooled. Their interpretation of these observations was that unpaired surface spins were coupling to the vortices trapped in the SQUID loop, forming a paramagnetic system that followed Curie’s law. Furthermore, they inferred an unpaired spin density of exactly $5 \times 10^{17}/\text{m}^2$. One year later, Bluhm et al. performed scanning SQUID measurements of the paramagnetism of Au rings deposited on a Si (with native oxide) substrate [56]. They measured a similar spin-1/2 density of $4 \times 10^{17}/\text{m}^2$ and, importantly, observed no paramagnetic behavior above the unmetalized substrate. This latter result implies that the spins exist only at the metal-insulator interface.

The surface spin model forms a key assumption of many subsequent analyses [45, 54, 57-59]. Moreover, recent experiments on anti-correlations of flux noise have confirmed the surface spin model [48, 49]. An unambiguous understanding of the mechanism by which the spins produce $1/f$ flux noise, however, has yet to be developed. Recent proposals include spin clusters [60], spin glasses [61], fractal spin clusters [62], and hyperfine interactions [63]; the models in Refs. [61, 62] suggest that $\alpha$ may differ from unity.
1.5.3 Proposed origin of spins

One leading theory of the origin of the unpaired surface spins is that they originate from Metal-Induced Gap States (MIGS), which can be understood as follows. From elementary quantum mechanics, the probability amplitude of a particle with energy $E$ approaching a step barrier of constant potential $V_0 > E$ will decay exponentially inside the barrier. In an analogous situation, we consider the interface of a metal and insulator (or semiconductor). An electron state with an energy in the conduction band of the metal, but in the band gap of the insulator, will similarly decay exponentially—with oscillations due to the periodicity of the lattice—in the insulator [64]. In the absence of disorder the MIGS are occupied by paired electrons (spin up and down) and not localized. However, if the interface is disordered—as it is expected to be in physical systems—the picture changes. Choi et al. ran numerical computations where they randomly varied the on-site energy within two atomic layers of the interface and found that this disorder created states with a large Hubbard $U_i$, the energy cost of double occupation [65]. They also found that MIGS with a large $U_i$ tended to localize spatially. Finally, they found that only a modest amount of disorder was sufficient to generate the experimentally observed unpaired spin density of $5 \times 10^{17}/m^2$.

1.5.4 Flux noise in qubits

Flux noise is a major source of dephasing [18] in superconducting flux [44, 46–49] and phase quantum bits (qubits) [42, 45, 50] and in the quantronium [43]. Furthermore, reported values of $A^2$ in qubits are remarkably similar to those in SQUIDs [44, 45, 48, 66–68], despite the fact that the qubits were typically only a few micrometres in size. Interestingly, the $1/f$-like noise has been identified in qubits across a bandwidth spanning at least 11 decades [66–68].
Chapter 2

Experimental measurement system and procedures

2.1 Overview

We begin with a description of the experimental characterization of flux noise in SQUIDs. The success of this endeavor relies on a robust experimental measurement system that satisfies a number of stringent requirements. First and foremost, because we seek to measure accurately the intrinsic noise in SQUIDs, which are themselves ultra-low-noise devices, every care must be taken to insure that the measurement system produces as little noise as possible. In this same thread, $1/f$-noise measurements at low frequencies are necessarily slow measurements, implying that the system must be ultra-stable over long time scales (on the order of hours, in our case). As noise originating from myriad sources—electronic and mechanical as well as temperature and magnetic field fluctuations, for example—can couple into the measurement, the design considerations are substantial.

We also seek to characterize the flux noise as a function of device temperature, which means that we need a refrigeration apparatus that can maintain a stable temperature from approximately 0.1 K up to 4 K. Standard dilution refrigerators are quite capable of reaching such temperatures, but we shall see that care is needed in order to achieve the required stabilities, which are better than 1 part in $10^4$.

For a number of reasons, it is also advantageous for the system to have the ability to cool multiple SQUIDs in a single cool-down and to measure each one individually. The length of the cool-down and warm-up procedures of our dilution refrigerator means that considerable time is saved compared with cooling each SQUID individually. Furthermore, cooling several SQUIDs simultaneously eliminates many potential confounding factors—trapped magnetic fields, for example—that could conceivably cause differences between measurements from different cool-downs.

In this Chapter we review the experimental design and implementation of our system as well as the validation, calibration, and measurement procedures.
2.2 Measurement overview

The principle behind our measurement technique is that a SQUID biased into the normal state generates fluctuations based on changes in its critical current $I_c$ and the flux $\Phi$ through its loop and the transfer coefficient $(\partial I_c/\partial \Phi)_b$. Naturally, these fluctuations are quite small and cannot be read directly using even the best room-temperature amplifier. Therefore, we use a second SQUID, operated in a standard flux-locked loop (FLL) configuration, to detect the fluctuations. The fluctuations of the measured SQUID are monitored for long periods of time, converted to spectral densities, and subsequently analyzed.

The basic configuration of our system is the circuit schematic illustrated in Fig. 2.1, where we have modified the design used by Wellstood [40], which was designed to measure a single SQUID, in order to accommodate multiple measured SQUIDs. In our circuit, several SQUIDs are connected in series with a small compensating resistor $R_c$, the input coil to the
2.2. MEASUREMENT OVERVIEW

readout SQUID, and a choke inductor, which prevents high-frequency cross-talk between the measured and readout SQUIDs. Appropriate wiring to the circuit allows us to inject currents across the compensating resistor and across each of the SQUIDs independently. To make a measurement, we begin by injecting a current $I_b$ through a particular SQUID and increasing it until the SQUID enters the normal state with voltage $V_{\text{SQUID}}$. At this point, most of the current bias $I_b$ passes through the measured SQUID, but a small fraction $V_{\text{SQUID}}/I_b R_c$ is shunted through the compensating resistor. If $I_b$ is increased by $\delta I_b$, a fraction $R_c/(R_d + R_c)$ of $\delta I_b$ passes through the compensating resistor, where $R_d$ is the dynamic resistance of the measured SQUID, which depends on the bias current. For reasons we discuss later, $R_c \ll R_d$ by design. If the shunted current that flows around the big loop exceeds the critical current of any other SQUID in the circuit, that SQUID will be driven normal, thereby adding its own intrinsic noise. To prevent this, a current $I_r$ is passed through the compensating resistor, as in Fig. 2.1, until the voltage across $R_c$ equals that across the SQUID: $V_{\text{SQUID}} = I_r R_c$. In this scenario, no net current flows around the loop, ensuring that SQUIDs not under measurement remain superconducting and contribute no noise.

Under these conditions, a variation in the critical current of the measured SQUID will redistribute the bias currents to the circuit. For instance, if the critical current decreases, a fraction of $I_b$ will be diverted around the big loop, through $R_c$ and the input coil, thereby introducing a flux offset in the readout SQUID. The FLL, which is modulated above 100 kHz, rapidly cancels this offset by changing the current to the feedback coil of the readout SQUID (Fig. 2.1) until the flux offset is canceled. A voltage $V_{\text{FLL}}$ proportional to the feedback current and, correspondingly, to the current circulating in the loop is read at the output of the FLL. For small variations in the measured SQUID, the response is linear so that $V_{\text{FLL}}$ is also proportional to the critical current of the SQUID. We now discuss the conversion factors as well as our choice of circuit and device parameters, chosen to optimize sensitivity to the fluctuations.

Operated in this configuration, any single measurement is insufficient to discern the origin of the fluctuations that we measure. The measured noise could originate from intrinsic critical current noise within the junctions; flux noise in the SQUID, which, in turn, modulates the critical current of the SQUID; or noise in the readout SQUID or the measurement and bias circuitry. Therefore, to verify our SQUID noise measurements, we must first verify that the noise of our bias and readout circuitry is negligible. Next, in order to distinguish between critical current and flux noise, we utilize the property of the SQUID whereby its sensitivity to flux $(\partial I_c/\partial \Phi)$ varies as a function of its flux bias [see Sec. 1.3.1 and Fig. 1.2(b)]. By varying the flux bias and corresponding sensitivity of the SQUID, we can verify which of the various noise spectra represent flux noise, a concept which we explain in more detail in Sec. 2.5.

2.2.1 Optimal circuit and device parameters

In general, we choose parameters to maximize the change in $V_{\text{FLL}}$ for a given change in critical current or flux in the measured SQUID. First, we discuss the value of the compensat-
2.2. MEASUREMENT OVERVIEW

Figure 2.2: Bias configuration and operating points of measured SQUID. (a) Simplified schematic of the typical biasing of a measured SQUID. (b) Current-voltage characteristics of a SQUID (red) for a small change in critical current $\delta I_c$. The load line for $R_c$ (blue) is calculated as follows. For zero current flowing through $R_c$, $V_{Rc} = 0$ and the current through the SQUID, $I_{\text{SQUID}}$, is maximum: $I_{\text{SQUID}} = I_t$. For $I_{\text{SQUID}} = 0$, the total current flows through $R_c$ and $V_{Rc} = I_t R_c$. The steady-state solution occurs when $V_{\text{SQUID}} = V_{Rc}$, which is the intersection of the two curves. Therefore, $\delta I_c$ causes the operating points to change little in voltage, but maximally in current.
ing resistor. In Fig. 2.2(a) we show a simplified schematic of the typical biasing of a SQUID, omitting the superconducting SQUIDs; here, \( I_t = I_r + I_b \) is the total current applied to the circuit. Ultimately, we seek to maximize the change in current through the input coil for a given change in SQUID critical current, \( I_c \). In this familiar circuit, we recognize that if \( R_c \gg R_d \), the SQUID is effectively current biased at the relevant low frequencies. That is, if the critical current of the SQUID decreases slightly, \( V_{\text{SQUID}} \) will increase by \( \delta V \) and the current through \( R_c \) will increase by only \( \delta V/R_c \), which, by construct, is small. However, if \( R_c \ll R_d \), the SQUID is effectively voltage biased. We show this situation in Fig. 2.2(b), using the concept of load lines because the SQUID is a nonlinear circuit element. In this scenario, a decrease \( \Delta I_c \) in the \( I_c \) of the SQUID corresponds to a maximum increase of current through \( R_c \) and, correspondingly, through the input coil of the readout SQUID \( \Delta I_{\text{loop}} \),

\[
\Delta I_{\text{loop}} = \frac{1}{1 + R_c/R_d} \Delta I_c. \tag{2.1}
\]

For typical values of \( R_d \sim 10 \, \Omega \), \( R_c \) should be on the order of 1 \( \Omega \). From Eq. (2.1), we see that values of \( R_c \) much less than \( R_d/10 \) do not offer much additional signal and can, in fact, degrade the performance of the system if \( R_c \) is too low. We also see that for \( R_c \ll R_d \), \( \Delta I_{\text{loop}} \approx \Delta I_c \) and, therefore, throughout this thesis we will often use \( \Delta I_{\text{loop}} \) and \( \Delta I_c \) interchangeably. We note that this analysis is true only if \( R_c \) is the dominant resistance in the loop when all SQUIDs are superconducting.

With a small value of \( R_c \), we ensure that fluctuations in the readout SQUID generate maximum current fluctuations through the input coil of the readout SQUID. In turn, we maximize the flux coupled into the readout SQUID by using the largest feasible mutual inductance for the input coil \( M_i \). In general, \( M_i \sim 10 \, \text{nH} \), or \( M_i^{-1} \sim 0.1 \, \mu\text{A}/\Phi_0 \). Conversely, we aim to make the mutual inductance of the feedback coil \( M_f \) small (\( \sim 0.5 \, \text{nH} \)) so that the feedback current is relatively large. Finally, the feedback resistor \( R_{\text{feedback}} \), which sets the conversion factor between the feedback current and \( V_{\text{FLL}} \), is chosen to be large (\( \gtrsim 100 \, \text{k}\Omega \)). Thus, \( V_{\text{FLL}} \) is related to \( I_{\text{loop}} \) via

\[
V_{\text{FLL}} = R_{\text{feedback}} \left( \frac{M_i}{M_f} \right) I_{\text{loop}}, \tag{2.2}
\]

and the FLL has a transresistance on the order of \( 10^6 \, \text{V/A} \).

It is also possible to optimize the junction parameters of the SQUIDs. By varying the value of the junction shunt resistance \( R \) and critical current of the junctions, we can modify the level of white noise and flux sensitivity, respectively. At high enough frequencies—typically, above 10 to 100 Hz when the SQUID is biased at maximum flux sensitivity—the white noise generated by the shunt resistors dominates both the flux and critical current \( 1/f \) noises of the SQUID. As a current, the white noise magnitude scales as \( R^{-1} \) [69], which indicates that we can reduce the white noise level by increasing \( R \). Because \( \partial I_c/\partial \Phi \) is roughly proportional to \( I_c \) (Sec. 1.3.1), we can also increase the sensitivity of our measurement to flux in the measured SQUID. These two criteria suggest that one should increase \( R \) and
2.3. IMPLEMENTATION

$I_c$ as much as possible in order to maximize the relative current variations due to flux noise. However, our measurements require that the SQUIDs are non-hysteretic, that is, $\beta_c \equiv 2\pi I_0 R^2 C / \Phi_0 < 1 \ [8, 9]$, so that $R$ and $I_0$ cannot both be increased without bound. Here, $I_0 = I_c / 2$, $R$, and $C$ are the average junction critical current, resistance, and capacitance, respectively. For fixed critical current density $j_c$ and constant capacitance per unit area $C/A$, $I_0$ is typically varied by changing the junction area $A$ so that $I_0 R^2 C \propto A^2 R^2$. Assuming that one designs a SQUID for fixed $\beta_c < 1$, this constraint also implies a fixed product $I_0 R$. Because the white noise spectral density decreases only as $R^{-1}$ but the flux sensitivity increases as $(\partial I_0 / \partial \Phi)^2 \propto I_0^2$, we see that it is most advantageous to increase $I_0$ as much as possible, even at the cost of decreasing $R$. In our experiments $I_0$ is typically between 5 and 10 $\mu$A and $R$ is around 20 $\Omega$.

2.3 Implementation

We now discuss the specifics of how we implement the measurement system, shown schematically in Fig. 2.3. The SQUIDs to be measured, which are typically fabricated onto a single chip, are first mounted and wire-bonded onto a circuit board in one of the chambers of a copper sample box (Fig. 2.4). Flux bias is provided via either an on-chip coil or a coil mounted in the circuit board, directly below the chip. The compensating resistor, which is simply a short segment of resistance wire, is also mounted to the circuit board. The readout SQUID is located in another chamber on the opposite side of the box in order to prevent cross-talk between SQUIDs. Intermediate chambers contain the choke inductor and a cold transformer on the voltage line of the readout SQUID, which is used both to step up the voltage and to present a resistance to the room-temperature amplifier that minimizes the noise temperature. Once the sample is mounted into the box, a lipped Cu lid is attached. For magnetic shielding, the inner surfaces of both the box and lid are electroplated with approximately 200 $\mu$m of Pb.

The sample box is then mounted onto the cold finger of an Oxford Kelvinox 300 dilution refrigerator. In addition to the Pb plating of the sample box, further attenuation and shielding of external magnetic fields are provided by two concentric cans of Pb and Cryoperm. Electrical signals are carried via three 4-conductor cables with Reichenbach connectors for the various biasing of the measured SQUIDs and three SubMiniature version A (SMA) coaxial cables for the current bias, voltage readout, and flux feedback current of the readout SQUID. All lines are rf-filtered using standard copper powder filters (CPF) and low-pass filtered using discrete filters. Bias currents are provided via lithium-ion battery-powered boxes using one of two circuit designs, shown in Fig. 2.5. The current biases for $I_b$ and $I_c$ typically need to provide only tens of microamps and use the “passive” design [Fig. 2.5(a)]. The current bias for the flux bias of the measured SQUIDs, however, can require in excess of 10 mA, so that we use an “active” source [Fig. 2.5(a)]. Because noise in the bias boxes can cause noise in the measurement, we have taken great care to design sources with negligible low-frequency noise. Both designs rely on the extremely flat discharge voltage of Li-ion batteries, which allows us
Figure 2.3: Schematic of measurement system. The sample is mounted to the cold finger of a dilution refrigerator and is surrounded by a combination of Pb and Cryoperm shields for magnetic shielding. All lines are low-pass filtered using copper powder filters (CPF) and discrete filters. The typical configuration of bias boxes and readout equipment is shown. The shaded portion is connected only when measuring $\partial I/\partial \Phi$, not when acquiring noise data.
2.3. IMPLEMENTATION

Figure 2.4: Photo of sample box. Six chambers separate the various portions of the measurement circuit. The circuit board holding the measured SQUIDs is on the left. On the right is the board holding the readout SQUID. The chambers in the middle hold the choke inductor and cold transformer, wound around molypermalloy powder (MPP) cores, which work well at cryogenic temperatures.

Figure 2.5: Circuit schematics of current supply boxes. (a) “Passive” circuit design, type A. Li-ion battery sources a current controlled by a 10-turn potentiometer. This design is useful for providing currents $\lesssim 100 \, \mu$A. (b) “Active” circuit design, type B. The Li-ion battery and 10-turn potentiometer provide a stable voltage source input to the op-amp (ISL28134). $R_{\text{set}}$ sets the current output range, which, for this op-amp, has a maximum of 60 mA. Since the op-amp sources the current, the discharge on the Li-ion battery is minimal.
2.4. CALIBRATION AND VALIDATION

to use them as simple, yet stable, voltage references. For the readout and control electronics
of the FLL, we have used two different commercial packages available from Easy SQUID and
STAR Cryoelectronics, respectively. Both packages come with amplification units, which we
attach directly on top of the fridge. For additional isolation and electric shielding, the fridge
and the electronics mentioned thus far are located within a shielded room, consisting of an
electrically continuous copper sheet (Fig. 2.3).

The FLL preamplification box interfaces with the control electronics, external to the
shielded room, via a filtered D-subminiature DB9 cable. The FLL control electronics in
turn output \( V_{\text{FLL}} \) and also connect to a computer, which provides control over the FLL
parameters. The output \( V_{\text{FLL}} \) is passed to a oscilloscope, which is used during the biasing
procedure, and to an Agilent 35670A signal analyzer. It is also passed to a lock-in amplifier,
which is used to measure \( \partial I / \partial \Phi \), the measurement of which we describe in Section 2.5.

Data are acquired using the signal analyzer and sent to a computer for analysis and
storage. Under normal operation, the signal analyzer acquires data and computes the spectral
density, which is stored in a buffer that the computer can access. However, we have found
that it is much more advantageous to use the signal analyzer to capture the raw time series of
the signal. A single time capture allows us to compute multiple FFTs with differing numbers
of averages to form a single, stitched spectrum (Appendix A). In a typical measurement,
the sampling rate and length of time capture, which defines the buffer size, are programmed
into the signal analyzer via the computer interface and a custom LabVIEW program (Signal
Analyzer Controls in Fig. 2.6). The signal analyzer then fills the buffer and uploads the
time series to the LabVIEW program, where we can save it to disk, together with a header
file with details of the measurement, and compute the spectral density. For more advanced
plotting and analysis functions, we use a suite of Matlab programs that we have written for
this purpose.

