Abstract

Single-Photon Detection, Kinetic Inductance, and Non-Equilibrium Dynamics in Niobium and Niobium Nitride Superconducting Nanowires

Anthony Joseph Annunziata

2010

This thesis is a study of superconducting niobium and niobium nitride nanowires used as single optical and near-infrared photon detectors. The nanowires are biased in the zero-voltage state with a current just below their critical current and at a temperature well below their critical temperature. In this state, an absorbed photon induces localized heating at the point of absorption. This suppresses the critical current in that location, creating a resistive region in the nanowire. The resistive region can grow under Joule heating and can self-reset to the zero-voltage state without the dc bias current being reduced.

This study is twofold. First, niobium is investigated as an alternate detector material to niobium nitride. This study compares the performance niobium nanowire detectors of several geometries and fabricated in two different ways to the performance of niobium nitride nanowire detectors. Niobium detectors are found to have longer reset times and are more difficult to bias in a regime where they self-reset to the zero voltage

state after detecting a photon. This makes niobium a less suitable material than niobium nitride for these detectors.

In the second part of this study, the reset dynamics of these detectors are studied. Thermal relaxation is studied using a combination of experiments and numerical simulations. It is found that the thermal relaxation time for a niobium nanowire depends significantly on the amount of energy dissipated into the hotspot during the detection event. This energy depends on the bias current and on the kinetic inductance of the nanowire. The kinetic inductance is proportional to the length; thus a shorter nanowire will have a shorter thermal relaxation time, and a shorter reset time. Using this theoretical framework, the difference in reset time between niobium and niobium nitride nanowire detectors is explained. The temperature and current dependence of the kinetic inductance of niobium and niobium nitride nanowires is also investigated.

Single-Photon Detection, Kinetic Inductance, and Non-Equilibrium Dynamics in Niobium and Niobium Nitride Superconducting Nanowires

A Dissertation

Presented to the Faculty of the Graduate School

of

Yale University

in Candidacy for the Degree of

Doctor of Philosophy

by

Anthony Joseph Annunziata

Dissertation Director: Professor Daniel E. Prober

May 2010

© 2010 by Anthony Joseph Annunziata.

All rights reserved.

Preface

I enjoy discussing science with nonscientists. I especially like explaining what I have been working on for the past 5 years. Depending on the background of the audience, I sometimes begin by explaining why I think physics is a simple subject, despite the popular perception. So I will begin here in a similar fashion, which will hopefully lend some insight into my experience researching and writing this dissertation.

Physics is the quantitative study of the laws that govern the physical world. In applied physics, we try to use these laws to make useful things. Physics is an exact and mathematically rigorous science. Despite its broad scope and sometimes complicated mathematical nature, however, the basic goal of physics is to explain the most things with the fewest ideas. In this way, physics is a very simple way of looking at the world, much simpler in fact than most other things. Take rocket science as an example. The motion of rockets is described by the same equation as describes a block sliding down a plane, which everyone learns about in high school. It is just more difficult to use the equation to predict how to build and fly a rocket than it is to use the equation to predict how the block will slide down a plane.

Complexity is a good thing in many cases. It makes life interesting and exciting, like in art, or music, or cooking. In physics (and all of science) a complex explanation of some phenomenon is sometimes necessary, but it is not desirable. Simplicity is desired because it enables a wide variety of phenomena to be explained using only a few concepts. The simplicity of physics makes theories universal, meaning they can be applied to situations we haven't yet studied. In applied physics, we use this universality to make new devices and materials that are based on the same fundamental concepts that underpin existing devices and materials. Thus, the simplicity of physics enables applied physicists to create a wide array of useful things, such as a single-photon detector made from a superconducting nanowire that can reset very quickly. These detectors are the subject of this dissertation, but they are based on many of the same concepts as other superconducting detectors and their operation is explained by the same basic theories as are used to explain a large variety of other phenomena.

–Anthony J. Annunziata,

May 2010, New Haven, CT

Acknowledgements

I must begin by thanking God that this thesis is complete. A little grace from above is instrumental in completing a PhD, and I have certainly received my share of it over the last 5 years. Amen.

I thank my parents for their love, guidance and support. I dedicate this thesis to them. From helping me move into my first apartment to giving me advice on my post-PhD career plans, they have been essential in enabling this thesis.

I thank my country and its citizens for supporting my work through research grants and through a personal fellowship from the National Science Foundation. The United States of America is the greatest place in the world to do science. May she always remain that way and may my small contribution help strengthen and defend her.

I thank my advisor, Professor Daniel Prober. Dan once remarked that he was less a rocket scientist than a professor of “fussy details”. In five years I have come to realize the truth of this. In the course of my training, Dan and I have had many spirited discussions. He has challenged me as well as supported me. For all of this and more, I owe Dan a great debt of gratitude.

I thank Professor Aviad Frydman. With his help I began this project, and through his many visits from Israel he helped me push it all through to completion. Along the way, he taught me many important life lessons, including the importance of being careful when necessary but not necessarily to always be careful, as well as the importance of keeping good notes in one’s lab book.

I thank Luigi Frunzio, with whom I share a Neapolitan heritage and an interest in nanofabrication. His tireless and spirited guidance has been a tremendous help and an inspiration over the years.

I thank Mike Rooks, with whom I worked at IBM and later here at Yale. As I write this, I wonder if I might still be working on device fabrication if Mike hadn't offered me his 20+ years of expertise in electron beam nanolithography. He makes it look easy, but it is not.

I thank my thesis committee. Professor Robert Schoelkopf and Professor Charles Ahn have provided very helpful and inspiring guidance over the years. I also thank Professor Robert Grober who graciously stepped in for Rob at my thesis defense.

I thank my colleagues. In the course of this work, I received the help, advice, and support of many. To me, Daniel Santavicca is a fixture of the Prober lab. Before me he came, and after me he stands (now as a post-doc). He has been a tremendous help to me over the years, a sounding board for new and sometimes odd ideas, a voice of reason, and a friend. Joel Chudow has been a loyal and extremely helpful colleague and friend with many useful insights on physics and life. Matt Reese was a tireless and pedagogical mentor my first year, and Veronica Savu made the office bloom. Andrew Mack was a dutiful assistant to me during my third year – he began a numerical modeling project that I later worked on with Orlando Quranta. Orlando worked with me during part of my fourth year. The model we developed then led to a significant part of this thesis. I also thank everyone else in the Prober lab, on the fourth floor, and beyond who has helped me over the years.

I thank Lauren Priolo, my brothers David and Nicholas Annunziata, my grandmother Rita Annunziata, who is the other scientist in the family, and Robert Barbieri, Elise Panza, and Giovanni Zinn for advising, supporting, criticizing, teaching, warning, inspiring, listening, and reminding me of what is important in life.

Table of Contents

List of Figures

List of Tables

List of Symbols and Abbreviations

All units are in SI unless specifically stated otherwise. The approximate values quoted here were used in all calculations unless otherwise stated.

SNSPD	Superconducting nanowire single-photon detector
SEM	Scanning electron microscope
FET	Field effect transistor
Nb	Niobium
NbN	Niobium nitride
NbTiN	Niobium titanium nitride
InGaAs	Indium Gallium Arsenide
PICA	Picosecond imaging circuit analysis
PPM	Pulse position modulation
QKD	Quantum key distribution
VAP	Vortex-antivortex pair
IPA	Isopropyl alcohol
PMMA	Polymethymethacrylate (electron resist)
RF	Radio frequency
PCB	Printed circuit board
CB-CPW	Conductor-backed coplanar waveguide
min	Minimum value of a quantity
max	maximum value of a quantity
h	Planck's constant: 6.626×10^{-34} J·s
$\hbar$	Planck's constant: $h/2\pi$
c	Speed of light: 2.998×10^{8} m/s
k_B	Boltzmann's constant: 1.38×10^{-23} J/K

f	Frequency
ω	Radial frequency
λ	Wavelength
n	An integer number
m_e	Mass of an electron: 9.1×10^{-31} kg
e	Charge of an electron: 1.6×10^{-19} C
ε_o	Permittivity of free space: 8.85×10^{-12} $m^{-3}kg^{-1}$ s^2 C^2
ε	Relative dielectric constant: material dependent
μ_o	Permeability of free space: 1.26×10^{-6} $m\cdot kg\cdot C^{-2}$
τ_t	Bias current transfer time: $\tau_t = L_K/(R_L+R_d)$
τ_r	Bias current return time: $\tau_r = L_K/R_L$
τ_e	Electron-electron elastic scattering time (Drude model)
$\tau_{e\text{-}e}$	Electron-electron inelastic scattering time
$\tau_{e\text{-}ph}$	Electron-phonon inelastic scattering time
$\tau_{ph\text{-}e}$	Phonon-electron inelastic scattering time: $\tau_{ph\text{-}e}(T) = \tau_{e\text{-}ph}(T)$
$\tau_{ph\text{-}e}$	Phonon-electron inelastic scattering time
τ_{esc}	Phonon escape time (from metal into substrate)
τ_c	Instantaneous cooling time of a hotspot in an SNSPD: $\tau_c = \tau_c(T_e)$
$\langle\tau_c\rangle$	Average cooling time of a hotspot in an SNSPD
τ_D	Diffusion time
$\Delta\tau$	SNSPD timing jitter for photon detection
E	Energy
E_{dis}	Energy dissipated
T_c	Superconducting critical temperature: midpoint of transition
H_{c1}	Lower critical field of a type-II superconductor
T_o	Operating temperature

T_e	Electron temperature
T_{ph}	Phonon temperature
Δ	Superconducting energy gap
ξ	Superconducting coherence length
ξ_o	Pippard coherence length
λ_M	Magnetic penetration depth
λ_o	Magnetic penetration depth at zero temperature
I_c	Superconducting critical current
I_{co}	Critical current at the operating temperature T_o
I_{latch}	Latching current
I_r	Return current; transition from normal to superconducting state at $T < T_c$
I_d	Device (nanowire) current
I_b	Bias current
I_s	A supercurrent
J_s	A supercurrent density
j_c	Superconducting critical current density
j_b	Bias current density (uniform): $j_b = I_b/A_{cs}$
j_d	Device current density
ℓ_e	Electron mean free path in a normal metal
R_S	Shunt resistor; dc-coupled to device
R_A	Amplifier input resistance; ac-coupled via C_{BT}
R_L	Load resistance of the readout circuit: $R_L = R_S \| R_A$
R_d	Device (nanowire) resistance: $R_d = R_d(t)$
ρ_d	Device (nanowire) resistivity: $\rho_d = \rho_d(x,y,t)$
R_n	Device (nanowire) full resistance in normal state
ρ_n	Device (nanowire) full resistivity in normal state
L_K	Kinetic inductance of device

L_M	Magnetic inductance of device
L_{BT}	Bias tee dc-coupling inductance
C_{BT}	Bias tee ac-coupling capacitance
C_s	Shunt capacitance used in measuring kinetic inductance
l_d	Device (nanowire) length
l_{Th}	Thouless length
l_2	Length scale involved in electron-electron scattering
w_d	Device (nanowire) width
d_d	Device (nanowire) thickness
A_{cs}	Device (nanowire) cross-sectional area: $A_{cs} = w_d \cdot d_d$
A_d	Device (nanowire meander) detection area
σ	Device (nanowire meander) fill factor; active area = σA_d
n_{Nb}	Index of refraction of niobium: $\approx 2.71 - 2.95i$
n_{sapp}	Index of refraction of sapphire: 1.78
n_{air}	Index of refraction of air: 1.00
a	Thin film photon absorption probability
σ_e	Conductivity of normal electrons
D_e	Diffusion constant for electrons
D_{ph}	Diffusion constant (effective) for phonons
C_e	Electron heat capacity per unit volume
C_{ph}	Phonon heat capacity per unit volume
Ψ	Superconducting order parameter
$\lvert\psi\rvert$	Magnitude of superconducting order parameter; $\lvert\psi\rvert^2 = n_s$
ϕ	Phase of superconducting order parameter
A	Vector potential
v_s	Cooper pair velocity
γ	Phase difference along a nanowire

k	Superconducting wave vector
k_c	Superconducting critical wave vector
p_c	Superconducting critical momentum
V	Voltage
V_L	Voltage across R_L
V_d	Voltage across device (nanowire)
C	SNSPD count rate
Θ	SNSPD dark count rate
Σ	Blackbody power spectral density per unit area per unit steradian
ε	Vortex polarizability
Γ_{VAP}	Vortex-antivortex pair unbinding attempt frequency
Γ_{VH}	Vortex hopping attempt frequency
U_{VAP}	Vortex-antivortex unbinding potential energy barrier
U_{VH}	Vortex hopping potential energy barrier
Θ_{VAP}	Dark count rate due to vortex-antivortex unbinding
Θ_{VH}	Dark count rate due to vortex hopping

Chapter 1

Introduction to Single-Photon Detection

1.1 Overview of this dissertation

In chapter 1, a historical overview of the concepts and methods of single-photon detection is given. This discussion begins with the discovery of the photoelectric effect, traces the development of single-photon detection technologies to the present day, and summarizes by giving an overview of the present state of the art, including the superconducting nanowire single-photon detector (SNSPD), which is the subject of this dissertation. Next, a description of each important performance parameter for a single-photon detector is given. In the final part of chapter 1, several important applications of single-photon detection are described. These applications motivate the ongoing development of improved detector technologies.

In chapter 2, a description of the theory and operation of SNSPDs is given. This begins with an overview of device operation and readout. This is followed by a survey of important theoretical concepts. This survey provides the foundation for understanding the specific goals of this thesis. These goals are 1) to study niobium (Nb) as an alternate to niobium nitride (NbN) as a material for SNSPDs, 2) to understand what physical processes govern photon detection and reset in Nb, and 3) to understand the nature of kinetic inductance in Nb and NbN nanowires, including its functional dependence on device parameters as well temperature and current. This second chapter is meant to orient the reader and prepare for a detailed discussion of the results of experiments meant to compare the performance of Nb and NbN SNSPDs.

In chapter 3, a detailed description of the techniques and procedures for fabricating Nb SNSPDs is given. This includes a discussion of methods to optimize the quality of Nb films from which SNSPDs are patterned, as well as work done to optimize the lithography and etching process. Fabrication of Nb SNSPDs was done at Yale and at IBM T. J. Watson Research Center in Yorktown Heights, NY. Later in chapter 3, a description of the fabrication techniques and procedures for the NbN devices is given. These devices were fabricated by collaborators at the University of Salerno in Salerno, Italy, and at CNR – Istituto di Cibernetica in Pozzuoli, Italy. Included in this chapter is a summary of the normal state device characterization, which was used to develop the fabrication process and to screen the best devices from those with defects.

In chapter 4, a description of the experimental apparatus and methods is given. This includes an overview of the measurement setup for dc characterization, photon detection characterization, and kinetic inductance measurements. After this, a description of the design and construction of the home-built cryogenic insert used in these measurements is given. Finally, the microwave readout electronics and optical excitation subsystems are described in detail. This chapter is fairly detailed since this thesis is the first time this measurement apparatus has been used in the Prober laboratory at Yale and will thus be useful for future students to reference.

In chapter 5, Nb and NbN SNSPD devices of several geometries are compared. First, the dc characterization of the superconducting properties of the devices is discussed. This is used to further screen out defective devices. Next, the detection performance of good devices is compared. Comparison to the relevant theoretical expectations developed in chapter 2 is made wherever clear predictions exist. Where the theory is unclear, the experimental results are discussed in relation to fundamental theoretical concepts in as much detail as possible. Chapter 5 is concluded with a summary and explanation as to why NbN SNSPDs are more desirable than Nb SNSPDs for most applications. A major component of this discussion is the non-ideal phenomenon of latching, which is discussed theoretically in chapter 2 and examined in detail in chapter 6.

In chapter 6, the reset dynamics of the device are studied. First, kinetic inductance, which is one of the most important physical concepts in SNSPDs, is explored. In a properly resetting device, kinetic inductance determines the reset time. Measurements of the kinetic inductance in Nb and NbN nanowires are presented. These highlight the temperature and current dependence. These results are partially explained using the theory presented in chapter 2. Next, the concept of latching is explained by studying the thermal relaxation of the out-of-equilibrium electron system of the nanowires using both experimental and theoretical results for Nb and NbN devices. A theoretical model developed in chapter 2 is used to show rigorously why latching occurs in Nb SNSPDs. This is compared to results from the literature studying latching in NbN devices. Finally, a summary of various methods to mitigate latching in SNSPD devices is given.

1.2 History of single optical photon detection

Photons are the quantized units of energy in an electromagnetic wave of a given frequency. For a frequency f, the energy of a photon is $E = hf$, where h is Planck's constant. Thus, the total energy in a single Fourier mode (characterized by the mode frequency f and the polarization) of electromagnetic radiation is equal to nhf where n is the number of photons in that mode. Although the photoelectric effect is sometimes considered the first proof of the existence of single-photons, in fact it is only proof that some form of energy quantization is present in the interaction between matter and electromagnetic waves (Lamb 1969). As he admits in the title of his famous 1905 paper, Einstein's theoretical explanation of the photoelectric effect based on the absorption of light quanta (the term photon was not introduced until later) was a phenomenological assumption (Einstein 1905). It took the experiments of Compton, concurrent with the formalization of quantum theory by Planck, Heisenberg, Dirac, Born, and others, to establish the quantum nature of electromagnetic radiation and with it the existence of the photon.

Once it was shown that the photoelectric effect was, in fact, due to the absorption of single photons of light, the photoelectric effect was exploited as the basis for the first single-photon detector. Here, the phrase "single-photon detector", refers to a device that is capable of producing a measurable response to the absorption of a single photon of light. The photomultiplier tube was the first such device. Invented in 1935 at RCA in Harrison, NJ, (Iams 1935) the photomultiplier tube utilized the photoelectric effect in conjunction with electron cascade amplification based on secondary emission to detect photons with a wavelength of 800 nm. The device was packaged in a vacuum tube. A photograph of the first photomultiplier tube is seen in Figure 1.1. Photomultipliers advanced rapidly after 1935, and are a very mature technology today for detecting single photons of wavelengths shorter than approximately 1 μm. At longer wavelengths, the lower energy of the incident photons makes initiating the cascade process difficult.

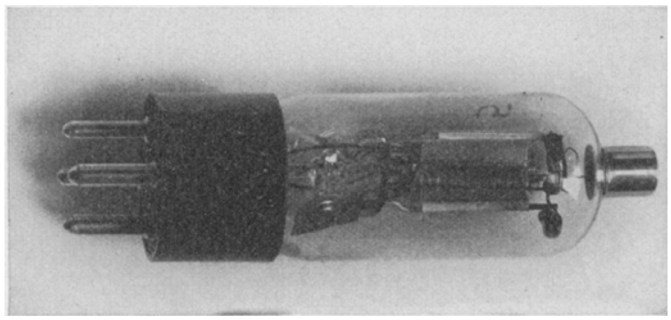

Figure 1.1: Image of first photomultiplier tube, from (Iams 1935).

Although photomultipliers remain a key technology for single-photon counting applications today, other technologies for detecting single photons, particularly at infrared wavelengths, have also been developed. A solid-state semiconductor analog of a photomultiplier tube exists in the form of an avalanche photodiode. In avalanche photodiodes, the photo-excited electron emission and cascade amplification occurs within a semiconductor, instead of within a vacuum as in a photomultiplier tube. Avalanche photodiodes for detection of single photons in the visible and near infrared wavelengths were an outgrowth of studying the avalanche breakdown process in silicon diodes. The first examples of single-photon detection in semiconductor diodes were reported by Haeker *et al*. (1971). This work was based on previous work done by Conradt *et al*. (1969) studying photoemission processes in germanium diodes. They are based on reverse-biasing a p-n junction to a very high voltage. In this case, the absorption of a single photon in or near the high field region of a p-n junction (biased near its breakdown voltage) can excite a single electron that accelerates and excites other electrons, leading to a cascade breakdown and a measurable current through the junction. Avalanche photodiodes have been developed since the early 1970s into a mature technology for detecting single near infrared and visible photons of wavelengths less than approximately 1.5 μm.

In addition to avalanche photodiodes, newer techniques for detecting single photons in semiconductor devices using quantum wells and quantum dots have also been demonstrated. In these detectors, a photon excites an electron-hole pair; the electron or hole is then spatially confined by a quantum dot or well structure that forms the gate of a field effect transistor (FET). The trapped charge can modulate the gate voltage, and therefore the conductance, of the FET. The conductance of a sufficiently small FET is sensitive to the charge induced by the absorption of a single photon. Pioneering work in quantum dot/well single-photon detectors was done by Shields *et al*. (2000) and has developed steadily in the decade since (Rowe 2006, Blakesley 2005).

In addition to photomultiplier tubes and semiconductor-based technologies, a number of technologies based on superconducting devices have been developed for detecting single optical and near infrared photons. These include superconducting tunnel junction detectors, transition-edge sensors, and superconducting nanowire detectors. The last is the subject of this thesis.

The use of superconducting tunnel junctions for detecting single optical photons was an outgrowth of their development as very sensitive x-ray detectors. Pioneering work was done by Chi *et al*. (1981) and significant early development work at Yale was completed by Wilson *et al*. (2000) whose studies were conducted in the same research group as the present work. A superconducting tunnel junction consists of two superconducting electrodes separated by a thin insulating layer. The detectors are operated well below the critical temperature of the metal that comprises the electrodes, so that nearly all electrons within those metals are condensed into Cooper Pairs. At temperatures far below the critical temperature, there is no current because there are no normal (unbound) electrons. The Cooper pairs can form a supercurrent for voltages $eV < 2\Delta$ where Δ is the superconducting energy gap, however a magnetic field is used to suppress the superconducting critical current of the junction to zero, which prevents any supercurrent from flowing. This allows the junctions to be biased with a finite voltage. When a single photon is absorbed in one of the superconducting electrodes of a junction biased in this way (or within an absorber connected to the electrodes by a short diffusive channel), it breaks many Cooper pairs, creating a large number of normal electrons. These normal electrons diffuse across the junction under voltage bias, inducing a normal current transient that can be measured with a low noise amplifier. It has been shown that these detectors are sensitive to UV and optical photons, and may well be sensitive even to mid- and far-infrared single photons (Prober 2007). Furthermore, by integrating the current flowing during a detection event to find the total number of normal electrons produced, the energy of the incident photon can be determined. Although they are presently used for x-ray detection, superconducting tunnel junction detectors for optical and infrared photons have mostly been replaced by a second type of superconducting optical photon detector.

The second type of single-photon detector based on a superconducting element is the transition edge sensor. This is also referred to as a transition edge calorimeter when used to determine the energy of single photons. The use of transition edge sensors for detecting visible and infrared photons was an outgrowth of their use as sensitive power detectors of continuous (multi-photon) infrared radiation, and from their use as single-photon detectors for higher energy photons and other single particles. Transition edge sensors with sensitivity to single optical and infrared photons were first demonstrated by Cabrera *et al*. (1998) and were a direct extension of the technology developed by Irwin *et al*. (1995) for energy resolving particle detectors for high energy physics experiments. These devices are based on a strip of superconducting metal that is biased with a voltage such that the strip is held within the middle of the superconducting-normal metal phase transition. Thus, the resistance of the strip is finite, but less than the full normal state value. When biased in this transition region, the resistance of the strip is extremely sensitive to changes in temperature. An absorbed photon heats up the strip slightly, causing its resistance to increase. This increase in resistance can be measured using a low noise amplifier. These types of detectors have seen use in astronomy and quantum information applications. Development of transition edge sensors continues at present, with work concentrated at the National Institute of Standards and Technology for quantum information applications and NASA Goddard Space Flight Center for astronomy applications. Significant new advances in transition edge sensor technology have also recently been made by Santavicca *et al*. (2010) in the same lab as this thesis work. This work extends the sensitivity of a transition edge calorimeter into the mid-infrared, with the hope of extending their sensitivity into the far infrared for direct single THz photon sensing applications in astronomy.

In addition to transition edge sensors and superconducting tunnel junction detectors, a third type of superconducting single-photon detector has been developed recently. This is based on a current biased, fully superconducting nanowire. These detectors are known as superconducting nanowire single-photon detectors (SNSPDs). They offer extremely high performance for visible and near-infrared single-photon detection. SNSPDs were first discussed by Kadin *et al*. (1996) and were first demonstrated by Semenov (2001). SNSPDs have advanced rapidly and are now beginning to be used in applications. SNSPDs are the subject of this thesis.

1.3 General properties of single optical photon detectors

The performance of a single-photon detector is characterized by several important parameters: 1) detection efficiency, 2) energy and number resolution, 3) count rate, 4) dark count rate, and 5) jitter. In this section, these parameters are formally defined.

1.3.1 Detection efficiency

The most fundamental property of a single-photon detector is that an individual photon absorbed by the detector will trigger some measurable event. Single-photon detectors must be able to accurately detect individual, randomly arriving photons; this is known as a “single shot” detector. This is distinct from a detector that can detect a single-photon only when the output signal is averaged over many cycles of a periodic input pulse that, on average, consists of a single-photon.

The sensitivity of a single-photon detector is given by the detection efficiency. There is some variation in the literature as to how detection efficiency is defined. In this thesis, it is defined as the number of individually measured events that are triggered by single-photons divided by the total number of photons incident on the detector. Thus, a detection efficiency of 0.5 means that, on average, half of all photons that arrive at a detector are measured. For nearly all single-photon detector technologies, the detection efficiency depends on wavelength and is generally lower for less energetic photons than for higher energy photons.

1.3.2 Energy and number resolution

Energy resolution refers to the ability of a detector to resolve the energy of a detected photon. Most single-photon detectors do not have single shot energy resolution. A simple reason for this is that the photon energy is generally much less than the total energy dissipated in the detection event; that is, a photon is simply the initiator of a far more energetic cascade within the detector that tends to blur the energy of the initiating photon. Cascade detectors are inherently non-linear: the output signal does not scale with the energy of the incident photon but is either constant or changes only slightly when photons of different energy are absorbed.

A few types of single-photon detectors, including the transition edge sensor and superconducting tunnel junction detector, can accurately determine the energy of a single incident photon. For monochromatic incident photons, these detectors can also be used to determine the number of photons that are absorbed at the same time (or more precisely, within a span of time that is shorter than the detector response time). For detectors with accurate intrinsic energy resolution, the output signal is typically proportional to the energy or number of photons absorbed. Thus, these detectors have a linear response. Figure 1.2 illustrates the difference in output signals between linear and non linear detectors. Although one report (Semenov 2008) has suggested that SNSPDs may be able to resolve the energy of photons in the near infrared (wavelengths of 800-1200 nm), this report has not been corroborated. Furthermore, there has been no report that SNSPDs have an energy or number resolving capability for visible photons.