2.4 Calibration and validation

To validate the system, we must first measure certain parameters in order to calibrate
conversion factors and then verify that the readout and bias electronics are sufficiently low-noise.
First, we seek to calibrate the output of the FLL; that is, we want to know \( \partial I_{\text{loop}} / \partial V_{\text{FLL}} \).
From Eq. (2.2), this means we need to measure \( M_i, M_f, \) and \( R_{\text{feedback}} \). There is a trick,
however, that allows a direct calibration of \( \partial V_{\text{FLL}} / \partial \Phi_a \), where \( \Phi_a \) is the flux in the readout
SQUID. Then,

\[
\frac{\partial I_{\text{loop}}}{\partial V_{\text{FLL}}} = \left( \frac{\partial I_{\text{loop}}}{\partial \Phi_a} \right) \left( \frac{\partial \Phi_a}{\partial V_{\text{FLL}}} \right) = M_i^{-1} \left( \frac{\partial V_{\text{FLL}}}{\partial \Phi_a} \right)^{-1}.
\]

To measure \( \partial V_{\text{FLL}} / \partial \Phi_a \), we begin by locking the FLL with zero current in the big loop and we
record the \( V_{\text{FLL}} \). Next, we unlock the FLL, inject a current into the input coil corresponding
2.4. CALIBRATION AND VALIDATION

Figure 2.6: Screenshot of LabVIEW time capture acquisition program. For a typical measurement, ‘Signal Analyzer Controls’ set the sampling frequency and buffer size of the signal analyzer. ‘Run Time Capture’ instructs the signal analyzer to start filling the buffer. Once the time capture finishes and is downloaded, we can compute the FFT in LabVIEW with parameters set in ‘LabVIEW FFT Controls’. A ‘Notes’ section allows the user to specify details of the measurement; these notes, along with the parameters of ‘Signal Analyzer Controls’ are saved to the specified file and path. Several different time captures can be queued to run sequentially.
to approximately one flux quantum, and lock the loop.\(^1\) With the loop locked, we now reduce the bias current back to zero and again record \(V_{\text{FLL}}\). The difference in recorded voltages \(\Delta V_{\text{FLL}}\) corresponds to one \(\Phi_0\) in the readout SQUID. Therefore, \(\partial V_{\text{FLL}} / \partial \Phi_a = \Delta V_{\text{FLL}} / \Phi_0\), which is about 0.5 V/\(\Phi_0\) for our system. This method is described in detail in Sec. 2.10a of Ref. 40. The mutual inductance of the input coil \(M_i\) can be readily measured by injecting a known bias current into the circuit—again, across the compensating resistor if SQUIDs are in the circuit—and measuring the response in \(V_{\text{FLL}}\). For \(M_i \approx 10\) nH, as in our system, \(\partial I_{\text{loop}} / \partial V_{\text{FLL}} \approx 0.4 \mu\text{A/V}\).

Next, we calibrate \(V_{\text{FLL}}\) to changes in the measured SQUID. Changes in \(I_c\) are related to \(V_{\text{FLL}}\) via

\[
\frac{\partial I_c}{\partial V_{\text{FLL}}} = \left( \frac{\partial I_c}{\partial I_{\text{loop}}} \right) \left( \frac{\partial I_{\text{loop}}}{\partial V_{\text{FLL}}} \right) \approx \frac{\partial I_{\text{loop}}}{\partial V_{\text{FLL}}} \tag{2.4}
\]

by Eq. (2.1) for \(R_c \ll R_d\). For our applications, Eq. (2.4) is sufficiently accurate—10% for \(R_c = R_d/10\)—as we are not focused on highly accurate measurements of critical current noise. Note that this approximation does not affect the accuracy of flux noise measurements as we shall see.

To scale \(V_{\text{FLL}}\) to an equivalent flux noise, we seek to measure the change \(\Delta V_{\text{FLL}}\) in response to a known applied flux in the measured SQUID \(\Delta \Phi_m\). In order to do so, we first measure the mutual inductance \(M_m\) between the flux bias coil (which is either fabricated on-chip or mounted in the circuit board to which the chip is mounted) and the measured SQUID. We readily accomplish this by injecting a known current \(I_\Phi\) into the coil and monitoring the response of the SQUID, which has periodicity \(\Phi_0\) in flux. For example, the difference in flux, achieved by a difference in current \(\Delta I_\Phi\) in the coil, that separates maxima in SQUID critical current is exactly \(\Phi_0\). Therefore, the mutual inductance is simply \(M_m = \Phi_0 / \Delta I_\Phi\). In principle, only one period is sufficient for this measurement, but we typically measure over five or more and perform a linear regression for better accuracy. In addition, we have found that measuring the flux bias corresponding to a minimum SQUID critical current, rather than a maximum, is more precise because the \(I_c - \Phi\) characteristics are much sharper at \((n + \frac{1}{2})\Phi_0\) than at \(n\Phi_0\). Finally, to measure \(\partial V_{\text{FLL}} / \partial \Phi_m\) we inject a known current, converted to a flux in the measured SQUID via \(M_m\), and measure the response in \(V_{\text{FLL}}\). This calibration, which must be performed before each flux noise measurement, is described in greater detail in the next Section.

After establishing the relevant conversion factors, we characterize the noise of the measurement system. With the measured SQUIDs in the superconducting state, the major sources of noise are (i) the amplification and readout electronics of the FLL, (ii) the intrinsic critical current and flux noise of the readout SQUID, and (iii) the Nyquist noise of the compensating resistor [21]. While it is possible to measure the noise from (i) and (ii) independently, the characterization process is complicated, lengthy, and involves several cooldowns.

\(^1\)We can do this even with the SQUIDs bonded into the loop by injecting a current \(I_r\) across the compensating resistor. Since the SQUIDs are superconducting, the current flows through the lowest resistance path, that is, all the current flows through the SQUIDs and input coil.
to 4 K. A much simpler method exists that sets an upper bound on the combined noise of (i) and (ii). With accurate knowledge of the compensating resistance $R_c$ and temperature $T$, we can calculate the Nyquist current noise spectral density as $S_{I,R_c} = 4k_B T/R_c$. By comparing the measured and predicted noises, we can estimate the noise added by the readout. In addition, we can confirm the calibration of the system. We find that our electronics add approximately $3 \left( \text{pA}^2/\text{Hz} \right)$ of white noise, which is negligible compared to $S_{I,R_c}$ for $T \gtrsim 0.3$ K and $R_c \approx 0.5 \Omega$. Above this temperature, agreement between measurement and prediction is better than approximately 10%. This measurement also yields the $1/f$ noise of (i) and (ii); (iii) contributes no $1/f$ noise. We found that the Easy SQUID readout system exhibited significantly more $1/f$ noise than the Cryoelectronics system, which was a problem when measuring low-noise SQUIDs.

Finally, we remark that a fourth contribution to the noise in the system originates in the inherent instability of the temperature of the devices. Through the temperature dependence of the critical currents of the Josephson junctions, these temperature fluctuations induce fluctuations in the critical current of the SQUID. We find this noise source to be significant in our devices for $T \gtrsim T_c/3$. This effect is described in detail in Chap. 4.

### 2.5 Measurement procedure

The SQUID is variably sensitive to small changes in flux depending on its flux bias, ranging from zero sensitivity when biased at $n\Phi_0$ to maximum sensitivity at nominally $(n \pm 1/4)\Phi_0$. To measure flux noise, clearly we must bias the SQUID at a point where it is sensitive to changes in flux. However, as discussed in Chapter 1, the total noise generated by the SQUID includes both critical current and flux noise. Furthermore, the magnitude of the contribution of asymmetric critical current variations varies as the flux sensitivity. Therefore, if the critical current noise is significant compared to the flux noise, its contribution is very difficult to subtract accurately from the total noise. If the critical current noise is much smaller than the flux noise, however, we can safely neglect its contribution. Therefore, we first characterize the critical current noise of the SQUID to ensure that it is negligible and then we measure at maximum flux sensitivity.

To measure the critical current noise of a SQUID we first bias it into the voltage state—typically $V_{\text{SQUID}} \sim 5 \mu\text{V}$—and null the induced loop current $I_{\text{loop}}$, with an appropriate current through the compensating resistor as discussed in Sec. 2.2. At this point, we vary the flux bias $\Phi_m$, adjusting the current bias $I_b$ as the critical current changes—to the point where $I_c$ is maximum and $\partial I_c/\partial \Phi_m = 0$, which corresponds to a flux bias of $n\Phi_0$. In this situation, a symmetric variation in the critical currents of the junctions causes $I_c$ to vary, which, in turn, induces a current $I_{\text{loop}}$ into the input coil of the readout SQUID as in Eq. (2.1). We acquire a time capture of $V_{\text{FLL}}$, sampled at approximately 1 kHz, for between 15 minutes to 1 hour, which is sufficiently long to characterize the noise at low frequencies (down to approximately $10^{-1}$ Hz to $10^{-2}$ Hz). From the time capture, we compute the spectral density, which we scale to an equivalent current in the loop via Eq. (2.3). At high frequencies, the
spectrum is flat due to the white noise generated by the resistive shunts of the junctions. If
the magnitude is large enough, the $1/f$ noise from symmetric critical current fluctuations
will cause the observed spectral density to increase at low frequencies.

Once the critical current noise is characterized, we next seek to bias the SQUID at the
point of maximum sensitivity to flux in order to measure its flux noise. In other words, we
vary $\Phi_m$ until $\partial I_{\text{loop}}/\partial \Phi_m$ is maximized. To measure $\partial I_{\text{loop}}/\partial \Phi_m$, we first use the internal
oscillator of a Stanford Research Systems SR830 lock-in amplifier and a custom voltage-to-
current box to generate an oscillating current of known magnitude. As shown in Fig. 2.3,
the oscillating current is summed with the static flux bias current and injected into the flux
bias coil of the measured SQUID. Using the measured value of $M_m$, we calculate the rms
magnitude of oscillating flux $\Phi_{m,\text{osc}}$ applied to the measured SQUID. The oscillating flux
causes $I_c$ and, correspondingly, $V_{\text{FLL}}$ to modulate. We use the lock-in to
demodulate $V_{\text{FLL}}$ and compute the rms magnitude of the voltage oscillations $V_{\text{FLL},\text{osc}}$, from which we readily
calculate

$$
\frac{\partial I_{\text{loop}}}{\partial \Phi_m} = \frac{V_{\text{FLL,osc}}}{\Phi_{m,\text{osc}}} \left( \frac{\partial I_{\text{loop}}}{\partial V_{\text{FLL}}} \right).
$$

(2.5)

To ensure that the response is linear, it is important that $\Phi_{m,\text{osc}} \ll \Phi_0$; in our measurements,
$\Phi_{m,\text{osc}}/\Phi_0$ was typically between $10^{-2}$ and $10^{-4}$. Depending on the value of $\beta_L$, the flux bias
corresponding to the maximum value of $\partial I_{\text{loop}}/\partial \Phi_m$ can vary from approximately $(n \pm \frac{1}{2})\Phi_0$
for SQUIDs with $\beta_L \approx 1$ to very near $(n \pm \frac{1}{2})\Phi_0$ for SQUIDs with $\beta_L \ll 1$ [70].

With $\Phi_m$ set such that $\partial I_{\text{loop}}/\partial \Phi_m$ is maximized, we voltage bias the SQUID and null $I_{\text{loop}}$
in the manner previously described, acquire a 1-hr time capture, and compute the spectral
density. At this flux bias, $I_c$ is maximally sensitive to both asymmetric critical current and
flux noise. By comparing this spectrum to that acquired at $n\Phi_0$, we can readily determine
the noise added by these two sources.

While we cannot directly distinguish asymmetric critical current noise from flux noise,
we can make a simple approximate argument to relate the former to symmetric critical
current noise. First, suppose that each junction has a noise power spectrum $S_{I_0}$. Symmetric
variations in the critical currents of the junctions will directly change the critical current of
the SQUID so that $S_{I_0,\text{sym}} = 2S_{I_0}$. Asymmetric critical current variations do not directly
change the SQUID critical current, but they can indirectly change the critical current by
inducing a flux. The asymmetric variations—and, therefore, the circulating current—have a
spectral density $S_{I_0}/2$ and couple flux via the self-inductance of the SQUID,

$$
S_{I_0,\text{asym}} = \frac{1}{2} S_{I_0} L^2 \left( \frac{\partial I_c}{\partial \Phi} \right)^2.
$$

(2.6)

Because $|\partial I_c/\partial \Phi|_{\text{max}} \sim I_c/\Phi_0$ and $\beta_L = LI_c/\Phi_0$, we see that

$$
S_{I_0,\text{asym}} \lesssim \frac{1}{2} S_{I_0} \left( \frac{LI_c}{\Phi_0} \right)^2 = \frac{1}{2} S_{I_0} \beta_L^2 \lesssim S_{I_0,\text{sym}},
$$

(2.7)
Figure 2.7: Current noise spectra for measured SQUID in the superconducting state and in the voltage state at $0 \Phi_0$ and $\frac{1}{4} \Phi_0$. At low frequencies, the presence of any $1/f$ critical current noise is obscured by the $1/f$ noise of the readout SQUID and electronics: the spectrum acquired with the voltage biased SQUID at $0 \Phi_0$ (red) is indistinguishable from the $1/f$ noise of the readout SQUID and electronics (blue). When the measured SQUID is biased at $\frac{1}{4} \Phi_0$, it becomes sensitive to flux noise and the spectrum considerably increases at low frequencies. The flattening of all spectra at high frequencies is due to the onset of white noise from the shunt resistors.
so that critical current noise due to asymmetric variations will never be much larger than that due to symmetric variations. Therefore, a large difference in noise between the two spectra acquired at $n\Phi_0$ and $(n+\frac{1}{4})\Phi_0$ for fixed $n$ must be due to the flux noise. In Fig. 2.7 we plot the observed current noise—computed from the time series as described in Appendix A—for a measured SQUID in the superconducting state as well as in the voltage state at $n\Phi_0$ and $(n+\frac{1}{4})\Phi_0$. At $(n+\frac{1}{4})\Phi_0$ the SQUID is nominally\(^2\) maximally sensitive to flux and the current noise at low frequencies increases significantly due to the contribution of flux noise.

As a further verification, we can confirm that the observed noise at low frequencies scales as an equivalent flux. To do this, we reduce $\partial I_{\text{loop}}/\partial \Phi_m$ to, say, half its maximum and again acquire data. If the observed noise is indeed due to flux variations, its magnitude referenced as a current in the input coil should scale as $(\partial I_{\text{loop}}/\partial \Phi_m)^2$, that is, the power spectral density at a particular frequency should be reduced by a factor of four compared to the spectrum acquired at $|\partial I_{\text{loop}}/\partial \Phi_m|_{\text{max}}$. Of course, the white noise from the shunts, which dominates at high frequency, does not scale as a flux, so the “knee” between white current noise and $1/f$ flux noise will decrease in frequency, a fact that highlights the importance of acquiring data where $\partial I_{\text{loop}}/\partial \Phi_m$ is a maximum. We have plotted in Fig. 2.8 two noise spectra scaled to equivalent flux in the measured SQUID. The spectra were acquired at $(2n + 1)\Phi_0/4$

\(^2\)As discussed previously, the value of the flux bias that maximizes the transfer factor $\partial I_{\text{loop}}/\partial \Phi_m$ is not always $(n \pm \frac{1}{4})\Phi_0$ and can vary depending on the value of $\beta_L$. Nonetheless, for brevity we will often refer to the flux bias that maximizes the transfer factor as $\frac{1}{4} \Phi_0$. 

Figure 2.8: Equivalent flux noise spectrum showing scaling with $\partial I_{\text{loop}}/\partial \Phi_m$. The $1/f^\alpha$ flux noise dominates at low frequencies so that the spectrum is independent of $\partial I_{\text{loop}}/\partial \Phi_m$. At high frequencies, white current noise from the resistive shunts of the SQUID dominates and the spectrum scales as $(\partial I_{\text{loop}}/\partial \Phi_m)^{-2}$. Here, the fit parameters $A^2$ and $\alpha$ agree within 6% and 1%, respectively.
and \((2n + 1)\Phi_0/8\), where the corresponding values of \(\partial I_{\text{loop}}/\partial \Phi_m\) were 15.5 \(\mu\text{A}/\Phi_0\) and 11.0 \(\mu\text{A}/\Phi_0\), respectively. As expected, the flux noise portion of the spectrum indeed scales as an equivalent flux in the measured SQUID. Note that our fitting routine for all spectra can be found in Appendix B.

**Acknowledgments**

This chapter would not be complete without the explicit mention of Andrew Fefferman, Sean O’Kelley, and Jeffrey Birenbaum, all of whose contributions were instrumental to the successful development of the experimental measurement system. In particular, Andrew’s expertise and experience with dilution refrigerators was key for all of us to learn to operate a device that was created when I was learning multiplication. Sean’s background in electronics helped us all to make intuitive sense of the readout circuit as well as to account for the various stages of electronics amplification and noise in a systematic and comprehensible way. Finally, Jeff, who was just a first-year graduate student at the time we developed the system, was enormously helpful putting together the nuts and bolts, as it were. Of his innumerable contributions, I am most thankful for his obliging and painstaking assembly of the much-hated Reichenbach cables. They have withstood the test of time, and for that we are thankful.
Chapter 3

Scaling of flux noise with temperature

3.1 Introduction

A crucial question is whether the independent-spin model describes the flux noise dynamics correctly or if spin-spin interactions are significant. Sendelbach et al. [60] inferred the existence of clusters from measurements of inductance fluctuations in SQUIDs, suggesting a non-negligible spin-spin interaction. The independent-spin model was used in an analytical calculation of the mean-square value of the flux noise $\langle \Phi^2 \rangle$ as a function of loop geometry [45, 57]. Measurements of decoherence in flux qubits [47] were in agreement with the predicted scaling while measurements of decoherence in phase qubits instead showed a scaling with loop inductance [50].

In this Chapter we report flux noise measurements as a function of temperature for 10 dc SQUIDs with systematically varied geometries. We find that both $S_\Phi(1 \text{ Hz})$ and the slope $\alpha$ vary systematically with temperature. Remarkably, for a given SQUID $S_\Phi(f)$ pivots about a nearly fixed frequency as the temperature is changed. Values of $\langle \Phi^2 \rangle$ inferred from our measurements of $S_\Phi(f)$ deviate markedly from the predicted scaling with loop geometry. Furthermore, although there is no evident temperature dependence in the prediction of $\langle \Phi^2 \rangle$ over the range of temperatures investigated, pivoting causes the inferred values of $\langle \Phi^2 \rangle$ to vary over several orders of magnitude as we change the temperature. We are unable to reconcile our data with the independent-spin model.

We measured $S_\Phi(f)$ in a total of 20 dc SQUIDs, connected in series in sets of five on four chips, fabricated on oxidized Si wafers using a trilayer Nb-AlOx-Nb junction technology. Very generously, our collaborators Vlad Bolkhovsky, Danielle Braje, George Fitch, Matthew Neeley, and Will Oliver at MIT Lincoln Labs and Gene Hilton, Sherry Cho, and Kent Irwin at NIST, Boulder, fabricated to our design a large number of high-quality chips for us to measure. Of these, we measured four chips, where two of the chips were made at MIT-LL and two at NIST. Since the two chips from each institution showed very similar results, we present data on only one from each, labeled I (LL) and II (NIST). The SQUIDs were patterned in a square geometry with outer widths $2R$ and loop linewidths $W$ listed in Table 3.1, along with their estimated loop inductances. Each junction was resistively shunted to eliminate
3.2. METHODS

Table 3.1: Dimensions and inductances of SQUIDs I and II.

<table>
<thead>
<tr>
<th></th>
<th>R (µm)</th>
<th>W (µm)</th>
<th>R/W</th>
<th>L (pH)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.1</td>
<td>12</td>
<td>0.5</td>
<td>24</td>
<td>80</td>
</tr>
<tr>
<td>I.2</td>
<td>6</td>
<td>0.5</td>
<td>12</td>
<td>33</td>
</tr>
<tr>
<td>I.3</td>
<td>3</td>
<td>0.5</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>I.4</td>
<td>1.5</td>
<td>0.5</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>I.5</td>
<td>1.5</td>
<td>0.5</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>II.1</td>
<td>265</td>
<td>240</td>
<td>1.1</td>
<td>120</td>
</tr>
<tr>
<td>II.2</td>
<td>145</td>
<td>120</td>
<td>1.2</td>
<td>98</td>
</tr>
<tr>
<td>II.3</td>
<td>85</td>
<td>60</td>
<td>1.3</td>
<td>92</td>
</tr>
<tr>
<td>II.4</td>
<td>55</td>
<td>30</td>
<td>1.8</td>
<td>96</td>
</tr>
<tr>
<td>II.5</td>
<td>40</td>
<td>15</td>
<td>2.7</td>
<td>106</td>
</tr>
</tbody>
</table>

hysteresis on its current-voltage characteristic.

3.2 Methods

As shown schematically in Fig. 3.1, the five SQUIDs on a given chip were connected in series with an off-chip compensating resistor $R_c = 0.46 \ \Omega$, the input coil of a readout SQUID operated in a flux-locked loop (FLL), and a choke to suppress high-frequency currents [71]. The entire circuit was enclosed in two superconducting shields and a cryoperm shield to reduce ambient magnetic field fluctuations and the Earth’s static magnetic field. The magnetic flux in the SQUIDs was established by means of a current in a common coil. To measure the flux noise in a given SQUID, we biased it at a voltage of 2.5 or 5 µV with a current $I_b$ and canceled the quasistatic current induced into the input coil and remaining SQUIDs with a compensating current $I_r$ in $R_c$. Fluctuations $\delta \Phi$ in $\Phi$ of the measured SQUID induced a current $\delta \Phi(dI/d\Phi)$ in the input coil of the readout SQUID. For all data represented here, we biased the SQUID where $dI/d\Phi$—determined by measuring the response to a small oscillating flux—was a maximum. The temperature $T$, measured using calibrated Ge (mixing chamber) and RuOx (sample box) resistance thermometers, was stabilized with feedback from the Ge thermometer to better than 1 part in $10^4$ during data acquisition [72]. Nyquist noise from $R_c$ yielded a temperature within ±5% of that of the thermometers.

We acquired a time-series of the voltage fluctuations for 1 hour, computed the spectral density and converted it to $S_\Phi(f)$. We performed a least squares fit to $S_\Phi(f) = A^2/(f/1 \ Hz)^\alpha + C^2$, representing the flux $1/f^\alpha$ noise and the white noise from the resistive shunts, to obtain $A$ and $\alpha$. To confirm that the inferred value of $A^2$ was independent of $dI/d\Phi$, we measured the noise at other values of $\Phi$ and $dI/d\Phi$. In addition, we made measurements at $dI/d\Phi = 0$, enabling us to determine the critical current $1/f$ noise. We verified that the power spectrum of the white noise from the resistive shunts dominated those from $R_c$ and the readout electronics, and that, with a flux bias of $n\Phi_0$, it remained white at frequencies down to the $1/f$ knee of the readout electronics with a magnitude in
3.3 Results

As representative data, in Fig. 3.2(a) we show raw power spectra for SQUID II.3. The spectra are $1/f$-like, with a slope that flattens at higher frequencies as the white noise from the shunt resistors becomes significant. At low frequencies ($f \lesssim 10^{-1}$ Hz) and high temperatures ($T \gtrsim 1.2$ K), fluctuations in the critical current are significant, thereby increasing the slope of the spectra [72]. Consequently, in our fits to the spectra, we disregard noise from this region.

The fitted values of $A^2$ and $\alpha$ are plotted in Fig. 3.3 for all five SQUIDs on both chips. We see immediately that the systematic trends with $T$ for each chip differ between the two chips. For chip I, $A^2$ increases as $T$ is lowered to 0.1 K, with an upturn for $T < 0.15$ K, and increases monotonically with $R$. In contrast, for chip II $A^2$ becomes independent of $T$ for $T \lesssim 0.6$ K, a result reminiscent of the findings of Wellstood, et al. [38]. With regard to the slope, for both chips $\alpha$ increases dramatically as $T$ is lowered from 4.0 K to 0.1 K. For chip I, the value of $\alpha$ increases with $R$. Similarly low values of $\alpha$, 0.4–0.5 at 4 K, have been observed by other authors [53, 73]. Although we have no explanation for the disparate dependencies of $A^2$ and $\alpha$ on $T$, we remark that for chip I $W$ is constant at 0.5 $\mu$m while for chip II $W$ varies from 15 to 240 $\mu$m. A conceivable explanation is that interactions between spins give rise to spatial correlations.

The progressive increase in $\alpha$ as $T$ is lowered is illustrated vividly in Figs. 3.2(b,c), where...
Figure 3.2: Power spectra $S_\Phi(f)$. (a) As-acquired flux noise spectral densities for SQUID II.3 at 11 temperatures. Inset shows data at 0.1 K with corresponding fit to $A^2/(f/1 \text{ Hz})^\alpha + C^2$ (dashed line). Fits of $S_\Phi(f)$ for (b) SQUID I.1 at 10 temperatures and (c) SQUID II.3 at 11 temperatures.
Figure 3.3: Temperature dependence of $1/f$ flux noise for 10 SQUIDs. (a) and (b) Fit coefficients $A^2$ for SQUIDs I and II vs. $T$. (c) and (d) Fit coefficients $\alpha$ for SQUIDs I and II vs. $T$. Confidence in fits is $\pm 10\%$ in (a) and (b); $\pm 0.03$ ($T \leq 1.1$ K) and $\pm 0.05$ ($T > 1.1$ K) in (c); and $\pm 0.03$ ($T \leq 1.1$ K) and $\pm 0.07$ ($T > 1.1$ K) in (d).
Figure 3.4: Fitted power spectra $S_{\phi}(f)$ for SQUIDs I at 10 temperatures.
Figure 3.5: Fitted power spectra $S_{\Phi}(f)$ for SQUIDs II at 11 temperatures.
we plot $A^2/(f/1\ \text{Hz})^\alpha$ for (b) SQUID I.1 and (c) II.3 for all $T$ at which we obtained data. Ignoring the two highest spectra in Fig. 3.2(b)—corresponding to $T < 0.15$ K with upturns in the data in Fig. 3.3(a)—quite unexpectedly and remarkably we observe that each set of spectra pivots around an approximately constant frequency $f_p$: $f_p = 76 \pm 20$ Hz in (a) and $f_p = 6.2 \pm 0.5$ Hz in (b). (Figures 3.4 and 3.5 show similar plots for all 10 SQUIDs.) The plots in Fig. 3.2 imply that, although one traditionally defines the noise magnitude as $A^2 \equiv S_\Phi(1\ \text{Hz})$, in fact $A^2$ is an incomplete characterization in the absence of a knowledge of $\alpha$.

We next estimate the mean square flux noise $\langle \Phi^2 \rangle$. For a circular loop of outer radius $R$ and linewidth $W$, in the limit $R \gg W$, Bialczak et al. [45] showed that

$$\langle \Phi^2 \rangle \approx \frac{2\mu_0^2}{3}\mu_B^2\sigma R W \left[ \ln\left(\frac{2bW}{\lambda^2}\right) \right] + 0.27$$

(3.1)

for uncorrelated spins with uniform surface density $\sigma$ and magnetic moment $\mu_B$. Here, $\mu_0$ is the vacuum permeability, $b$ is the film thickness, and $\lambda$ is the penetration depth. For the square geometry of our devices, there is a multiplicative correction of order unity, which we neglect. Because we cannot directly measure $\langle \Phi^2 \rangle$, we use

$$\langle \Phi^2 \rangle = \int_{f_1}^{f_2} S_\Phi(f) df$$

(3.2)

to relate $\langle \Phi^2 \rangle$ to our measurements of $S_\Phi(f)$. Here, we extrapolate $S_\Phi(f)$, which we measure typically over several decades, to $f_1$ and $f_2$. We set $f_1 = 10^{-4}$ Hz, approximately the lowest frequency to which flux noise has been measured. The value of $f_2$ is more difficult to estimate as it is poorly understood. Since Bylander et al. [66] and Slichter et al. [68] reported flux noise up to measurement-limited frequencies of 20 MHz and 1 GHz, respectively, we use $f_2 = 10^9$ Hz for our analysis. Separate measurements on Al SQUIDs, fabricated in exactly the same way in the same equipment as the qubits in Ref. [68], yielded average low-temperature values remarkably similar to those observed in SQUIDs I and II: $A^2 = 2.1\ (\mu\Phi_0)^2/\text{Hz}$ and $\alpha = 0.72$. These results suggest that details of the processing do not greatly alter the characteristics of flux noise. In particular, we expect $f_2$ to be comparable in SQUIDs and qubits alike.

For all 10 SQUIDs, Fig. 3.6 shows $\langle \Phi^2_{\text{inf}} \rangle$, inferred using Eq. (3.2) and the data plotted in Fig. 3.3, vs. $T$. For comparison, the dashed lines indicate the value of $\langle \Phi^2_{\text{calc}} \rangle$ for each SQUID calculated using Eq. (3.1), the values of $R$ and $W$ from Table 3.1, $\sigma = 5 \times 10^{17}$ m$^{-2}$, $\lambda = 39$ nm and $b = 150$ nm (I), 200 nm (II). While $\langle \Phi^2_{\text{calc}} \rangle$ is independent of $T$ for $T \ll T_c$, the values of $\langle \Phi^2_{\text{inf}} \rangle$ increase strongly with increasing $T$. At the lowest $T$, values of $\langle \Phi^2_{\text{inf}} \rangle$ exceed those of $\langle \Phi^2_{\text{calc}} \rangle$ by one to two orders of magnitude and at the highest $T$ by three to five orders of magnitude. Clearly, the rapid calculated increase of $\langle \Phi^2_{\text{inf}} \rangle$ with $T$ is inevitable given the behavior in Fig. 3.2.

As a further test of the theory, we investigated the scaling of $S_\Phi(f)$ with SQUID dimensions to compare with the predictions of Eq. (3.1). For brevity, we present results only for
3.3. RESULTS

Figure 3.6: Mean-square flux noise, inferred from measured power spectra using Eq. (3.2), vs. $T$. (a) Chip I and (b) chip II. Horizontal dashed lines represent the predictions of Eqs. (3.1) and (3.2) for the values of $R$ and $W$ listed in Table 3.1.

SQUIDs I, for which $W$ is constant, so that $\langle \Phi^2_{calc} \rangle \propto R$. Figure 3.7(a) shows $S_\Phi(54 \, \text{Hz})$ vs. $R$ for the four distinct geometries of SQUIDs I. Here, we chose the average value of $f_1$, 54 Hz, to minimize the effect of the $T$-dependence of the spectra on the geometric scaling. To a reasonable approximation, apart from the data for SQUIDs I.1 and I.2 at the two lowest temperatures, $S_\Phi(54 \, \text{Hz})$ scales linearly with $R$. Figure 3.7(b) shows $\langle \Phi^2_{inf} \rangle$ vs. $R$ for 10 temperatures, which we see is approximately independent of $R$, increases with $T$ and has a magnitude substantially higher than the prediction.

We now discuss potential sources of error in $\langle \Phi^2_{inf} \rangle$ that would give rise to a strong temperature dependence. A $T$-dependent spin density is precluded by the Curie-Law behavior of the spin paramagnetism [54, 56]. We also note that, for the observed values of $\alpha$, $\langle \Phi^2_{inf} \rangle$ changes by less than 1% as $f_1$ is varied from zero to $10^{-2} \, \text{Hz}$. With regard to $f_2$, one could argue that the value of $f_2 = 1 \, \text{GHz}$, observed at 50 mK [68], becomes lower as $T$ increases. However, to reduce the highest-temperature value of $\langle \Phi^2_{inf} \rangle$ in Fig. 3.7 to its lowest temperature value by lowering $f_2$ would require $f_2$ to be roughly 700 kHz for SQUIDs I and just 8 kHz for SQUIDs II. Such values of $f_2$ are improbably low; it is difficult to imagine a noise process that extends to 1 GHz at 0.1 K, yet to less than 1 MHz at 4 K. Regarding $\alpha$, high-frequency measurements in qubits have sampled the flux noise spectrum from 0.2 to 20 MHz [66] as well as from $10^{-2}$ to 1 Hz, 1 to 20 MHz, and 0.7 to 1 GHz [68]. These measurements each align with a spectrum of constant $\alpha$ ($\alpha = 0.9$ [66] and $\alpha = 0.57$ [68]) and value of $S_\Phi(1 \, \text{Hz})$ typical of those observed in SQUIDs. These measurements suggest that, while $\alpha$ can vary between devices, within a single device it is frequency-independent over a very wide bandwidth. Finally, errors in our fits of $\alpha$ will inevitably lead to errors in $\langle \Phi^2_{inf} \rangle$. For instance, at low (high) temperatures decreasing $\alpha$ by 0.03 (0.07) increases $\langle \Phi^2_{inf} \rangle$ by a
3.3. RESULTS

Figure 3.7: Scaling of noise power with $R$ for SQUIDs I for 10 temperatures from 0.1 K to 4.0 K. (a) $S_{\phi}(54 \, \text{Hz})$ vs. $R$. Dashed line indicates scaling with $R$, with arbitrary magnitude. (b) $\langle \Phi^2 \rangle$ vs. $R$. Dashed line represents prediction of Eq. (3.1).
factor of roughly 1.6 (4). These errors, while not negligible, cannot account for the strong temperature dependence and discrepancies of several orders of magnitude.

Given the difficulty of reconciling our data with the predictions of a model based on \( N \) single, uncorrelated spins—for which \( \langle \Phi^2 \rangle \propto N \mu_B^2 \)—we consider the possibility that the spins form clusters [60] following an argument raised by Alexander Shnirman. We assume the clusters to be uncorrelated with each other and to contain an average of \( Z \) spins producing a magnetic moment \( \mu_c \), where \( Z \) may depend on \( T \). There are three scenarios. (i) Ferromagnetic: \( \mu_c = Z \mu_B \) and \( \langle \Phi^2 \rangle \propto (N/Z)(Z \mu_B)^2 = NZ \mu_B^2 \); (ii) random (glassy): \( \mu_c = Z^{1/2} \mu_B \) and \( \langle \Phi^2 \rangle \propto (N/Z)(Z^{1/2} \mu_B)^2 = N \mu_B^2 \); and (iii) antiferromagnetic: \( \mu_c \approx \mu_B \) and \( \langle \Phi^2 \rangle \propto (N/Z)\mu_B^2 \). As \( Z \) grows, the rate at which a cluster reverses its magnetic moment is assumed to decrease rapidly [74]. At a given \( T \), such a distribution of reversal rates offers a natural explanation for the 13-decade range in lifetimes required for \( S_\Phi(f) \) to range from \( 10^{-4} \) to \( 10^9 \) Hz [24]. Furthermore, a plausible scenario for the spectral pivoting as \( T \) is lowered is a progressive growth in \( Z \), reducing the number of small clusters generating noise above \( f_p \) and increasing the number of larger clusters generating noise below \( f_p \)—thereby increasing \( \alpha \). This simplistic picture has implications for the temperature dependence of each of the three cases. As \( T \) is lowered, \( \langle \Phi^2 \rangle \) increases for ferromagnetic clusters, remains constant for random clusters, and decreases for antiferromagnetic clusters.

Finally, the noise behavior must be compatible with the observed spin paramagnetism [54, 56], which follows a Curie \( 1/T \) scaling. We note that the contribution of a cluster with magnetic moment \( \mu_c \) to the linear magnetic susceptibility scales as \( \mu_c^2/k_B T \). On the other hand, once \( \mu_c B > k_B T \) the susceptibility ceases to be linear in magnetic field, the magnetization of the cluster saturates and Curie scaling is violated. Consequently, since Curie behavior was observed at \( T = 50 \) mK and \( B = 10 \) mT, we find \( \mu_c < 7 \mu_B \). Thus, \( Z < 7 \) for the ferromagnetic case, \( Z < 7^2 \) for the random case and \( Z \) is unrestricted for the antiferromagnetic case. We also point out that the areal single spin density of \( 5 \times 10^{17} \) m\(^{-2} \) derived from both noise and paramagnetism experiments is unchanged for glassy clusters. Needless to say, a detailed understanding of the interactions between both the spins that form such clusters and between the clusters themselves will be required to develop a credible explanation of flux noise.

**Acknowledgments**

This Chapter represents the fruits of labor from a large number of contributors that spanned several years. The SQUID fabrication by our collaborators Vlad Bolkhovsky, Danielle Braje, George Fitch, Matthew Neeley, and Will Oliver at MIT Lincoln Labs and Gene Hilton, Sherry Cho, and Kent Irwin at NIST, Boulder, was unquestionably essential. In addition, Fred Wellstood and Alexander Shnirman were helpful in developing a fuller understanding of our results.
Chapter 4

Low-frequency critical current noise in Josephson junctions induced by temperature fluctuations

4.1 Introduction

At low frequencies $f$ the critical current $I_0$ of a Josephson tunnel junction fluctuates with time, yielding a power spectrum $S_{I_0}(f)$ that scales approximately as $1/f$ [28, 29, 31, 32, 37, 41, 75, 76]. It is generally accepted that the noise arises from the uncorrelated trapping and subsequent release of electrons in the tunnel barrier [27–34]; a trapped electron decreases the junction transparency locally, reducing the critical current. The superposition of the resulting random telegraph signals produces $1/f$ noise [24]. This noise has long been a limitation on the resolution of SQUIDs, particularly in high transition temperature ($T_c$) devices. Current bias reversal schemes, however, can reduce the effect of the noise drastically [29, 37]. Furthermore, critical current noise may eventually become a limitation to the decoherence time of superconducting qubits such as flux and phase qubits, the transmon and the quantronium [41, 42, 77, 78]. In the case of the flux qubit, for example, critical current noise—unlike flux noise—causes dephasing even at the degeneracy point, since it directly modulates the energy level splitting.

In 2004 Van Harlingen et al. [41] compiled a number of published critical current noise data obtained from ex-situ fabricated Josephson tunnel junctions and found that the scaled noise power at 1 Hz, $AS_{I_0}(1\ \text{Hz})/I_0^2$, where $A$ is the junction area, was relatively constant—consistent with the charge trap model. Furthermore, it was reported that $S_{I_0}(1\ \text{Hz})$ scaled quadratically [75, 79] with temperature $T$. Subsequently, significantly lower values of $S_{I_0}(1\ \text{Hz})$ were reported in in-situ fabricated junctions, and it was found that $S_{I_0}(1\ \text{Hz})$ scaled linearly with $T$ [80–85]. The scaling with temperature is important both to understand the underlying charge-trapping mechanism and to calculate dephasing in qubits. Currently, the magnitude of the critical current noise is too small to be measured directly at typical qubit operating temperatures of tens of millikelvin, and must be extrapolated from higher temperatures [41].
4.2. METHODS

In this Chapter we show that even small temperature fluctuations, which modulate \( I_0(T) \) via its temperature dependence, can contribute significantly to low-frequency critical current noise in addition to the intrinsic charge trapping mechanism. Moreover, because both temperature stability and the sensitivity of the critical current to changes in temperature vary as functions of temperature, this mechanism can distort the spectrum and temperature dependence of the observed critical current noise. For the particular case of our dilution refrigerator, we show that, for low-noise Al-AlOx-Al junctions, the temperature-induced noise can dominate for \( T \gtrsim T_c/3 \).