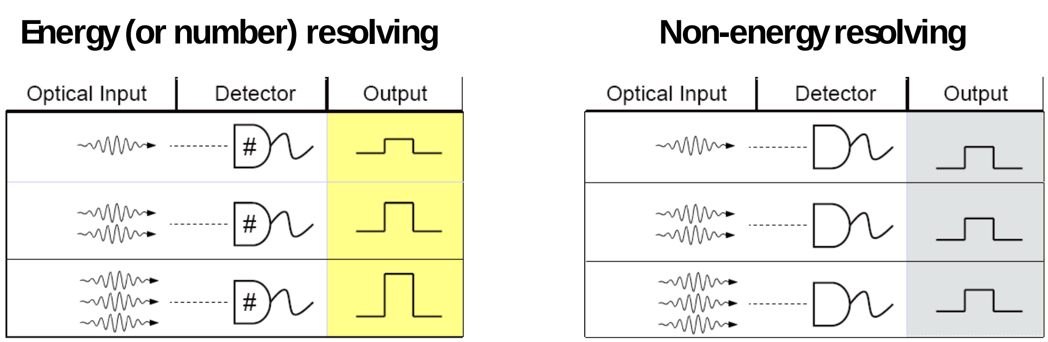

Figure 1.2: There are two categories of single-photon detectors, energy-resolving (linear response) and non-energy resolving (non-linear or cascade response), (Nam 2004).

1.3.3 Count rate

The count rate of a single-photon detector is the number of single-photons detected per unit time. The maximum count rate is specified by the maximum number of photons that a detector can count per unit time with a given detection efficiency. For example, if a detector has a detection efficiency of 0.2 and a 10 MHz maximum count rate, it can detect 10^7 photons per second where the probability of detecting any individual photon incident on the detector is 0.2. If photons arrive at a higher rate than 10 MHz, the detector will typically miss some (not be triggered) such that its average detection efficiency is less than 0.2. For nearly all detectors, at sufficiently high incident photon arrival rates, the detector will reach saturation and will not be able to detect any additional single-photons. The details of this saturation and the resulting tradeoff between count rate and detection efficiency depend on the detector technology. Typically, the count rate is set by the reset time or "dead time" of a detector. This is defined as the amount of time required for the detector to return to its initial quiescent state after a photon is detected. If a second photon arrives during this dead time, the photon will either not be detected at all, or might be detected, but with probability significantly less than the probability of being detected if the detector was fully reset.

1.3.4 Dark count rate

Dark counts are spurious, measured events that are indistinguishable from real detection events caused by incident single-photons. They can be randomly occurring, or might follow the detection of an actual photon as the device resets. They are the chief source of noise that leads to uncertainty in the measured single-photon arrival rate. For a single detection event, this means there is a probability that the count is a false count, and not the result of an incident photon. Dark counts are specified by a dark count rate, which is the number of spurious counts a detector measures per unit time. The probability that a single detection event is a dark count is simply equal to: (dark count rate)/[(photon arrival rate × detection efficiency) + (dark count rate)]. Thus, dark counts are especially a concern when the photon arrival rate or the detection efficiency is low.

1.3.5 Jitter

Jitter is the uncertainty in the time that a photon is detected. It is defined as the full width at half maximum of the distribution of delay times between when a photon is incident on the detector and when an output signal is first detected. Therefore, it is a measure of the precision with which a detector can localize the arrival of a photon in time. The degree to which this delay is Gaussian distributed depends on the source of the jitter, which depends on the type of detector.

1.3.6 Single-photon detector technologies compared

In Table 1.1 a comparison of current state of the art single-photon detector technologies is presented. The performance of the best infrared single-photon detectors from each category is compared for each of the performance criteria outlined in sections 1.3.1 – 1.3.5. Overall, three major observations can be made. First, the best semiconductor detectors have lower detection efficiency at near infrared wavelengths than the best superconducting detectors. They also have a significantly shorter cutoff wavelength. Figure 1.3 shows the detection efficiency of the best Hamamatsu photomultiplier tubes. As can be seen, this cutoff is also very sharp. In this example, the cutoff is due to the bandgap of InGaAs, which is approximately 0.74 eV (Nahory 1975). Since InGaAs has one of the smallest bandgaps of any compound semiconductor that can be reliably manufactured, it is necessary to move to technologies that do not use semiconducting absorbers in order to detect longer wavelength photons. Superconducting detectors have no such band-gap limited cutoff, since the superconducting energy gap is ~1 meV. The second major observation is that SNSPDs are much faster than any other type of detector. They exhibit count rates that are at least an order of magnitude higher than semiconductor detectors, and jitter that is approximately an order of magnitude less. Thus, SNSPDs are particularly attractive for fast infrared photon counting. Finally, the drawback to superconducting detectors is that they must be operated at cryogenic temperatures. However, with the continued development of cryogen-free cooling technology, operating temperature should not be a major roadblock to using superconducting detectors in high performance photon counting applications in the future.

Detector Technology	Detection Efficiency (λ)	Energy Resolution	Max. Count Rate	Dark Count Rate	Jitter	Max. λ	Operating Temp.
InGaAs photodiode[1]	0.1 (1.55 μm)	Yes	100 MHz	16 kHz	55 ps	1.7 μm	240 K
Photomultiplier Tube[2] (Hamamatsu R5509-73)	0.01 (1.55 μm)	No	9 MHz	160 kHz	150 ps	1.7 μm	200 K
Superconducting tunnel junction[3]	0.6 (1.3 μm)	Yes	0.005 MHz	0	N/R	> 2 μm	0.25 K
Superconducting transition edge sensor (NIST)[4]	0.95* (1.55 μm)	Yes	0.1 MHz	0	100 ns	> 2 μm	0.1 K
NbN SNSPD[5]	0.57* (1.55 μm)	No	300 MHz	100 Hz**	30 ps	5.0 μm	2 K
NbTi SNSPD[6]	0.02 (0.65 μm)	No	300 MHz	N/R	N/R	N/R	2K
Nb SNSPD[7]	0.01 (1.55 μm)	No	150 MHz	100 Hz	100 ps	> 1.5 μm	2K

Table 1.1: A comparison of the best of each type of infrared single-photon detector technology. Data from: [1]Dixon, et al. (2008), [2]Hammamatsu (2009), [3]G'oltsman *et al.* (2005), [4]Lita, *et al*. (2008), [5]Rosfjord, *et al*. (2006), [6]Dorenbos *et al.* (2008), [7]from this thesis work. For more information, including information on the performance of visible single-photon detectors, see Hadfield *et al*. (2009). Notes: *These measurements are for a detector within an optical cavity to maximize the coupling efficiency; **In this thesis, the measured dark count rate for NbN SNSPDs is quoted in Table 5.2 to be significantly less than the value listed here from the MIT work; this is because in this thesis, dark counts are reported for a slightly lower value of the bias current than in the MIT work; N/R: not reported.

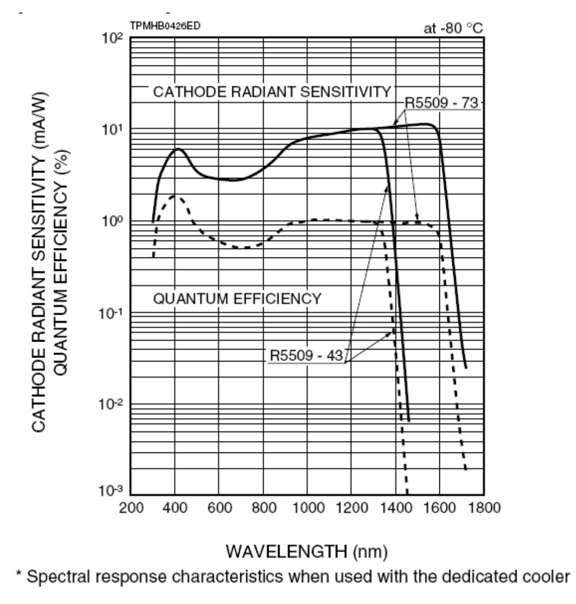

Figure 1.3: Detection efficiency versus wavelength for Hamamatsu photomultiplier tubes. Here, "quantum efficiency" is defined in the same way as "detection efficiency" is defined in this thesis. From Hamamatsu (2009).

1.4 Applications of visible and near-infrared single-photon detectors

Historically, single-photon detectors have seen a variety of applications, ranging from precise characterization light emitting devices, to scintillation detectors in particle physics experiments. This section focuses on emerging applications for single-photon detectors. These new applications require significantly higher performance detectors than have been available in the past. Thus, they are the main driver of single-photon detector development. Below is a survey that, while not exhaustive, is representative of the diversity of applications requiring high performance near-infrared single-photon detectors. A lack of sufficiently high performance detectors is a significant bottleneck to the continued development of these technologies.

1.4.1 Picosecond Imaging Circuit Analysis

The initial motivation for this thesis work was to develop a single-photon detector for picosecond imaging circuit analysis (PICA). PICA is an analysis technique that uses time-resolved measurements of single-photon emission from integrated circuits to locate and analyze defects (McManus 2000). It was developed at IBM and is used in fabrication development for advanced CMOS circuitry. The technique measures the photoemission that occurs from individual CMOS transistors when in the on and off states as well as when they switch. By characterizing this photoemission and measuring how it varies in time within the clock cycle and spatially across a chip, then a map of process defects can be composed. Such a map of an entire chip is seen in Figure 1.4 for one sample point within the clock cycle. Since clock cycles are now typically ~300 ps or less, timing resolution of less than ~30 ps is necessary for accurate sampling. Thus, very low jitter single-photon detectors are required. SNSPD-type detectors are a prime candidate. PICA analysis using SNSPDs has been demonstrated by Korneev *et al.* (2003).

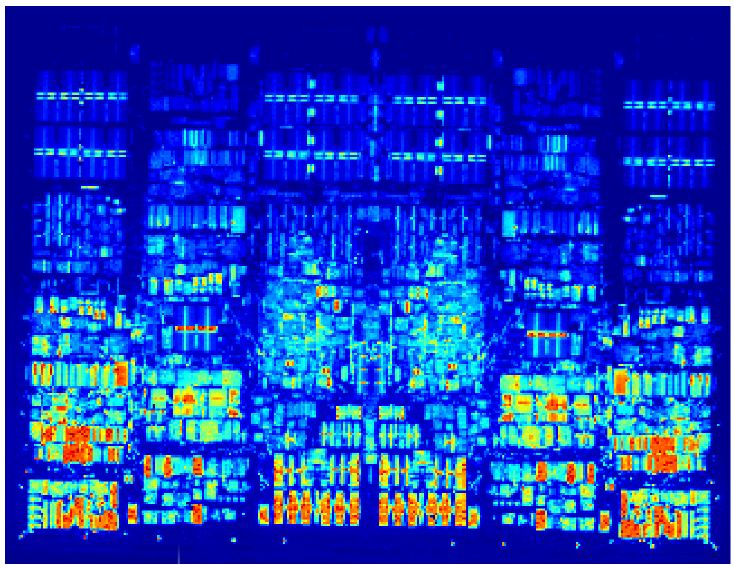

Figure 1.4: A false-color image of photoemission intensity across a chip at a single point within the clock cycle compiled using picosecond imaging circuit analysis (Polonksy, 2005).

1.4.2 Other Applications

Several other prominent applications for high performance single-photon detectors exist, including in communications and single-photon spectroscopy. For communications applications, single-photons can be used to transmit data via the arrival time of individual photons using a technique referred to as Pulse Position Modulation (PPM). If a detector can resolve the single-photon arrival time accurately, then very high data rates can be achieved with very low power using PPM. Single-photon PPM is ideal in long distance communication applications, in particular with space-based receivers. Robinson *et al.* (2006) have demonstrated data transmission at ~Gbit/s rates at a wavelength of 1550 nm using NbN SNSPD detectors at MIT Lincoln Laboratory, where development continues.

Another type of communication application of single-photon detectors is in Quantum Key Distribution (QKD). In QKD, quantum entanglement of the polarization of photon pairs, or another quantum state of the photon, is used to transmit an encryption key using a quantum algorithm. This quantum algorithm makes any interception of the key immediately apparent to the sender. Thus, QKD facilitates unbreakable encrypted data transmission. QKD has been demonstrated using meander-geometry SNSPD by Jaspan *et al.* (2006) at the National Institute of Standards and Technology, where development is ongoing.

In addition to communications applications, single molecule fluorescence measurements have been demonstrated using SNSPDs by Stevens *et al.* (2006). In this application, the long wavelength sensitivity as well as the timing accuracy of SNSPDs is potentially very attractive. SNSPDs may also be ideal candidates for use in lidar, similar to radar but with visible/near infrared light used to precisely locate and determine the distance to objects.

Chapter 2

Device Theory and Operation

2.1 Superconducting Nanowire Single-Photon Detector Operation

The development of superconducting nanowires single-photon detectors (SNSPDs) is based on previous work developing superconducting hot electron bolometers, although the detection mechanism sin these devices are quite different. A hot electron bolometer consists of a short superconducting micro- or nanowire that is biased on its resistive transition, so as to make a stable resistive hotspot with resistance that depends very strongly on the hotspot temperature. Power added by an input signal increases the temperature of the hotspot slightly, increasing the resistance and therefore the device voltage (Floet 1999). This modulation of the device voltage can occur from dc to ~50 ps timescales depending on the duration or frequency of the input signal, but these fast devices are not sensitive to single-photons of energy ≤ 1 eV (Santavicca 2009, Reese 2006). Transition edge sensors are similar to hot electron bolometers but can detect single-photons; however, they operate on much longer timescales than bolometers and require a much lower operating temperature (Santavicca 2010). In contrast to hot electron bolometers and transition edge sensors, which both have stable hotspots, in SNSPDs, detection is based on the single-photon-induced creation of a transient resistive hotspot, with a lifetime $\sim$ 1 ns, within a current biased, fully superconducting nanowire (Semenov 2001). Thus, SNSPDs are extremely non-linear detectors, providing a much larger resistance change than a hot electron bolometer or transition edge sensor for a small input energy, and therefore develop a much larger voltage signal. Like transition edge sensors, they can detect single-photons, but they operate on timescales that are as fast as hot electron bolometers. However, unlike hot electron bolometers and transition edge sensors, SNSPDs are not energy resolving.

A single-photon detection mechanism in superconducting nanowires was first proposed by Kadin *et al*. (1990 (1), 1990 (2)). Since then, there has been considerable discussion in the literature about the microscopic mechanism(s) responsible for photodetection in SNSPDs. A basic overview of the process of detection in an SNSPD is illustrated in Figure 2.1, and a typical readout circuit is shown in Figure 2.2. An SNSPD is usually patterned in a meander so as to form a pixel with an area $A_d > \lambda^2$ where λ is the wavelength of the photons that are to be detected. A Thevenin-equivalent circuit for the combined device and readout circuit is seen in Figure 2.3 in three stages; each stage corresponds to one of the three stages of the illustration of detection in Figure 2.1. The device has kinetic inductance, L_K, that is assumed constant for a specific device and proportional to the device length l_d (and therefore also to the device detection area, A_d) and a time dependent resistance, $R_d(t)$. The readout circuit consists of an inductively-coupled dc bias current, I_b, and a capacitively-coupled high-frequency amplifier with load resistance, R_L, that is typically equal to 50 Ω.

The detection process is as follows: (a) Initially, a nanopatterned strip of superconducting metal of typical width w_d ~100 nm and typical thickness d_d ~5 nm is biased with a current, I_b, whose value is just below the critical current at the operating temperature $I_{co} = I_c(T_o)$ where the operating temperature $T_o << T_c$. In this state, only a small amount of additional energy is required so that I_c is suppressed to a value less than I_b, creating a finite voltage. Thus, a single photon can induce a localized resistive hotspot. (b) This hotspot will expand quickly (in much less than a nanosecond) to encompass the width of the wire, at which point the device will develop a series resistance, R_d, that will cause Joule heating from the bias current. This Joule heating will amplify the size of the hotspot substantially and lead to a multiplication of the energy of the initial photon. Since the maximum hotspot resistance $R_{d,max} >> R_L$, most of the bias current will transfer to the load and the peak signal will simply be $V_{L,max} = I_b \cdot R_L$. The time scale of the transfer is $\tau_t = L_K/(R_L+R_d)$ where R_d is a function of time. (c) As R_d becomes large, the current that is shunted into the load reduces the Joule heating within the device, allowing the hotspot to cool and the zero-resistance state to be restored on a time scale set by thermal relaxation. Once the hotspot returns to near zero-resistance, the current slowly begins to transfer back into the device with a return time constant $\tau_\mathrm{r} = L_K/R_L$. Once the bias current has transferred back into the nanowire, the device is reset and ready to detect another photon.

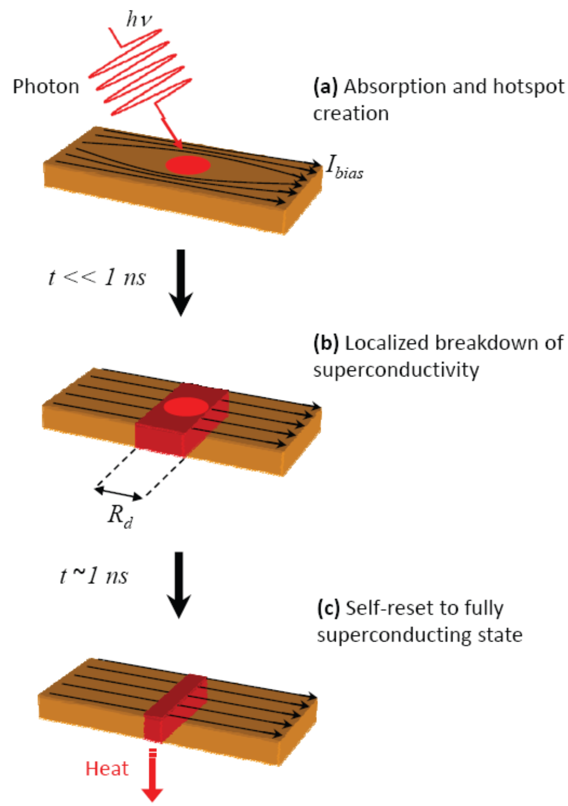

Figure 2.1: Overview of photon detection in an SNSPD. (a) $R_d = 0$ in equilibrium, as well as just after the photon is absorbed; (b) the device has become resistive due to the absorption of a photon and the spreading of the hotspot, which reduces I_c to a value below I_b, after which R_d quickly increases to a large value; (c) the hotspot resistance has returned to nearly zero and temperature is returning to equilibrium. Adapted from Semenov, *et al*. (2001).

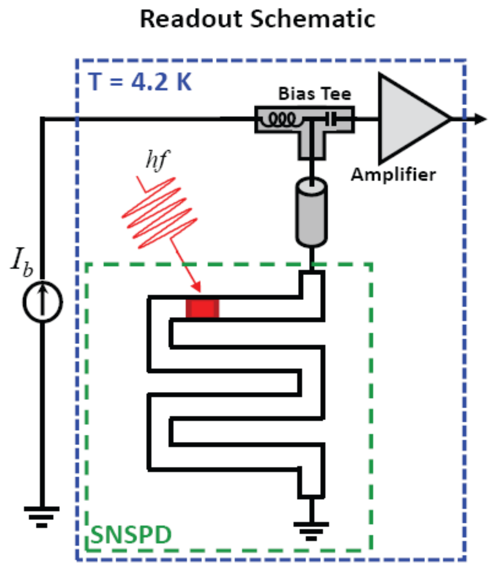

Figure 2.2: Overview of the readout circuit for an SNSPD. The device is represented as a meandered strip with a resistive hotspot that occupies only a small fraction of the total nanowire. The device is biased via an inductively-coupled dc bias current. The output voltage is measured by a capacitively-coupled high frequency amplifier with an input resistance, R_L, that is typically equal to 50 Ω. This readout circuit is explained in greater detail in chapter 4.

2.2 Scattering and Energy Sharing Processes in Metals at Low Temperature

In SNSPDs, energy is added directly to out-of-equilibrium electrons (non-equilibrium quasiparticles, as the elementary excitations from the superconducting state are known).[1] This energy must exit the nanowire through the substrate. While the energy transfer from the photon to the environment will not be traced in its entirety, it is important to understand the fundamental interactions that define how energy is shared among electrons, phonons, and the substrate. Once the energy enters the substrate, it is

[1] In this thesis, the word “electron” and the word “quasiparticle” will be taken to be equivalent, since effects related to charge imbalance or other phenomena where quasiparticles differ in physical character from electrons are believed not to be relevant.

effectively gone from the nanowire system. In the nanowires studied in this thesis, the electrons and phonons within the nanowire are separately out of equilibrium with the substrate. An electron temperature, T_e, and a phonon temperature, T_{ph}, is defined. The substrate temperature, T_o, is in general less than T_e and T_{ph}, and is equal to the operating temperature of the cryostat. The time scales of the interactions that govern energy transfer between these three systems are now discussed.

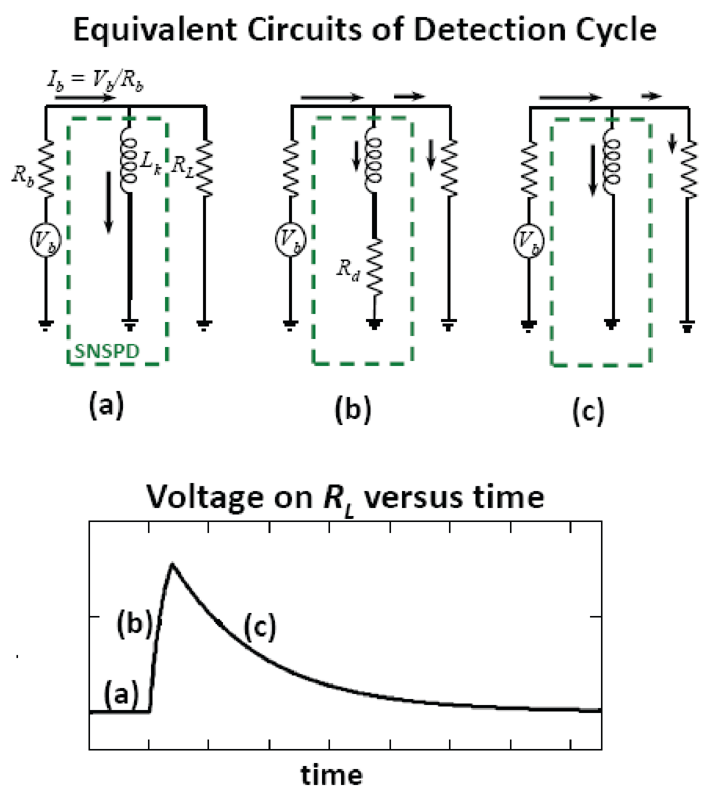

Figure 2.3: Equivalent circuit and typical output voltage pulse for an SNSPD. (a) $R_d = 0$ in equilibrium, as well as just after the photon is absorbed; (b) the device has become resistive due to the absorption of a photon and the spreading of the hotspot, which reduces I_c to a value below I_b after which R_d quickly increases to a large value and most of the bias current transfers into the load with a time scale of $\tau_t = L_K/(R_L+R_d)$, where R_d is a function of time; (c) the hotspot resistance has returned to zero (or nearly zero), and the current is returning slowly to the device with a time constant $\tau_r = L_K/R_L$.

2.2.1 Ginzburg-Landau Time

The Ginzburg-Landau time is the time scale of intrinsic superconducting fluctuations. It is also the time scale over which superconductivity can break down into the normal state or be restored from the normal state at temperatures close to T_c. Thus, it is related to the coherence length, $\xi(T_e)$, which is the characteristic length scale over which the superconducting order parameter can vary. The Ginzburg-Landau time is given by(Tinkham 1996):

$$\tau_{GL} = \frac{0.731}{D_e} \cdot \frac{\xi_o \ell_e}{1 - T_e/T_c} \tag{2.1}$$

where D_e is the diffusion constant for electrons, ξ_o is the Pippard coherence length, and ℓ_e is the electron mean free path. For both Nb and NbN, τ_{GL} is < 1 ps for nearly all temperatures (in the range of validity of the Ginzburg-Landau theory) except extremely close to the critical temperature.

2.2.2 Electron-Electron Inelastic Scattering

The electron-electron inelastic scattering is the time scale over which an excited electron shares its energy with another electron. Thus, it is the time scale over which a quasiparticle hotspot in a superconductor thermalizes. In disordered superconducting films such as Nb and NbN, a simple approximation for the electron-electron scattering time exists if the film is two dimensional with respect to two characteristic length scales,

$$l_{Th} = \left(\frac{\hbar D_e}{k_B T_e} \right)^{1/2} \qquad , \tag{2.2}$$

$$l_2 = \left(\frac{2\pi\left(\hbar / e^2\right)}{\rho} \right) l_{th}^2 \left(\frac{1}{\ln\left(\frac{\pi\left(\hbar / e^2\right)}{R_\square} \right)} \right). \tag{2.3}$$

If the thickness of the film is less than the Thouless length, l_{Th}, as well as a second length scale, l_2, that is proportional to l_{Th}^2 as well as to the ratio of the resistance quantum to the film resistivity, then the electron-electron inelastic time can be approximated by

$$\tau_{e-e} = \frac{1}{4 \times 10^7 R_\square T_e}, \tag{2.4}$$

where $R_\square$ is the sheet resistance, and ρ is the resistivity of the film (Santhanam 1987). Using typical parameters for Nb (D_e = 1 cm^2/s, $R_\square$ = 100 Ω/□, ρ = 75 μΩ-cm, and T_e = 4 K) these lengths are: $l_{Th} \approx$ 14 nm and $l_2 \approx$ 1.35 μm. For NbN, using typical parameters (D_e = 0.25 cm^2/s, $R_\square$ = 1000 Ω/□, ρ = 500 μW-cm, and T_e = 10 K) these lengths are: $l_{Th} \approx$ 4.4 nm and $l_2 \approx$ 20.3 nm. Typical Nb film thicknesses are ~ 7.5 nm; typical NbN films are ~ 5 nm. Thus, equation (2.4) is applicable to Nb and NbN. In this case, $\tau_{e-e} \approx$ 63 ps for Nb and $\tau_{e-e} \approx$ 2.5 ps for NbN at T_e = 4 K and 10 K, respectively.

2.2.3 Electron-Phonon Inelastic Scattering

The electron-phonon time is the characteristic timescale over which an electron can relax by scattering its energy into a phonon mode within the lattice of the film. The electron-phonon time is difficult to calculate theoretically for disordered metallic films such as Nb and NbN. The temperature dependence of the electron-phonon time follows an inverse power law with temperature. The power may range from three in a clean metal to less than two in a very disordered metal such as NbN. Experimentally, it is found that:

$$\tau_{e-ph} \approx 700 ps \times \left(\frac{6.5K}{T_e} \right)^2 \quad (2.5)$$

for typical thin Nb films (Santavicca 2009, Gershenzon 1990),

$$\tau_{e-ph} \approx 2ns \times \left(\frac{6.5K}{T_e} \right)^2 \quad (2.6)$$

for ultra thin (7.5 nm) Nb films used in this thesis,[2]

$$\tau_{e-ph} \approx 10 ps \left(\frac{10K}{T_e} \right)^{1.5} \quad (2.7)$$

for typical NbN (Ptitsina 1997).

[2] This is an estimate based on measurements in Santavicca (2009), but using parameters from the films studied in this thesis. Films used for the best performing Nb SNSPDs studied in this thesis had a resistivity that was significantly higher, and a critical temperature that was significantly lower, than the films studied in Santavicca (2009).

2.2.4 Phonon Boundary Scattering

Phonon reflection at the boundary (scattering) lengthens the escape time for phonons to exit the metal film and enter the substrate. This reflection is due to the finite speed of sound and to a lattice mismatch at the boundary between the metal film and the substrate, which leads to a non-unity transmission coefficient for phonons. Although theoretical studies have been undertaken, it is also difficult to derive an accurate theoretical expression for the escape time from a thin, disordered film. The escape time can be expressed as:

$$\tau_{esc} \approx N \frac{2d_d}{v_s} \tag{2.8}$$

where d_d is the thickness of the film, v_s is the speed of sound in the film, and N is the average number of attempts to cross the boundary that a phonon makes. Experimentally, for NbN films the escape time is measured to be 30 ps for d_d = 5 nm (Gousev, Semenov 1994; Gousev, Gol'tsman 1994). This value is used in all calculations in this thesis. In Nb, τ_{esc} is difficult to measure because most experiments only measure the quantity ($\tau_{e\text{-}ph}$ + τ_{esc}), and in thin Nb, $\tau_{e\text{-}ph} >> \tau_{esc}$. In this work, it is assumed that τ_{esc} = 45 ps for a 7.5 nm film, however there are no instances when the exact value of τ_{esc} is significant for Nb SNSPDs.

2.2.5 Diffusion

The rate at which excited electrons move away from the source of the excitation in a disordered metal is given by diffusion. In one dimension, the time for an electron to diffuse across a length l is expressed simply as:

$$\tau_D \approx \frac{l^2}{D_e} \qquad (2.9)$$

This expression can be used as an approximation for the time to diffuse from the center of a nanowire of width 100 nm to the edge (l = 50 nm). For Nb, $D_e \approx 1$ cm^2/s, yielding τ_D = 25 ps; in NbN, $D_e \approx 0.25$ cm^2/s, yielding τ_D = 100 ps (Reese 2006, Santavicca 2009, Ptitsina 1997).

2.2.6 Summary of Relevant Timescales in Nb and NbN

Table 2.1 contains a summary of the relevant timescales for Nb and NbN thin films. Calculations are for typical disordered 7.5 nm thick Nb films and 5 nm thick NbN films.

Parameter	**Nb** (d_d = 7.5 nm)	**NbN** (d_d = 5 nm)
τ_{GL}	~0.5 ps	~0.05 ps
$\tau_{e\text{-}e}$	63 ps	2.5 ps
$\tau_{e\text{-}ph}$	2 ns (T_e = 6.5 K)	10 ps (T_e = 10 K)
τ_{esc}	~45 ps	30 ps
τ_D	25 ps	100 ps

Table 2.1: Calculated values of the relevant timescales in Nb and NbN thin films.

2.3 Kinetic Inductance

In this section, the concept of kinetic inductance is introduced and compared to magnetic inductance. Next, an expression for the kinetic inductance is derived for a normal metal using the Drude model, and for a superconductor using the Ginzburg Landau theory. Finally, this expression is used to make predictions of the kinetic inductance in Nb and NbN SNSPD devices as a function of temperature and current.

2.3.1 Overview of Kinetic and Magnetic Self Inductance

Magnetic self-inductance is associated with the energy stored in a magnetic field. According to Faraday's law of induction, changes in magnetic flux can "induce" voltages that drive currents that oppose that change in the magnetic flux. The current persists until the difference in energy stored in the magnetic field between the initial and final values of the flux is depleted by the dissipative current that is induced. (In a perfect conductor, the current will not dissipate any energy and therefore will persist forever, so the total flux will never change.) Thus, inductance can be seen to represent a kind of magnetic flux "inertia", whose analogy to Newton's law of inertia is clear from Lenz's law; namely, that an induced current is always in such a direction as to oppose the motion or change causing it. It is instructive to derive the equation that relates induced voltage to magnetic self inductance. First, consider the definition of the magnetic self inductance,

$$L_M = \frac{\Phi}{I_d} \tag{2.10}$$

where Φ is the flux induced by a current I_d. Solving for Φ and differentiating both sides,

$$\frac{d\Phi}{dt} = L_M \frac{\partial I_d}{\partial t} + I_d \frac{\partial L_M}{\partial t}. \tag{2.11}$$

Then, by Faraday's law:

$$V = L_M \frac{\partial I_d}{\partial t} + I_d \frac{\partial L_M}{\partial t}. \tag{2.12}$$

where the second term is usually zero since L_M is typically (but not *necessarily*) constant in time. This is a completely general result that is simply an extension of Faraday's law. It applies any time there is a current to generate a magnetic field. For a straight metal wire with length l_d and with a circular cross section of radius r, the magnetic self inductance is given approximately by:

$$L_M \approx \frac{\mu_o}{2\pi} l_d \left[\ln\left(\frac{2l_d}{r}\right) - 1 \right] \tag{2.13}$$

where $\mu_0 = 4\pi \times 10^{-7}$ H/m (Miller 2002, but originally Rosa 1908).[3] As will be shown, this logarithmic dependence means that for very high aspect ratio superconducting wires, magnetic inductance will be small compared to kinetic inductance. In general, because of this logarithmic dependence, a typical rule of thumb is that $L_M \sim 1$ nH/mm for a very wide range of aspect ratios.

Kinetic inductance has no relation to Faraday's law of induction. It is referred to as an inductance purely because if a sinusoidal drive is assumed, an equation analogous to (2.12) can be written, with L_K replacing L_M. Kinetic "inductance" is due to Newton's

[3] This calculation is for a circular cross section at frequencies greater than dc, without a dielectric or ground plane nearby, and so is only an approximation for a nanostrip on a dielectric substrate. As will be shown, the magnetic inductance of a superconducting nanowire is much less than the kinetic inductance, and these details are largely irrelevant for this thesis work. In addition, the typical meander geometries of SNSPDs will reduce the magnetic inductance to values that are less than equation (2.13) predicts for a straight wire.

law of inertia. Charge carriers have mass, which means that it takes time for a finite electric field to accelerate them to a certain velocity, and if an opposing electric field is applied, it will take time to change their velocity. Thus, a better name for kinetic inductance would be "current inertia" or perhaps "charged inertia." Furthermore, kinetic inductance is not a special property of a superconductor, nor is it a purely quantum mechanical phenomenon.

2.3.2 Kinetic Inductance in a Normal Metal

To derive the kinetic inductance in a normal metal, begin with the Drude model of transport (Ashcroft 1976):

$$\vec{J} = \sigma_e \vec{E} \tag{2.14}$$

$$= \frac{n_e e^2 \tau_e}{m_e \left(1 - i\omega\tau_e\right)} \vec{E}\,, \tag{2.15}$$

where J is the current density, E is the electric field, σ_e is the complex conductivity of electrons, ω is the angular frequency of the oscillating field, τ_e is the elastic scattering time for an electron in a metal and n_e is the electron (charge carrier) density, e is the charge of an electron, and m_e is the mass of an electron. In Ohmic transport, typically $\omega\tau_e$ << 1, and so the real component of the conductivity is dominant and Ohm's law can be easily derived. Letting $\omega\tau_e$ >> 1, as is the case at very high frequencies, this instead simplifies to give a purely imaginary impedance:

$$\vec{E} = i\omega\left(\frac{m_e}{n_e e^2}\right)\vec{J}\,. \tag{2.16}$$

Or, solving for the voltage that results by having this field across some finite length:

$$V = \int i\omega \left(\frac{m_e}{n_e e^2} \right) \overline{J} \bullet \overline{dl} \,. \tag{2.17}$$

Assuming J and n_e are uniform across the length, for a nanowire of length l_d and cross sectional area A_{cs} carrying current I_d this simplifies to:

$$V = i\omega \left(\frac{m_e}{n_e e^2} \right) \left(\frac{l_d}{A_{cs}} \right) I_d \,. \tag{2.18}$$

From this equation, the kinetic inductance, L_K, is defined such that:

$$V = i\omega L_K I_d \,, \tag{2.19}$$

so that the impedance is

$$Z = \frac{V}{I_d} = i\omega L_K \,, \tag{2.20}$$

which is an inductive response with equivalent "inductance" of L_K. Assuming a sinusoidal signal, and recasting Z and I into phasors, an equation analogous to (2.12) is obtained:

$$V = L_K \frac{dI_d}{dt} \,. \tag{2.21}$$

Note that this equation does not contain any time derivative of L_K. This is different from a magnetic inductance, where an explicit time dependence of L_M can contribute to the voltage across the wire. The magnetic inductance is typically set by the geometry of a circuit loop or by the diameter of a wire, in either case independent of the current. Also note that the reason kinetic inductance is only a small component of the impedance of a normal metal is that usually $\omega\tau_e \geq 1$ only at very high frequencies (THz or greater), so the resistive impedance dominates for almost all relevant transport measurement frequencies.

2.3.3 Kinetic Inductance of a One-dimensional Superconductor

In a narrow superconducting strip where the strip width, w_d, and thickness, d_d, are less than the magnetic penetration depth, λ, and the coherence length, ξ, the supercurrent density is uniform within the wire. In this case, to obtain the correct expression for L_K, it is only required to assume that $\tau_e \rightarrow \infty$ such that $\omega\tau_e >> 1$ even as $\omega \rightarrow 0$, and then to replace m_e, n_e, and e with the Cooper pair mass ($2m_e$), density (n_s), and charge ($2e$), respectively, in equation (2.16):

$$\overline{E} = i\omega\left(\frac{m_e}{2n_s e^2}\right)\overline{J}\ , \tag{2.22}$$

after which the above derivation applies and the resulting expression for the kinetic inductance in a superconductor is just:

$$L_K = \left(\frac{m_e}{2n_s e^2}\right)\left(\frac{l_d}{A_{cs}}\right). \tag{2.23}$$

In a superconductor, however, n_s depends on temperature and current. The dependence on temperature is given simply by the temperature dependence of n_s: $n_s(T) = n_s(0)(1-T/T_c)$. To determine the current dependence, the effect of a finite current on n_s

must be considered. It is instructive to begin with the Ginzburg-Landau theory of superconductivity. Note that in this derivation, as in all calculations in this thesis, SI units will be used.[4] Although the Ginzburg-Landau theory is technically only valid near T_c, empirically, the critical current in Nb and NbN nanowires depends somewhat strongly on temperature down to approximately half of the critical temperature (see section 5.2), which suggests that this derivation should be valid in the temperature range of $T_c/2 < T < T_c$. In general, it will be assumed that each parameter (e.g., ξ) will have some temperature dependence ($\xi = \xi(T)$) not explicitly noted in this calculation. This calculation will assume constant electron temperature and current which changes adiabatically with respect to all relevant relaxation times for quasiparticles and Cooper pairs. If R_d is the resistance of the hotspot and R_n is the total normal state resistance of the superconductor, then there may be some small fractional discrepancy of order (R_d/R_n) between what this calculation predicts and the actual L_K versus time for an SNSPD whose current changes due to the formation of a resistive hotspot. First, consider $d_d \sim |r|$ such that the superconducting order parameter can be expressed simply as:

$$\Psi(\vec{r}) \approx |\psi| e^{i\phi(\vec{r})}, \tag{2.24}$$

where ϕ is the phase and $|\psi|^2 = n_s$. Then, from the Ginzburg-Landau theory (and fundamental quantum mechanics), the following expression (in SI units) for the superconducting current density can be used:

[4]It is conventional to express the Ginzburg-Landau equations in CGS units. This is done in most textbook examples, in particular Tinkham (1996) and many papers in the literature. Thus, at the end of these derivations, the result is reported in both SI and CGS units and is clearly labeled.

$$\vec{J}_s = \frac{e}{m_e}|\psi|^2\left(\hbar\nabla\phi(\vec{r}) - 2e\vec{A}\right) \tag{2.25}$$

$$= 2e|\psi|^2\frac{\hbar k}{2m_e} = 2e|\psi|^2 v_s\,, \tag{2.26}$$

where A is the vector potential and v_s is the velocity of Cooper pairs (Tinkham 1996). If the cross-sectional area of the nanowire, A_{cs}, is such that $A_{cs} << \lambda_M^2$ where λ_M is the magnetic field penetration depth of the superconductor, then the contribution of the screening field in the calculation of free energy can be ignored. Then, minimizing the free energy with respect to the supercurrent velocity as done by Tinkham (1996), the following result is obtained:

$$|\psi|^2 = n_0\left(1 - \left(\frac{\xi(T)m_e v_S}{\hbar}\right)^2\right), \tag{2.27}$$

so that

$$\vec{J}_S = 2e|\psi|^2 v_S \tag{2.28}$$

$$= 2en_0\left(1 - \left(\frac{\xi(T)m_e v_S}{\hbar}\right)^2\right)v_S\,, \tag{2.29}$$

where n_0 is the Cooper pair density at $I = T = 0$. This can be expressed more intuitively using:

$$v_s = \frac{\hbar k}{2m_e} \tag{2.30}$$

and

$$\gamma = kl_d\,, \tag{2.31}$$

where k is the magnitude of the superconducting wave vector and γ is the total phase difference along the nanowire. Taking γ to be the independent variable, the following equation for the supercurrent can be derived:

$$I_s = \frac{e\hbar}{m_e}\left(\frac{A_{cs}}{l_d}\right)n_0\left(\gamma - \frac{\xi^2\gamma^3}{l_d^2}\right). \tag{2.32}$$

The phase, γ, can be related to L_K using the first Josephson relation, which is a general result of quantum mechanics that applies when the time dependence of the driving potential that comprises the Hamiltonian is slow compared to the Cooper pair and quasiparticle relaxation times (the usual adiabatic approximation of time-dependent quantum mechanics).[5] Then, the Hamiltonian is:

$$H = T + U(t)\,, \tag{2.33}$$

$$U(t) \approx U\,, \tag{2.34}$$

and thus:

$$\frac{2eV}{\hbar} = \frac{d\gamma}{dt}\,, \tag{2.35}$$

[5] This is true for an ac voltage at ≈100MHz, the highest frequency used in this work to measure L_K. It is also true for the frequencies associated with the dynamics of I_d after a photon is detected, for all the devices tested in this thesis.

which is just the first Josephson relation where γ is the gauge-invariant phase difference between the points where the voltage is measured. The voltage, V, can be related to L_K via the circuit expression:[6]

$$V = L_K \frac{dI_s}{dt} \ . \tag{2.36}$$

Thus,

$$L_K = \frac{\dfrac{\hbar}{2e}\dfrac{d\gamma}{dt}}{\dfrac{dI_s}{dt}} \tag{2.37}$$

$$= \frac{m_e}{2e^2}\left(\frac{l_d}{A_{cs}}\right)\left(\frac{1}{n_0\left(1-\dfrac{3\gamma^2\xi^2}{l_d^2}\right)}\right) \tag{2.38}$$

Substituting for the magnetic penetration depth,

$$\lambda^2(T) = \frac{m_e}{\mu_o n_s(T) e^2} , \tag{2.39}$$

the following result for the kinetic inductance is obtained:

[6] In the case of an adiabatic drive signal as assumed above, the voltage is due to a temporal gradient of the phase of the order parameter Ψ. The local instantaneous electric field may be finite but, averaged over a cycle of the driving signal, it is equal to zero; thus there is no dc voltage. Then according to any relaxation-time model of the response of the charge carriers, the ac current and ac voltage will be out of phase by exactly $\pi/2$, which suggests a purely "inductive impedance" of the superconductor. Thus, the kinetic inductance results from the frequency-independent phase lag between the components of the ac probe signal I and V induced by the purely inertial response of the charge carriers to a sinusoidal drive.

$$L_K = \frac{\mu_o \lambda^2}{2}\left(\frac{l_d}{A_{cs}}\right)\left(\frac{1}{\left(1-3k^2\xi^2\right)}\right). \tag{2.40}$$

Now, at $I = I_c$, it is the case that $k = k_c$, a critical wave number, which corresponds to:

$$\frac{p_c}{\hbar} = \frac{m_e v_c}{\hbar}, \tag{2.41}$$

with

$$p_c = \frac{\hbar}{\sqrt{3}\xi}, \tag{2.42}$$

which is the critical momentum of the superconductor. Combining these equations, it follows that:

$$k_c \xi = \frac{1}{\sqrt{3}}, \tag{2.43}$$

so that

$$L_K\left(I_c\right) \to \infty. \tag{2.44}$$

Thus, the kinetic inductance of a nanowire should diverge at the critical current, I_c. In order to solve for the current dependence explicitly, the equation for *I(k)* must be inverted so as to express *k(I)*. This can be easily done numerically. A plot of a prediction for $L_K(I_d)$ for a NbN nanowire with l_d = 105 μm, w_d = 130 nm, d_d = 5 nm, T_c = 10 K, and I_c = 26.2 μA is seen in Figure 2.4

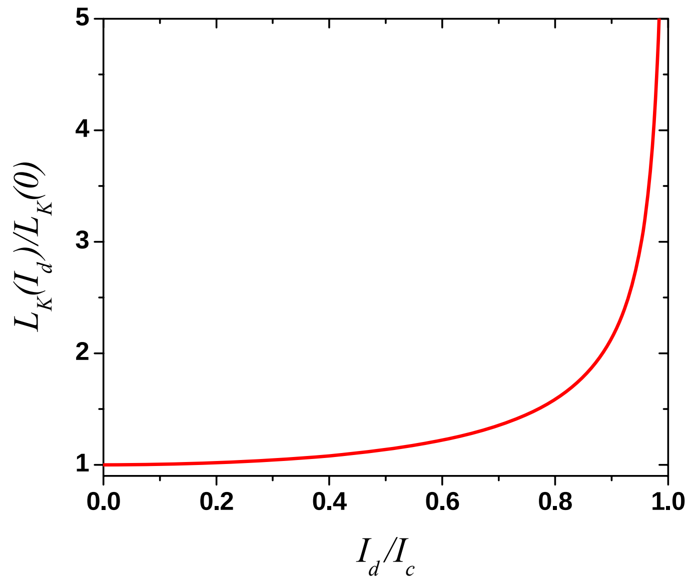

Figure 2.4: Theoretical prediction for the kinetic inductance of a NbN nanowire for l_d = 105 μm, w_d = 130 nm, d_d = 5 nm, T_c = 10 K, and I_c = 26.2 μA.

To obtain an analytical approximation for the current dependence of L_K for device currents near the critical current, a Taylor series expansion of *I(k)* near I_c can be used. First, express the critical current in terms of parameters that have been defined:

$$I_c(T) = \frac{4}{3}\frac{en_e(T)\hbar}{m_e}\frac{A_{cs}}{\xi}. \tag{2.45}$$

Then, taking the following result from Tinkham (1996):

$$\frac{I}{I_c} = \frac{3\sqrt{3}}{2}\left(k\xi - k^3\xi^3\right), \tag{2.46}$$

and expanding it in a Taylor series, the following expression is obtained:

$$\frac{I}{I_c} = 1 + \frac{1}{I_c}\left(\frac{dI}{dk}\right)_{k_c}\left(k - k_c\right) + \frac{1}{I_c}\frac{1}{2}\left(\frac{d^2 I}{dk^2}\right)_{k_c}\left(k - k_c\right)^2 + 0\left(k^3\right). \tag{2.47}$$

The linear term → 0, since $I = I_c$ is a local maximum. In this case,

$$\frac{I}{I_c} = 1 + \frac{1}{I_c}\frac{1}{2}\left(\frac{d^2 I}{dk^2}\right)_{k_C}\left(k - k_c\right)^2 + 0\left(k^3\right). \tag{2.48}$$

Solving for the second derivative and substituting, it follows that:

$$\frac{I}{I_c} \approx 1 - \frac{9}{2}\xi^2\left(k - k_c\right)^2. \tag{2.49}$$

Defining $\Delta k = (k - k_c)$, this equation can be inverted to solve for Δk:

$$\Delta k \approx \frac{\sqrt{2}}{3\xi}\left(1 - \frac{I}{I_c}\right)^{\frac{1}{2}}. \tag{2.50}$$

Now, returning to equation (2.40) for L_K, for k near k_C it follows that:

$$1 - 3k^2\xi^2 \approx 1 - 3\xi^2\left(k_c - \Delta k\right)^2, \tag{2.51}$$

$$\approx \frac{6}{\sqrt{3}}\xi\Delta k. \tag{2.52}$$

Then, combining with the above, the following expression for $L_K(I)$ for $I_b \approx I_c$ is obtained:

$$L_K \approx \mu_o \lambda^2 \left(\frac{l_d}{A_{cs}}\right)\left(\frac{1}{2\left(\frac{2}{3}\left(1 - \frac{I}{I_c}\right)\right)^{\frac{1}{2}}}\right), \text{ (SI)} \tag{2.53}$$

$$L_K \approx \lambda^2\left(\frac{\pi}{c^2}\right)\left(\frac{l_d}{A_{cs}}\right)\left(\frac{1}{2\left(\frac{2}{3}\left(1-\frac{I}{I_c}\right)\right)^{\frac{1}{2}}}\right). \quad \text{(CGS)} \tag{2.54}$$

Therefore:

$$\frac{L_K(I)}{L_K(0)} \approx \left(\frac{1}{2\left(\frac{2}{3}\left(1-\frac{I}{I_c}\right)\right)^{\frac{1}{2}}}\right). \tag{2.55}$$

Although this was derived somewhat differently in this thesis, it matches the result from Anlage (1989). An expression can also be derived for the current dependence of L_K at low values of the bias current. Beginning with the expressions (2.40) and (2.46) derived above:

$$L_K = \mu_o\lambda^2\left(\frac{l_d}{A_{cs}}\right)\left(\frac{1}{\left(1-3k^2\xi^2\right)}\right), \tag{2.40}$$

$$\frac{I}{I_c} = \frac{3\sqrt{3}}{2}\left(\xi k - k^3\xi^3\right), \tag{2.46}$$

when $I << I_c$ it follows that I/I_c can be approximated as:

$$\frac{I}{I_c} \approx \frac{3\sqrt{3}}{2}\left(\xi k\right). \tag{2.56}$$

Solving for ξk and substituting into the expression for L_K and simplifying, the following equation is obtained:

$$L_K \approx \mu_o \lambda^2 \left(\frac{l_d}{A_{cs}} \right) \left(\frac{1}{\left(1 - \frac{4}{9} \frac{I^2}{I_c^2} \right)} \right). \tag{2.57}$$

Using the binomial expansion for $(4/9)(I/I_c)^2 << 1$, this becomes:

$$L_K(I) \approx \mu_o \lambda^2 \left(\frac{l_d}{A_{cs}} \right) \left(1 + \frac{4}{9} \frac{I^2}{I_c^2} \right), \quad \text{(SI)} \tag{2.58}$$

$$L_K(I) \approx \lambda^2 \left(\frac{\pi}{c^2} \right) \left(\frac{l_d}{A_{cs}} \right) \left(1 + \frac{4}{9} \frac{I^2}{I_c^2} \right). \quad \text{(CGS)} \tag{2.59}$$

Therefore:

$$\frac{L_K(I)}{L_K(0)} \approx \left(1 + \frac{4}{9} \frac{I^2}{I_c^2} \right), \tag{2.60}$$

This, again, matches the result from Anlage (1989), which was derived somewhat differently.

2.3.4 Estimates of Magnetic and Kinetic Inductance in a Nanowire

In this section, typical values for the magnetic inductance and the kinetic inductance of a nanowire will be estimated. For zero current and at zero temperature (where kinetic inductance is smallest), the following equations will be used:

$$L_M \approx \frac{\mu_o}{2\pi} l_d \left[\ln\left(\frac{2l_d}{r}\right) - 1 \right], \tag{2.13}$$

$$L_K \approx \frac{\mu_o \lambda_o^2}{2} \left(\frac{l_d}{A_{cs}} \right), \tag{2.61}$$

where (2.61) is equivalent to (2.23). Using $l_d = 10^{-4}$ m, $\lambda_o = 2\times10^{-7}$ m, $A_{cs} = 5 \times 10^{-16}$ m^2, and $r = (A_{cs}/r)^{-1/2}$, which are typical values for an SNSPD fabricated from a disordered superconducting film, $L_M = 1.7 \times 10^{-10}$ H and $L_K = 2.3 \times 10^{-8}$ H. Thus, the magnetic inductance is negligible in all cases for SNSPDs. Finally, it is instructive to calculate a typical kinetic inductance per unit length for comparable Nb and NbN nanowires at finite temperatures (from equation 2.61 with $\lambda(T_o) \rightarrow \lambda_o$). For $T_o = 1.7$ K, $A_{cs} = 7.5 \times 10^{-16}$ (Nb) and $A_{cs} = 5 \times 10^{-16}$ (NbN), $\lambda(T_o) = 3\times10^{-7}$ m (Nb) and $\lambda(T_o) = 5\times10^{-7}$ m (NbN), the inductance for Nb is $L_{K,Nb} = 0.27$ nH/μm and for NbN is $L_{K,NbN} = 1.2$ nH/μm.