4.2 Methods

In our experiment, the temperature control is typical of that for temperature-stabilized measurements above the base temperature of a dilution refrigerator. We raise the temperature of the mixing chamber (MXC) by passing a current through a resistor thermally anchored to it. The magnitude of the current is controlled by a digital proportional-integral-derivative (PID) feedback loop updated at 10 Hz, where the thermometer—germanium, in our case—is close to the resistor. The silicon chip on which the junctions are grown is attached to a fiberglass carrier mounted inside a copper box. The box is bolted to the bottom of a cold finger attached to the MXC. The thermal mass of the sample and its thermal conductivity to the MXC, both of which can vary with temperature, define a thermal filter time constant \( \tau_{\text{chip}} \), so that temperature fluctuations at frequencies well above \( 1/2\pi\tau_{\text{chip}} \) are filtered out while those well below are transmitted to the junctions under investigation. In our system, \( \tau_{\text{chip}} \) is of the order of a few seconds. We note that Nyquist noise generated in the resistive shunts of the junctions is likely to dominate the intrinsic critical current noise at frequencies above typically 0.1 to 1 Hz, so that accurate measurements of critical current noise require frequencies as low as \( 10^{-3} \) Hz.

For our measurements we used a SQUID, rather than a single junction, simply because we became aware of the temperature-induced critical current noise in the course of investigating low-frequency noise in SQUIDs. The SQUID in which we perform the following measurements was fabricated in a joint effort between Allison Dove, Gus Olson, Zach Yosovits, and Jim Eckstein, who deposited the washer of the SQUID, and Chris Nugroho, Vlad Orlyanchik, and Dale Van Harlingen, who fabricated the junctions of the SQUID. In all of the experiments reported here, we flux bias the SQUID at \( n\Phi_0 \) so that the flux noise vanishes (\( n \) is an integer, \( \Phi_0 \equiv h/2e \) is the flux quantum). Thus, at low frequencies the two junctions behave as a single junction with critical current \( I_0 \) given by the sum of the individual critical currents and an area given by the sum of the two areas. The SQUID loop was fabricated from Nb and was completed with shadow-deposited Al-AlOx-Al junctions that were resistively shunted to remove hysteresis in the current-voltage characteristics. For each junction, the critical current was approximately 5 \( \mu \)A, and the junction area was 0.5 \( \mu \text{m}^2 \), corresponding to a self-capacitance \( C \) of approximately 50 fF. The shunt resistance was \( R = 18 \) \( \Omega \), yielding a hysteresis parameter \( \beta_c \equiv 2\pi I_0 R^2 C/\Phi_0 \approx 0.25 \) so that the junctions were over-damped.
4.3 Results

Following a suggestion by Chris Nugroho, to characterize the temperature dependence of the critical current, in Fig. 4.1(a) we plot $I_0(T)/I_0(0)$ vs. $T$ for the measured SQUID,
Figure 4.2: Temperature and critical current noise for three settings of temperature controller. Typical 20-minute-time traces for the measured (a) temperature and (b) critical current variations at 0.4 K.

compared with the Ambegaokar-Baratoff [86] prediction

\[
I_0(T) = \frac{\pi \Delta(T)}{2eR_N} \tanh \left[ \frac{\Delta(T)}{2k_B T} \right].
\]  

(4.1)

Here, we normalize the energy gap [87] \( \Delta(T) \) to \( \Delta(0) = 1.76 k_B T_c \), where \( T_c = 1.27 \) K is the measured transition temperature and \( R_N \) is the resistance of an unshunted junction at voltages above \( 2\Delta(T)/e \). We take the measured value of the bias current at a SQUID voltage of 1.25 \( \mu \)V as \( I_0 \). The measured temperature dependence is in excellent agreement with the prediction. To determine \( dI_0(T)/dT \) we increased the temperature by a small amount \( \delta T \), typically 10 mK, and measured the change in the current in the input coil. Figure 4.1(b) shows that the measured and predicted values of \( dI_0(T)/I_0(0)dT \) vs. \( T \) are in good agreement. As a check on our methodology, we showed that with the measured SQUID in the superconducting state the temperature change \( \delta T \) produced no discernible change in the output of the flux-locked loop.

To characterize the role of temperature stability in critical current noise measurements, we first stabilized the temperature of the SQUID at 0.4 K using the PID feedback loop. Once the temperature had stabilized, at 2 Hz we sampled simultaneously the measured temperature and the output from the flux-locked loop for three different sets of PID parameters, which yielded different amplitudes of temperature noise. Importantly, our temperature readings are electronically unfiltered in the sense that we disable the low-pass filters typically used in commercial temperature controllers. For example, the LakeShore Model 370 AC Resistance Bridge, used in our measurements, by default applies a 10-s linear moving average to the
output temperature reading. While the actual temperature noise remains unchanged, the filter reduces the magnitude and distorts the spectral and temporal response of the measured noise.

In Fig. 4.2(a) we plot a typical 20-minute time trace of the temperature fluctuations for three settings of the PID feedback loop, A, B, and C. The noise evidently increases from A to C. The standard deviations $\sigma_T$ of the temperature fluctuations are 8, 13, and 26 $\mu$K for A, B, and C, respectively. Figure 4.2(b) shows the corresponding time series for the critical current variations. In run C, the critical current is clearly highly correlated with the temperature. In run A, however, the level of correlation is not discernible.

To quantify the correlation between temperature and current variations, in Fig. 4.3 we plot the normalized cross-correlation function $R_{TI_0}(\tau)$ between the temperature and critical current for each time series shown in Fig. 4.2 vs. lag time $\tau$. For run C, we observe a maximum in $R_{TI_0}(\tau)$ of 0.88, indicating a high degree of correlation, consistent with our observations in Fig. 4.2. As the temperature noise decreases, the maximum in $R_{TI_0}(\tau)$ drops correspondingly to 0.54 and 0.13 for runs B and A, respectively. The lag time at which the peak in $R_{TI_0}(\tau)$ occurs is related to the low-pass temperature filter discussed previously.

Whereas $R_{TI_0}(\tau)$ characterizes correlations in the time domain, we are particularly interested in frequency-dependent correlations. In Fig. 4.4(a) we plot the normalized cross-spectral density (coherence) $S_{TI_0}^2/S_T S_{I_0}$ for each run; $S_{TI_0}$ is the cross-spectrum of temperature and critical current. As $f$ increases from $10^{-3}$ Hz, the coherence retains its maximum value until about $10^{-2}$ Hz, above which it decreases, becoming essentially zero for $f > 0.1$ Hz. The near-unity coherence in run C (blue) for $f < 10^{-2}$ Hz indicates that the temperature and critical current variations are almost completely correlated, so that most of the critical
4.3. RESULTS

current fluctuations at these frequencies originate from temperature fluctuations. Runs B (red) and A (black) exhibit a smaller maximum coherence, consistent with our analysis in the time domain.

In Fig. 4.4(b), for run C we plot the measured current noise $S_{I_0}(f)$ and the temperature noise, scaled to a current noise: $S_{I_0,T}(f) = [dI_0(T)/dT]^2 S_T(f)$. As expected, the spectra converge at low frequencies. As the frequency increases, however, the spectra no longer coincide and show significant differences, which we explain by two effects. First, because of the low-pass temperature filter discussed earlier, temperature variations at the MXC for $f > 1/2\pi \tau_{\text{chip}} \approx 0.05$ Hz are strongly attenuated when they reach the SQUID. Second, Nyquist current noise generated by the resistive shunts in the SQUID dominates the current noise generated by temperature variations for $f > 0.2$ Hz.

We remark that $S_T(f)$ in our system scales approximately as $1/f^{1.7}$ between roughly $10^{-1}$ and $10^{-3}$ Hz. Because the power spectrum of a quantity with a linear drift scales as $1/f^2$, we investigated the effect of drift in our system. By computing a linear regression of each measured time trace and the corresponding power of that linear drift, we find that drifts account for at most 0.2% of the total noise power. Because of the roll-off from the low-pass temperature filter, $S_{I_0}(f)$ scales as $1/f^{2.2}$ for $0.01$ Hz < $f$ < $0.1$ Hz. The measured slope of $-2.2$—too large to have originated from charge trapping, which is expected to scale approximately as $1/f$ over a wide frequency range—is apparent only in long measurements.
4.3. RESULTS

Figure 4.5: Critical current noise caused by charge-trapping compared to that due to temperature variations. For $S_{I_0}(1 \text{ Hz})$ we use values from Eqs. (4.2) and (4.3), whereas $S_{I_0,T}(1 \text{ Hz})$ is calculated for Al and Nb assuming $S_T(1 \text{ Hz})/(10^{-6}T)^2 = 2 \text{ Hz}^{-1}$ (a), $4 \text{ Hz}^{-1}$ (b), and $8 \text{ Hz}^{-1}$ (c). Here, $I_0 = 1 \mu\text{A}$ and $A = 1 \mu\text{m}$.

taken to sufficiently low frequencies. Thus, it is conceivable that a short measurement yielding $S_{I_0}(f)$, dominated at low frequencies by temperature variations, could produce a spectrum that mimics the expected $1/f$ noise in a narrow frequency range near the “knee”, the intersection with the white noise.

To investigate the effect of temperature noise at different temperatures, in Fig. 4.5 we compare typical experimental values of $S_{I_0}(1 \text{ Hz})$ to $S_{I_0,T}(1 \text{ Hz}) = (dI_0/dT)^2S_T(1 \text{ Hz})$, calculated for Al and Nb for three values of $S_T(1 \text{ Hz})$ representing experimentally realistic levels of temperature stability. As before, $(dI_0/dT)^2$ is derived from the Ambegaokar-Baratoff temperature dependence of $I_0(T)$. From our empirical observation that $\sigma_T(T)$ for a fixed set of PID parameters scales roughly as $T$, we choose $S_T(1 \text{ Hz})/(10^{-6}T)^2 = 2 \text{ Hz}^{-1}$, $4 \text{ Hz}^{-1}$ (corresponding to run C) and $8 \text{ Hz}^{-1}$. The long-dashed curve corresponds to critical current noise [41]

$$S_{I_0}(1 \text{ Hz}) = 144 \times 10^{-12}I_0^2 (T/4.2) \text{K}^2/(A/\mu\text{m}^2) \text{Hz},$$

(4.2)

while the short-dashed curve corresponds to the lower values recently achieved in both Al and Nb junctions [35]

$$S_{I_0}(1 \text{ Hz}) = 0.1 \times 10^{-12}I_0^2 (T/1) \text{K}/(A/\mu\text{m}^2) \text{Hz},$$

(4.3)

Qualitatively, at high enough temperatures the effect of the temperature noise dominates that of charge-trapping, even for the highest temperature stability. For junctions with noise
given by Eq. (4.2), the temperature-induced noise is larger in both Al and Nb junctions only for \( T > 0.75 T_c \). However, lower noise junctions obeying Eq. (4.3) are particularly difficult to measure, since temperature noise dominates at temperatures above \( T_c/4 \) in Nb junctions and \( T_c/3 \) in Al junctions. Consequently, as fabrication technologies improve and critical current noise is lowered, measurements of the noise become increasingly susceptible to temperature fluctuations.

There are several strategies to reduce the relative effect of temperature variations on measurements of critical current noise. (1) One can increase \( \tau_{\text{chip}} \) by decreasing the thermal conductivity to the MXC. While this strategy readily removes temperature-induced critical current fluctuations at frequencies above, say, \( 10^{-3} \) Hz, the corresponding time to stabilize the junction temperature at a given value would be correspondingly long, several hours. (2) One can measure the noise in two or more matched junctions in physical proximity. Low-frequency temperature variations produce highly correlated changes in the critical currents of the junctions, whereas noise caused by charge-trapping is uncorrelated between the junctions. Thus, removing the correlated noise leaves the intrinsic charge trapping noise. (3) In the case of unshunted junctions, one can deduce critical current noise from fluctuations in \( R_N \), since \( I_0 \) and \( R_N \) are subject to precisely the same fluctuations in the barrier transparency [30, 35, 80, 82, 88]: at a given temperature, \( S_{I_0}(f)/I_0^2 = S_{R_N}(f)/R_N^2 \). Compared to the critical current, the resistance is considerably less sensitive to temperature fluctuations. Finally, we note that, since \( S_{I_0} \) scales as \( I_0^2 \), one could increase \( S_{I_0} \) simply by increasing \( I_0 \). However, since \( S_{I_0;T} \) scales as \( (dI_0/dT)^2 \) and hence also as \( I_0^2 \), this approach does not diminish the relative magnitude of the temperature noise.

In conclusion, we have shown that temperature variations inherent in cryogenic systems, caused by imperfect temperature control, can contribute significantly to critical current noise, adding to the noise generated by charge trapping in the barrier. This contribution is temperature dependent, potentially confounding studies of the temperature dependence of the intrinsic critical current noise. Measurements on junctions with very low values of critical current noise using cryostats with poor temperature stability are particularly prone to this problem. For our measurement system with Al-AlOx-Al junctions, temperature fluctuations dominate the intrinsic critical current variations for \( T > T_c/3 \).

Acknowledgments

As previously mentioned, this work was possible due to the joint fabrication of the SQUID by Jim Eckstein’s and Dale Van Harlingen’s groups. Allison Dove, Gus Olson, and Zach Yoscovits deposited the washer of the SQUID and Chris Nugroho and Vlad Orlyanchik fabricated the junctions of the SQUID. I also wish to thank Chris Nugroho and Jeff Birenbaum for stimulating discussions that ultimately lead to our understanding of this phenomenon.
Chapter 5

Mean square flux noise in SQUIDs and qubits: numerical calculations

5.1 Introduction

To characterize the noise magnitude for a SQUID or qubit loop, various authors have calculated the total mean square flux noise \( \langle \Phi^2 \rangle \) for the loop assuming models with independent surface spins with a magnetic moment \( \mu_B \) (the Bohr magneton) that are uniformly distributed on the loop surface \([45, 55, 57]\). With certain assumptions regarding the frequency distribution of the spins, one can estimate the spectral density \( S_\Phi(1 \text{ Hz}) \). The first such model, which computed \( \langle \Phi^2 \rangle \) numerically, used existing measurements of \( S_\Phi(1 \text{ Hz}) \) to deduce an areal spin density \( \sigma = 5 \times 10^{17} \text{ m}^{-2} \) \([55]\), a number corroborated in subsequent measurements of paramagnetism in SQUIDs \([54]\) and Au rings \([56]\). This model also reproduced the relatively weak scaling of noise with geometry. The numerical method used in the model, however, is computationally slow and has poor accuracy and spatial resolution, making it unsuitable for studying qubit-sized loops. Subsequently, Bialczak et al. \([45]\) derived an analytic result for circular thin-film loops in the limit that the outer radius \( R \) was much greater than the loop linewidth \( W \).

These authors found that the dominant scaling was \( \langle \Phi^2 \rangle \propto R/W \), implying that loops with \( W \) approaching \( R \) should yield low-noise devices. However, loops with \( R \sim W \) violate the assumptions used to derive the scaling, so that this analytic model cannot be used to predict \( \langle \Phi^2 \rangle \) accurately. Furthermore, neither model accounts for the contribution of spins located on the edges of the superconducting film. In qubits, the linewidth and thickness and, consequently, the number of surface and edge spins, are often of the same order of magnitude so that the noise due to these edge spins may well be significant.

Accurate calculations of \( \langle \Phi^2 \rangle \) are important for a number of reasons. Among them, the ability to estimate the flux noise generated by a loop is invaluable in designing devices, particularly qubits, where direct measurements of flux noise are time consuming and must be carried out at dilution refrigerator temperatures. Furthermore, to improve our understanding of the mechanism of flux noise, it is essential to have an accurate means of translating the
5.2. METHODS

magnetic fields produced by fluctuating spins on a loop into the fluctuating flux threading the loop.

In this Chapter, we introduce a numerical method for calculating $\langle \Phi^2 \rangle$ in thin-film superconducting loops of arbitrary geometries, and present several results for square loops. We begin in Section 5.2 by reviewing the two existing methods and discussing their limitations in greater depth. We continue by explaining the framework for our numerical technique of calculating $\langle \Phi^2 \rangle$ that is computationally efficient, accurate, and valid for arbitrary geometries. Finally, we develop the model for the specific case of square loops, which are common among both SQUIDs and qubits. We continue in Section 5.3 by computing $\langle \Phi^2 \rangle$ for seven different scenarios where the loop dimensions and physical parameters are systematically varied, and discuss the results. We conclude in Section 5.4 with a brief summary.

5.2 Methods

The general method of finding $\langle \Phi^2 \rangle$ in a SQUID or qubit (hereafter referred to as “loop”) involves calculating the flux change $\Phi_s(r)$ in the loop due to the reversal of a single spin, located at $r$ and with magnetic moment of orientation $\hat{\mu}$, and summing $\Phi_s^2(r)$ for each spin on the loop surface. Provided that the number of spins is large, the spins are uniformly distributed on the surface and $\hat{\mu}$ is randomly oriented, the summation involves merely an integration of $\Phi_s^2(r)$ over the surfaces and angles. For certain loop geometries, $\Phi_s$ is known analytically for arbitrary $r$ and can be readily integrated. One can also compute $\Phi_s(r)$ numerically. With existing numerical methods, the computation must be repeated for each position of $r$ on the surface and also for three orthogonal spatial orientations. Furthermore, a separate computation is necessary for each particular loop geometry. Given this complexity, an analytic solution is highly desirable. However, such solutions exist only for very restricted geometries; to calculate $\langle \Phi^2 \rangle$ in arbitrary geometries, one must do so numerically.

5.2.1 Previous methods

Koch et al. [55] calculated $\Phi_s$ using the superconducting version of FastHenry [89, 90] to compute numerically the mutual inductance $M(x, y)$ between the device loop and a small test loop of area $a \approx (0.1 \ \mu m)^2$, representing the spin dipole moment, centred at $(x, y)$. Under these conditions, $\Phi_s(x, y) = M(x, y)\mu_B/a$. This calculation was repeated a large number of times as $x$ and $y$ were varied to cover the entire SQUID surface. While effective, this method of calculating $\langle \Phi^2 \rangle$ suffers from a number of drawbacks. (i) Because of the relatively large area of the test loop, the spatial resolution is limited to roughly $a^{1/2} \approx 0.1 \ \mu m$. This low resolution limits the accuracy of the computed contribution of the surface spins to structures larger than this and prohibits the calculation of the effect of edge spins, as film thicknesses are typically on the order of 0.1 $\mu m$. (ii) Because of the fixed direction of the vector area of the test loop, computing $\Phi_s$ for a spin of arbitrary direction requires one to compute the mutual inductance in the $x$, $y$ and $z$ directions separately. (iii) Because an accurate result
5.2. METHODS

for \( \langle \Phi^2 \rangle \) requires a dense sampling of dipole locations on the surface of the device loop, and a separate FastHenry calculation is required for each dipole location and for each direction \( \mu \) at that location, this method is computationally intensive.

The analytical method of Bialczak et al. [45] is radically different. Starting with the principle of reciprocity, the authors showed that the flux coupled into a loop by a dipole of magnetic moment \( \mu \) located at \( r \) is simply \( \Phi_s = B(r) \cdot \mu / I \), where \( B(r) \) is the magnetic field at \( r \) due to a circulating current \( I \) in the loop. With this insight, the calculation of \( \langle \Phi^2 \rangle \) is reduced simply to calculating \( B(r) \). Typically, this is accomplished by first calculating the distribution of \( I \) within the loop, then calculating the magnetic field \( B(r) \) using the Biot-Savart law or Ampere’s law. Finally, if one assumes randomly distributed and oriented spins,

\[
\langle \Phi^2 \rangle = \sum \Phi_s^2 = N\langle (B(r) \cdot \mu)^2 \rangle / I^2 = \frac{N\mu^2}{3I^2} \langle B^2(r) \rangle.
\]  

(5.1)

Thus, one needs only to calculate the mean square magnetic field \( \langle B^2 \rangle \) over the surface. Here, \( N \) is the number of spins on the surface of the washer, which is generally assumed to be large enough to justify a continuum approximation.