2.4 Detection Efficiency Limitations

The detection efficiency is the probability that a photon will form a resistive hotspot if it is incident on the area of the detector. There are generally two categories of effects that reduce this probability. First, there are effects that are external to the nanowire, which primarily affect the coupling of the photon to the SNSPD. Second, there are effects that are internal to the nanowire, which reduce the probability that a resistive hotspot will form even if the photon is absorbed. In this section, external limits to the detection efficiency are first discussed, followed by internal limits that prevent a resistive hotspot from forming.

2.4.1 External Limits to the Detection Efficiency

A photon that is incident on the area of the detector must first be absorbed by the superconductor before it can generate a hotspot. The probability of absorption depends on the optical properties of the thin film as well as the geometry of the nanowire. In this discussion, the probability that an incident photon will be absorbed is calculated.

There are generally two methods for coupling electromagnetic waves to a nanowire: with an absorbing pad of area > λ^2, or via an antenna with poles of length ~$\lambda/2$. Antenna coupling is needed if the nanowire itself is much shorter than λ, and is also useful if strong polarization dependence is acceptable or desired. At optical frequencies, however, lithography constraints and resistive losses make high efficiency antenna coupling difficult (Santavicca 2009). In SNSPD detectors, the nanowire is typically patterned into a pixel of area > λ^2 that approximates an absorbing pad. An example is seen in Figure 2.5; the light colored material is the metal (Nb), while the sapphire substrate is darker. Besides being a non-resonant design with large area, the absorption probability in this meander geometry also has a relatively strong polarization dependence (Anant 2008) which is useful in some applications.

For optical modes with the electric field polarized in the preferential direction (parallel to the meander strips), the maximum possible absorption probability for a single photon per unit area is a function of the fill factor and the absorption probability of the unpatterned film. The fill factor, σ, is the fraction of the area of the detector, A_d, that is actually covered by the nanowire strips and can therefore absorb photons. For all detectors studied in this thesis, $\sigma = 0.5$. Since the films used to pattern SNSPDs are metallic (reflective) and very thin, with d_d ~5 nm, the absorption probability for a single photon in the unpatterned film, α, will be much less than unity. The absorption probability per for the SNSPDs studied in this thesis is therefore equal to $\alpha\sigma$.

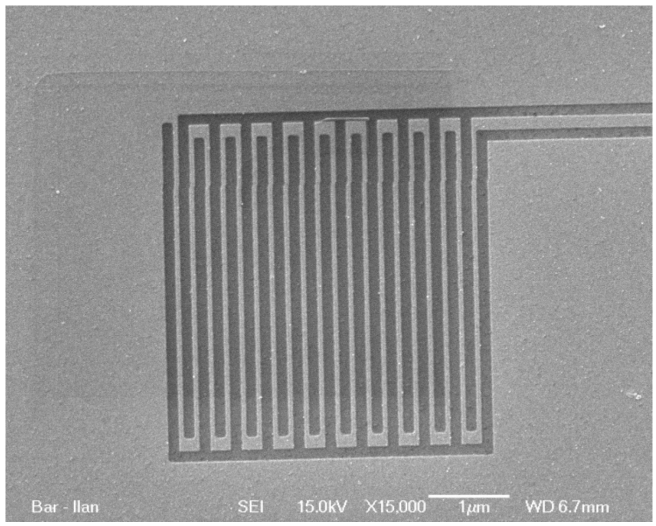

Figure 2.5: Scanning electron micrograph of a meandered Nb SNSPD with detection area A_d = 16 μm^2. The light colored material is Nb, while the sapphire substrate is darker. The shift is due to a stitching error, which was corrected in later fabrication iterations. An explanation of stitching errors is contained in chapter 3.

Using the results of Fowles (1989), α can be computed for a continuous sheet of Nb using the complex indices of refraction for air, niobium, and sapphire. Sapphire is the substrate for all samples. Using $n_{Nb} \approx 2.70 - 3.0i$ and n_{sapp} = 1.75, and n_{air} = 1.00 (Carrol 1982 and references therein)[7] and multiplying the calculated values of α by σ = 0.5, the

[7] This value of n_{NB} is for a wavelength of 690 nm; n_{Nb} changes little (~10%) for wavelengths from 470 nm through 690 nm so it is assumed that this calculation is equally accurate throughout the visible range. It is worth noting that this calculation is only an approximation because the granularity and thinness of the niobium films tested in this thesis probably alters the value of n_{Nb} from the value quoted here, which is for

probability that a photon will be absorbed within an SNSPD of thickness d_d is plotted versus d_d in Figure 2.6. This would be equal to the detection efficiency if every photon that was absorbed were to create a resistive hotspot. This is not generally the case, however, therefore this plot should be interpreted as the maximum possible detection efficiency for a Nb SNSPD of thickness d_d. Furthermore, if non-preferentially polarized photons are incident on the detector, the detection efficiency will be further reduced.[8]

In NbN SNSPDs, the absorption probability has been calculated for $d_d = 5$ nm and $\sigma = 0.5$ by Yang (2009) to be approximately 24% and measured by Anant (2008) to be 21% for a wavelength of 1550 nm polarized parallel to the meander strips. In Yang (2009), it was further shown that by adding a simple half-wavelength resonant cavity and an antireflection coating, the probability of absorption in NbN SNSPDs was increased from 0.21 to 0.5. It is likely that a similar enhancement could be attained for Nb or any other SNSPD. In this thesis, all NbN devices are patterned from ~5 nm thick films. Nb devices range in thickness from ~7.5 nm to 14 nm.

---

bulk niobium. In the literature (e.g., several references within Carrol 1982), the absorption in niobium for infrared wavelengths (> 700 nm) may depend more strongly on the wavelength because of band structure effects in a perfectly crystalline bulk sample, however in the granular, disordered thin films used in this thesis, it is believed that this is less likely to be a significant factor through wavelengths of 1550 nm.

[8] In the case of photons polarized perpendicular to the meander strips, the absorption probability is reduced by approximately a factor of 2.1, as measured by Anant (2008). For randomly polarized incident photons, if an average polarization angle of $\pi/4$ is assumed, and the meander structure is assumed to act like a wire grid polarizer as found by Anant (2008), then the absorption probability would be reduced by a factor of approximately $1+1.1\cdot\cos(\pi/4) = 1.78$

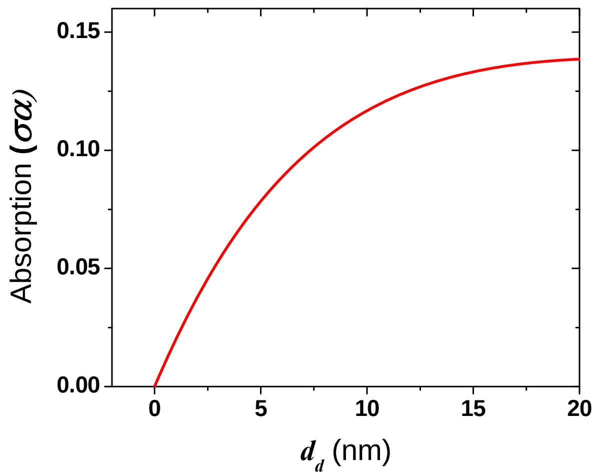

Figure 2.6: Probability of absorbing a photon ($\sigma\alpha$) as a function of SNSPD film thickness for a photon wavelength of 690 nm and fill factor $\sigma = 0.5$.

2.4.2 Internal Limits to the Detection Efficiency

A low absorption probability is a significant limit to the detection efficiency of an SNSPD. This absorption probability can be significantly increased, however, by incorporating the SNSPD into an optical resonant cavity. Once a photon is absorbed, however, it still may not create a resistive hotspot. It has been shown that the probability that an absorbed photon will form a resistive hotspot depends on the bias current, the operating temperature, and the energy of the photon to be detected (Semenov 2003, Korneev 2004, Engel 2004, Semenov 2008, Kerman 2007). This will be analyzed in detail in chapter 5 for measurements of Nb and NbN SNSPDs.

Despite experimental measurements from several groups showing the dependence of the detection efficiency on bias current, temperature, and photon energy, the initial formation of the resistive hotpsot is still not completely understood. Several models for the formation mechanism that are consistent with these experimental results have been presented (Kadin 1996, Semenov 2001, Semenov 2003, Engel 2004, Semenov 2008, Semenov 2009). In discussing these models, it is useful to distinguish between two stages of hotspot growth. The "formation stage" of the hotspot begins when a photon is first absorbed. This occurs before there is any Joule heating from the bias current. At the end of the formation stage, a finite device resistance has developed: $R_d > 0$. The second stage is the "heating stage", which begins when R_d becomes finite. The growth stage is dominated by Joule heating from the bias current. The evolution of the hotspot under Joule heating is discussed in section 2.5 and chapter 6.

The hotspot formation stage begins when a photon is absorbed in the nanowire. According to quantum mechanics, a photon will couple to a single electron, elevating the energy of that electron by $hf \sim 1$ eV, where f is the frequency of the absorbed photon and h is Planck's constant. When this occurs, a thermalization process via electron-electron interactions occurs on a timescale given by $\tau_{e\text{-}e}$ (see section 2.2.2). This allows the energy hf to be spread and break many Cooper pairs creating an expanding hotspot of excess quasiparticles. [9] An illustration of this hotspot is seen in Figure 2.7. Whether this hotspot will have enough energy to suppress I_c in that section of the nanowire to a value below I_b will depend on I_b, I_c, the operating temperature, and the energy of the photon. While

[9] In the absence of any energy loss to the substrate or to phonons in the metal film, the number of extra quasiparticles created is simply $hf/\Delta(T_o)$ where $\Delta(T_o) \approx \Delta(0) = 1.76k_bT_c$. Since thermalization occurs on a much shorter time scale ($\tau_{e\text{-}e}$) than the electron-phonon interaction ($\tau_{e\text{-}ph}$), this is a reasonable approximation.

experimental studies agree , there is still debate in the literature as to which microscopic theory best describes the hotspot formation stage.

There are three general qualitative microscopic models of the hotspot formation process. At the time of this writing, there is no conclusive experimental evidence to determine which of the three is the correct microscopic physical picture. Furthermore, there are no accurate quantitative models of the hotspot formation process. Such a model would be able to predict the dependence of the detection efficiency on both current and temperature. The qualitative models are:

1) Hot quasiparticles spread quickly across the width of the wire so that the critical current density, j_c, is suppressed across the entire wire until the uniform bias current density, j_b, exceeds j_c in a certain location where the temperature is highest. After j_c is exceeded in a certain location, Joule heating causes the normal region to expand.

2) The suppression of j_c and the temperature rise is localized in a hotspot with radius significantly less than the width of the wire. This pushes the supercurrent outside the hotspot, leading to a momentary non-uniform device current density $j_d(x,y)$ that increases in the regions adjacent to the hotspot until j_c is exceeded in the side regions (referred to as "sidewalks" in the literature) that have not been heated. This model is espoused by Semenov *et al.* (2001, 2003) and Engel *et al.* (2004). An illustration of hotspot formation under this model is seen in Figure 2.8.

3) The critical current density is not suppressed below the value of j_b anywhere in the hotspot after it thermalizes initially, but a temperature fluctuation combined with the increase in temperature due to heating from the photon can cause j_c to momentarily dip

below j_b. This will nucleate either (a) a phase slip of the superconducting order parameter (a one-dimensional model) (see Tinkham 1996) or (b) the unbinding of vortex-antivortex pairs (a two-dimensional model) (Semenov 2008). Either (a) or (b) will lead to a localized resistance that can be sustained and expanded by Joule heating.

Both models (1) and (2) would suggest a rigid cutoff for detection at some threshold current, for a given photon energy and operating temperature. However, this is never seen in SNSPDs. In fact, resistive hotspot formation exists even at very low currents, and the cutoff with current is exponential but not particularly sharp. This can be interpreted as evidence of the validity of model (3), particularly at low values of the bias current. However, for sufficiently high photon energies, the detection efficiency does not depend strongly on the bias current, suggesting that models (1) or (2) may be valid for high energy photons. Semenov (2008) suggests that model (2) is valid for photon wavelengths larger than approximately 800 nm, while model (3, two-dimensional) is valid for longer wavelengths. The evidence for model (2) over model (1), however, seems based mostly on early modeling work that assumed this was the picture. In fact, it is difficult to determine if model (1) or (2) is more appropriate.

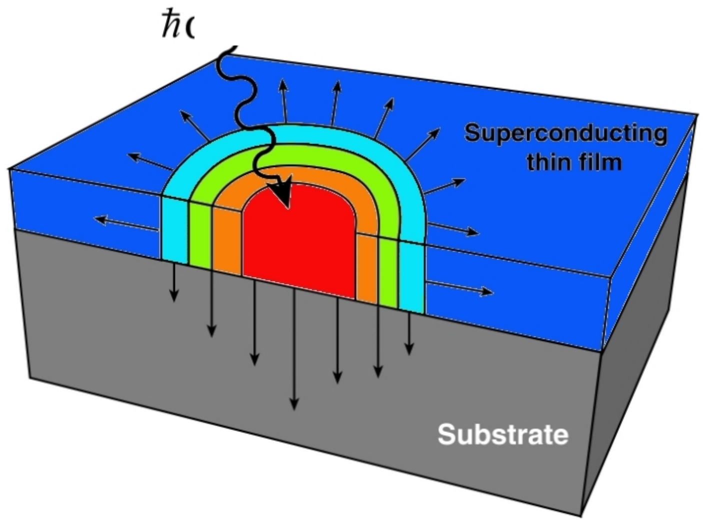

Figure 2.7: Illustration of a photon-induced hotspot in a superconducting strip. Here, red is the highest temperature region, with the temperature cooling with distance from the center of the hotspot. In the blue region, the temperature is equal to the operating temperature, T_o. From Quaranta (2008).

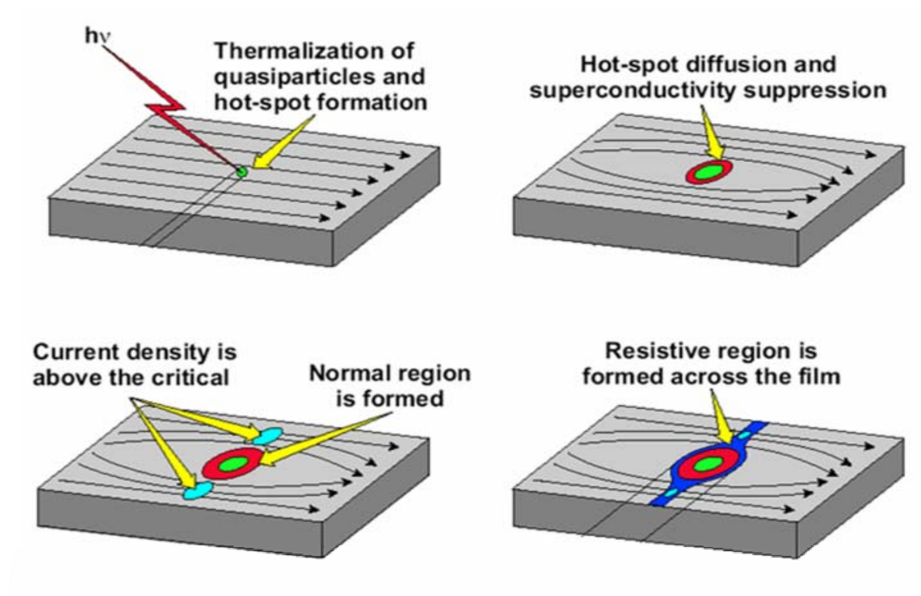

Figure 2.8: Illustration of hotspot formation under model (2). From Semenov (2001).

2.5 Limits to the Photon Energy and Number Resolution

In practical operation, SNSPDs are not sensitive to the energy of an absorbed photon or the number of photons that are absorbed in a given detection event.[10] This is because the photon initiates a cascade of heating where the total energy dissipated is much greater than the energy of the photon.[11] Typically, the voltage signal measured across the readout circuit load has a maximum value of $V_{L,max} \approx I_b \times R_L$, with the overall shape of the pulse set by the readout circuit (to be explained further in section 2.6). This is the case because the hotspot resistance is usually much greater than R_L, so almost all of the bias current is shunted into the load during a detection event. The size of the hotspot depends on the energy dissipated into it; the energy comes from the photon itself ($E_o = hf$) as well as Joule heating via the bias current. The total amount of energy dissipated by Joule heating is approximately $E_{dis} \approx \frac{1}{2} L_K \cdot I_b^2$. The Joule energy is usually much larger than the energy of a photon, so that changes in the energy of the photon or the number of photons absorbed (number of hotspots in series within a long nanowire) only influence the maximum signal size slightly. Furthermore, in practice any small dependence of the signal on the photon energy is washed out by noise in the measurement circuitry.

---

[10] Semenov (2008) reports a mechanism for intrinsic energy resolution for near infrared photons (0.8-1.2 eV) though these measurements have not yet been corroborated by others.

[11] This amplification of the energy of the photon is qualitatively analogous to the type of amplification that occurs in a secondary electron cascade in a photomultiplier tube and in the avalanche effect in a semiconductor photodiode. In all three cases, the cascade nature of detection washes out any sensitivity to the original energy of the perturbing photon. Furthermore, Kadin (1990) shows that there is an exact analog between avalanche photodiodes and superconducting photodetectors.

2.6 Count Rate and Reset

In this section, the reset of an SNSPD after a photon is detected is discussed. First, the timescales that define this reset are introduced. This is followed by the explanation of a simple model for electro-thermal reset that works well in transition edge sensors. Although this model fails to capture the full dynamics of SNSPDs, it is instructive to first study it. Finally, an accurate model of the electrothermal dynamics in SNSPDs is introduced. This model is based on work in the literature as well as work that was done as part of this thesis.

2.6.1 Kinetic Inductance Limit

As explained in section 2.1, in a properly functioning SNSPD (non-ideal functioning is discussed next) the reset time of the detector is determined by the current return time, $\tau_r = L_K/R_L$, which sets the rate of return of the current from the load back into the device once the resistive hotspot has cooled enough to transition back to the fully superconducting state (Kerman 2006). As explained in section 2.2, the bias current must be very close to the critical current in order to achieve high detection efficiency. Thus, after the hotspot has returned to the zero-resistance state, the device will only be ready to detect another photon with high probability after the device current has returned to within ~$0.95I_b$; this takes a reset time of approximately $3\tau_r$, giving a maximum "high efficiency" count rate, C, of approximately $(3\tau_r + 3\tau_t)^{-1} \approx (3\tau_r)^{-1}$, since $\tau_t = L_K/(R_d(t) + R_L)$ and $R_d >> R_L$ for almost the entire time the hotspot is resistive. If photons arrive at a higher rate than this, the detection efficiency of the SNSPD will decrease from the optimum value discussed in section 2.4 because the device current will not have enough time to return to near I_c before another photon arrives.

An obvious way to decrease the reset time, $3\tau_r$, and therefore increase the count rate, is to decrease L_K or increase R_L so that τ_r decreases. Decreasing L_K can be accomplished by decreasing the total length of the nanowire, however this also reduces the detection area: $A_d \propto l_d \propto L_K$. It can also be accomplished by using another material with lower intrinsic inductivity, such as Nb instead of NbN. This was a significant motivation behind this thesis work (Annunziata 2006, 2009). Decreasing L_K has the added advantage of decreasing the total Joule power dissipated during the detection event, which is important and is discussed in detail in the next section. Increasing R_L can be accomplished simply by using a series resistor between the device and the readout circuit input, as in Yang (2007) and Kerman (2009). However, there is a limit to how short the reset time can be. If τ_r is made too short by either reducing L_K or increasing R_L, an SNSPD will not reset at all. Instead, it will latch into a finite voltage state, where the resistive hotspot is sustained by Joule self-heating from the bias current. In this state, the SNSPD is insensitive to photons. This phenomenon is referred to as "latching" in the literature (Yang 2007, Annunziata 2009, Kerman 2009). In general, latching occurs when the current return time, τ_r, is too fast compared to the cooling time of the resistive hotspot, τ_c. Studies of latching in Nb and NbN have been a significant component of this thesis research and will be discussed in detail in chapter 6.

2.6.2 Thermal Relaxation Limit

As discussed above, if τ_r is too fast, the SNSPD will no longer self-reset but will latch into a stable, finite resistive state where it is insensitive to photons. The minimum value of τ_r at which a device will still reset properly depends on the properties of that device. These properties include the parameters of the material, which include the thermal time constants discussed in section 2.2, the resistivity, and the inductance, as well as the device geometry. Latching also depends on the readout circuit load and the bias current. Thus, latching is a complicated phenomenon that must be modeled based on all of these parameters. A model for the dynamics of SNSPDs that accurately reproduces latching is discussed in section 2.6.4.

Thus far, the device operation as outlined in section 2.1 assumes that the cooling time for the hotspot, once most of the current has been shunted into the readout circuit load and Joule heating has nearly ceased, is much shorter than τ_r. This cooling time, τ_c, is not always much less than τ_r, however. When τ_r is reduced to a value less than τ_c, latching occurs. Thus, in order to reduce the reset time (= 3 τ_r) and therefore increase the count rate, the cooling time, τ_c, must be reduced. The cooling time depends not only on the energy dissipated in the hotspot, $E_{dis} \approx ½L_K \cdot I_b^2$, but also on intrinsic material parameters. As is clear in the table in section 2.2.6, the important relaxation times associated with cooling are the electron-phonon time, $\tau_{e\text{-}ph}$, the phonon escape time, τ_{esc}, and the diffusion time, τ_D. NbN has a much shorter electron-phonon time than Nb (Table 2.1), which decreases τ_c. However, it also has larger kinetic inductance (calculated in section 2.3.4 and measured in chapter 7) and in general also a higher critical current, which necessitates a higher bias current. Thus, significantly more Joule energy is dissipated in NbN SNSPDs. A greater amount of dissipated energy increases τ_c. Thus, it is not obvious that a NbN SNSPD would have a shorter reset time for self-resetting operation than a comparable Nb SNSPD, although in general it does. This is discussed in chapter 5.

Since the current return time, τ_r, and with it the reset time ($=3\tau_r$), cannot be reduced to values less than τ_c without having the device latch, a thorough understanding of what determines τ_c is necessary. This is investigated in the next two sections.

2.6.3 The Hot Electron Bolometer as a Model of Thermal Relaxation in an SNSPD

In this section, a simple model that is useful for a hot electron bolometer is used to determine the input parameters of τ_c. A hot electron bolometer is a type of transition edge radiation sensor (see section 1.2) that couples power directly into the electron system of a superconducting microbridge.[12] This microbridge is dc voltage biased in the middle of its superconducting transition, such that the device resistance, R_d, is in a dc-stable, intermediate range between zero resistance and the full normal state resistance: $0 < R_d < R_n$. In this "intermediate" state, the electron temperature is approximately equal to the superconducting critical temperature, $T_e \approx T_c$. A small change in the power absorbed by the electrons will heat them up slightly, creating a large increase in R_d. If a large output bandwidth is desired, as is often the case, the change is resistance is read out by

[12] Dissertations on Nb hot electron bolometers have been written by Reese (2006) and Santavicca (2009).

measuring the voltage with an amplifier with input impedance matched to the approximate resistance of the bridge in the intermediate state, ~50 Ω. In this case, the device can respond to changes in the input power that are as fast as the rate at which the hot electrons cool. The cooling time is set by the electron-phonon and phonon-escape times. Since absorbed power only raises the temperature of the electron system slightly, it is always the case that $T_e \approx T_c$. Thus, $\tau_{e\text{-}ph} \approx \tau_{e\text{-}ph}(T_c)$. Since the phonon escape time, τ_{esc}, is approximately independent of temperature, the total the cooling time is simply $\tau_c = \tau_{e\text{-}ph}(T_c) + \tau_{esc} \approx \tau_{e\text{-}ph}(T_c)$, since $\approx \tau_{e\text{-}ph} >> \tau_{esc}$ in Nb. In NbN, $\tau_c \approx \tau_{e\text{-}ph}(T_c) + \tau_{esc}$ since $\tau_{e\text{-}ph}(T_c) \sim \tau_{esc}$ (Gousev 1994).

In an SNSPD, the temperature change that results from a photon-created hotspot is large because the change in Joule heating is much greater in an SNSPD than in a hot electron bolometer. In an SNSPD, the electron hotspot ranges in temperature from T_o before a photon is absorbed, to greater than T_c while the hotspot is resistive and back to T_o as the hotspot returns to equilibrium. In this case, $\tau_{e\text{-}ph}(T_e)$ changes in time and therefore an accurate analytical expression for the average cooling time, τ_c, is difficult to obtain. Thus, a bolometer model that treats $\tau_{e\text{-}ph}$ as a constant, while instructive, is not accurate for an SNSPD. To determine an accurate model, the full temperature-dependent dynamics need to be simulated. This is the topic of the next section.

2.6.4 A Model of the Hotspot Dynamics in SNSPDs

Models for NbN SNSPDs have been presented by Semenov *et al.* (1995), Il'in *et al.* (2000), Semenov *et al*. (2009) and Kerman *et al*. (2009), however only Kerman *et al*. (2009) studies the phenomenon of latching in detail. The analysis in Kerman *et al*. (2009) uses a phenomenological model of the heating and cooling of the photon-created resistive hotspot to determine if a device will latch. According to Kerman *et al*. (2009), electrothermal feedback creates a situation where a resistive hotspot is either stable (finite dc resistance) or unstable in the steady state, depending on the dc bias current, I_b, and τ_r. Stable hotspots are not the desired case for a properly operating SNSPD. Solutions to the model in Kerman *et al*. (2009) are obtained analytically by determining the stability of the hotspot under arbitrarily small sinusoidal perturbations. This type of small-signal analysis does not model the actual time-dependent formation and evolution of the hotspot. The predictions of the model in Kerman *et al*. (2009) were fit to data collected from NbN SNSPDs. By varying several of the phenomenological parameters of the model, good agreement between the model predictions and experimental data was obtained. However, some of the phenomenological parameters used in Kerman *et al*. (2009), notably the "hotspot temperature stabilization time," are not clearly connected to the microscopic physical processes – including electron-phonon scattering, phonon-escape and electron diffusion – that govern energy relaxation in superconducting thin films. These physical processes have been instrumental in understanding other non-equilibrium superconducting devices such as hot electron bolometers and transition edge sensors (Gershenzon 1990, Burke 1999). Furthermore, in Kerman *et al*. (2009), the values of other standard parameters that are extracted by fitting, such as the thermal conductivity, appear to be quite different from those of independent measurements(Gershenzon 1990, Gousev 1994 (1), Gousev 1997 (2), Burke 1999, Annunziata 2006, Leoni 2006, Kerman 2007). For example, the results based on fitting in Kerman *et al*. (2009) imply a value of the thermal conductivity in NbN of approximately 0.0017 W/K-m. Direct measurements of the thermal conductivity in NbN, and calculations found in previous work by the author of Kerman *et al*. (2009) based on the Wiedemann-Franz law have obtained $\kappa \approx 0.16$ W/K-m (Kerman 2007). Still, the model in Kerman *et al*. (2009) provides guidance that is useful in understanding trends within the data presented.

In this thesis, a model is developed to analyze latching in Nb SNSPDs based on the fundamental, microscopic physical processes that are well-known to govern thermal relaxation in superconducting thin films and nanowires. Similarly to Kerman *et al*. (2009), it is found that in steady state, a photon-created hotspot can stabilize to either a finite resistance or cool back to zero resistance, depending on I_b, L_K, and R_L . Unlike in Kerman *et al*. (2009), the explanation given in this work of what determines whether or not a hotspot will latch to a steady state finite resistance is based on examining the full dynamics of the electrothermal feedback under a large temperature perturbation. This perturbation comes from the photon and from the initial heating that occurs as the hotspot is forming, which is not accounted for in Kerman *et al*. (2009). We find that this initial heating is essential to determining whether a Nb device will latch or self-reset. The predictions of the present model agree well with measurements of Nb SNSPDs without using any free parameters to fit the measured data. Furthermore, unlike in Kerman *et al*.

(2009), all of the parameters use in the present model are based on independent measurements.

The model used in this thesis was adapted from Semenov *et al*. (1995) to simulate the dynamics of Nb SNSPD detectors. This model assumes that the electron system has a temperature $T_e(t)$, that the phonon system has a temperature $T_{ph}(t)$, and that the substrate is at a constant temperature, T_o. In accordance with the results in section 2.2 of this thesis, the electron-electron internal thermalization time for electrons excited by an absorbed photon, $\tau_{e\text{-}e}$, is assumed to be much shorter than the energy sharing time between electrons and phonons ($\tau_{e\text{-}ph}(T_e)$, $\tau_{ph\text{-}e}(T_{ph}) = (\tau_{e\text{-}ph}(T_{ph}))$, and the time for phonons to escape into the substrate (τ_{esc}). Thus, $\tau_{e\text{-}e}$ is neglected in this model. In this "two temperature" model, the temperature of the electron and phonon subsystems is spatially dependent in two dimensions (the model is for an ultrathin strip of superconductor, which describes the geometry of an SNSPD); thus, $T_e = T_e(x,y,t)$ and $T_{ph} = T_{ph}(x,y,t)$. The spatial diffusion of heat within the electron system is modeled with an electron diffusion constant D_e, and within the phonon system it is modeled with an effective diffusion constant for phonons of D_{ph}.[13] The flow of heat between the electrons and phonons is modeled using $\tau_{e\text{-}ph}$ and $\tau_{ph\text{-}e}$ as coupling parameters. By energy conservation, $\tau_{ph\text{-}e}(T_{ph}) = (C_{ph}(T_{ph})/C_e(T_e))\,\tau_{e\text{-}ph}(T_e)$ where C_e is the temperature-dependent electron heat capacity per unit volume and C_{ph} is the temperature-dependent phonon heat capacity per unit volume (Semenov 1995). The heat added to the electron system via an absorbed photon is

[13] Phonons are not normally thought to diffuse in a three-dimensional crystal because they are bosons and scattering with impurities is rare. In these ultrathin films however, where the thickness is much less than the phonon wavelength, surface scattering is prevalent, allowing a diffusion constant for phonons moving laterally (in-plane) to be defined. In practice, the predictions of this model are rather insensitive to D_e, even when it is effectively zero.

modeled by including the influx of heat from a photon of energy hf that acts to raise the temperature locally such that $\Delta T_e(x=0, y=0, t=0) = hf/(d_d C_e dxdy)$, where $dxdy$ is the photon absorption area and d_d is the film thickness. Joule heating per unit area is modeled using the spatially dependent device current density, $j_d(x,y,t)$, and the temperature and current-density dependent device resistivity, $\rho_d(T_e, j_d)$. Cooling (energy loss to the substrate) is modeled using τ_{esc} as the coupling parameter between the phonon system and the substrate, which is modeled as a thermal reservoir with temperature T_o. An illustration of this model is seen in Figure 2.9.

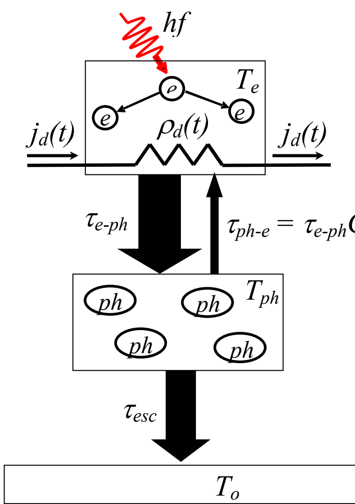

Figure 2.9: An illustration of the two-temperature model used to simulate the dynamics of the hotspot in Nb SNSPDs. In this illustration, "*ph*" indicates a phonon, "*e*" indicates an electron (quasiparticle). Adapted from Quaranta (2008).

This model can be summarized using two coupled differential heat flow equations, one for the electron temperature and the other for the phonon temperature:

$$\dot{T}_e = -\frac{1}{\tau_{e-ph}(T_e)}\left(T_e - T_{ph}\right) + \frac{1}{C_e(T_e)} j_d(x,y,t)^2 \rho_d(x,y,t) + D_e \nabla^2 T_e, \tag{2.62}$$

$$\dot{T}_{ph} = \frac{1}{\tau_{e-ph}(T_e)} \frac{C_e(T_e)}{C_{ph}(T_{ph})} \left(T_e - T_{ph}\right) - \frac{1}{\tau_{esc}} \left(T_{ph} - T_o\right) + D_{ph} \nabla^2 T_{ph} . \tag{2.63}$$

where $T_e = T_e(x,y,t)$ and $T_{ph} = T_{ph}(x,y,t)$. The spatial distribution of the current density, $j_d(x,y,t)$, is determined by the resistivity of the device, $\rho_d(x,y,t)$, which depends on the spatially-dependent temperature and current density in the superconductor . The total current flowing through the device, $I_d(t)$, is determined by the readout circuit. The readout circuit used in this model is the same one illustrated in Figure 2.3, which is a good approximation of the actual measurement circuit (to be discussed in chapter 4). Thus, the equation that governs the total device current is:

$$\dot{I}_d(t) = -\frac{1}{L_K} \left(R_d(t) + R_L \right) I_d(t) + R_L \frac{I_b}{L_K} . \tag{2.64}$$

which is obtained using Kirchhoff's laws. Equation (2.64) can be seen as a type of dynamic boundary condition on j_d. If the coordinate system is oriented such that the positive x-axis is along the length of the nanowire in the direction of current flow, the total device current is given by:

$$I_d(t) = I_d(x,t) = d_d \int_0^{w_d} j_d(x,y,t) dy . \tag{2.65}$$

which, by conservation of charge, must be equal at all points (x). The local resistivity, $\rho_d(T_e(x,y,t), j_d(x,y,t))$, will depend on the whether the point (x,y) in the material is in the superconducting or normal conducting state. In this model, the resistivity will be approximated as:

$$\rho_d(x,y,t) = \rho_o \left[1 - \left[H\left(T_c - T_e(x,y,t)\right) \cdot H\left(j_c - j_d(x,y,t)\right) \right] \right] \tag{2.66}$$

where H is the Heaviside step function and ρ_o is the normal state resistivity of the film. This assumes that the superconducting transition is infinitesimally narrow, and thus there is no intermediate state in this model, unlike in the case of the hot electron bolometer. This also simplifies the calculation of R_d, the total device resistance, since only those sections of the strip at point (x) that are normal for all values of (y) at (x) will contribute to R_d. Thus,

$$R_d(t) = \frac{l_{norm}}{d_d w_d} \rho_o . \tag{2.67}$$

Finally, an important quantity to calculate is the critical current of the device as a function of time, $I_c(t)$. Initially, an absorbed photon triggers the formation of a resistive hotspot which suppresses the critical current in that section of wire to zero. The device critical current is defined as the minimum critical current along the length of the nanowire. It is given by:

$$I_c(t) = \min_x \left[d_d \int_0^{w_d} j_c\left(T_e\left(x,y,t\right)\right) dy \right]. \tag{2.68}$$

The location along the length (x) where I_c is a minimum is also the location where T_e is a maximum. As solutions to this model show, the temperature across a resistive hotspot is fairly uniform. Furthermore, since the critical current is only determined by temperature, once the hotspot begins to cool to below the critical temperature, $I_c(t)$ becomes a measure of the thermal relaxation of the hotspot back to equilibrium. Thus, the time scale over which the critical current returns to its equilibrium value at $T_e = T_o$ is the hotspot cooling time, τ_c.[14]

---

[14] Solving for $I_c(t)$ and using it as a probe of thermal relaxation is unique to the simulations done for this thesis. The advantages of this method will be apparent when examining the results of the simulations.

To investigate the thermal relaxation further, the model defined by equations (2.62-2.64) must be solved for $T_e(x,y,t)$ and $T_{ph}(x,y,t)$. From these electron and phonon temperature functions, the critical current, $I_c(t)$, the device voltage, $V_d(t) = I_d(t)R_d(t)$, as well as other quantities, can be solved for. Equations (2.62-2.64) are coupled, non-linear, and non-separable. Thus, the most accurate solutions necessitate a numerical approach. As part of this thesis work, a numerical solution to these equations has been implemented in Matlab. This computation utilizes established solution algorithms that are part of the Matlab Partial Differential Equation Toolbox as well as customized algorithms that account for the normal-superconductor phase transition. The device was represented by a two dimensional grid with longitudinal grid spacing, Δx and transverse grid spacing, Δy. In all simulations, the grid spacing was less than one tenth of the dimension of the strip in that direction. At each grid point, the electron and phonon temperatures were defined. From these temperatures, all other temperature-dependent quantities are defined for each grid point. When a volume-dependent quantity such as the heat capacity is calculated, it is calculated over the volume of the cell centered at the grid point (x,y) where the volume of the cell is equal to $d_d \Delta x \Delta y$. In these simulations, consistent with measurements in Nb devices, the Ginzburg-Landau expression for the temperature dependence of the critical current was used:

$$j_c(T_e) = j_c(0)\left(1 - \frac{T_e}{T_c}\right)^{3/2}, \tag{2.69}$$

with $j_c(0)$ determined for each device based on measurements (Tinkham 1996). The absorption of a photon is simulated by increasing the temperature of one grid point such that :

$$T_e(x_o, y_o) = T_o + \frac{hf}{C_e(T_o)\cdot(\Delta x \Delta y d_d)}, \tag{2.70}$$

where (x_o,y_o) is the grid point where the photon is absorbed. In this way, the full dynamics of the hotspot evolution are simulated. This is not the case in some analytical models of NbN SNSPDs, namely Kerman (2009), where a perturbative method is used to model the stability of a preexisting hotspot.

Numerical solutions to this two-temperature model can accurately reproduce the signal measured when a single photon is detected. In Figure 2.10, the simulated (dashed, color curves) and measured (solid curves) current through the load resistance is plotted as a function of time for two different values of the bias current, I_b = 5 μA and I_b = 8.1 μA. This data is for a single-photon detection event in a Nb SNSPD with L_K = 235 nH, R_L = 50 Ω, I_c = 8.2 μA, T_c = 4.5 K, and T_o = 1.7 K. The curves for the lower bias, I_b = 5 μA, show a properly resetting device. This is referred to as "self-reset" and is the ideal operating mode of an SNSPD. The three labeled stages correspond to those in Figure 2.3 and those discussed in section 2.1. The curves for the higher bias current, I_b = 8.1 μA, show an example of latching. In general, there is some current, I_{latch}, that when the SNSPD is biased with a dc current $I_b > I_{latch}$, it will not self-reset but will rather latch to a state with finite resistance. Note that the measured latching pulse in Figure 2.10 has more noise because it is a single-shot measurement, while the measured self-resetting pulse is an average. The measured decay of $I_L(t)$ in the latching case for t > 5 ns is due to ac coupling of the amplifier and will be explained in chapter 4.

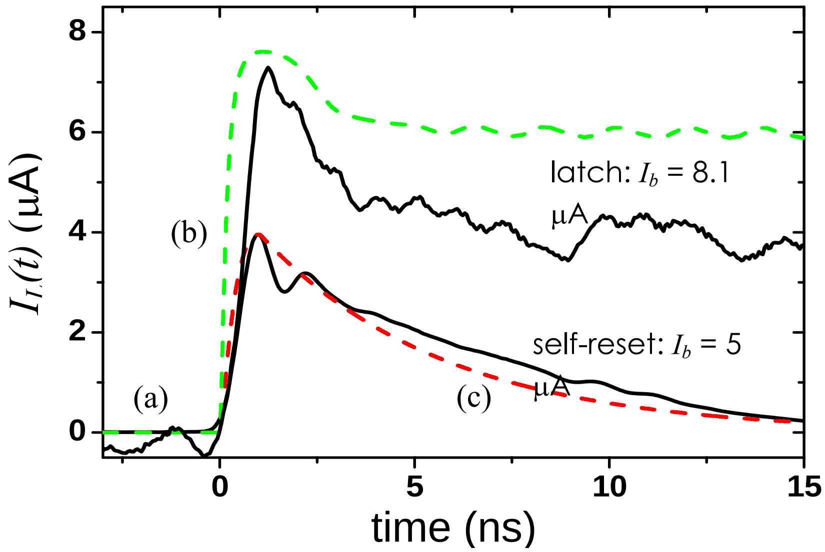

Figure 2.10: Single-photon pulses in a Nb SNSPD with L_K = 235 nH, R_L = 50 Ω, I_c = 8.2 μA, and T_c = 4.5 K, and T_o = 1.7 K. The plot shows measured (solid lines) and simulated (dashed, colored lines) output pulses, $I_L(t)$, with the self-resetting case labeled by the three regimes of operation: (a) the device is in equilibrium with $R_d = 0$; (b) a photon has been absorbed, the hotspot is growing and the current is transferring into the load; (c) the hotspot resistance has returned to zero, and the current is returning to the device with a time constant $\tau_r = L_K/R_L$.

The simulation can predict $I_L(t)$ as well as $I_d(t)$, $I_c(t)$, and $R_d(t)$, which are instrumental to understanding thermal relaxation and latching in SNSPDs. In Figure 2.11, plots of $I_d(t)$, $I_c(t)$, and $R_d(t)$ are shown for the same device as in Figure 2.10 with I_b = 5 μA. In Figure 2.12, the effect of increasing the bias current is shown using simulations of $I_d(t)$ and $I_c(t)$, for the same Nb SNSPD as in Figures 2.10 and 2.11. The bias currents simulated are: I_b = 5 μA, 7 μA, 8.1 μA. For this device and readout load, I_{latch} = 7.1 μA. Figure 2.12 clearly shows the crossover from self-resetting operation to latching. Latching occurs when the cooling of the hotspot and the current return is such that the trajectory of $I_c(t)$, which is set by τ_c, never rises above $I_d(t)$, which is set by τ_r. In this case, the device remains in the resistive state. The crossover from self-resetting operation to latching occurs when $\tau_c \sim \tau_r$ (blue curves, I_b = 7 μA) as is apparent from the graph, however, the simulation predicts that the increase of $I_c(t)$ as the hotspot cools is not a simple exponential function. Thus, it is not governed by a single time constant, τ_c, but by a time constant that increases as time progresses and the hotspot cools (as $T_e(t) \rightarrow T_o$).[15] Thus, $\tau_c = \tau_c(T_e)$. This is due to the inverse temperature dependence of $\tau_{e\text{-}ph}$. It is possible that a similar effect in NbN SNSPDs may explain the surprisingly large thermal relaxation time reported by Kerman (2009) and Ejrnaes (2009). Furthermore, because of the time dependence of $R_d(t)$, the return of the device current is also not a purely exponential decay when R_d is finite. Thus, simulations are needed to predict the exact trajectories of $I_c(t)$ and $I_d(t)$, particularly in the important region when the hotspot first begins to cool below T_c, which is when I_c first begins to increase from zero.

[15] This thesis work does not claim that fitting to a stretched exponential is the most physically relevant way to characterize thermal relaxation in superconducting nanowire, however it is possible that more studies may indeed show this. Stretched exponential decays are those that are $\sim exp[-(t/\tau_c)^\beta]$ with β as the stretching parameter. They are common and have physical meaning in the case of optical energy relaxation in semiconductors or anywhere else where there are relaxation processes that either consist of a distribution of relaxation times, or where relaxation is via multiple scattering events that successively reduce energy, with a longer characteristic relaxation time for each successive relaxation step. See, e.g., Pavesi (1993).

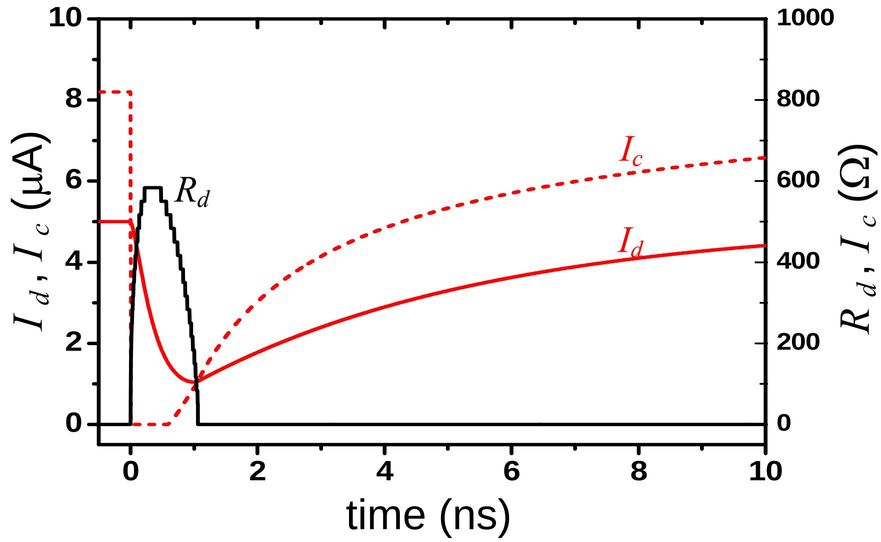

Figure 2.11: Simulations of $I_d(t)$, $I_c(t)$, and $R_d(t)$ for the same Nb SNSPD as in Figure 2.10 with $I_b = 5$ μA .

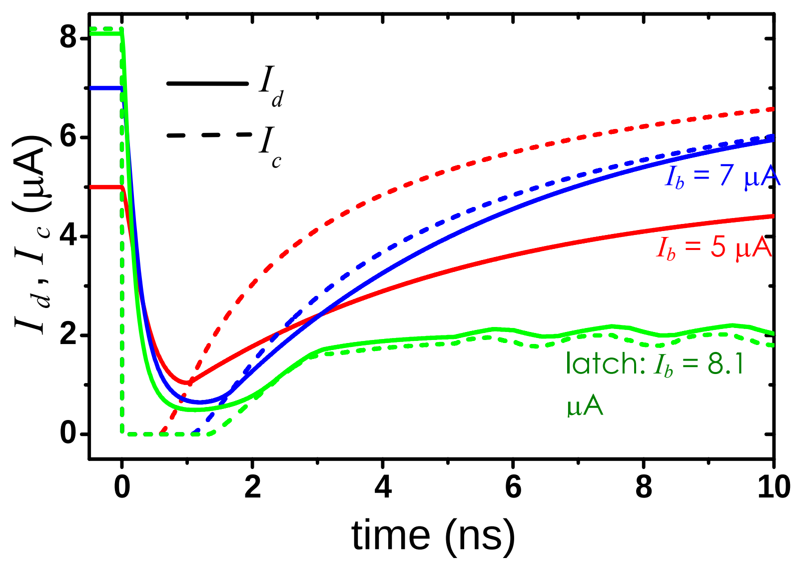

Figure 2.12: Simulations of $I_d(t)$ and $I_c(t)$, for the same Nb SNSPD as in Figure 2.10 for several bias currents: $I_b = 5$ μA, 7 μA, 8.1 μA. This shows the crossover from self-resetting operation to latching. Latching occurs when the heating of the hotspot is such that the trajectory of $I_c(t)$ never rises above $I_d(t)$, and so the device remains in the resistive state. The crossover point occur when $\tau_c \sim \tau_r$, however, the increase of $I_c(t)$ is not a simple exponential function, while the increase of $I_d(t)$ is, making a direct quantitative comparison of the time scales difficult.

Simulations show that the total energy dissipated in the hotspot, E_{dis} (before R_d begins to decrease), can be very large, as discussed in section 2.5. Some typical values for E_{dis} are given below. It is found that $E_{dis} = hf + ½L_K(I_b^2 - I_{min}^2)$, where hf is the energy of the absorbed photon and I_{min} is the minimum value of I_d. Typically $I_{min} << I_b$ and $hf <<$ $½L_K(I_b^2 - I_{min}^2)$ (see section 2.5), so $E_{dis} \approx ½L_K I_b^2$. The simulations show that a larger value of E_{diss} leads to a larger and higher temperature hotspot. A higher temperature increases the average value of τ_c, lengthening the time for thermal relaxation back to the base temperature, T_o. Since E_{dis} is approximately independent of R_L, R_L can be reduced to increase τ_r without affecting τ_c and therefore avoid latching. However, it is desired that τ_r be as short as possible so that the reset time ($3\tau_r$) is as short as possible. Thus, it is desirable to have τ_r be as to close to (for speed) but just above (to prevent latching) approximately the average value of τ_c. Thus, it is desirable to reduce the average value of τ_c: $\langle\tau_c\rangle$. In Fig. 2.13, simulations of $R_d(t)$ and $I_c(t)$ are plotted for Nb SNSPDs using R_L = 25 Ω. As E_{diss} is increased by increasing L_K (device C → device B) or by increasing I_b (device A from I_b = 5 μA → 8.19 μA), $\langle\tau_c\rangle$ also increases. In this plot, it is seen that τ_{th} increases from 2.2 to 3.13 to 3.8 to 5.0 ns as E_{dis} is increased from 5 to 12 to 26 to 85 eV (hf = 2.6 eV, λ = 470 nm). For comparison, for a hot electron bolometer where heating is small, measurements of τ_c in Nb are often made using mixing techniques (Santavicca 2009). These measurements are performed at T_c such that $\langle\tau_c\rangle = \tau_c(T_c) = \tau_{e\text{-}ph}(T_c) = 2.0$ ns. Reducing L_K is advantageous because it reduces $\langle\tau_c\rangle$ by reducing E_{dis}. An important caveat is that when L_K is reduced for a Nb SNSPD, R_L must also be reduced. This is because $\tau_r = L_K/R_L$ can only be reduced (approximately) proportional to the amount that

$\langle\tau_c\rangle$ is reduced and $\langle\tau_c\rangle$ scales sub linearly with L_K because some of the extra energy goes into making the hotpsot geometrically larger as well as hotter. Reducing R_L reduces the output signal: $V_{L,max} \approx I_b \times R_L$ (see section 2.5)

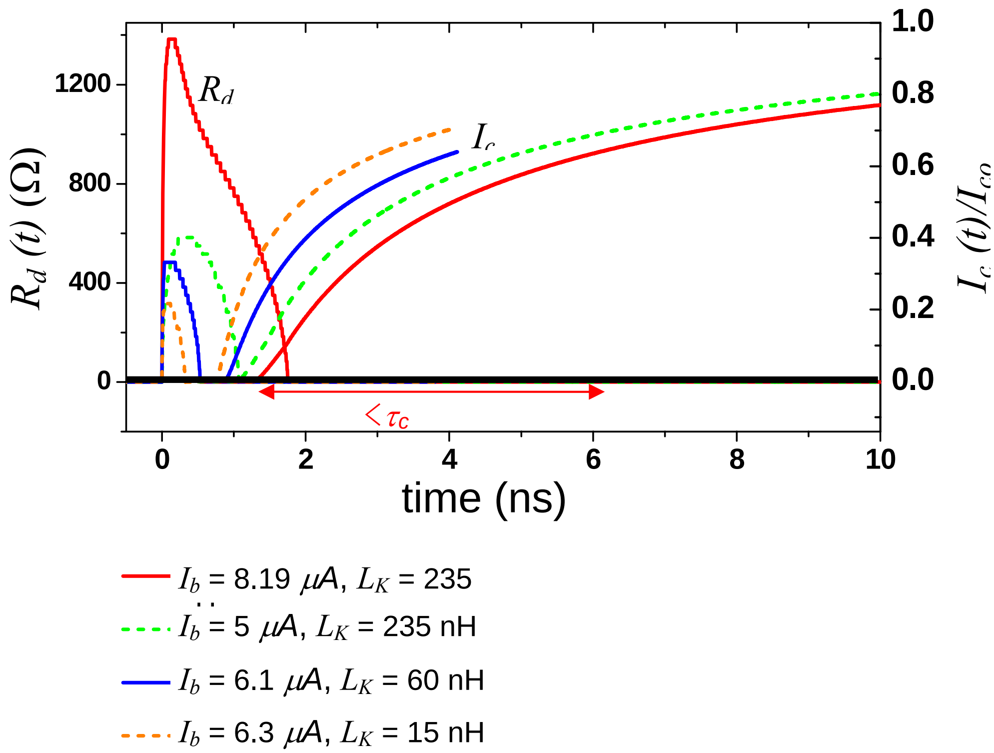

Figure 2.13: Simulations for Nb devices of various lengths ($l_d \sim L_K$) showing the dependence of $R_d(t)$ and $I_c(t)$ on I_b and L_K for $R_L = 25\ \Omega$; $\langle\tau_c\rangle$ is noted for the device with $L_K = 235$ nH with $I_b = 8.19$ μA where it is equal to approximately 5 ns. Note that the device with $L_K = 235$ nH is the same device as in Figures 2.10-2.12.