In this approach, the current distribution and resulting magnetic field can be calculated analytically in thin-film, circular loops. The magnetic field, integrated over the surface of the loop, yields [45]

\[
\langle \Phi^2 \rangle \simeq \frac{2\mu_0^3}{3} \mu_B^2 \sigma \frac{R}{W} \left[ \frac{\ln(2bW/\lambda^2)}{2\pi} + 0.27 \right],
\]  

(5.2)

where \( \mu_0 \) is the vacuum permeability, \( b \) is the film thickness, and \( \lambda \) is the penetration depth. This result is valid only for restricted geometries, namely \( R \gg W \gg \lambda \) and \( \lambda \sim b \). The leading order scaling, \( \langle \Phi^2 \rangle \propto R/W \), implies that loops with \( W \) approaching \( R \) would yield low-noise devices. However, this limit violates the assumptions used in the derivation of Eq. (5.2), so it cannot be used accurately to predict \( \langle \Phi^2 \rangle \). Under the stated restrictions, this analytical method represents an important step forward. However, many practical devices may not satisfy the geometric criteria for Eq. (5.2). For example, the use of Eq. (5.2) may be inaccurate in qubits in which \( R \) is not much greater than \( W \). Indeed, the very definition of \( R \) becomes ambiguous when \( R \approx W \) since the radius could vary by an order of magnitude depending on whether it is measured to the inside or outside of the linewidth. In addition, even under the stated restrictions Eq. (5.2) has further limitations. First, the magnetic field is calculated from the analytic current distribution [3] using Ampere’s law and the infinite sheet approximation, which breaks down when the current density changes on length scales comparable to the film thickness, as is often the case near the edges of the loop. Since \( \langle \Phi^2 \rangle \propto \langle B^2 \rangle \), these differences can introduce significant errors. Furthermore, the infinite sheet approximation neglects the perpendicular component of the magnetic field and, therefore, spins with magnetic moments pointed normal to the plane of the loop. Second, the magnetic field is calculated only on the upper surface of the loop, and neglects the potential...
5.2. METHODS

5.2.2 Numerical technique

We describe a numerical technique for efficiently and accurately calculating $B(\mathbf{r})$ and, using reciprocity and Eq. (5.1), $\langle \Phi^2 \rangle$. Although this technique is in principle applicable to arbitrary geometry, we confine our discussion to the widely used square-loop geometry, an example of which is shown in Fig. 5.1. Before describing the algorithm in detail, we summarize it as follows. Given a particular loop geometry, we first cut a slit into the loop and define two ports in the system along each of the slits (see Fig. 5.1). Next, we discretize the loop into a nonuniform mesh with InductEx [91, 92], which exports the mesh to FastHenry. To drive a circulating current, FastHenry establishes an alternating voltage difference between the ports and computes the current density $\mathbf{J}$ at positions defined by the

Figure 5.1: Current density in square loop with dimensions $R = 2 \, \mu m$ and $W = 1 \, \mu m$. Streamlines and colour contours represent the magnitude of the current density $|\mathbf{J}|$. The slit ($1 \, \mu m \leq x \leq 2 \, \mu m$, $-0.25 \, \mu m \leq y \leq 0.25 \, \mu m$) is required to force the current—sourced and sunk at opposite ports—to circulate around the entire loop.

contribution of edge spins. This may be particularly significant for qubits in which $W$ is comparable to $b$. 
mesh. Because the computed current density depends strongly on the particular mesh chosen, we iterate several times, using the computed current density and InductEx to optimize the mesh between iterations. Using the resultant current density, which we normalize such that \( I = 1 \ \mu A \), along with the Biot-Savart law, we next calculate \( \mathbf{B}(\mathbf{r}) \) over the surface and edges of the washer. Finally, \( \langle \mathbf{B}^2 \rangle \) is computed and converted to \( \langle \Phi^2 \rangle \) as discussed previously.

**Calculation of current density**

A precise calculation of the current density in the loop is essential, as it forms the basis for the rest of the calculation of \( \langle \Phi^2 \rangle \). To perform the bulk of this computation, we employ the 64-bit superconducting version of FastHenry [93], which computes the current distribution as an intermediary step to its designed purpose of inductance extraction. To perform this computation, a preprocessor in FastHenry first approximates a given thin-film structure by discretizing it into a large, two-dimensional network of rectangular prisms (hereafter referred to simply as “segments”) of height equivalent to the film thickness, each carrying an unknown current along its length that is uniform across its width and thickness. The FastHenry processor subsequently calculates the current density in each segment. The collection of current densities in the segments forms the numerical approximation to the two-dimensional current density. This discretization and solution process is described in detail in [89].

Unfortunately, this preprocessor suffers from a critical shortcoming in our application: the widths of all segments must be identical. This constraint forces one to use an excessively large number of segments to ensure high spatial resolution of the current density near the edges of the film, where the current density in superconductors typically changes most rapidly. For many geometries, the number of segments needed to compute an accurate current density would be far too large to be computationally feasible. For example, a relatively narrow 10-\( \mu \)m linewidth could be discretized into ten 1-\( \mu \)m-wide segments. However, this limits the spatial resolution of the current distribution to 1 \( \mu \)m, which is insufficient to capture the rapidly varying nature of the current density near the edges of the film. Simply using more segments that are narrower rapidly becomes computationally intractable. It can be shown that the computation time scales roughly as \( O(N_{\text{seg}}^3) \), where \( N_{\text{seg}} \) is the number of segments; furthermore, \( N_{\text{seg}} \) scales roughly as the inverse of the square of the segment width. Clearly, one needs a discretization algorithm that implements segments of different widths. Wide segments are appropriate near the middle of the loop linewidth, where the current density varies slowly, while narrow segments are appropriate near the edges, where the current density changes rapidly.

To implement segments with different widths, we use InductEx to replace the FastHenry preprocessor. InductEx allows one to specify the desired width of each segment and subsequently generates a discretized geometry file that can be directly imported by FastHenry, bypassing its own internal preprocessor. Figure 2(a) shows the computed current density using variable-width segments in a cross section across the linewidth of the geometry in Fig. 5.1. We see that narrow segments near the inner and outer edges accurately capture the rapidly changing current density, while wider segments in the middle of the linewidth save
5.2. METHODS

Figure 5.2: Current density across linewidth and corresponding magnetic field. (a) Computed and analytic current density for $-2 \mu m \leq x \leq -1 \mu m$ and $y = 0 \mu m$ showing nonuniform segment widths. Note that segments are narrower near edges than in the middle region. (b) Cross-sectional view of linewidth with magnetic field generated by computed current in (a).
computation time \[94\]. We note that in certain geometries the segment widths can vary by three orders of magnitude, highlighting the absolute necessity of this technique.

Because the ideal widths of the segments are not known precisely until the current density is well characterized, we iterate the solution of the current density several times, adjusting the segment widths between iterations. A fixed algorithm generates the initial widths of the segments, which are used in the first iteration to solve for current density. Subsequently, another algorithm analyzes the current density, optimizes the segment widths and, if necessary, increases the number of segments. The current density is recomputed using the optimized segment widths. This technique typically converges within three iterations to an optimal set of segment widths, which in turn leads to a well characterized current density. Figure 2(a) also clearly illustrates the inaccuracy of the analytic method as \( R \) approaches \( W \). In this situation, the current density tends to accumulate at the edge of the linewidth that borders the hole, as the numerical simulations show. The analytic current density, however, is always assumed to be symmetric about the centre of the linewidth.

**Calculation of magnetic field**

To calculate the magnetic field at an arbitrary position \( \mathbf{r} \), we sum the contributions of all the current-carrying segments using the Biot-Savart law,

\[
\mathbf{B}(\mathbf{r}) = \sum_{n=1}^{\text{seg}} \frac{\mu_0}{4\pi} \int_{V_n} \frac{(\mathbf{J}_n dV) \times \mathbf{s}}{s^3}. \tag{5.3}
\]

Here, \( V_n \) is the volume of the \( n \)th segment and \( \mathbf{s} \) is the vector pointing from the differential volume element \( dV \) to \( \mathbf{r} \). Because the current density within each segment is uniform, we can simplify the integral as

\[
\int_{V_n} \frac{(\mathbf{J}_n dV) \times \mathbf{s}}{s^3} = \mathbf{J}_n \times \int_{V_n} \frac{s dV}{s^3}. \tag{5.4}
\]

While in principle we could perform this integration numerically, in practice this is undesirable because of computational complexity and numerical instabilities in calculating the field on the surface of a segment, where the denominator in the integrand approaches zero. Instead, for a rectangular prism, we solve Eq. (5.4) analytically as

\[
\int_{V_n} \frac{s dV}{s^3} = \Theta(s) \bigg|_{V_n}, \tag{5.5}
\]

where the auxiliary function \( \Theta(s) \), which we calculate as

\[
\Theta_i(s) = s_j - s_i \arctan \left( \frac{s_j}{s_i} \right) + s_i \arctan \left( \frac{s_s s_k}{s_i \xi} \right) - s_k \log(2(s_j + \xi)) - s_j \log(2(s_k + \xi)), \tag{5.6}
\]
is evaluated at the eight corners of the segment. Here, \( \xi \equiv (s_i^2 + s_j^2 + s_k^2)^{1/2} \) and \( i, j, \) and \( k \) are cyclic permutations of \( x, y, \) and \( z \). Using Eq. (5.3)–Eq. (5.6), which are straightforward to parallelize to run on multi-core processors, we can rapidly and accurately compute at arbitrary positions the magnetic field generated by the circulating current in the loop.

Corresponding to the current density in Fig. 5.2(a), Fig. 5.2(b) shows a vector plot of the magnetic field calculated using the above method. Several features are noteworthy. First, we make the obvious point that, due to symmetry, the magnitude of the magnetic field at the top surface is equal to that at the bottom surface. Consequently, in a physical device we expect spin reversals at both the exposed surface and the interface with the substrate to contribute equally to flux noise. Second, we see that the perpendicular component of the magnetic field is significant near the film edges where the current changes rapidly. Third, the magnetic field is strongest at the edges of the film, implying that any edge-spin reversal couples strongly to the loop. Finally, as with the current density, the magnetic field is much larger on the inner edge of the loop.

**Calculation of \( \langle \Phi^2 \rangle \)**

To compute \( \langle \Phi^2 \rangle \) in the loop we need only calculate \( \langle B^2 \rangle \). We ignore the effect of the slit, which we introduced solely to drive a circulating current. Since the current density near the ports is locally distorted, we compute \( \langle B^2 \rangle \) on the surface of one quarter of the superconductor opposite the slit—specifically, on the loop surface for which \( x < 0 \) and \( y < 0 \)—which, by symmetry, is equal to \( \langle B^2 \rangle \) for the entire loop (without the slit). To calculate \( \langle B^2 \rangle \), we first compute \( B(\mathbf{r}) \) at a nonuniformly sampled set of points on the surface of the loop corresponding to the nonuniformity of the segment lengths and widths. Next, \( B(\mathbf{r}) \) is squared and averaged, yielding \( \langle B^2 \rangle \) which is converted to \( \langle \Phi^2 \rangle \) using Eq. (5.1). For both top and bottom surface spins and for edge spins, we have used an areal spin density \( \sigma = 2.5 \times 10^{17} \text{ m}^{-2} \), one-half of the “canonical” spin density, which was inferred assuming spins on only one surface.

**Estimation of numerical accuracy**

Potential errors in our calculation of \( \langle \Phi^2 \rangle \) can be divided into three categories: (i) the error associated with the discretization prepared for InductEx and FastHenry, (ii) the error within the InductEx/FastHenry calculation and (iii) the error in calculating \( \langle \Phi^2 \rangle \) from the current distribution.

To estimate (i) we calculated \( \langle \Phi^2 \rangle \) for several fixed geometries while varying the number of segments in the mesh as well as the number of iterations as described in Section subsec:calcJ. We found that doubling our default number of segments across the linewidth changed \( \langle \Phi^2 \rangle \) by less than 5\% and doubling the default number of iterations changed \( \langle \Phi^2 \rangle \) by less than 1\%. Furthermore, \( \langle \Phi^2 \rangle \) asymptotes to a constant value for each geometry examined as the number of segments increases. With regard to (ii), the current density in superconducting integrated circuit structures in the micrometre range has a strong influence on inductance—both through magnetic and kinetic terms—and recent comparisons of measured inductances
with those computed with InductEx showed agreement to within 3% [95], even for complex structures with ground plane holes [96]. We therefore estimate (ii) to be small. Finally, because we calculated \( \mathbf{B}(\mathbf{r}) \) analytically for a single current-carrying segment and \( \langle \mathbf{B}^2 \rangle \) is simply a straightforward spatial average of \( \mathbf{B}(\mathbf{r}) \), we conclude that (iii) is negligible. We therefore conclude that our method is well converged and that the accuracy of \( \langle \Phi^2 \rangle \) is roughly \( \pm 5\% \) for our choice of parameters. By increasing the number of segments, one can increase the accuracy at the cost of computation time.

5.3 Results and discussion

Using the technique described in Section 5.2, we compute and plot \( \langle \Phi^2 \rangle \) for several scenarios in which the loop parameters are individually and systematically varied. Unless specified otherwise, we use \( b = 150 \, \text{nm} \) and \( \lambda = 90 \, \text{nm} \). Here, the choice \( \lambda = 90 \, \text{nm} \) is based on microwave surface impedance measurements of the penetration depth versus film thickness in Nb thin films for a film thickness of 150 nm [97]. We later consider the effect of varying \( \lambda \) over a broad experimentally relevant range. We compute the total mean square flux noise \( \langle \Phi^2_{\text{tot}} \rangle \), the sum of the contributions from all surface spins \( \langle \Phi^2_{\text{surf}} \rangle \)—that is, surface spins with \( \mathbf{\mu} \) pointed in-plane with the loop \( \langle \Phi^2_{\parallel} \rangle \) and surface spins with \( \mathbf{\mu} \) pointed perpendicularly to the loop \( \langle \Phi^2_{\perp} \rangle \)—and all edge spins with both in-plane and perpendicular spins \( \langle \Phi^2_{\text{edge}} \rangle \):

\[
\langle \Phi^2_{\text{tot}} \rangle = \langle \Phi^2_{\text{surf}} \rangle + \langle \Phi^2_{\text{edge}} \rangle = \langle \Phi^2_{\parallel} \rangle + \langle \Phi^2_{\perp} \rangle + \langle \Phi^2_{\text{edge}} \rangle.
\] (5.7)

We compare these results with the analytical predictions of Eq. (5.2), which we multiply by a factor \( 4/\pi \) to convert from a circular to a square geometry. We note that while the analytic derivation assumes \( R \gg W \), to explore its limitations we plot Eq. (5.2) even as \( R \) approaches \( W \). In this limit, because of the ambiguity in the definition of \( R \), we use three definitions of \( R \) when evaluating Eq. (5.2): the radius measured to the inner edge \( (R_{\text{ins}}) \), midpoint \( (R_{\text{mid}}) \), and outer edge \( (R) \) of the linewidth.

5.3.1 Role of loop radius

In Fig. 5.3 we plot the various components of \( \langle \Phi^2 \rangle \) versus \( R_{\text{ins}} \) for loops of constant linewidth \( W = 1 \) and 10 \( \mu \text{m} \), where Eq. (5.2) simplifies to \( \langle \Phi^2 \rangle \propto R \). Indeed, in the limit \( R \gg W \) our simulations confirm this scaling for all components of \( \langle \Phi^2 \rangle \), which can be understood as follows. For fixed \( W \) and \( R \gg W \), each side of the loop can be approximated as an isolated line, remote from any other superconductors. As such, the current distribution—and, therefore, the resulting magnetic field magnitude and coupling strength to spins on the surface—is symmetric across the linewidth and independent of \( R \). Consequently, for fixed \( W \), \( \langle \Phi^2 \rangle \) is proportional to the number of spins and, therefore, the circumference of the loop, \( 8R \). Naturally, this scaling breaks down as \( R \) approaches \( W \), where the current distribution
5.3. RESULTS AND DISCUSSION

is no longer symmetric and independent of $R$ and the contribution from the corners of the loop is no longer negligible. In this limit, the value of Eq. (5.2) depends greatly on the particular definition of $R$. Defining the radius as $R$ or $R_{\text{mid}}$ and plotting $\langle \Phi^2 \rangle$ versus $R_{\text{ins}}$, Eq. (5.2) predicts that $\langle \Phi^2 \rangle$ levels off to a constant value as $R - W = R_{\text{ins}}$ approaches zero. Our numerical computations also show $\langle \Phi^2 \rangle$ levelling off to an approximately constant value. On the other hand, defining the radius as $R_{\text{ins}}$, Eq. (5.2) predicts that $\langle \Phi^2 \rangle$ decreases without bound, contradicting our simulations. The trend of $\langle \Phi^2 \rangle$ to a constant value implies that lowering flux noise by reducing the size of the hole in the loop becomes much less effective once $R \approx W$.

We now comment on the magnitudes of each of the components of $\langle \Phi^2 \rangle$ in Fig. 5.3. Because Eq. (5.2) includes only the contribution of the parallel component of spins on the surface, we might expect Eq. (5.2) and $\langle \Phi_\parallel^2 \rangle$ to agree well for $R \gg W$. However, in this limit for both figures 5.3(a) and 5.3(b) the analytic result exceeds $\langle \Phi_\parallel^2 \rangle$ by approximately 70% and 40%, respectively. In addition, the contributions of corner effects are small and we can typically reconcile the numerical and analytical results to within a few percent by accounting for the following errors and discrepancies. (i) Our numerical current density and analytical current density used in [45] agree well in the middle of the linewidth but diverge by up to a factor of 2 near the edges [see figures 5.4(a) and 5.4(b)]. (ii) As discussed in Section subsec:prevmeth, the Ampere-law approximation loses accuracy when the current distribution varies rapidly [see figures 5.4(c) and 5.4(d)]. (iii) The analytic current distribution used to

![Figure 5.3: Numerically computed and analytic $\langle \Phi^2 \rangle$ versus $R_{\text{ins}}$ and $R/W$ for constant $W$. Blue lines represent Eq. (5.2) evaluated for $R$ (- - -), $R_{\text{mid}}$ (---), and $R_{\text{ins}}$ (--).](image-url)
Figure 5.4: Comparison of numerical and analytic results. (a) and (b) Current densities and (c) and (d) calculated surface magnetic fields across the loop linewidth in the centre ($y = 0$). Here, $R = 250 \, \mu m$ and $W = 10 \, \mu m$. 
derive Eq. (5.2) is not integrated all the way to the edges of the linewidth. (iv) The logarithm term in Eq. (5.2) originates from a series expansion of a more accurate functional form. Together, the magnitudes of the errors (i-iv) are 20%, 64%, 25% and 17%, respectively, for $R = 25 \mu m$ and $W = 1 \mu m$ [see Fig. 5.3(a)]; for $R = 250 \mu m$ and $W = 10 \mu m$, the errors are 15%, 34%, 18% and 6%, respectively [see Fig. 5.3(b)]. As $R$ approaches $W$, the contributions of the loop corners and the increasing current density around the hole become significant and the analytic formula becomes progressively less accurate.