In addition to L_K and I_b, $\langle\tau_c\rangle$ also depends on material parameters. In particular, $\langle\tau_c\rangle$ decreases when $\tau_{e\text{-}ph}$ and the phonon escape time, τ_{es}, are decreased and when D_e is increased. In Nb, $\langle\tau_c\rangle$ is significantly larger than in NbN for similar sized detectors, even though NbN has larger L_K and I_{co}. This is because NbN has a much stronger electron-phonon interaction than Nb, which more than compensates for the deficiency of having lower diffusivity and generally larger values of E_{dis} (see section 2.2). The model suggests that in a Nb SNSPD, electron-phonon relaxation sets the ultimate limit on the minimum cooling time: $\langle\tau_c\rangle_{\min} \approx \tau_{e\text{-}ph}(T_c) + \tau_{es} \approx \tau_{e\text{-}ph}(T_c)$. This limit would be recovered in the case of very small heating, i.e. when $L_K \rightarrow 0$ so that $E_{dis} \rightarrow 0$. In this case, the bolometer result is recovered. In NbN, $\langle\tau_c\rangle_{\min} \approx \tau_{e\text{-}ph}(T_c) + \tau_{es} \approx 40$ ps, but to operate in this regime, both L_K and R_L would need to be impractically small to avoid latching.

The results of this model will be applied to analyze latching further in chapter 5, where measurements of latching are reported and analyzed.

2.7 Dark Counts

Dark counts are spurious voltage pulses that do not result from the intended absorption of photons. There are two major sources of dark counts: dark counts that are a result of external stimuli, referred to as externally sourced dark counts, and dark counts that are a result of internal stimuli, referred to as internally sourced dark counts. Externaly sourced dark counts can be completely eliminated by proper filtering and shielding; internally sourced dark counts are intrinsic to the device and therefore cannot be easily eliminated.

2.7.1 Dark Counts from External Sources

External dark counts are either due to the SNSPD absorbing spurious photons that are not intended to be absorbed or from electrical noise from the readout circuitry. Photons that are not intentionally directed at the detector may still be absorbed and create a resistive hotspot. These photons may be due to the coupling of stray light into the illuminating optics or to blackbody radiation. Electrical noise may be due to fluctuations in the dc current bias, coupling of room temperature Johnson noise via the readout line, back-action from the amplifier, or electrical reflections due to impedance mismatches in the transmission lines that are used to readout the device. These electrical issues are discussed in chapter 4 and can be eliminated through proper experimental design. Likewise, stray light coupling can also be eliminated. For black body radiation, it is essential that the detection area of the device only be exposed to blackbody radiation from a low temperature emitter (~4 K). As will be shown next, there are very few visible and near infrared photons in a 4 K blackbody radiation spectrum. In a room temperature blackbody radiation spectrum, there are many more photons. An SNSPD is not sensitive to mid-infrared or longer wavelength radiation. Thus, only stray photons from visible room light would lead to the formation of spurious resistive hotspots. These can be eliminated by simply utilizing a dark measurement environment

The number of blackbody photons incident on an SNSPD can be calculated from Planck's law of blackbody emission:

$$\Sigma(f,T) = \frac{2hf^3}{c^2}\frac{1}{e^{\frac{hf}{kT}}-1}, \tag{2.71}$$

where Σ is the power spectral density per unit area per solid angle that the emitter subtends with respect to the surface of the detector, T is the temperature of the emitter, and f is the frequency of the power emitted (Planck 1901). In order to calculate the number of photons per second, it is necessary to integrate equation (2.71) from f_{min} to infinity. Then, assuming the sources of black body emission are the walls of the cryogenic vessel in which the detector is housed, which are at a uniform temperature T, the solid angle that the sources subtend is simply 2π. Then, for a detector of area A_d, the power due to photons with a frequency greater than f_{min} incident on the detector per unit time is given by:

$$\int_{f_{\min}}^{\infty} \frac{4\pi h f^3 A_d}{c^2} \frac{1}{e^{\frac{hf}{kT}} - 1} df \quad . \tag{2.72}$$

By dividing this result by hf_{min}, the number of photons with a wavelength of 2 μm equivalent to the blackbody energy incident on the device in one second can be determined. For T = 4.2 K and A_d = 25 μm^2, the number of equivalent 2 μm photons in this band is approximately 1 per second. If T = 300 K, however, there will be approximately 2.5×10^4 equivalent 2 μm photons incident on the device per second. Therefore, in a properly filtered, dark, 4.2 K experimental environment, externally sourced dark counts are not significant.

2.7.2 Internally Sourced Dark Counts

In practice, the dark count rate of an SNSPD is not negligible when biased with a current just below the critical current even in a dark, well filtered measurement environment. Typical dark count rates for a well filtered, dark measurement environment are 100-1000 dark counts per second for SNSPDs with $I_b > 0.95 I_{co}$. These dark counts are entirely from internal sources. As will be shown, a voltage pulse that results from a dark count is identical to a voltage pulse that is instigated by a photon. From this, it can be inferred that a dark count hotspot experiences a similar amount of Joule heating, along with a similar reset process, to a photon-induced hotspot. Furthermore, the functional form of the dark count rate, Θ, versus bias current is similar to the detection efficiency versus bias current for photons. Thus, whatever mechanism causes a dark counts does so in a process that is similar to how photons are detected.

There have been several studies of dark counts in NbN SNSPDs (Kitaygorsky 2005, Engel 2006, Kitaygorsky 2007, Bell 2007). These studies, as well as studies done for this thesis, suggest that the mechanism for dark counts is that of local resistance fluctuations, which lead to temporary resistive states in the superconducting nanowire when it is biased very close to the critical current. These short-lived resistive states are amplified by Joule heating, undergoing the same electro-thermal dynamics as described in section 2.6 for photon-initiated hotspots. Thus, a model for determining the dark count rate, Θ, need only predict the frequency of occurrence of these temporary resistive states.

The character of resistance fluctuations in nanowires depends strongly on dimensionality. The dimensionality of a superconductor is determined by comparing the two characteristic length scales of superconductivity, the coherence length, ξ, and the magnetic penetration depth, λ_M, to the transverse dimensions of the superconducting strip, w_d and d_d. If ξ and λ_M are each shorter than both w_d and d_d, the nanowire is a one dimensional superconductor. If ξ and λ_M are only shorter than either w_d or d_d then the superconductor is two-dimensional. In Nb, typical values for these parameters at low temperature ($T_o < \sim T_c/3$) in the films tested in this thesis are estimated to be: $\lambda_m(T_o) \approx$ 400 nm[16], $\xi(T_o) \approx 6$ nm based on measurements by Santavicca (2010).[17] Near the critical temperature, these values increase to approximately: $\lambda_m(0.9T_c) \approx 450$ nm[18], $\xi(0.9T_c) \sim 19$ nm[19]. In NbN, estimated values for these parameters at low temperature ($T_o < \sim T_c/3$) in the films tested in this thesis are: $\lambda_m(T_o) \approx 700$ nm[20], $\xi(T_o) \approx 3$ nm.[21] Near the critical temperature, these values increase to approximately: $\lambda_m(0.9T_c) \approx 800$ nm[22], $\xi(0.9T_c) \sim 10$

---

[16] Santavicca measures $\xi(T_o) \approx 11$ nm for Nb films that are approximately a factor of 2.6 less resistive than the Nb films typically used in this thesis. The value presented here is obtained by scaling Santavicca's measurement by $(2.6)^{1/2}$ since the coherence length for a superconductor in the dirty limits is proportional to the mean free path for electrons (Tinkham 1996).

[17] This value of the penetration depth for Nb is substantially higher than is typically reported. This is due to the very disordered nature of Nb films used to make SNSPDs in this thesis. As is explained in chapter 5, very dirty Nb is necessary to make SNSPDs with high detection efficiency. Cleaner Nb measured for this thesis had $\lambda_M(T_o) \approx 200$ nm.

[18] Based on the temperature dependence of the kinetic inductance for Nb.

[19] This is calculated from the value at T_o using the Ginzburg-Landau temperature dependence (Tinkham 1996).

[20] This is calculated based on measurements of the kinetic inductance of NbN SNSPDs, using equation (2.61).

[21] This value of the penetration depth for NbN is substantially higher than reported in much of the literature for NbN SNSPDs (see, e.g., Kerman 2006, Yang 2009), were typically, $\lambda_M(T_o) \approx 400$ nm. This is attributed to the substantially greater resistance, and therefore reduced mean free path, of the NbN studied in this thesis.

[22] Based on the temperature dependence of the kinetic inductance NbN.

nm[23]. For Nb SNSPDs, $w_d \approx 100$ nm, $d_d \approx 7.5$ nm, and for NbN nanowires, $w_d \approx 130$ nm, $d_d \approx 5$ nm. Based on these values, neither Nb nor NbN SNSPDs are perfectly one or two dimensional superconductors. Instead, SNSPDs fabricated from both materials are quasi one-dimensional, with $\lambda_m > w_d$, d_d, but $\xi \sim d_d$ and $\xi < w_d$.

Bartolf, *et al*. (2010) along with Engel, *et al*. (2006) and Bell, *et al*. (2007) discuss this quasi-two dimensionality in detail. It is argued that even though NbN nanowires are nearly one-dimensional, vortices should be able to exist within the strip because $\xi < w_d$, giving the resistance fluctuations a two dimensional character. In two dimensions, resistance fluctuations are due to the flow of vortices. Vortex motion at high bias currents and low temperatures ($T_e << T_c$) may result from either unbinding of vortex-antivortex pairs (Engel 2006) or from vortices that enter the film from the edges of the strip, where the magnetic self-field is greater than H_{c1} for bias currents close to the critical current (Kogan 1994, 2007). This latter effect is referred to as "vortex hopping." Both the unbinding of vortices and vortex hopping are processes that require activation over a potential barrier whose height is given by the difference in free energy between the superconducting state (where the vortex is bound or absent from the film) and the resistive state (where vortices are moving). Calculations that predict the rate of resistance fluctuations are now presented based on calculating the probability that this energy barrier will be exceeded.

The following results are taken from the work of Bartolf *et al*. (2010) and many references therein, as well as from Mooij, *et al*. (1984). In order for a vortex-antivortex

[23] This is calculated from the value at T_o using the Ginzburg-Landau temperature dependence (Tinkham 1996).

pair (VAP) to unbind, a bias current dependent energy barrier, $U_{VAP}(I_b, T)$, must be overcome. The barrier is given by:

$$U_{VAP}\left(I_b, T\right) = \frac{A(I_b, T)}{\varepsilon} \cdot \left[\ln\left(\frac{2.6 I_c(T)}{I_b} \right) - 1 + \frac{I_b}{2.6 I_c(T)} \right], \tag{2.73}$$

where

$$A\left(I_b, T\right) = \frac{h^2}{8\pi e^2} \cdot \frac{d_d}{\lambda_M (T)^2}, \tag{2.74}$$

and where ε is the average polarizability of a VAP. As can be seen, the presence of a current reduces the barrier height. At bias currents near the critical current, thermal fluctuations may cause the barrier to be exceeded occasionally. The probability of this occurring is given by:

$$\Theta_{VAP}\left(I_b, T\right) = e^{-\frac{U_{VAP}(I_b, T)}{k_B T}}. \tag{2.75}$$

The total dark count rate due to the unbinding of VAPs is given by:

$$\Theta_{VAP}\left(I_b, T\right) = \Gamma_{VAP} e^{-\frac{U_{VAP}(I_b, T)}{k_B T}}, \tag{2.76}$$

where Γ_{VAP} is a proportionality constant which gives the attempt rate for overcoming the barrier and includes the details of the geometry of the nanowire. A similar physical situation exists for the case of vortex hopping. In the case of vortices entering the film from the edges, the energy barrier is given by:

$$U_{VH}\left(I_b,T\right)=\frac{h^2}{8\pi e^2\mu_o}\cdot\frac{d_d}{2\lambda_M(T)^2}\left[\ln\left(\frac{2w_d}{\pi\xi(T)}\cdot\frac{1}{\sqrt{1+\left(\frac{I_b}{I_B(T)}\right)^2}}\right)-\frac{I_b}{I_B(T)}\cdot\left[\arctan\left(\frac{I_B(T)}{I_b}\right)-\frac{\pi\xi(T)}{2w_d}\right]\right], \quad (2.77)$$

where

$$I_B(T)=\frac{h}{4e\mu_o I_b}\cdot\frac{d_d}{2\lambda_M(T)^2}. \quad (2.78)$$

As with VAP unbinding, the presence of a current reduces the barrier height so that at high bias thermal fluctuations may cause the barrier to be exceeded. The probability of this occurring is given by:

$$\Theta_{VH}\left(I_b,T\right)=e^{-\frac{U_{VH}(I_b,T)}{k_BT}}. \quad (2.79)$$

The total dark count rate due to vortex hopping is given by:

$$\Theta_{VH}\left(I_b,T\right)=\Gamma_{VH}I_b e^{-\frac{U_{VH}(I_b,T)}{k_BT}}, \quad (2.80)$$

where the attempt rate depends on the bias current, with Γ_{VH} as a proportionality constant which includes the details of the geometry of the nanowire. These equations have been used by Bartolf (2010) to accurately model dark counts in NbN SNSPDs across a wide range of bias currents. In that work, equations (2.75) and (2.78) were fit to data by taking Γ_{VAP} and Γ_{VH} as well as T as free parameters. Modeling the experimental measurements in this way allowed an accurate fit to the data, however Bartolf was unable to distinguish which model, that of vortex-antivortex unbinding or of vortex hopping, was a better fit.

Studies of dark counts have also been done using similar models by Engel (2006), Bell (2007), and Kitaygorsky (2005, 2007) that generally show good agreement with experimental measurements, but again rely on allowing one or more parameters to vary in order to fit the data.

2.8 Jitter

Jitter is the standard deviation in the delay between when a photon is incident on the detector and when an output signal is first detected. Sources of jitter fall into two categories: readout jitter and intrinsic jitter. Readout jitter is simply the uncertainty due to noise and other variation within the measurement electronics. Although readout jitter is a significant concern in practical applications, from a device physics standpoint it is easily mitigated. Intrinsic jitter is jitter that, like internally sourced dark counts, is inherent in the operation of the device. Measurements of the jitter in NbN SNSPDs place an upper bound of approximately 30 ps on the intrinsic jitter (Yang 2009). The cause of this intrinsic jitter, however, is unknown. It may be due to variation in where the photon is absorbed along the width of the nanowire strip. If absorption occurs closer to an edge, it may take a different amount of time for the hotspot to form than if absorption occurs in the center of the strip. The majority of the jitter is not due to variation in where the photon lands along the length, as the difference in propagation delay added by even a ~1 mm delay line (longer than most nanowires that make up SNSPDs) is only 3 ps. In Nb SNSPDs, the total jitter due to both the readout electronics and intrinsic sources is measured and discussed in chapter 5.

Chapter 3

Device Fabrication

3.1 Fabrication Overview

The Nb devices studied in this thesis were fabricated in part at Yale and in part at IBM T. J. Watson Research Center in Yorktown Heights, NY. The NbN devices studied were fabricated by collaborators at the University of Salerno in Salerno, Italy, and at Istituto di Cibernetica in Pozzuoli, Italy. At Yale, Nb film growth and ion beam etching was completed. SEM imaging was done at both locations. In this chapter, the Nb superconducting nanowire single-photon detector (SNSPD) fabrication process is explained in detail. An overview of the NbN SNSPD fabrication process is also given. This process has been explained by Leoni, *et al*. (2006). The chapter concludes with a description of the screening method used to choose the best devices. In all, over a thousand individual Nb SNSPD devices were fabricated for this thesis work. Out of these, a large majority of devices were defective in some way, rendering them unusable. In many instances, these defects were not obvious until electrical testing was performed after most of the fabrication of a particular run of devices was completed.

The Nb SNSPDs were fabricated using electron beam lithography in a process based on subtractive methods using reactive ion etching. A brief overview of the major Nb SNSPD fabrication steps is illustrated in Figure 3.1, where a cross section of a device in various stages of the fabrication process is shown. (a) The Nb film is grown using sputter deposition of Nb onto a sapphire substrate in an ultra high vacuum chamber. (b) Poly(methyl methacrylate) (PMMA) is spin-coated onto the Nb film, and patterned using electron beam lithography. (c) The Nb film is etched using reactive ion etching. (d) The device consists of patterned Nb on a sapphire substrate where all PMMA has been removed using acetone. (e) The patterned device is thinned with an argon ion beam. In order for Nb devices to have high detection efficiency, the Nb film is initially deposited thicker than needed, and is thinned after all patterning is complete. This thinning increases the resistivity of the film and is necessary for high detection efficiency in Nb SNSPDs, as will be shown in section 5.3. In the following sections, each step of the fabrication process is explained in detail.

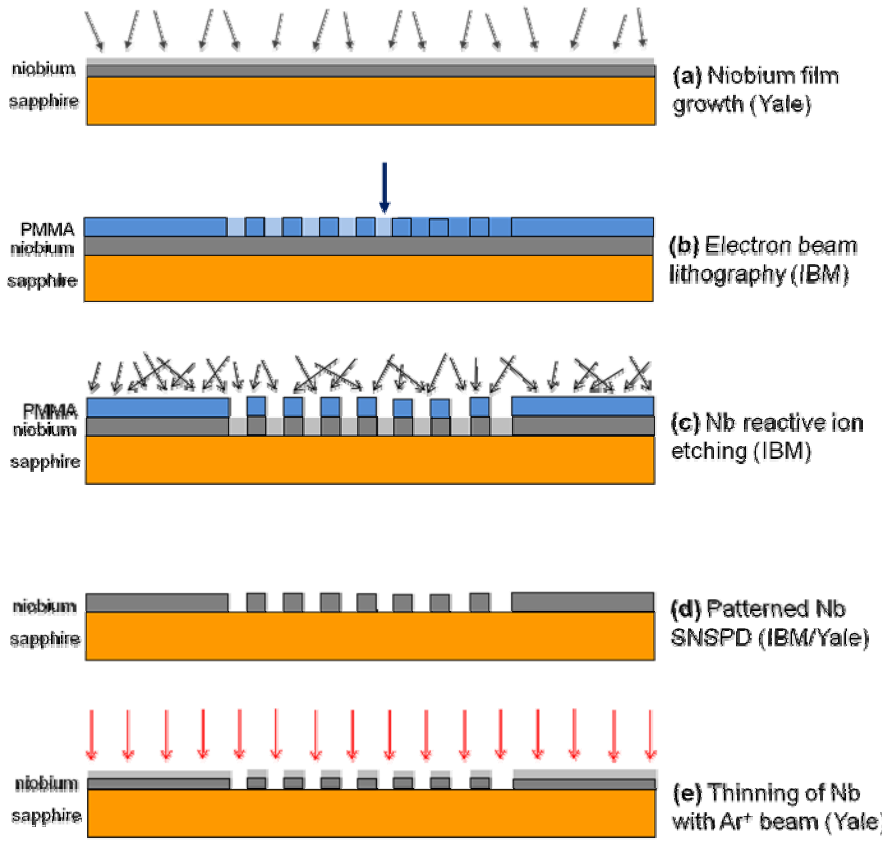

Figure 3.1: Overview of the fabrication process for Nb SNSPDs. A cross section is shown: (a) Sputtering of Nb onto a sapphire substrate; the lighter region shows the full thickness of the film once growth is complete. (b) Electron beam lithography using PMMA. Electron beam lithography is a serial process; each feature is written one at a time. The lighter area shows he part of the PMMA that has been exposed but not yet developed. (c) Reactive ion etching of Nb; the light area shows the part of the Nb hat will be completely etched away. (d) Patterned device. (e) Thinning using an argon ion beam; here, the lightened area shows the reduction from the full thickness, typically 14 nm.

3.2 Nb Film Growth and Optimization

Nb films were grown using dc magnetron sputtering in an ultra high vacuum chamber. Sputtering is a proven method of depositing high quality superconducting thin films of Nb (Reese 2006).

Other methods such as electron beam evaporation that are available at Yale for growing Nb films could have been used, but these deposition systems generally have higher base pressures and were not optimized for growing very thin (< 20 nm) films. For a discussion and analysis of this, see (Reese 2006).[24] For the fabrication of SNSPDs, ultra thin Nb films are required, with thicknesses ~10 nm or less. In this case, a high vacuum and a well characterized deposition process are necessities.

3.2.1 Sputter Deposition

Sputtering is a process of material deposition whereby a plasma of a chemically inert gas (typically Ar) is focused by a magnetic field so that the Ar ions (Ar^{+}) impinge upon the surface of a target that consists of the material to be deposited. The energetic ions strike the surface, knocking molecular-sized pieces of the target material into the vacuum. The plasma of argon ions and deposition material is focused toward the

[24] It was shown by Reese (2006) that the Plassys electron beam deposition system at Yale deposited poorer Nb quality films than the sputtering process when using an additive fabrication procedure (lift-off with a PMMA/MMA bilayer). This was attributed to outgassing of the electron beam resists (PMMA/MMA) because of heating from the evaporated Nb impinging upon and heating the resist. Since in a subtractive process there is no resist present during film growth, this contamination may not occur and it may be possible to grow high quality Nb films in the Plassys. The sputtering system available for this work however still had a significantly lower base pressure.

substrate using a magnetic field. The Ar^+ ions do not significantly affect the crystal structure of the film that is grown because they do not bond with Nb atoms in the lattice (which are in covalent bonds). In this fashion, sputtering is a high purity, highly controllable but non-directional method of depositing a high quality polycrystalline film.

3.2.2 Overview of the Kurt J. Lesker Sputter Deposition System

The deposition system used for sputtering Nb was manufactured by the Kurt J. Lesker Company in 1989 and was originally purchased for growing high temperature superconductor films. The system has been modified from that purpose by past researchers and is presently used to grow thin films of Nb, Al, SiO, Ti, and Au. The system is a multi-target sputtering and evaporation system with an in situ ion beam, heat lamps for substrate heating, a load-locked vacuum chamber, and a mass spectrometer for measuring the concentration of background contaminants. The machine has four stations within the deposition chamber, with a rotating sample holding arm to move between stations without breaking vacuum. In one station, Nb can be sputtered from two different 2” Torus magnetron sputter guns, a proprietary design from K. J. Lesker, Co. with Advance Energy 1.0 kW and 1.5 kW dc power sources for the plasma. In another station, Al and SiO can be evaporated from R. D. Mathis thermal sources powered by Lambda LT-821 and Varian high current power sources, respectively. In a third station, Al, Au, and Ti are sputtered from another cluster of 4 magnetron sputtering guns. In the fourth station, there is a 3 cm Kaufman Commonwealth ion gun with a Ta grid, which is used for ion beam etching. In the fabrication process for Nb SNSPDs, typically only one of the

2” Nb sputtering sources and the ion gun were utilized. The vacuum system consists of two CTI CryoTorr 8 cryopumps, one each for the load lock and deposition chamber. These pumps are able to maintain a stable ~8×10^{-9} Torr base pressure when the chamber has been properly sealed to prevent air leaks and baked out to eliminate water vapor. A detailed description of this system as well as the major use and maintenance issues is given by Reese (2006).[25]

3.2.3 Procedure and Parameters Used

In this section, the procedure for growth of ultra thin Nb in the Kurt J. Lesker sputter deposition machine (hereafter, “the Lesker system”) is outlined. This procedure was used for all Nb devices presented in this thesis. In most cases, the steps of this procedure were decided upon after considering and testing several alternative methods.

Substrate preparation: A 2” sapphire substrate cut so that the surface intersects the R-plane $[1\bar{1}02]$ of the sapphire crystal was used. These were obtained in a prime grade, single-side polished preparation from CrysTec, Inc. Using R-plane sapphire results in a good lattice match to [110] Nb. This orientation of Nb grows reliably even on a room temperature substrate and results in a high quality superconductor when deposition conditions are such that the Nb has internal compressive strain (Celaschi 1985, Durbin

[25] In the course of this thesis work, the author, along with Dr. Luigi Frunzio, has made many repairs and upgrades to this system as well. This was the cause of a very significant amount of time delay in fabricating devices. Future users are hereby cautioned to use and maintain this machine with great care.

1982). This generally requires a very clean substrate, a deposition chamber with very high vacuum, and a fast sputter deposition rate. The procedure for sputter deposition was:

1) Ultrasonically agitate in acetone for 60 seconds.

2) Without drying, immediately transfer to isopropyl alcohol and ultrasonically agitate for 60 seconds.

3) Blow dry with nitrogen.

4) Load substrate into sample holder of the Lesker system with the polished side down as quickly as possible; use a nitrogen blow gun to clean to ensure no particulates collect on the wafer before closing loadlock.

5) Standard procedures for pumping down are used. Prepare the substrate surface by etching off surface contaminants from the sapphire substrate with the Ar^+ ion beam with a beam current density of 6.7 A/m^2 for 15 seconds.

Nb sputter deposition: Nb was sputtered using gun E of the Lesker system with the substrate at room temperature. Sputtering power was 1.73×10^5 W/m^2 over a 2" sputter target, with 1.3×10^{-3} Torr Ar pressure added to a base pressure of less than 1×10^{-8} Torr in the deposition chamber. This results in a deposition rate of approximately 1.3 nm/s, which results in high quality films (Reese 2006).

Post-deposition handling: For all devices tested in this thesis, after deposition the device was transported to IBM T. J. Watson research center for electron beam lithography and reactive ion etching.

3.2.4 Effects of Base Pressure, Substrate Temperature, and Substrate Type

The base pressure of the Lesker system deposition chamber, typically 8×10^{-9} Torr, was not a significant factor in limiting the quality of ultra thin Nb films grown using the above procedure. If the base pressure had been a factor of 10 greater, then oxidation within and between grains of the Nb film during deposition could limit the quality of the crystal grown. For a base pressure of 1×10^{-7} Torr with oxygen as the only residual gas, a rule of thumb is that a monolayer of oxygen molecules will be impinged upon the surface of the substrate once every 60 s. Thus, significant oxygen may be incorporated into the crystal. For studies of the importance of base pressure, see Reese (2006).

The substrate type and temperature play a significant role in determining the quality of the superconducting film grown on it. R-plane sapphire substrates at room temperature were used for all Nb devices fabricated for this thesis. This was chosen both for high quality as well as reliability. A heated sapphire substrate was shown to result in higher quality films by Celaschi *et al.* (1985). A study of the effect of substrate type and temperature on the quality of Nb was undertaken at the start of this thesis work.[26] The results of this study are seen in Table 3.1 below. For these measurements, a Nb film thickness of approximately 15 nm was used for all measurements. Measurements were of bare, unpatterned Nb films with dimensions of approximately 2 mm × 10 mm. As can be seen, growing on R-plane sapphire with an elevated substrate temperature results in the highest quality film. However, heated film growths in the Lesker system were not

---

[26] The study of substrates was undertaken in conjunction with Matthew Reese and Prof. Aviad Frydman.

reliable; using the same nominal deposition parameters, several widely disparate values of T_c and $R_\square$ were obtained due to variation in the system that was not controllable. Data from the device with the highest value of T_c is reported here. thus, the second best process, growth on room temperature R-sapphire was used for this thesis work because it was far more reliable. Furthermore, in chapter 5 it is shown that low resistivity films are not necessarily desirable for high detection efficiency in Nb SNSPDs. However, a high critical temperature is always desired. Note that in the data shown in Table 3.1, the bare Si substrates were obtained by etching away the oxide layer with the ion beam and sputtering the Nb immediately afterward without breaking vacuum. The heated, bare Si substrate did not grow a film that had a superconducting transition above 1.6 K. It is believed that on bare, heated Si, a Nb-silicide film forms which is not superconducting.

Substrate	**Dep. Temp. (K)**	T_c **(K)**	**RRR**	$R_\square$ **(Ω/□)**
SiO_2 on Si	300 K	6.2	1.87	20
SiO_2 on Si	1000 K	6.6	1.8	20
Si	300 K	7.2	2.34	8
Si	1000 K	-	1.35	30
R-sapphire	300 K	7.35	2.62	6
R-sapphire	1000 K	7.8	4.1	3.5

Table 3.1: Measurements of 15 nm thick Nb films on various substrates and at both elevated (1000 K) and room temperatures. The heated deposition on pure silicon did not have a measurable superconducting transition down to 1.7 K. This can be attributed to the formation of a non-superconducting Nb-silicide compound at high temperature.

3.3 Lithography

Lithography is illustrated in Figure 3.1(b). Lithography is a process used to define planar geometric patterns using a source of radiation and a thin film of radiation-sensitive material. There are two types of high resolution lithography useful in fabricating devices with features of ~µm or less. Photolithography uses UV light and thin polymer films called photoresists that are sensitive to this light. By exposing parts of the photoresist to UV light, those parts undergo a photo-activated chemical reaction which, for the case of positive resists, makes them soluble in certain aqueous solutions, referred to as developers. For negative resists, the opposite occurs, whereby the exposed light hardens the resist, leaving other parts soluble. In this way, photoresists can be patterned into masks that are used to either shield other materials underneath from etching or to mask the substrate from metal deposited above the photoresist layer. Photolithography is useful for features that range from sizes of ~ 1 mm to < 1 µm using tools available at most academic cleanrooms, including at Yale. The resolution is limited by the wavelength of UV light as well as the alignment technology of the lithography tool.

If features less than ~ 1 µm are necessary, then either deep ultraviolet lithography or electron beam lithography is required. Deep ultraviolet lithography is not available in most research cleanrooms; it necessitates expensive and complicated equipment and photomask preparation. Rather, electron beam lithography is used for sub-micron lithography in most labs due to its flexibility and high resolution. Electron beam lithography uses a focused beam of electrons to expose certain areas of an electron resist. In a similar manner to photolithography, for positive electron resists (as used in this thesis), the exposed areas become soluble in certain solutions referred to as "developers."

Generally, electron beam lithography developers are different from those used in photolithography. In addition, the process by which an electron beam effects the exposed area of an electron beam resist is also different. When an electron beam impinges upon a thin, positive electron resist, it penetrates through to the substrate. Scattering with the substrate atoms create secondary electrons, many of which backscatter into the resist. These backscattered electrons break down the long polymer chains that make up the resist into smaller molecules, which have lower molecular weight. These lower molecular weight polymers are much more soluble in the developer than the high molecular weight polymers. For negative resists, the opposite occurs, whereby the exposed areas of the resist are hardened by the secondary electrons, leaving other parts soluble.

For fabrication of SNSPDs, the electron beam resist must be exposed in narrow strips ≈100 nm wide. Thus, the resolution requirement is 100 nm. These strips of exposed area must be very high aspect ratio, to form a high aspect ratio nanowire. The width of the exposed strip cannot vary by more than approximately +/-5% along the entire length in order for the nanowire to have uniform critical current. Thus, the uniformity requirement is +/- 5% over an area of ~100 μm^2. Furthermore, these exposed strips must meander back and forth, within 100 nm of each other so as to form a detector with fill factor $\sigma = 0.5$ so as to have high detection efficiency (see section 2.4). This means that the overlap in exposures due to nearby patterning must be taken into account when determining the dose, or the total charge deposited, in the electron resist. This overlap in exposure is referred to as proximity effect. Taken together, the requirements on resolution, uniformity, and proximity effect necessitate a high performance electron beam lithography system.

3.3.1 Overview of Electron Beam Lithography

In this section, an overview of electron beam lithography tools is given. In a thermal field emission source, the type typically used in modern electron beam lithography tools, the source of electrons is a sharp, heated tip that is biased with a high voltage. The heat assists the emission of electrons, and causes zirconium oxide to flow down the tip, providing a low work-function surface (Rooks 2010).The emitted electrons are further accelerated and focused using electromagnetic fields to collimate the electron beam and focus into a very small spot, typically of order a few nanometers.

The resolution and proximity effect of electron beam lithography depends on several factors. The wavelength of electrons is much smaller than the wavelength of UV light, so the resolution (minimum feature size) is not limited by the wavelength. It is limited by scattering processes within the substrate and electron resist that tend to create secondary electrons that scatter randomly, making the exposed area wider than the beam width. Resolution is also limited by fluctuations in the incident electron beam width and in the beam current as well as vibrations of the sample holder stage. The degree to which the electron beam spreads out within the electron resist depends on the energy of the electrons. This depends on the total accelerating voltage used to create the electron beam. For the Nb devices fabricated for this thesis, 100 kV was used for all lithography. This is the highest available voltage of the IBM system and results in the highest resolution of any electron beam lithography system if used in conjunction with a very precise beam control optics and sample mounting stage. The accelerating voltage also determines the degree of proximity effect. A higher accelerating voltage leads to a narrower exposure width, and thus a higher resolution. Higher voltage also leads to less secondary electron side-scattering, which can exposes areas adjacent to the intended target area. Thus, the proximity effect is less for higher voltages. To further compensate for proximity effect, particularly at the edges of a large pattern, two methods are employed. First, advanced software is used to calculate the dose distribution across a large pattern. Second, structures that are specifically designed to promote an even distribution of secondary electrons across a large pattern are exposed near and/or within the desired pattern. Both of these methods were used in fabricating Nb SNSPDs for this thesis work.[27] An example of these proximity effect correction structures is seen in Figure 3.2. The central nanowire meander is the SNSPD device, while the unconnected bordering strips are written purely to correct for the under dose that the perimeter of the SNSPD might receive if those border structures were absent.

[27] Although later fabrication runs without these proximity effect compensation structures proved that the structures are not necessary when the exposure is optimized.

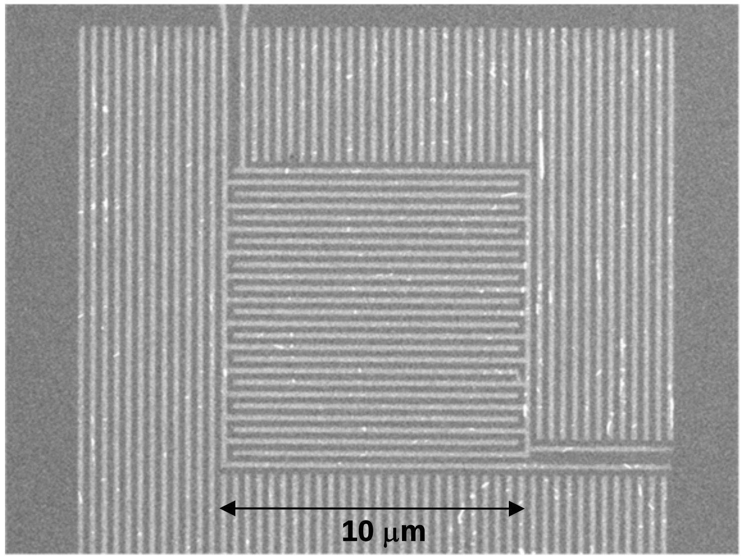

Figure 3.2: SEM image of a Nb SNSPD with line width of 200 nm patterned alongside unconnected proximity effect correction structures around the perimeter of the device. In this image, the dark areas are Nb and the light areas are the (charged) sapphire substrate. The bright spots visible in the image may be the result of charging of organic contaminants that settle on the surface of the film.

Uniformity (variation in the width of the nanowire strip) is determined by the stability of the beam current and the sample stage, the precision with which the electron beam can be positioned on the sample, and the development procedure. To achieve high uniformity typically requires high precision electron focusing and high precision stage control. In the applications where uniformity is important, stage controls with active vibration compensation and laser interferometric positioning are typically used. In addition, electromagnetically shielded, climate controlled cleanrooms with extremely stable foundations are often employed. In addition, a highly uniform developer consisting of isopropanol and water was used (Rooks 2002).

3.3.2 Overview of IBM Electron Beam Lithography Facilities

The lithography for this thesis work was performed at the electron beam lithography facilities at IBM T. J. Watson Research Center in Yorktown Heights, NY.[28] The electron beam lithography instrument utilized was a Leica VB6 machine with an accelerating voltage of 100 keV. This instrument is equipped with state of the art stage control with active vibration compensation and laser interferometric positioning. The instrument is located in a cleanroom and is maintained by experts in electron beam lithography.

3.3.3 Procedures and Parameters Used

In this section, the procedure for electron beam lithography using the Leica VB6 instrument at IBM T. J. Watson Research Center is outlined. This procedure was used for all Nb devices presented in this thesis. In most cases, the steps of this procedure were decided upon after considering and trying several alternative methods.

Spin coating of PMMA: The following procedure was used to spin coat 3% 950K PMMA in a solvent of anisole (PMMA A3) onto the previously grown Nb film on a 2” sapphire wafer. No primer or additional cleaning beside blow drying was utilized. It must be noted that a maximum baking temperature and time for the PMMA of 125^{o} C and 5 minutes was necessary to avoid oxidation of the ultrathin Nb film. At higher baking temperatures

---

[28] All electron beam lithography was performed by or under the oversight of Michael Rooks, formerly of IBM Research and now Associate Director of Facilities at the Yale Institute for Nanoscience and Quantum Engineering.

and/or longer baking times, increased sheet resistance and degradation of the superconducting properties of the Nb film were observed.

1) Blow clean with dry nitrogen.

2) Spin coat 950K PMMA A3 at 4000 4000 revolutions per minute for 60 seconds.

3) Examine the wafer for uniform coverage.

4) Immediately bake the spin-coated wafer at 125^{o} C for 5 minutes on a uniformly heated hotplate.

5) Load the wafer into the sample holder of the Leica VB6.

Electron Beam Exposure: The procedure for electron beam lithography is a standard IBM process developed by Dr. Rooks for defining high resolution features in PMMA. Resist was exposed at 100 kV, using 6 nA, with a typical spot size of ~7 nm with doses of 1 mC/cm^2 or approximately 63 $electrons/nm^2$ of PMMA. The required dose scales with the Bethe stopping power of electrons traveling through resist (Kim 2005). The critical dose for exposure at 100 kV is therefore 2.3 times higher than when exposing at 30 kV.

Development: The following procedure was used for developing the PMMA once it was exposed to the electron beam.

1) Develop by submersing the wafer in a mixture of 3 parts IPA:1 part water at 5^{o} C while ultrasonically agitating for 60 seconds.

2) Blow dry with nitrogen.

3) Examine the wafer under an optical microscope to examine the pattern.

3.4 Etching and Electrical Testing

Etching refers to the process of removing material from a wafer. There are several types of etching, including chemical etching, sometimes referred to as wet etching; reactive ion etching; and ion beam etching, sometimes referred to as ion milling. An important parameter in etching is the ratio of etch rates between one material and another. This is referred to as selectivity. Chemical etching utilizes a solution that dissolves or otherwise chemically reacts with a material to remove it. A chemical etch is usually isotropic, with selectivity depending completely on the chemical reactivity of one substance versus another. Reactive ion etching (RIEs) is also an isotropic, chemically reactive etch, but instead of using molecules in solution, the chemically active etchant molecules are ionized in a plasma within a vacuum chamber. The ionization promotes faster etch rates, greater selectivity, and much better process control. The third category of etching is ion beam milling. This is a purely kinetic process by which a beam of chemically inert ions impinges upon and erodes the surface of a material in a vacuum chamber. It is similar to sputtering in that a plasma is used to bombard and remove metal from a target. Ion beams are very directional. In fabricating Nb SNSPDs, RIE and ion beam etching are used. For RIE, a Unaxis 770 machine at IBM Watson Research Center is used. For ion etching, the ion gun in the Lesker system described in section 3.2 is used. A final ion etching step was necessary to fabricated devices with high detection efficiency. This is discussed in section 5.3 of this thesis.

3.4.1 Reactive Ion Etching Procedure

The following procedure was used for the etching of Nb after electron beam lithography and development.

1) First, clean the Unaxis 770 RIE chamber by etching with O_2 for at least 5 minutes.

2) Place the wafer with the pattern side up into the Unaxis RIE.

3) Etch for 3.5 minutes at 100 W and 30 mTorr of CF_4.

4) Remove wafer, examine under an optical microscope and/or check whether devices have been etched by measuring the resistance between two structures that should be isolated if the etch has fully penetrated the Nb film.

5) Once etching is complete, remove residual PMMA using acetone heated to 70^{o} C for 30 minutes followed by ultrasonically agitating in acetone at room temperature for 2 minutes.

Reactive ion etching with a Fluorine-based plasma is not particularly selective of the Nb compared to the PMMA. By using a low power for the reactive ion etch, the PMMA A3, which was approximately 100 nm thick, was sufficient as an etch mask for up to 14 nm of Nb. Use of low power prevents reflow of PMMA, which has a relatively low glass transition temperature of 110C (Rooks 2010). Etching for much longer than 3.5 minutes would result in the PMMA being completely removed and some of the Nb film being etched away from areas where it was desired to remain. A typical result of this lithography and reactive ion etching process is seen in 3.3. This test structure was used throughout the fabrication development.

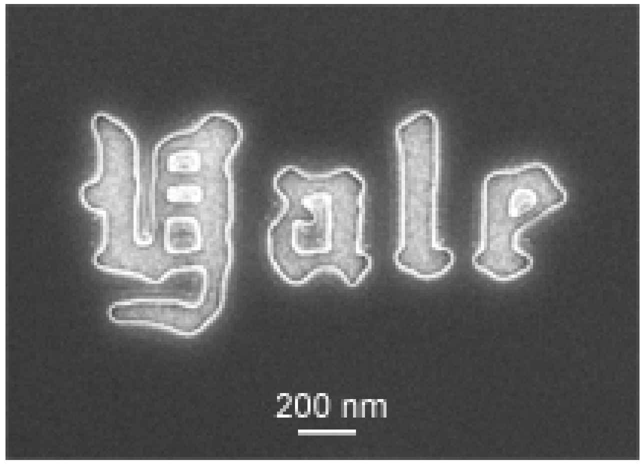

Figure 3.3: SEM image of etched Nb in a typical test pattern used in optimizing the lithography and etching process. Courtesy of M. Rooks.

3.4.2 Electrical Measurements and Dicing

After etching and cleaning is completed, the resistance of each device on the wafer is electrically tested using a dc, room temperature probe station. This is used to determine which devices have been fabricated successfully and which have defects that are apparent in their dc electrical resistance. Possible defects include partial and full short circuits and/or open circuits, which are due to defects in the patterning or etching. A map of the devices on the wafer is compiled, and then the wafer is respun with photoresist as a protective layer and then cut into individual chips. The process by which a wafer of multiple devices is cut into chips is called dicing. Dicing is done at Yale using a Microautomation (model number 1006) dicing saw with a diamond-tipped blade used to cut sapphire. The chips each have typically 4 devices. Those chips with 4 devices that have the best electrical resistance are chosen for further testing. Chips are measured one at a time at cryogenic temperatures in the measurement apparatus described in chapter 4.

3.4.3 Ar^+ Ion Beam Etching Procedure

It was found that in order for Nb SNSPDs to have high detection efficiency, a final etching step was necessary. It was found that devices that were initially patterned with thicker Nb films, and later thinned with an ion gun, had significantly higher detection efficiency than devices that were directly patterned from thin films. This is discussed in chapter 5. The following procedure was used for this etching step.

1) Choose a chip with defect-free devices after previous reactive ion etch and dicing.
2) Ultrasonically agitate the chip in acetone for 60 seconds.
3) Without drying, immediately transfer the chip into isopropyl alcohol and ultrasonically agitate for 60 seconds.
4) Blow dry with nitrogen.
5) Load chip into chip holder of the Lesker system with device side oriented away from backing plate; use nitrogen blow gun to ensure no particulates collect on wafer before closing load lock.
6) Standard procedures for pumping down are used. Etch with the Ar^+ ion beam with a beam current density of 6.7 A/m^2 for 30-50 seconds, depending on the original

thickness and the desired final thickness. It is typical to reduce the thickness from14 nm to 7.5 nm. The etch rate for Nb is roughly 0.13 nm/ second.

7) Remove from the Lesker system.

The argon ion etch is the last step of the fabrication process for high detection efficiency Nb SNSPDs.

3.5 NbN SNSPD Fabrication

The NbN SNSPDs tested in this thesis work were fabricated by collaborating groups at the University of Salerno in Salerno, Italy, and at CNR – Istituto di Cibernetica in Pozzuoli, Italy. The fabrication process is described by Leoni *et al.* (2006). In that process, ultrathin films of NbN are grown using dc magnetron sputtering on 10x10 mm^2 R-plane sapphire substrates heated to 400^o C in an environment of Ar and N_2, with N_2:Ar ratio of 1:2 at 3.4 mTorr total pressure. The lithography is a two step process. In the first step, a layer of PMMA is spin coated onto the NbN film, baked, exposed using a 100 keV electron beam lithography system, and developed. A 60 nm thick layer of Ti/Au is deposited using an electron beam evaporator, after which liftoff of the PMMA defines contact electrodes. For the NbN devices, the contact electrodes were a 50 Ω coplanar waveguide transmission line, tapered to transition between the microscopic device dimensions and the macroscopic transmission lines on the sample mounting board. At this point, a second layer of electron beam lithography is done. A layer of hydrogen silsesquioxane (HSQ), a negative electron beam resist, is spin coated on top of the NbN and contact pad layers, baked, exposed using a 100keV electron beam lithography system, and developed. A reactive ion etch based on a mixture of CHF_3 and SF_6 is used to etch the NbN into a nanowire meander.

3.6 Summary of Nb and NbN Devices Tested

In this thesis, over a thousand Nb SNSPDs were fabricated.[29] This consisted of 25 wafers, each with 16 to 25 chips, patterned in 4 × 4, 5 × 5, or 6 × 6 arrays. A schematic of a typical chip is seen in Figure 3.4. Each chip has 4 SNSPD devices with varying geometries. A schematic of a typical device is seen in Figure 3.5. The patterning was such that the 4 electrodes connecting to each of the devices on a chip would match up to 4 coplanar transmission lines on a custom designed printed circuit board on the cryogenic insert, with connections made via wirebonding.[30] The on-chip contact electrodes for Nb devices were not designed to be 50 Ω or any other specific transmission line impedance. A 50 Ω or other matched impedance is not necessary because the contacts are short compared to the wavelengths of even the fastest electrical dynamics of the device. To show this, consider that the on-chip contact electrodes are approximately 4 mm long. A general rule of thumb for high frequency design is that if the length of a component is less than $\sim\lambda/10$ where λ is the wavelength of interest, then the impedance that results from the distributed capacitance and inductance, which is the transmission line impedance, can be treated as a lumped element.[31] In this case, λ needs to be reduced by

[29] This excludes ion beam etching, which is the final step of the fabrication process and only done on those devices that were screened for fabrication defects by probing their electrical resistance.

[30] This printed circuit board is detailed in chapter 4.

[31] That is, one with no spatial extent, such as an ideal capacitor or inductor in a circuit model.

the square root of the effective dielectric constant of the transmission line. Assuming an effective dielectric constant of 3.2, which is the geometric mean between air and sapphire at GHz frequencies, the frequency that corresponds to wavelength that satisfies the equation, $\lambda/(1.8\cdot 10) = 4$ mm, is approximately 13.6 GHz, corresponding to a rise time of $(2\pi f)^{-1/2} \approx 12$ ps.[32] This is significantly faster than any electrical response time of the device. Furthermore, the capacitance and inductance of this short strip, even when treated as a lumped element, are very small, and in practice can be neglected at these frequencies. Also apparent from the illustration is a large (~100 μm) structure that is located in the upper right corner of the chip. This is used to measure the dc electrical properties of the two-dimensional Nb film. This structure was useful for making 4-probe measurements of the sheet resistance of the wide Nb films. This was used to determine whether the quality of the film itself was degraded by processing. In general, the quality of the film was impacted only slightly, with a critical temperature and sheet resistance within approximately 10% of the value for an unprocessed film.

[32] This equation comes from treating the maximum rise time as given by a low pass filter with a cutoff frequency, f, such that the transfer function through the filter is $V_{in}/(1-2\pi f\tau) = V_{out}/2$; this is the "3dB" cutoff point.

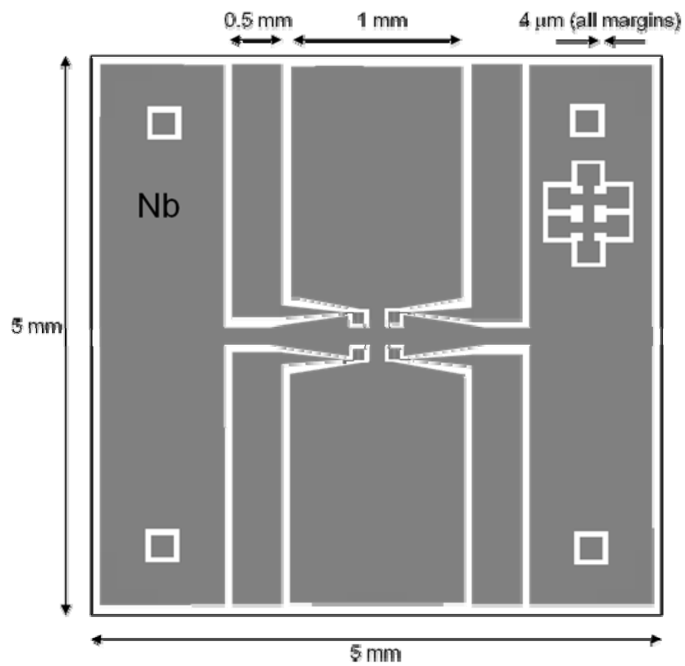

Figure 3.4: Schematic of a typical chip with 4 Nb SNSPD devices. The devices are represented by the small striped squares in the center of the chip, with contact electrodes connected to each. The white areas are the sapphire substrate. These areas are the ones where the PMMA over-layer is exposed and developed away, after which reactive ion etching removes the Nb film from the exposed areas.

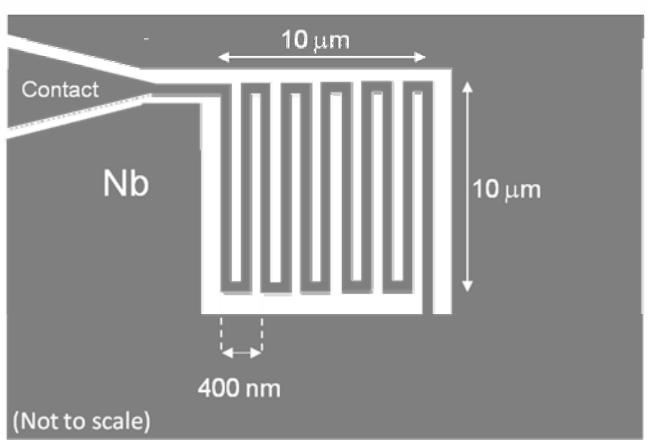

Figure 3.5: Schematic of a typical Nb SNSPD. The white areas are the sapphire substrate. These areas are the ones where the PMMA over layer is exposed and developed away, and then where reactive ion etching removes the Nb film.

Over the course of this thesis work, the superconducting properties of each Nb device on approximately 25 chips from 10 different wafers were tested. Many of the rest of the chips had devices with defects, such as resistance that was too high or too low. These types of defects were apparent before cooling to cryogenic temperatures. They were determined based on measurements of the room temperature dc electrical resistance in a probe station and based on SEM imaging. A schematic of the screening methodology is seen in Figure 3.6. During each step of the fabrication, devices with defects were discarded. While the fabrication yield from all fabrication runs taken in aggregate is quite low, the fabrication yield of a particular wafer of devices where the electron beam lithography and etching parameters were optimized correctly was > 90%. This fabrication yield is defined to be the number of devices with the expected dc room temperature resistance divided by the total number of devices on chip. Of the devices that had no fabrication defects, some still had internal defects that lead to the devices having reduced critical temperatures and/or reduced critical currents. These defects were only apparent when the devices were cooled down to cryogenic temperatures to test their superconducting properties (see section 5.2). Such defects included reduced I_c or T_c and are discussed in chapter 5. In Figure 3.7, images of typical Nb SNSPDs studied in this thesis are shown.