In addition to the contribution of parallel surface spins, Fig. 5.3 shows the contribution of perpendicular surface spins and edge spins, which are not included in Eq. (5.2). In both figures 5.3(a) and 5.3(b), we see that the contribution of perpendicular surface spins $\langle \Phi^2_{\parallel} \rangle$ scales roughly as $\langle \Phi^2_{\parallel} \rangle$ and is relatively small, typically 10% of $\langle \Phi^2_{\parallel} \rangle$ particularly for large $R_{\text{ins}}$. The edge-spin contribution, however, shows a different dependence. For large values of $R/W$, $\langle \Phi^2_{\text{edge}} \rangle$ scales as $R$, with a relative magnitude that, as expected, is smaller in the loop with the wider linewidth. We expect the relative contribution of $\langle \Phi^2_{\text{edge}} \rangle$ to decrease as $W$ increases because the ratio of the number of surface spins to edge spins increases. As $R/W$ decreases, the current density at the edge near the hole increases rapidly near $R \approx W$ so that $\langle \Phi^2_{\text{edge}} \rangle$ diverges from a scaling with $R$ and instead approaches a constant. In Fig. 5.3(a), $\langle \Phi^2_{\text{edge}} \rangle$ actually exceeds $\langle \Phi^2_{\text{surf}} \rangle$ by a factor of 1.9 at the lowest $R_{\text{ins}}$ so that edge spins are the dominant contribution to $\langle \Phi^2_{\text{tot}} \rangle$. This result is important for qubits, which typically have narrow linewidths ($W \lesssim 10 \mu b$) and for which $R$ is not necessarily much greater than $W$.

### 5.3.2 Role of loop linewidth for fixed hole size

In Fig. 5.5 we plot $\langle \Phi^2 \rangle$ versus $W$ for two constant hole sizes, $2(R-W) = 2 \mu m$ and $10 \mu m$. In the limit of large $R$, the majority of the circulating current accumulates near the edge bordering the hole with a distribution independent of $R$; therefore, we expect $\langle \Phi^2 \rangle$ asymptotically to become constant. We note that because $R/W$ approaches unity for increasing $W$, Eq. (5.2) coincidentally also predicts $\langle \Phi^2 \rangle$ to approach a constant using the definitions $R$ and $R_{\text{mid}}$. As $W$ approaches zero, Eq. (5.2) predicts that $\langle \Phi^2_{\text{surf}} \rangle$ will increase rapidly as $1/W$, which we observe in Fig. 5.5(b). In Fig. 5.5(a), however, $\langle \Phi^2_{\text{surf}} \rangle$ flattens off and even begins to decrease as $W$ decreases, because of two competing effects. On the one hand, the average current density—and thus the approximate average magnetic field and coupling strength of surface spins—across the linewidth scales as $1/W$, which tends to increase $\langle \Phi^2 \rangle$ as $W$ decreases. On the other hand, as $W$ approaches $\lambda = 90 \mu m$, the current distribution becomes increasingly uniform across the linewidth. Because the contribution to $\langle \Phi^2 \rangle$ scales as $B^2$, which in the Ampere-law approximation is proportional to $J^2$, this effect tends to decrease $\langle \Phi^2 \rangle$ as $W$ decreases. The peak in $\langle \Phi^2 \rangle$ for surface spins at approximately $W = 0.2 \mu m$ represents the point at which the latter effect begins to dominate the former for decreasing $R$. We note that the analytic current distribution does not show this behaviour because it assumes $W \gg \lambda$. Finally, we see that $\langle \Phi^2_{\text{edge}} \rangle$ scales as $1/W$ in both figures 5.5(a) and 5.5(b). As the thickness is constant, $\langle \Phi^2_{\text{edge}} \rangle$ is roughly proportional to the average current density at the edges, which for these geometries scales approximately as $1/W$. 
5.3. RESULTS AND DISCUSSION

5.3.3 Role of loop aspect ratio

In Fig. 5.6 we plot $\langle \Phi^2 \rangle$ versus $R/W$ for two constant values of $R$: 10 and 100 $\mu$m. According to Eq. (5.2), for this configuration we expect $\langle \Phi^2 \rangle$ to scale as $1/W$ with a small correction that depends logarithmically on $W$. We observe this scaling for the surface spins for $R/W \gtrsim 3$, below which the noise tends to drop rapidly with decreasing $R/W$.

5.3.4 Role of loop linewidth for fixed aspect ratio

In Fig. 5.7 we plot $\langle \Phi^2 \rangle$ versus $W$ for constant values of $R/W = 2$ and 20. For constant $R/W$, Eq. (5.2) predicts that $\langle \Phi^2 \parallel \rangle$ should increase logarithmically with $W$. Indeed, our simulations show that $\langle \Phi^2 \rangle$ for surface spins increases with $W$, although more rapidly than the analytic prediction. The contribution from edge spins, however, actually decreases slightly with $W$. It is notable that in both cases $\langle \Phi^2_{\text{tot}} \rangle$ changes only by a factor of about 2 for a 200-fold change in $W$, from 0.5 to 100 $\mu$m, confirming the dominance of the $R/W$ scaling.

5.3.5 Role of penetration depth

In Fig. 5.8 we investigate the role of the penetration depth by plotting $\langle \Phi^2 \rangle$ versus $\lambda$ for two fixed geometries. As changing $\lambda$ changes the current distribution only moderately, Eq. (5.2) predicts a logarithmic sensitivity in $\langle \Phi^2 \rangle$ for surface spins. In fact, our simulations
5.3. RESULTS AND DISCUSSION

Figure 5.6: \( \langle \Phi^2 \rangle \) versus \( R/W \) for two constant values of \( R \). Blue lines represent Eq. (5.2) evaluated for \( R \) (---), \( R_{\text{mid}} \) (---), and \( R_{\text{ins}} \) (---).

Figure 5.7: \( \langle \Phi^2 \rangle \) versus \( W \) for two constant values of \( R/W \). Blue lines represent Eq. (5.2) evaluated for \( R \) (---), \( R_{\text{mid}} \) (---), and \( R_{\text{ins}} \) (---).
Figure 5.8: $\langle \Phi^2 \rangle$ versus $\lambda$ for two fixed geometries. Blue lines represent Eq. (5.2) evaluated for $R$ (---), $R_{\text{mid}}$ (----), and $R_{\text{ins}}$ (--.--).

show very little change in $\langle \Phi^2_{\text{surf}} \rangle$ and $\langle \Phi^2_{\perp} \rangle$, which is related to the spatial derivative of the current density, increases rather significantly with $\lambda$. In addition, as $\lambda$ increases the peak current density near the edges ($J_{\text{edge}}$) decreases and the current density across the linewidth increases. The magnetic field is sensitive to currents within a distance on the order of the film thickness, which tends to average the effects of a rapidly changing current distribution. Therefore, the contribution of the edge spins, which is roughly proportional to $J_{\text{edge}}$, decreases significantly with increasing $\lambda$. These combined effects tend to make $\langle \Phi^2_{\text{tot}} \rangle$ relatively insensitive to changes in $\lambda$ over experimentally relevant values.

5.3.6 Role of edge spins

To consider in greater depth the contribution of edge spins, in Fig. 5.9 we replot Fig. 5.5(a)—$\langle \Phi^2 \rangle$ versus $W$ for constant hole size $2(R - W) = 2 \mu m$—where we now separate the contributions of spins on the inner edge, bordering the hole, from those on the outer edge. In the limit of large $W$ and small $R/W$, the majority of the current circulates closly around the hole. Therefore, we expect the contribution of the inner edge spins to dominate. As $W$ decreases and $R/W$ increases, the current density becomes increasingly symmetric across the linewidth. Correspondingly, the contribution of the outer edge spins increases rapidly and asymptotes to that of the inside edge spins.
5.3. RESULTS AND DISCUSSION

![Figure 5.9: Relative contributions to $\langle \Phi^2 \rangle$ of inside edge spins $\langle \Phi^2_{\text{ins}} \rangle$ and outside edge spins $\langle \Phi^2_{\text{out}} \rangle$ versus $R/W$ for two constant hole sizes. Blue lines represent Eq. (5.2) evaluated for $R$ (---), $R_{\text{mid}}$ (--.--), and $R_{\text{ins}}$ (---).]

5.3.7 Role of film thickness

Finally, we compute $\langle \Phi^2 \rangle$ versus the loop thickness $b$, plotting the result in Fig. 5.10 for $20 \text{ nm} \leq b \leq 200 \text{ nm}$. Equation Eq. (5.2) predicts that $\langle \Phi^2_{\text{surf}} \rangle$ increases logarithmically with $b$. Our simulations show, however, that $\langle \Phi^2_{\text{surf}} \rangle$ actually decreases with $b$, a discrepancy that originates in slight differences between the numerical and analytic current distributions. More surprisingly, the contribution of the edge spins increases dramatically with the thickness, approximately as $\langle \Phi^2_{\text{edge}} \rangle \propto b$. A priori, we might expect the opposite: that the current density, and thus the magnetic field, would scale approximately as $1/b$ so that $\langle \Phi^2_{\text{edge}} \rangle \propto B^2 N \propto J^2 b \propto 1/b$. Examining the numerical results, however, we find empirically that the peak current density near the edge ($J_{\text{edge}}$) scales roughly as $J_{\text{edge}} \propto b^{-1/2}$ and the magnetic field along the edge ($B_{\text{edge}}$) scales as $B_{\text{edge}} \propto J_{\text{edge}} b^{1/2} \propto \text{constant}$. Therefore, $\langle \Phi^2_{\text{edge}} \rangle \propto B_{\text{edge}}^2 b \propto b$, as we observe. This unexpected result has important implications for qubits, which are often made from thin-films with thicknesses of just tens of nanometres. Our numerical results show that, counter-intuitively, thinner films should indeed exhibit a lower flux noise.
5.4 Concluding remarks

Accurate calculations of $\langle \Phi^2 \rangle$ are essential to the development and improvement of both SQUIDs and qubits. We have developed a numerical method—suitable for calculating $\langle \Phi^2 \rangle$ for arbitrary loop geometries—that is robust, computationally efficient, and accurate. We compared our computations of $\langle \Phi^2 \rangle$ to a previous analytic calculation of $\langle \Phi^2 \rangle$ and find significant discrepancies, even in the limit $R \gg W$, where the two should agree. By accounting for inaccuracies arising from approximations used in deriving the analytic formula as well as differences in the analytic and numerically computed current distributions, we reconcile these discrepancies to within a few percent. We computed the effect of edge spins and find their contribution to be significant in narrow-linewidth ($W \lesssim 10b$) devices. As the film thickness is varied, we find that $\langle \Phi^2_{\text{surf}} \rangle$ decreases as $b$ increases, an opposite trend to that predicted by Eq. (5.2). Our calculations are particularly relevant for qubits, in which the linewidth is often narrow, $R$ is not necessarily much greater than $W$, and typical film thicknesses range between approximately 30 and 100 nm.

Recent experiments suggest that neighboring spins form clusters with a magnetic moment that may be at least several times greater than $\mu_B$ [58, 60]. For clusters small in spatial extent—that is, smaller than both the smallest dimension of the loop and the length scale on which the magnetic field changes—our calculations for $\langle \Phi^2 \rangle$ can be readily modified. As an example, suppose the clusters are composed of $Z$ single spins ($Z \in \{1, 2, 3, \ldots \}$ is fixed for all clusters) and are ferromagnetic, with a magnetic moment $\mu_z = Z \mu_B$ and areal density

![Graph](image_url)
\( \sigma_Z = \sigma/Z \). Assuming that the average areal density of spins is unchanged by clustering, for uncorrelated clusters one finds \( \Phi_s \propto \mu_Z \) and \( \langle \Phi^2 \rangle \propto \sigma_Z \mu_Z^2 \propto Z \). Similar, straightforward modifications can be made for antiferromagnetic clusters, in which \( \mu_Z \) is either unity (\( Z \) odd) or zero (\( Z \) even), and for “glassy” (random) clusters, in which \( \mu_Z = Z^{1/2} \mu_B \). For scenarios in which \( Z \) varies between clusters, \( \langle \Phi^2 \rangle \propto \sum_{Z=1}^{\infty} \sigma_Z \mu_Z^2 \). Furthermore, our method of numerically calculating \( \Phi_s(\mathbf{r}) \) remains valid even if interactions between spins or clusters are significant. However, in this scenario the relation \( \langle \Phi^2 \rangle = \sum \Phi_s^2 \) is no longer generally true; one requires knowledge of the interaction to sum the contributions from all spins and clusters properly.

Finally, we emphasize that our simulations—like all previous methods—yield \( \langle \Phi^2 \rangle \) whereas experiments determine the flux noise spectral density, \( S_\Phi(f) = A^2/(f/1 \text{ Hz})^\alpha \). The two quantities are related by \( \langle \Phi^2 \rangle = \int_{f_1}^{f_2} S_\Phi(f) df \). Clearly, a knowledge of the quantities \( f_1, f_2, A \) and \( \alpha \) is required to obtain an explicit prediction of the spectral density from \( \langle \Phi^2 \rangle \). In practice, however, one measures \( S_\Phi(f) \) over several decades of frequency and fits it to determine \( A \) and \( \alpha \). Furthermore, to an excellent approximation we can set \( f_1 = 0 \) provided \( \alpha \gtrsim 0.9 \), which generally appears to be the case [58]. Thus, if we assume \( \alpha \) remains constant over the entire frequency range, an explicit connection between \( \langle \Phi^2 \rangle \) and \( \int S_\Phi(f) df \) requires only a knowledge of \( f_2 \). The value of \( f_2 \), however, is currently the subject of debate.

Acknowledgments

This project sprouted out of an undergraduate project in 2009 undertaken by Ida Sognnaes, who made admirable progress on a problem that we quickly realized was far too complicated for a semester project. Subsequently, Keenan Pepper spent more nearly two years grappling with the issue of numerically calculating the current distribution in a superconducting washer. His work yielded several key insights, but we failed to obtain the numerical accuracy necessary to solve the problem.

Without question, our big break came after becoming aware of InductEx and working extensively with Coenrad Fourie, InductEx’s creator. I met Coenrad at a poster session at the Applied Superconductivity Conference in Portland in October of 2012, a point at which the project was stalled and the calculation seemed intractable. In just a few short months, Coenrad and I had largely resolved all numerical issues and had begun drafting the results of our work. We cannot thank Coenrad too much for his tireless efforts assisting us in this calculation.
Chapter 6

Dephasing in qubits due to flux noise

6.1 Introduction

The dynamics of superconducting quantum bits (qubits) [18]—broadly classified as charge qubits [98], flux qubits [99] and phase qubits [100]—can be characterized by two times: the relaxation time $T_1$ and the pure dephasing time $\tau_\phi$ [52], as discussed in Section 1.4.2. The time $T_1$ required for a qubit to relax from its first excited state to its ground state is determined by the strength of environmental fluctuations at a frequency corresponding to the energy level splitting $\nu$ of the two states. The decoherence time $T_2$, over which the phase of superpositions of two eigenstates becomes randomized, has two contributions: $1/T_2 = 1/(2T_1) + 1/\tau_\phi$. The pure dephasing time $\tau_\phi$ is limited by fluctuations in $\nu$, due predominantly to fluctuations in magnetic flux in the case of flux qubits.

Measurements of $\tau_\phi$ in flux qubits [44, 46–49] and phase qubits [45] have been used to infer the magnitude of the flux noise in these devices, under the assumption that the spectral density of the flux noise scaled as $1/f$. In this Chapter, we show theoretically that deviations in $\alpha$ from unity strongly impact $\tau_\phi$: the value of $\tau_\phi$ decreases markedly with decreasing $\alpha$. Additionally, we examine the influence of $\alpha$ and noise bandwidth on $\tau_\phi$, the ratio $\tau_\phi,E/\tau_\phi,R$ obtained in echo ($E$) and Ramsey ($R$) pulse sequences, and the functional dependence of the decay function.

6.2 Model

It is natural to ask what impact nonunity values of $\alpha$ have on the pure dephasing of flux qubits. Low-frequency flux noise modulates the energy splitting of the ground and first excited states of a flux qubit, $\hbar\nu = [\Delta^2 + \epsilon^2(\Phi)]^{1/2}$, via the bias energy $\epsilon(\Phi)$. The bias energy is the energy difference between the two states with persistent currents $\pm I_q$ when there is no tunneling between them ($\Delta = 0$) [101]. Here, $\epsilon = 2I_q(\Phi - \Phi_0/2)$ or equivalently, $I_q \equiv \frac{1}{2}(\partial \epsilon / \partial \Phi)$.

We define the sensitivity of the splitting to a change in $\Phi$ in terms of the longitudinal
6.2. MODEL

sensitivity of the qubit to flux noise,

\[ D_\Phi \equiv \frac{\partial \nu}{\partial \Phi} = \frac{1}{h} \left( \frac{\partial \epsilon}{\partial \Phi} \right) \epsilon / (\Delta^2 + \epsilon^2)^{1/2}. \]  

(6.1)

To first order, there is no dephasing from flux noise at the degeneracy point, \( \Phi = \Phi_0/2 \), where \( \epsilon \) vanishes. In this Chapter, however, we consider the limit \( \epsilon / \Delta \gg 1 \), far from the degeneracy point, at which \( \partial \nu / \partial \Phi = (1/h) \partial \epsilon / \partial \Phi = 2I_q/h \). We assume that, in this limit, \( 1/\tau_\phi \gg 1/2T_1 \) so that the measured dephasing time arises only from pure dephasing. We adopt the value \( D_\Phi = 10^{12} \) Hz/\( \Phi_0 \), corresponding to the typical value \([44, 46]\) \( I_q \approx 0.3 \) \( \mu \)A. Furthermore, based on the empirical observation that \( S_\Phi(1 \) Hz\) is relatively constant among a wide variety of SQUIDs, we assume that \( A = 1 \) \( \mu \Phi_0 \) Hz\(^{-1/2} \) regardless of the value of \( \alpha \).

The modulation of \( \nu \) by flux noise leads to an accumulation of phase error and thus to dephasing. The rate at which the dephasing occurs varies between different types of pulse sequences. For example, in a Ramsey sequence \([102]\) the qubit is excited by a microwave \( \pi/2 \) pulse from the ground state into a superposition of ground and excited states. After a time \( t \) another \( \pi/2 \) pulse is applied and the qubit state is measured. The results of many measurements with fixed \( t \) are averaged and \( t \) is varied from \( t \ll \tau_\phi \) to \( t \gg \tau_\phi \) to obtain the decay function \( g(t) \). Here, we define the dephasing time as \( g(\tau_\phi) \equiv 1/e \). To eliminate dephasing due to flux fluctuations between pulse sequences, one implements an echo sequence in which a \( \pi \) pulse is inserted midway between the two \( \pi/2 \) pulses \([103]\). In general, the echo sequence yields a dephasing time greater than that of the Ramsey: \( \tau_{\phi,E} > \tau_{\phi,R} \).

The sensitivity of the Ramsey and echo sequences to noise are described by the weighting function \( W(f, t) \) given by \([43, 100]\)

\[ W_R(f, t) = \frac{\sin^2(\pi ft)}{\pi^2 t^2}, \quad W_E(f, t) = \frac{\sin^4(\pi ft/2)}{\pi^2 (ft/2)^2}. \]  

(6.2)

For the Ramsey sequence with \( f \ll 1/t \), we see that \( W_R \approx 1 \), whereas for \( f \gg 1/t \), \( W_R \) falls as \( 1/f^2 \). Consequently, we expect the dominant contributions to the Ramsey dephasing time to arise from noise at frequencies \( f \lesssim 1/\tau_{\phi,R} \). In contrast, for the echo sequence \( W_E \) scales as \( f^2 \) for \( f \ll 1/t \) and as \( 1/f^2 \) for \( f \gg 1/t \). In this case, we expect the dominant contribution to the dephasing time to be from noise at frequencies \( f \approx 1/\tau_{\phi,E} \).

The decay function \( g(t) \) is calculated by ensemble averaging over the entire measurement time, yielding \([43, 100]\)

\[ g(t) = \exp \left[ -t^2 (2\pi D_\Phi)^2 \int_{f_1}^{f_2} df S_\Phi(f) W(f, t) \right]. \]  

(6.3)

Here, the symmetrized noise power is defined as

\[ S_\Phi(f) \equiv \frac{1}{2} \int dt \{ \langle \Phi(t) \Phi(0) \rangle + \langle \Phi(0) \Phi(t) \rangle \} e^{-2\pi ift}. \]  

(6.4)
6.3. Results

6.3.1 Dephasing times versus $\alpha$

As is evident from Eq. (6.3), a nonunity value of $\alpha$ will affect the integral in a complicated way. Figure 6.1 shows computed dephasing times for both sequences versus $\alpha$ for $f_2 \to \infty$ and $f_1 = 10^{-3}$, $10^{-1}$, $10^1$, and $10^3$ Hz. The effect of changing $\alpha$ is substantial: both $\tau_{\phi,R}$ and $\tau_{\phi,E}$ increase by an order of magnitude as $\alpha$ is varied from 0.6 to 0.9. By comparison, we find that an order of magnitude change in $A$ for a given value of $\alpha$ also changes $\tau_{\phi}$ by an order of magnitude. Figure 6.1 further shows that, because of its insensitivity to low-frequency noise, the echo sequence yields significantly longer dephasing times for all $\alpha$. Finally, while
6.3. RESULTS

\( \tau_{\phi,E} \) is insensitive to changes in \( f_1 \) for \( f_1 \ll 1/\tau_{\phi,E} \) (equivalently \( T \gg \tau_{\phi,E} \)), \( \tau_{\phi,R} \) becomes increasingly sensitive as \( \alpha \) increases.

6.3.2 Dephasing times versus cutoff frequencies

We now examine more quantitatively the sensitivity of \( \tau_{\phi} \) to changes in both \( f_1 \) and \( f_2 \) for various values of \( \alpha \). For the Ramsey sequence with \( f_2 \to \infty \), Fig. 6.2(a) shows \( \tau_{\phi,R} \), normalized to \( \tau_{\phi,R}(f_1 = 0.1 \text{ Hz}) \), versus \( f_1 \) for \( 0.6 \leq \alpha \leq 1.2 \). We again see that the sensitivity of \( \tau_{\phi,R} \) to \( f_1 \) increases with increasing \( \alpha \). Even so, for \( \alpha = 1.2 \), \( \tau_{\phi,R} \) changes by a factor of only 4 when \( f_1 \) is varied from 0.1 to \( 10^4 \) Hz.

To explore the effect of \( f_2 \) on \( \tau_{\phi,R} \), we fix \( f_1 = 1 \) Hz and vary \( f_2 \), plotting \( \tau_{\phi,R}(f_2)/\tau_{\phi,R}(f_2 \to \infty) \) for \( 0.6 \leq \alpha \leq 1.2 \) [Fig. 6.2(b)]. We see that the sensitivity of \( \tau_{\phi,R} \) to \( f_2 \) increases for decreasing \( \alpha \). Furthermore, Fig. 6.2(b) shows that \( \tau_{\phi,R} \) is insensitive to the particular value of \( f_2 \) for \( f_2 \gg 1/\tau_{\phi,R}(f_2 \to \infty) \), simply because the Ramsey sequence is insensitive to noise for \( f \gg 1/\tau_{\phi,R} \). However, as \( f_2 \) decreases through \( 1/\tau_{\phi,R}(f_2 \to \infty) \) a non-negligible amount of noise to which the qubit is sensitive is effectively eliminated, thereby reducing the total integrated noise and increasing \( \tau_{\phi,R}(f_2) \). This effect is greater for small \( \alpha \), where \( S_\Phi \) decreases with \( f \) more slowly and contributes to dephasing out to a higher frequency.

We perform a similar analysis of the sensitivity of \( \tau_{\phi,E} \) to the value of \( f_2 \). In Fig. 6.3 we plot \( \tau_{\phi,E}(f_2)/\tau_{\phi,E}(f_2 \to \infty) \) versus \( f_2 \) for \( f_1 = 1 \) Hz. As with the Ramsey sequence, we find that \( \tau_{\phi,E} \) is insensitive to \( f_2 \) for \( f_2 \gg 1/\tau_{\phi,E}(f_2 \to \infty) \). Indeed, since \( \tau_{\phi,E} \) is dominated by noise at \( f \approx 1/\tau_{\phi,E} \), this result as we expect. Also in analogy with the Ramsey sequence, \( \tau_{\phi,E} \) is more sensitive to \( f_2 \) for small \( \alpha \). Unlike the Ramsey sequence, however, where the dephasing is sensitive to frequencies over a large bandwidth (\( f_1 \) to \( 1/\tau_{\phi,R} \)), the echo sequence is sensitive to noise only in a narrow bandwidth around \( 1/\tau_{\phi,E} \), making \( \tau_{\phi,E} \) much more sensitive to changes in \( f_2 \) for \( f_2 \approx 1/\tau_{\phi,E} \). Here, \( \tau_{\phi,E} \) increases by an order of magnitude for a two-order-of-magnitude decrease in \( f_2 \).

6.3.3 The ratio \( \tau_{\phi,E}/\tau_{\phi,R} \)

Since the value \( D_\delta \) can vary significantly between flux qubits, we consider the ratio \( \tau_{\phi,E}/\tau_{\phi,R} \), which has the advantage of being rather insensitive to the precise values of both \( A \) and \( D_\delta \). We compute these times using Eq. (4), which shows that the decay function \( g(t) \) depends only on the product \( AD_\delta \). To explore the dependence of the ratio on \( \alpha \), we compute \( \tau_{\phi,E}/\tau_{\phi,R} \) versus \( \alpha \) for \( f_2 \to \infty \) (equivalent to \( f_2 \gg 1/\tau_\phi \)) and \( f_1 = 10^{-1}, 10^1, \text{ and } 10^3 \) Hz. Furthermore, for each value of \( f_1 \) we perform the calculation for \( AD_\delta/(10^6 \text{ Hz}^{1/2}) = 0.2, 1, \text{ and } 5 \). The results are shown in Fig. 6.4.

We first examine the dependence on \( \alpha \). As \( \alpha \) increases, noise at frequencies much greater than 1 Hz falls quickly, so that \( \tau_{\phi,E} \) increases rapidly. Conversely, noise at low frequencies near 1 Hz changes little as \( \alpha \) changes. The Ramsey dephasing time is sensitive to a large noise bandwidth where a significant contribution comes from frequencies near \( f_1 \). Therefore,
Figure 6.2: Normalized Ramsey dephasing times for $0.6 \leq \alpha \leq 1.2$ in steps of 0.1. (a) $\tau_{\phi,R}(f_1)/\tau_{\phi,R}(f_1 = 0.1 \text{ Hz})$ versus $f_1$ for $f_2 \to \infty$ and (b) $\tau_{\phi,R}(f_2)/\tau_{\phi,R}(f_2 \to \infty)$ versus $f_2$. The colored dots in (b) are placed at $f_2 = 1/\tau_{\phi,R}(f_2 \to \infty)$, above which $\tau_{\phi,R}$ displays no dependence on $f_2$ (see text).
Figure 6.3: Computed values of $\tau_{\phi,E}(f_2)/\tau_{\phi,E}(f_2 \to \infty)$ vs $f_2$. Lower cutoff frequency $f_1 = 1$ Hz and $0.6 \leq \alpha \leq 1.2$ in steps of 0.1. The colored dots are placed at $f_2 = 1/\tau_{\phi,E}(f_2 \to \infty)$, above which $\tau_{\phi,E}$ displays no dependence on $f_2$ (see text).
Figure 6.4: Ratio $\tau_{\phi,E}/\tau_{\phi,R}$ vs $\alpha$. Lower cutoff frequency $f_1 = 10^{-1}$, $10^1$, and $10^3$ Hz and $f_2 \to \infty$. The thin upper, heavy middle, and thin lower lines correspond to $AD_{\phi}/(10^6 \text{ Hz}^{1/2}) = 0.2$, 1, and 5, respectively.
6.3. RESULTS

as \( \alpha \) increases we expect \( \tau_{\phi,R} \) to increase less rapidly than \( \tau_{\phi,E} \), explaining the increasing trend of \( \tau_{\phi,E}/\tau_{\phi,R} \).

For small, fixed values of \( \alpha \), changing the value of \( f_1 \) changes the ratio only slowly because both \( \tau_{\phi,E} \) and \( \tau_{\phi,R} \) are limited by noise at \( f \gg f_1 \). As the value of \( \alpha \) increases, however, an increasing contribution to dephasing in the Ramsey sequence arises from lower frequencies \( f \approx f_1 \). Therefore, for large, fixed values of \( \alpha \), increasing \( f_1 \) has the effect of removing a significant noise contribution, thereby increasing \( \tau_{\phi,R} \) and decreasing the ratio \( \tau_{\phi,E}/\tau_{\phi,R} \). We remark that since \( f_1 \) is an experimentally variable parameter, measuring \( \tau_{\phi,E}/\tau_{\phi,R} \) for several different measurement times may shed light on the value of \( \alpha \).

Finally, we see that the \( \tau_{\phi,E}/\tau_{\phi,R} \) is moderately sensitive to the product \( AD_\phi \) only for \( \alpha \gtrsim 1 \). However, additional calculations show that, for \( f_2 \lesssim 1/\tau_\phi \), the ratio becomes extremely sensitive to the particular value of \( AD_\phi \).

6.3.4 Dependence of decay function on \( \alpha \) and ultraviolet cutoff frequency

The decay function is of particular interest experimentally, since it can be measured directly. In general, the decay function of \( T_1 \)-limited processes is a simple exponential, that is \( g(t) = \exp(-t/T_1) \). However, the decay function of pure dephasing processes is more complicated and can be characterized as \( g(t) \equiv \exp(-\chi(t)) \), where \( \chi(t) \) can contain terms that are higher order in \( t \).

Here, we examine the functional dependence of \( \chi(t) \) for both pulse sequences. In each case, we find that \( \chi(t) \propto t^n \), where \( n \) can take two values \( (\gamma_1 \text{ and } \gamma_2) \) within a single sequence, separated by a characteristic time set by \( 1/f_2 \): \( \chi(t \ll 1/f_2) \propto t^{\gamma_1} \) and \( \chi(t \gg 1/f_2) \propto t^{\gamma_2} \). For the Ramsey sequence, \( \gamma_1 = 2 \) and \( \gamma_2 = 1 + \alpha \) for \( \alpha \leq 1 \) and \( \gamma_2 = 2 \) for \( \alpha > 1 \). For the echo sequence, \( \gamma_1 = 4 \) and \( \gamma_2 = 1 + \alpha \). These results reveal two experimentally relevant insights. First, for \( t \gg 1/f_2 \), \( \gamma_2 \) depends on \( \alpha \). Thus, if \( \tau_\phi \gg 1/f_2 \), a careful fit of the experimentally observed decay envelope may shed light on the value of \( \alpha \). Second, the functional form of \( g(t) \) can reveal information about \( f_2 \). For example, if one does not observe that \( \chi(t) \propto t^4 \) in an echo experiment, \( f_2 \) must be as high as \( 1/\tau_{\phi,E} \), establishing an important lower bound on the bandwidth of the flux noise.

Figure 6.5 emphasizes the above statements, showing \( g(t) \) plotted for both sequences for \( \alpha = 0.6 \) and 1.2, and for \( f_2 \) both above and below \( 1/\tau_\phi(f_2 \to \infty) \), thereby showing both \( \gamma_1 \) and \( \gamma_2 \) dependence. In Figs. 6.5(a) and 6.5(b) we plot \( g(t) \) for the Ramsey sequence with \( \alpha = 0.6 \) and 1.2. We note that difference between the functional dependencies of the two traces in Fig. 6.5(a) is slight, and would be nearly impossible to measure experimentally. In Fig. 6.5(b) there is no functional difference since \( \gamma_1 = \gamma_2 \). Figures 6.5(c) and 6.5(d) show \( g(t) \) for the echo sequence for \( \alpha = 0.6 \) and 1.2. The difference is more dramatic since \( \gamma_1 = 4 \) is so large. In this case, such a difference might be experimentally observable.
Figure 6.5: Computed decay function $g(t)$ versus $t/\tau_\phi$ for (a) and (b) Ramsey sequences and (c) and (d) echo sequences with $\alpha = 0.6$ and 1.2. In the red trace, $f_2 \gg 1/\tau_\phi (f_2 \to \infty)$; in the blue trace $f_2 \ll 1/\tau_\phi (f_2 \to \infty)$. 

\[
\chi(t) \propto f_2 \ll 1/\tau_\phi (f_2 \to \infty) \quad \text{and} \quad \chi(t) \propto f_2 \gg 1/\tau_\phi (f_2 \to \infty)
\]
6.3. RESULTS

Figure 6.6: Computed dephasing times $\tau_\phi$ vs $\alpha$ for $f_1 = 1$ Hz and $f_2 \to \infty$. (a) Ramsey and (b) echo pulse sequences for fixed $S_\Phi(f_0)$, where the pivot frequency $f_0 = 10^{-2}, 1, 10^2,$ and $10^4$ Hz.

6.3.5 $S_\Phi(f)$ pivoting about $f_0 \neq 1$ Hz as $\alpha$ is varied

As mentioned previously, there is no a priori reason to hold $S_\Phi(1$ Hz) fixed as $\alpha$ is varied; the choice is based on the empirical observation that values of $S_\Phi(1$ Hz) are relatively uniform across a wide variety of devices and measured $\alpha$. In fact, as we shall see in Chapter 3, there is experimental evidence that pivoting in SQUIDs occurs at frequencies above 1 Hz, on the order of 10 to 100 Hz, an empirical observation that has been corroborated by very recent theory [104].

To explore the sensitivity of our calculations to this assumption, we calculated the dephasing times for both sequences versus $\alpha$ for fixed $S_\Phi(f_0)$, where $f_0 = 10^{-2}, 1, 10^2,$ and $10^4$ Hz. Conceptually, the spectra can be imagined as pivoting as $\alpha$ changes about a fixed spectral density $S_\Phi(f_0)$ at frequency $f_0$. In order to normalize the magnitude of each set of curves corresponding to a particular $f_0$, we choose as a convention that $S_\Phi(1$ Hz) = $A^2$ when $\alpha = 1$, regardless of the value of $f_0$, that is $S_\Phi(f) = (A^2/f_0)(f/f_0)^{-\alpha}$. This convention is based loosely on empirical observation; it does not significantly change the dependence of $\tau_\phi$ on $\alpha$, but merely sets the absolute scale.
The results of these calculations, plotted in Fig. 6.6, show a dramatic effect, both qualitatively and quantitatively, on the dependence of $\tau_{\phi,R}$ and $\tau_{\phi,E}$ on $\alpha$. For the Ramsey sequence [Fig. 6.6(a)], the general trend of increasing $\tau_{\phi,R}$ is significantly altered as $f_0$ increases and even becomes nonmonotonic for $f_0 = 10^4$ Hz. In addition, for small values of $\alpha$, $\tau_{\phi,R}$ increases dramatically as $f_0$ increases. We note that, because of our normalization condition, the curves intersect at $\alpha = 1$. Calculations for the echo sequence are shown in Fig. 6.6(b), which shows a similar dependence of $\tau_{\phi,E}$ on $f_0$. For both sequences, the dependence of $\tau_{\phi}$ on $\alpha$ is minimal for $f_0 = 10^4$ Hz, the highest computed $f_0$. This dependence is easily understood for the echo sequence, which is sensitive only to noise at $f \approx 1/\tau_{\phi,E}$. As $f_0$ approaches $1/\tau_{\phi,E}$, the effect of $\alpha$ eventually becomes negligible. In fact, if $f_0$ were to exceed $1/\tau_{\phi,E}$, the trend in $\alpha$ would actually reverse. The Ramsey sequence, however, is sensitive to a larger noise bandwidth and has a correspondingly more complicated dependence, exhibited by its nonmonotonic behavior for large values of $f_0$.

6.4 Concluding remarks

In conclusion, we have presented calculations showing that the predicted dephasing times $\tau_{\phi,R}$ and $\tau_{\phi,E}$ of a flux qubit are very sensitive to the value of $\alpha$. As the value of $\alpha$ increases, both $\tau_{\phi,R}$ and $\tau_{\phi,E}$ increase dramatically—by an order of magnitude in some cases. Since experimentally inferred values of $S_\Phi(1\text{ Hz})$ from qubit measurements have generally assumed that $\alpha = 1$, a nonunity value of $\alpha$ can introduce a significant error into the inferred value of $A$. Furthermore, we have shown that while the lower cutoff frequency $f_1$ (set by the total measurement time) does not significantly affect $\tau_{\phi}$, the upper frequency cutoff $f_2$ can significantly change $\tau_{\phi}$ in a manner dependent on the value of $\alpha$, particularly for the echo sequence. Moreover, we have shown that by examining the directly measurable ratio $\tau_{\phi,E}/\tau_{\phi,R}$ and the dephasing function $g(t)$, experimentalists may have a probe into the values of $\alpha$ and $f_2$. Finally, the frequency at which the flux noise spectra pivot can dramatically affect the sensitivity of $\tau_{\phi}$ to $\alpha$.

Most importantly, these results demonstrate that lowering the flux noise amplitude is not the only method of increasing qubit dephasing times. With a more detailed understanding of what sets $\alpha$ experimentally—for, example, the geometry of the qubit washer—it may be possible to increase dephasing times substantially by raising the value of $\alpha$. Finally, we note that with straightforward modification our formalism could be used to calculate dephasing times from critical current noise and charge noise for the case $\alpha \neq 1$.

Acknowledgments

Because the content of this Chapter was entirely theoretical in nature, we graciously acknowledge the essential support of our theoretician collaborators Clemens Müller, Gerd Schöen, and Alexander Shnirman. In addition to introducing us to the appropriate approach to modeling dephasing in qubits, they provided detailed guidance and revisions of our work.
Chapter 7

Concluding remarks

Prior to the work and developments presented in this thesis, we can summarize the understanding of magnetic flux noise as follows. SQUIDs and superconducting qubits exhibit an effective flux noise with power spectral density $S_{\Phi}(f)$ that scales with frequency $f$ approximately as $S_{\Phi}(f) \propto 1/f^\alpha$, where $0.4 \lesssim \alpha \lesssim 1.1$. The noise is generated locally by the superconducting loop and is not due to external fields or parameter variations. In SQUIDs, flux noise is the dominant noise source at low frequencies. In flux-sensitive qubits, flux noise is a dominant source of dephasing and inhibits the development of quantum computing. Remarkably and inexplicably, the measured magnitude of flux noise in modern devices is nearly always $S_{\Phi}(1\text{ Hz}) \approx 1 (\mu\Phi_0)^2$/Hz, quite independent of temperature, loop geometry, fabrication materials and processes, and the particular value of $\alpha$. Numerical and analytic models assuming independent surface spins predict a particular scaling of the noise magnitude with geometry, but this scaling had not been systematically and experimentally studied.

7.1 Summary of results

From our extensive and systematic measurements, we now have a much more detailed empirical understanding of flux noise. Through our subsequent analysis, we showed that these empirical observations do not fit existing microscopic models and therefore support a more complex picture, thereby motivating further theoretical study. We will now review some of our principal conclusions and suggest some promising avenues for future research.

Perhaps the most astonishing trend we observed in our measurements was the spectral pivoting of $S_{\Phi}(f)$. In this phenomenon, the slope of the noise spectrum $\alpha$ dramatically increased—from about 0.4 to 0.8—as the temperature was lowered. The noise magnitude $S_{\Phi}(1\text{ Hz})$ varied in a particular way so as to keep $S_{\Phi}(f_0)$ fixed. Here, the pivot frequency $f_0$ was constant for each device and was typically between 1 and 100 Hz. The ramifications of this effect are numerous. First, it suggests that the historical choice of defining the noise magnitude as $S_{\Phi}(1\text{ Hz})$ was fortuitous; if, say, $S_{\Phi}(1\text{ MHz})$ were reported instead of $S_{\Phi}(1\text{ Hz})$, then the “noise magnitude” would be seen to decrease considerably with decreasing temperature. This observation implies that using $S_{\Phi}(1\text{ Hz})$ as the sole parameter for quantifying
the noise magnitude is an insufficient and potentially misleading metric. Perhaps a better convention is to report the spectral density at the pivot frequency, $S_\Phi(f_0)$, along with the corresponding value of $\alpha$.

Second, we can infer from spectral pivoting that the mean square flux noise $\langle \Phi^2 \rangle$ cannot be temperature-independent as predicted by existing models of independent surface spins. This discrepancy has altered our understanding of possible correlations between spins. We attempted to reconcile the discrepancy between measurements and predictions, which sometimes exceeded five orders of magnitude, by hypothesizing that spins lock together to form noninteracting clusters that can vary in size with temperature. This extension to the surface spin model leads to further inconsistencies. Therefore, we conclude the presence of significant and complex correlations between spins. The mechanism governing these interactions, however, is not yet understood.

We also measured $S_\Phi(f)$ and inferred $\langle \Phi^2 \rangle$ as a function of SQUID loop geometry. Again, the scaling was in disagreement with the predictions of models of noninteracting surface spins. This observation can be taken as further evidence for the existence of complex interactions between spins. Furthermore, we observed that $\alpha$ scaled with loop geometry, an effect for which we have no model. It is conceivable that long-range interactions or structures can form over length scales on the order of $\mu$m.

Two existing methods assuming noninteracting surface spins calculate $\langle \Phi^2 \rangle$ numerically and analytically. While the development of each represented a milestone in our understanding of flux noise, both methods suffer from significant drawbacks. In particular, the numerical method suffers from poor accuracy and significant computational intensity. The analytic method is valid only for a restricted set of geometries and cannot calculate the contribution of edge spins and certain types of surface spins. Our numerical method, however, is capable of computing $\langle \Phi^2 \rangle$ from all spins for arbitrary loop geometries, is very accurate, and is computationally efficient. Using this method, we performed a detailed comparison of our results to the analytic formula in a number of scenarios. We demonstrated significant discrepancies even in the limit where the two methods should agree, and traced this discrepancy to several approximations used in the derivation of the analytic result. Importantly, we computed the effect of edge spins and found their contribution to be significant in narrow-linewidth devices, which is a regime particularly relevant for qubits.

The notion that $\alpha$ in a single device can vary significantly with temperature raises the question of how dephasing in qubits is affected by its particular value. Small values of $\alpha$ mean that the noise spectral density decreases relatively slowly with increasing frequency. Because qubits are sensitive to frequencies up to the GHz range, one naturally concludes that low values of $\alpha$ lead to shorter dephasing times $\tau_\phi$. We extended existing models, which assumed a unity value of $\alpha$, to quantify the relationship between $\alpha$ and $\tau_\phi$. Indeed, we find that dephasing times in both echo and Ramsey pulse sequences increase by an order of magnitude as $\alpha$ increases from 0.6 to 0.9, holding $S_\Phi(1 \text{ Hz})$ fixed. This trend is significant and implies an alternate route to increasing dephasing times by raising the value of $\alpha$. To do this, however, requires a more detailed understanding of what sets $\alpha$ experimentally.
7.2 Future directions

There are three general directions that we consider promising avenues of future research. In fact, we feel strongly that all three should be pursued simultaneously, as results from one direction are likely to facilitate progress in the other two.

First, we feel that developing a comprehensive theory that captures the microscopic dynamics of individual spins is critical to the ultimate reduction of the deleterious effects of flux noise. For instance, it is still not known whether spins are spatially confined and randomly reverse direction or if they move randomly across the surface—thereby changing their flux coupling to the SQUID loop—as in the spin-diffusion model. The activation of spins, be it thermal or quantum, is not known. As previously mentioned, we have strong evidence that the spins are correlated in a complex way, but we do not yet have a detailed model for this interaction. It could, perhaps, be RKKY\cite{105,107} or dipole-dipole in nature or some combination of the two. To guide theory is a growing body of experimental measurements. A successful theory will have to explain a 13-decade noise bandwidth, a mechanism by which $\alpha$ can double with decreasing the temperature, and how the spectra can pivot about a fixed frequency. The goal is lofty, but the results of several recent theories are extremely promising.

Second, it is becoming increasingly important to measure flux noise at frequencies between 1 kHz and 1 GHz. SQUID measurements are generally limited to an upper frequency of 10 Hz to 1 kHz due to the onset of white noise from the shunt resistor. Because flux noise is inherently a wide-bandwidth problem, measurements on SQUIDs shed light on only a fraction of the problem. From a combination of techniques, measurements in qubits reveal the flux noise spectrum over a much wider bandwidth and to much higher frequencies. Performing these wide-bandwidth measurements as, say, the geometry of the qubit loop is varied would provide invaluable data and help to guide the theory. It should be be noted, however, that qubit measurements are limited to low temperatures, generally below 0.1 K.

Third, an extremely interesting avenue of research has spawned from unpublished results from Robert McDermott’s group, who showed that the addition of SiNx capping layers to an Al SQUID washer—both above and below—resulted in a significant decrease in $S_\Phi(1 \text{ Hz})$. If one accepts the picture that unpaired spins exist at the surfaces of the superconductor, then it is not surprising that the addition of capping layers could affect the noise properties. Nevertheless, it is surprising that the deposition of an amorphous insulator should lower the noise. In our own lab, we have repeated this experiment with similar results. We have also measured SQUIDs capped with a number of other materials—NbNx, Al$_2$O$_3$, Au—with varying results. At this point it is clear that capping layers can affect both the noise magnitude and the slope. However, the picture for how this occurs is still unclear. Systematic flux noise measurements of SQUIDs fabricated with tightly controlled parameters are likely to be essential to solving this puzzle.

With the growing importance and urgency of understanding and reducing flux noise, a correspondingly growing number of resources and talented people are being devoted to the problem. We are excited for the field and are hopeful that significant advances will reward dedicated efforts.
Bibliography


[98] Y. Nakamura, C. D. Chen, and J. S. Tsai. “Spectroscopy of Energy-Level Splitting be-


Appendix A

Calculating a stitched spectrum from time series

Many existing programs and scripts exist that compute the fast Fourier transform (FFT) and corresponding spectral density of an input time series. However, for a number of reasons we have found it advantageous to compute several FFT’s from the time capture and to “stitch” them together to represent the spectrum. In this Appendix, we review our algorithm to compute the FFT’s and to stitch them into a single spectrum.

Given an input time series, sampled at a frequency $1/\Delta t$ and consisting of $N$ points of amplitude $y_j$ ($j \in \{1,2,3,\ldots,N\}$), the FFT is defined as

$$Y_k = \sum_{j=1}^{N} y_j \omega_N^{(j-1)(k-1)},$$  \hspace{1cm} \text{(A.1)}$$

where $\omega_N = \exp[-2\pi i/N]$ is an $N$th root of unity. In general, the FFT algorithm works best when $N$ is a power of 2; if it is not, the time series is usually padded with zeros until its length is a power of 2. Because $y_j \in \mathbb{R}$, $Y_k$ is symmetric ($Y_k = Y_N^{\ast} = Y_{N-k+2}^{\ast}$ for $k \geq 2$) and the power spectral density $S_k$ can be written as

$$S_k = \begin{cases} 
\frac{\Delta t}{N}|Y_1|^2 & \text{if } k = 1 \\
2 \frac{\Delta t}{N}|Y_k|^2 & \text{if } k \in \{2,3,4,\ldots,N/2+1\},
\end{cases}$$  \hspace{1cm} \text{(A.2)}$$

which is defined at frequencies

$$f_k = (k-1)/N\Delta t,$$  \hspace{1cm} \text{(A.3)}$$

where $k \in \{1,2,3,\ldots,N/2+1\}$. The factor of 2 in Eq. (A.2) comes from summing the symmetric components of $Y_k$.

While it is possible to compute a single FFT of a time capture with $N_{tc}$ points, this technique generally leads to a spectrum that appears extremely noisy. That is, the difference in spectral densities at adjacent frequencies is of the order of the spectral density itself,
regardless of the length of the time capture. This property can be understood as follows. At any particular frequency, $S_k$ is the sum of the squares of the real and an imaginary components of $Y_k$, which are independent for random noise. Therefore, $S_k$ follows a Chi-squared distribution with two degrees of freedom ($S_k \sim \chi^2_2$). A property of the $\chi^2_2$ distribution is that the variance $\sigma^2$ is equal to its mean $\mu$ squared. Therefore, $\sigma^2(S_k) = \mu^2(S_k)$ so that the spread of spectral densities is relatively large and the spectrum appears noisy. A longer time capture increases $N_{tc}$, which in turn increases the frequency resolution—the width of the frequency bins $df = 1/N_{tc}\Delta t$—but does not decrease $\sigma^2(S_k)$.

To reduce the apparent noise in the spectrum, it is typical to introduce some form of averaging. A popular method is to split the time capture into $N_{avg}$ equivalent-length segments, computing the FFT for each segment and averaging the spectral densities to form a single averaged spectrum. In other words,

$$S_k = \frac{1}{N_{avg}} \sum_{n=1}^{N_{avg}} S_{k,n}, \quad (A.4)$$

where $S_k$ is now an averaged spectral density and $S_{k,n}$ is the spectral density at $f_k$ of the $n^{th}$ time segment. In this case, $S_k$ obeys the central limit theorem and, for sufficiently large $N_{avg}$, is normally distributed with a variance that scales as $1/N_{avg}$. Since

$$df = \frac{N_{avg}}{N_{tc}} \frac{1}{\Delta t}, \quad (A.5)$$

for fixed $N_{tc}$ we see that there is a tradeoff between fine frequency resolution and low variance of the spectral density. Because the lowest frequency of the spectrum (zero frequency excluded) is equal to $df$, the tradeoff is equivalently between a spectrum that extends low in frequency and one that is highly averaged.

To circumvent this tradeoff, we observe that, for our application, a fine frequency resolution is not necessary at high frequencies. Similarly, at low frequencies we can tolerate relatively few averages, which in turn gives us fine frequency resolution. This situation suggests a time capture split into a relatively small number $N_{avg,min}$ of long time series is ideal for computing the spectrum at low frequencies, and a time capture split into a relatively large number $N_{avg,max}$ of short time series is ideal for computing the spectrum at high frequencies. The two spectra corresponding to the different values of $N_{avg}$ can be “stitched” together at some intermediate frequency $f_{cut}$. That is, we define a new spectrum where the spectral densities below $f_{cut}$ are taken from the spectrum computed with $N_{avg,min}$ and the spectral densities above $f_{cut}$ are taken from the spectrum computed with $N_{avg,max}$. If $N_{avg,min}$ and $N_{avg,max}$ are significantly different, intermediate values of $N_{avg}$ can be used as well. The standard 1-hr time capture we acquire during noise measurements consists of $3200 \times 1024 = 3,276,800$ points, sampled at 1024 Hz. To compute the spectral density, we use $N_{avg} = [3, 12, 50, 200, 800, 3200]$ to generate six averaged spectra, which we stitch together at the frequencies $f_{cut} = [1/2^{10}, 1/2^8, 1/2^6, 1/2^4, 1/2^2] \times (512 \text{ Hz}/50) = [0.01, 0.04, 0.16, 0.64, 2.56] \text{ Hz}$. 


Figure A.1: Spectral densities of (a) white and (b) 1/f noise vs frequency, normalized to 1/dt. Spectra are computed from randomly generated time series with $N_{tc} = 2^{18}$ for varying $N_{avg}$. Colored dots are plotted to highlight the lowest frequency of each spectrum. In addition, a stitched spectrum with $N_{avg} = [4, 16, 64, 256]$ and $f_{cut} = (0.5/50) \times [1/2^6, 1/2^4, 1/2^2]$ is shown.

Figure A.1 shows the spectral density of white and 1/f noise computed in a variety of ways from two separate randomly generated time series. From the unstitched spectra we see that the variance in spectral density points decreases as $N_{avg}$ increases. Correspondingly, the lowest frequency to which the spectral density extends, represented by the colored dots, also increases with $N_{avg}$. However, the stitched spectrum, created by stitching five averaged spectra, extends to a low frequency while maintaining a low variance at high frequency points.

In Fig. A.2 we histogram the spectral densities of the white noise shown in Fig. A.1(a) to illustrate the effect that $N_{avg}$ has on the variance. As discussed above, the spectral densities generated from a single FFT are distributed following a $\chi^2_2$ distribution, as shown in Fig. A.2(a). As $N_{avg}$ increases, the distribution of spectral densities, which are now averaged from several FFTs, approaches a normal distribution with a standard deviation that scales as $N_{avg}^{-1/2}$ according to the central limit theorem. For $N_{avg} = 64$, the agreement with the normal distribution is good [Fig. A.2(d)].

1Technically, the correct statement is that the collection of values of spectral density $S_k$ at a fixed frequency $f_k$, taken from the FFTs of each of a large number of independent time series, follows a $\chi^2_2$ distribution. Instead, Fig. A.2(a) is a histogram of $S_k$ for all $k$ for a single time series. However, because the noise is white, the spectral densities of different frequencies are independent and each follow the identical distribution. Therefore, we histogram the spectral densities at all frequencies from a single time series rather than generate $2^{18}$ separate time series and FFTs to histogram the spectral densities of a single frequency!
Figure A.2: Histograms of white noise from Fig. A.1(a) for (a) $N_{\text{avg}} = 1$, (b) $N_{\text{avg}} = 4$, (c) $N_{\text{avg}} = 16$, and (d) $N_{\text{avg}} = 64$. The probability distributions fit (a) a $\chi^2$ distribution (solid red line) and (d) a normal distribution with standard deviation that scales as $N_{\text{avg}}^{-1/2}$ (solid red line).
Appendix B

Fitting spectra

In the majority of measurements of flux noise in SQUIDs, the flux noise $S_{\Phi}(f)$ is observed to scale with frequency as $1/f^\alpha$ with little or no deviation. As such, $S_{\Phi}(f)$ is well characterized by just two parameters: a noise magnitude and slope. For historical reasons, the spectral density at 1 Hz has become synonymous with the noise magnitude [37]. To determine the values of $\alpha$ and $S_{\Phi}(1 \text{ Hz})$, which we denote as $A^2$, we fit a measured spectrum $S_{\Phi,\text{meas}}$ to an assumed functional dependence. Here, $S_{\Phi,\text{meas}}$ is expressed in terms of equivalent flux in the measured SQUID. Because typical measurements of flux noise in SQUIDs include a non-negligible white noise contribution from the resistive shunts across the junctions, a third fitting term $C^2$, the magnitude of the white noise, is included in the functional dependence:

$$S_{\Phi,\text{meas}}(f) = S_{\Phi}(f) + C^2 = \frac{A^2}{(f/1 \text{ Hz})^\alpha} + C^2.$$  \hspace{1cm} (B.1)

Conceptually, the fitting process is quite straightforward. Nonetheless, there are a number of issues one must consider in order to implement a routine that programatically generates high quality fits, which we now review.

In general, a nonlinear curve fit of a function $g(x, \beta)$ to $N$ pairs of independent and dependent data points $(x_i, y_i)$ is accomplished via a nonlinear regression routine, which seeks to optimize the fit coefficients $\beta$ such that the sum of the squared residuals

$$\Lambda(\beta) = \sum_{i=1}^{N} [y_i - g(x_i, \beta)]^2$$ \hspace{1cm} (B.2)

is minimized. In our case, $g(x, \beta)$ is defined by Eq. (B.1), where the three components of $\beta$ are $A^2$, $\alpha$, and $C^2$, and $y$ and $x$ are defined by the spectral densities and frequencies, respectively, of the spectrum we are fitting. Unlike linear regression routines, nonlinear regression routines are iterative and can only guarantee that solutions are local, not global, minima. This fact implies that an initial guess of the fit coefficients must be provided to the routine. Furthermore, if the guess is poor the routine may converge to local minimum that may represent a poor fit.
To supply the routine with an accurate initial guess, we have developed a robust routine that works well with our flux noise data. We rely on the empirical observation that it is often easier to initially guess a single parameter at a time, holding the others fixed. We first guess $C^2$ to equal the spectral density at the highest frequency of the measured spectrum. Then, holding $C^2$ fixed, we perform a nonlinear regression for the values of $A^2$ and $\alpha$. Using these values for $C^2$, $A^2$, and $\alpha$ as the initial guesses, we perform a nonlinear regression for all three simultaneously, the solution of which we keep as our final fit coefficients.

The concept of minimizing $\Lambda(\beta)$ raises two more issues. First, because $S_\Phi(f) \propto 1/f^\alpha$, $y_i$ and $g(x_i, \beta)$ can vary several orders of magnitude. Therefore, the residuals from data at lower frequencies, which tend to be the largest, will provide the dominant contribution to $\Lambda(\beta)$. This, in turn, will cause the routine to yield a fit that is highly accurate at low frequencies at the cost of “accuracy” at high frequencies. In this situation, we refer to an accurate fit in the intuitive sense of a curve that visually approximates the measured spectrum when plotted on a log-log scale.

Second, because the frequency spacing $df$ is constant in an unstitched spectrum, the density of points increases linearly with frequency when plotted on a log scale; that is, the number of points per octave is proportional to its center frequency. This effect increases the influence of higher frequency data and overly favors an accurate fit at high frequencies.

Unfortunately, these two issues, which produce opposite effects, cancel each other only for spectra that vary as $1/f$. In this situation the spectral density and, correspondingly, the variance scale as $1/f$, while the density of the log frequency points $[\log(f_{i+1}) - \log(f_i)]^{-1}$ increases as $f$. Since our measured spectra rarely vary as $1/f$, we handle the previous two issues in the following manner. To ensure that the spacing between log frequency points is roughly constant $[\log(f_{i+1}) - \log(f_i) \approx \text{constant}]$, we use the stitched spectrum explained in Appendix A. Here, the number of averaged FFTs used to compute a particular frequency range of the stitched spectrum is roughly proportional to center frequency of the range. By Eq. (A.5), $df \propto N_{\text{avg}} \propto f$ so that $[\log(f_{i+1}) - \log(f_i) = \log(1 + df/f_i) \approx \text{constant}]$, which addresses the overrepresentation of high frequencies in an unstitched spectrum. Next, by fitting the logarithm of the spectra, we ensure that low frequencies are not overrepresented. That is, we perform a nonlinear regression such that

$$\Lambda'(\beta) = \sum_{i=1}^{N} [\log(y_i) - \log(g(x_i, \beta))]^2$$

is minimized. We have found that this method of fitting the logarithm of stitched spectra to be highly reliable and have used it for all fits presented in this thesis.