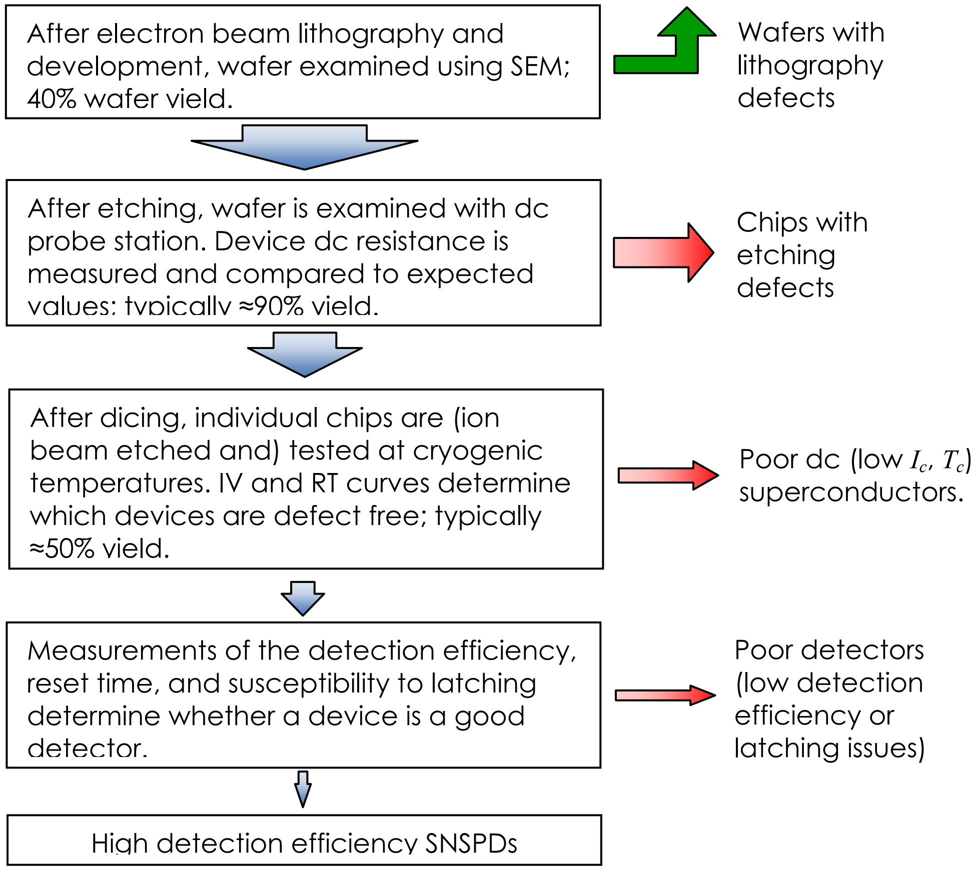

Figure 3.6: Schematic of screening steps used during the fabrication and characterization process for Nb SNSPDs. After the first step, wafers with lithography defects can be carefully cleaned and recycled without degradation of the Nb film. The final two screening steps are performed post-fabrication, and are discussed in detail in chapter 5.

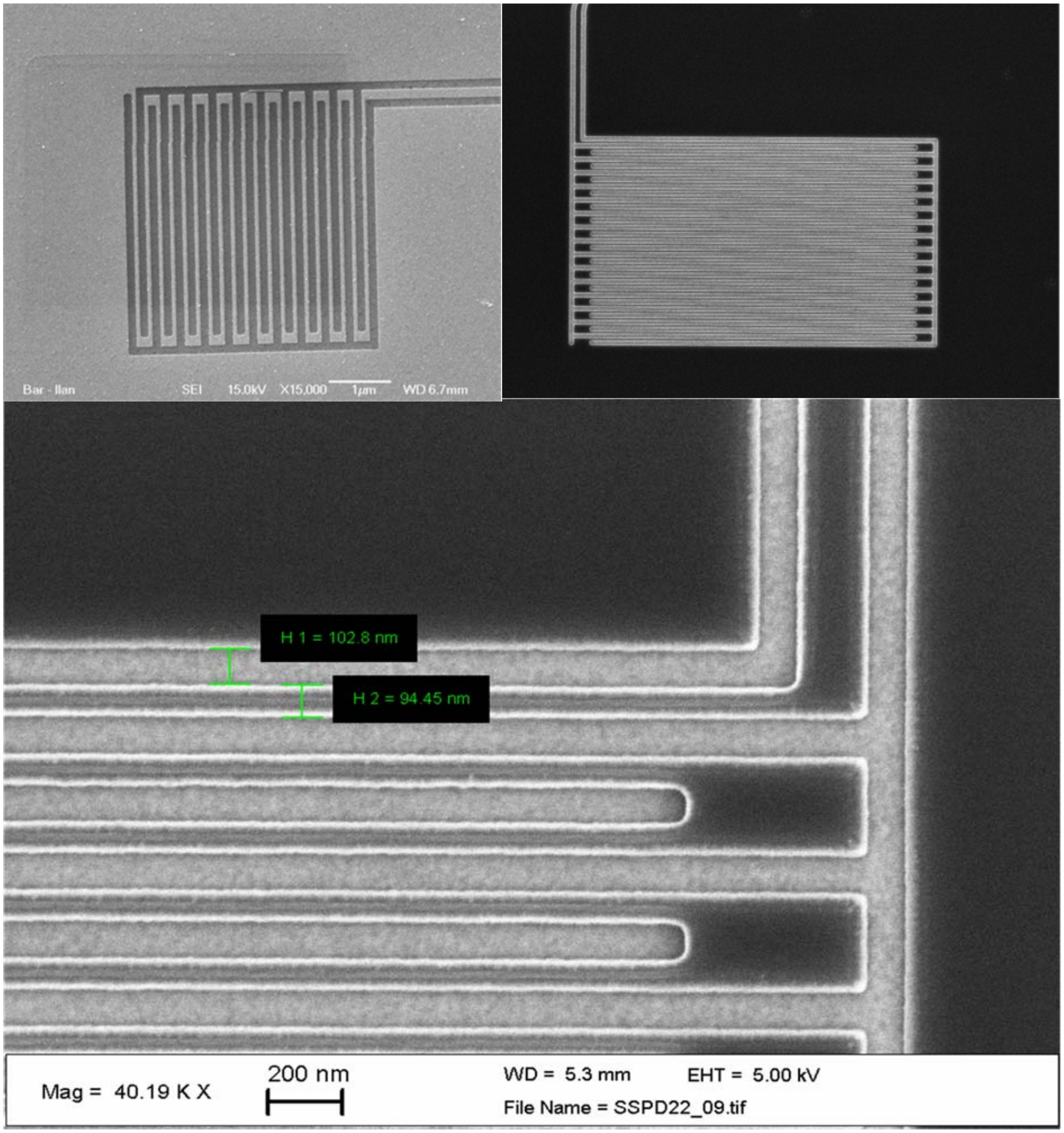

Figure 3.7: SEM images of typical Nb SNSPDs studied in this thesis. In the top two images, the Nb is light and the substrate is dark, while in the bottom image, the tone is reversed. In the top left image, a small stitching offset is apparent. This can occur if the electron beam field of view has a border within the device detection area. In later fabrication runs, this offset error was eliminated. In the bottom, a close up of a meander fabricated using the optimized process is shown with scale bars. As can be seen, line width uniformity is excellent, even without the use of proximity effect compensation structures.

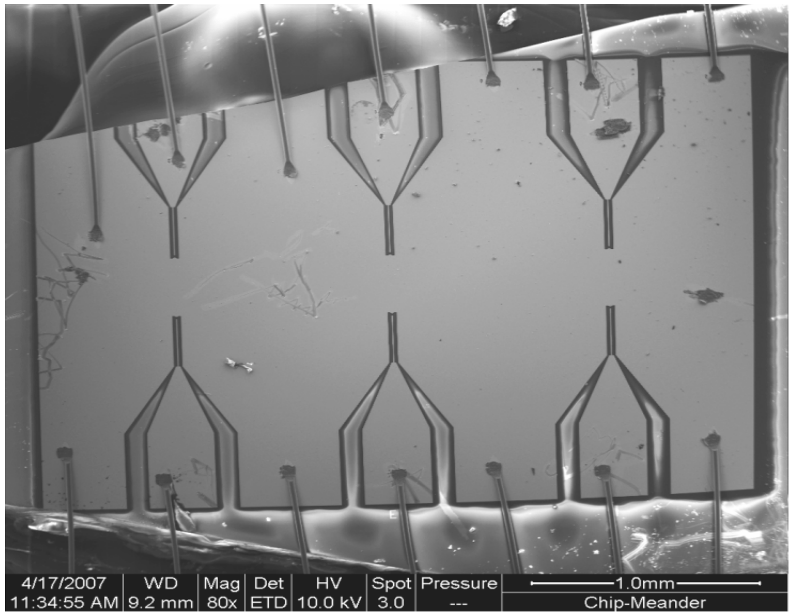

Figure 3.8: An optical microscope image of a typical NbN SNSPD chip with 6 devices. The wires that are visible are the wirebonds that make contact between the device contact electrodes and the sample board transmission lines. For the NbN devices, the contact electrodes were a 50 Ω coplanar waveguide transmission line, tapered to transition between the microscopic device dimensions and the macroscopic transmission lines on the sample mounting board.

Since the NbN SNSPDs were fabricated and tested first by collaborators, the devices received were already known to have acceptable superconducting properties. Two chips of NbN SNSPDs were tested. These chips each consisted of 6 devices of varying geometries. An optical microscope image of a NbN SNSPD chip is seen in Figure 3.8. An SEM image of a typical NbN device is seen in Figure 3.9.

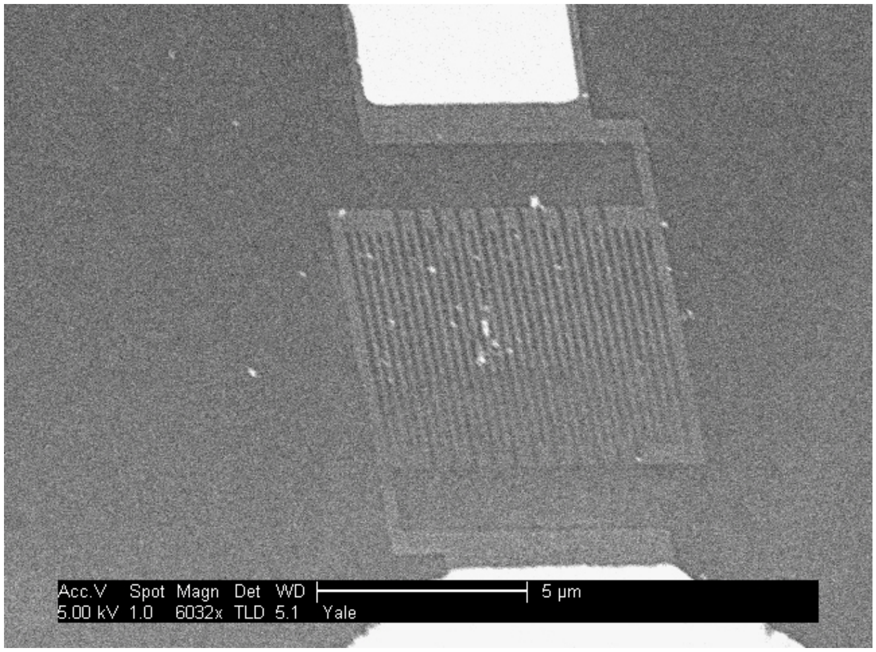

Figure 3.9: An SEM image of a typical NbN device. In this image, the NbN is light and the substrate is dark. The image is somewhat dull and distorted due to the difficulty of focusing the SEM because of substrate charging. In the NbN devices (fabricated by collaborators), the substrate near the devices is bare, while in Nb devices (fabricated at Yale and IBM), the substrate is mostly covered by Nb, which eliminates charging effects and allows accurate SEM imaging.

Chapter 4

Experimental Apparatus

4.1 Overview of Measurement Setup

The accurate characterization of single-photon sensors requires a carefully designed measurement setup. As in many nanoscale cryogenic physics experiments, special care must be taken to create a low temperature, optically dark, and electrically shielded environment where the device will not be significantly affected by external stimuli. To fully characterize a superconducting nanowire single-photon detector (SNSPD), an integrated system of dc, microwave, and optical components had to be developed. An overview of the measurement setup used for detector characterization is seen in Figure 4.1. There are four major parts: 1) (in red) laser diode sources and fiber optics used for exciting the detector in a highly controlled manner, 2) (in green) radio frequency (RF) readout electronics, which are used to measure and analyze fast voltage pulses, 3) (in purple) dc electronics, which are used for biasing as well as measuring the resistance versus temperature and current versus voltage characteristics of SNSPDs, and 4) (in black and yellow) the insert mechanical core along with the vacuum system and temperature regulation system, the cryostat, and electromagnetic shielding. Each of these parts of the measurement apparatus will be discussed in detail in the next sections of this chapter.

The cryogenic insert is the structure which contains the thermal, electrical and optical measurement components. The insert is illustrated in Figure 4.1; in this schematic, it includes all components within the cryostat. The sample is isolated from room light and electric fields by a copper inner vacuum (IVC) can that is part of this insert. The insert is discussed in detail in section 4.2. It is immersed in a liquid 4-helium cryostat. The cryostat consists of a glass dewar filled with liquid 4-helium (1.6 – 4.2 K) surrounded by a second, larger glass dewar filled with liquid nitrogen (77 K). The liquid helium bath is typically held at pressures much less than 1 atmosphere (typically a few Torr) using a Welch mechanical vacuum pump. This reduces the helium boiling point to approximately 1.6 K, which reduces the temperature of the liquid, and the lower part of the cryogenic insert, to this value. The cryostat is magnetically shielded with μ-metal[33] and aluminum and copper foil.

In the course of this thesis work, a completely new cryogenic insert was designed and constructed. This insert incorporated several novel features to facilitate data collection, including the use of a 6-channel RF-bandwidth switch at cryogenic temperatures, which enabled testing of multiple devices and with variable shunt resistances in the same cool-down. In addition to several novel features, the insert, measurement circuits, optics, and readout electronics incorporated many standard design features that are common in an experimental apparatus at low temperatures and high frequencies. In this chapter, many aspects of the measurement setup will be explained, with special emphasis on those features that are new or novel to the laboratory where this thesis work was completed.

[33] The term μ-metal refers to a nickel-iron alloy (approximately 75% nickel, 15% iron, plus copper and molybdenum) that has very high magnetic permeability (Jiles 1998).

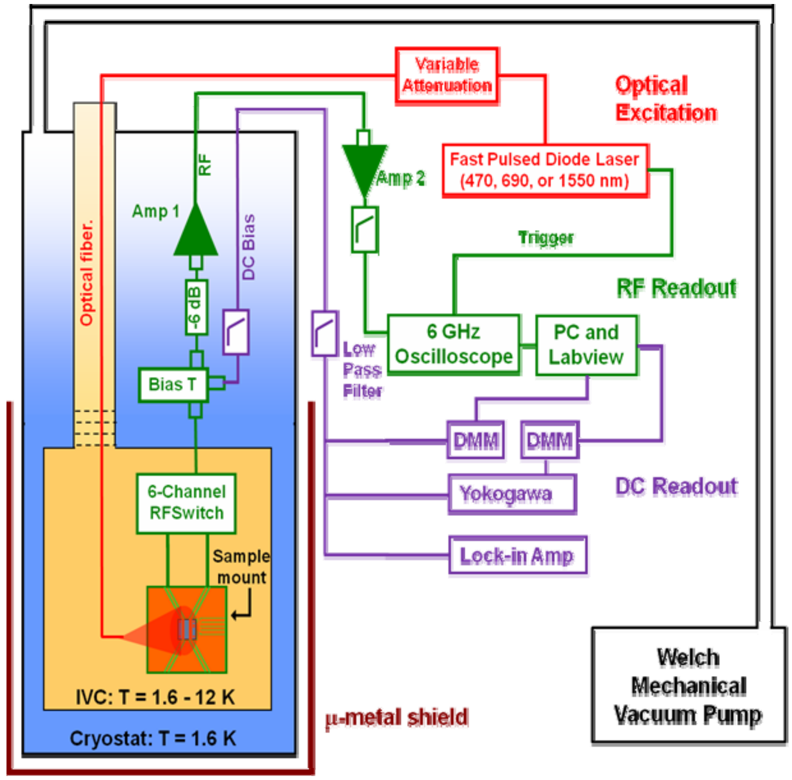

Figure 4.1: Overview of measurement setup for SNSPD dc and detection characterization. The liquid nitrogen dewar is not shown.

4.1.1 Measurement Circuit for DC and Photon Detection Studies

A schematic of the readout circuit used for dc characterization and photon detection measurements is shown in Figure 4.2. The SNSPD is wirebonded to a copper coplanar waveguide patterned on a printed circuit board (not shown, but discussed in section 4.2), feeding into a coaxial input of a remote controlled 6-channel switch (Radiall R573423600 with a bandwidth of 0-18 GHz). This switch enables connecting 4 devices and 2 additional loads in any parallel combination (one device channel and two load channels are shown here). RF and dc signals are coupled through the top wall of the copper inner vacuum can (IVC) from the switch common port using a glass bead vacuum feedthrough and are split using a bias tee (Minicircuits ZFBT-6GW with RF bandwidth of 0.1-6000 MHz)[34] held at 4.2 K in the helium bath.

The dc bias line is 0.086” outer diameter semi-rigid coaxial cable, filtered using Minicircuits lumped element low pass filters (LP1) and a home-built copper powder low pass filter (LP2, f_{cutoff} ~ 1 MHz) located in the helium bath.[35] DC current biasing for detection measurements is from a Yokogawa 7651 low noise voltage source in series with a large bias resistor (100 kΩ, 1 MΩ, or 10 MΩ) at room temperature. For current and voltage measurements, the same dc bias source is used, but the device voltage is measured using an HP 34401A digital voltmeter. When measuring resistance versus temperature, the current bias is sourced and the device voltage is measured by a Stanford Research SR850 lock-in amplifier.

RF amplification is accomplished using a cryogenic first stage amplifier (Amplitech APTC3-00050200-1500-P4 with bandwidth of 30-3000 MHz)[36] located in the helium bath with a 6 dB attenuator at the input and a semi-rigid coaxial delay line separating the device from the amplifier in order to mitigate the effects of reflections of the high frequency signal due to impedance mismatches. This amplifier was ordered

---

[34] This bias tee is not designed for cryogenic use; however, careful measurements of the frequency-dependent transmission through the dc and RF ports show that the Minicircuits ZFBT-6GW works well and reliably at temperatures down to 1.6 K through many thermal cycles of the cryogenic insert.

[35] For a discussion of the design and construction of copper powder filters and other types of dissipative filters useful in high frequency cryogenic measurements, see Santavicca (2009).

[36] This High Electron Mobility Transistor (HEMT)-based amplifier was custom designed by Amplitech, Inc. for these experiments and has the lowest noise of any commercially-available amplifier with this bandwidth.

specifically for this project. The second stage amplifier is at room temperature (several models used, depending on measurement; typical bandwidth of 100 – 8000 MHz or 1-1000 MHz). Lumped-element low pass filters (LP3) are used at the input to the oscilloscope, with cutoff frequencies dependent on the measurement being performed. High speed readout is done using a 50 Ω input impedance, 6 GHz, 20 gigasample per second real time oscilloscope (Agilent 54853).

In this simple circuit schematic, the optical pulse incident on the detector determines the device resistance, $R_d(t)$, by controlling a switch between a zero-resistance short to ground and a resistor with constant value of R_d. This simple electrical model of photon detection in an SNSPD is explained in section 2.1.[37] The photon source is a pulsed diode laser (discussed in section 4.2).

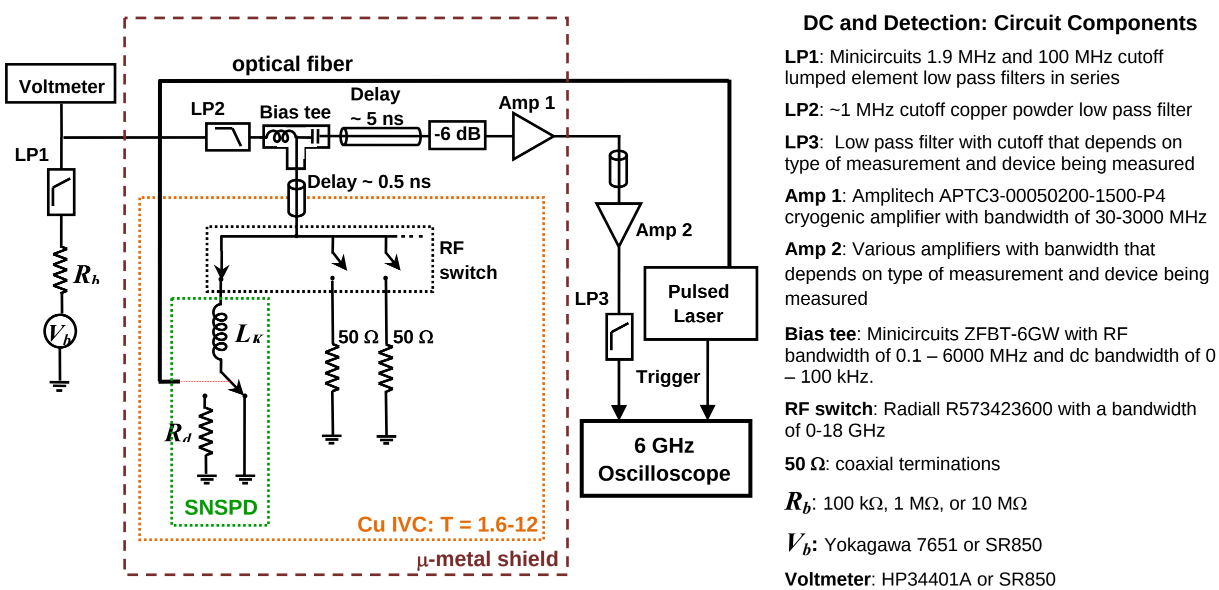

Figure 4.2: Circuit diagram for dc and photon detection measurements.

[37] As discussed in section 2.3, actual device operation is more complicated than this simple model.

4.1.2 Measurement Circuit for Kinetic Inductance Studies

A schematic of the readout circuit used for kinetic inductance measurmeents is shown in Figure 4.3. The dc bias component of the circuit as well as the bias tee, RF switch, and sample mounting printed circuit board is exactly the same as is used for dc and detection measurements. Now, however, a shunt capacitor, C_s, and a 50 Ω resistor are also in parallel with the device. When the frequency $f = (L_K C_s)^{-1/2}$, the reactive impedance will be zero and the circuit will look like a purely resistive 50 Ω termination, producing a minimum in the reflection coefficient. Thus, by measuring the frequency-dependent reflection coefficient with a known capacitance, the kinetic inductance may be determined. The capacitor is a 0203 form-factor NPO ceramic[38] chip capacitor that is soldered directly between the center conductor and ground plane of the coplanar waveguide that connects to the device. The resistor is a Farnel 0102 form-factor metal film chip resistor with nominal resistance of 47 Ω.[39] Here, the RF switch is only used to connect the resonant circuit to the readout line.

A probe signal is input to the coupled port of a directional coupler (type and bandwidth dependent on value of C_s and device being measured) from port 1 of a network analyzer (HP 8722D for high frequency measurements or HP 3589A for lower frequency measurements; dependent on value of C_s and device being measured). The

[38] NPO ceramic is known to be mechanically reliable and to have a relatively temperature-independent dielectric constant from room temperature to temperatures below 2 K. The capacitance of the chip capacitors used was measured from room temperature down to 4.2 K. In this range, there was no measurable change in the capacitance, to within the accuracy of the meter used, which was +/- 1 pF.

[39] These metal film resistors were measured to be relatively temperature independent from room temperature to temperatures below 2 K.

reflected signal is coupled through the directional coupler and amplified by a series of room temperature amplifiers and filtered using lumped-element low pass filters at the input to the oscilloscope. The bandwidth of the amplifiers and filters is dependent on the particular device being measured and value of C_s used. No optical pulse is used during a measurement of the kinetic inductance. The sample is kept in a completely dark environment within the vacuum can with the fiber blocked with a metal cover at the same temperature as the vacuum can.

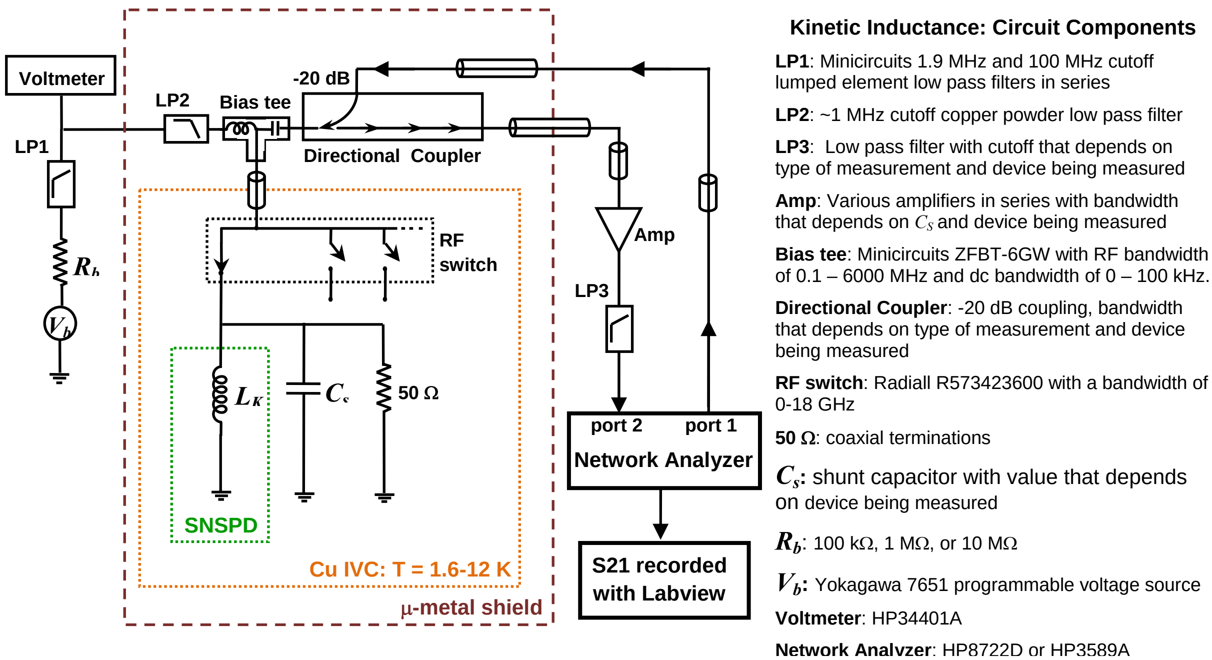

Figure 4.3: Circuit diagram for kinetic inductance measurements.

4.2 Cryogenic Insert

The cryogenic insert is the experimental apparatus that holds the sample under measurement at a stable, low temperature (typically ≤ 4.2 K) in a dark, shielded environment while making electrical and optical connection from room temperature to the cold device in order to probe its properties. It does this by being immersed in liquid helium at one end, where the sample is located. The other end is held at room temperature, where connections to room temperature electronics and optics are made. Since a large thermal gradient must be maintained across the length of the insert, it typically has a high aspect ratio and must be constructed of carefully chosen materials and electrical and optical components. In Figure 4.4 (a), an image of the complete insert used in this thesis work is seen.

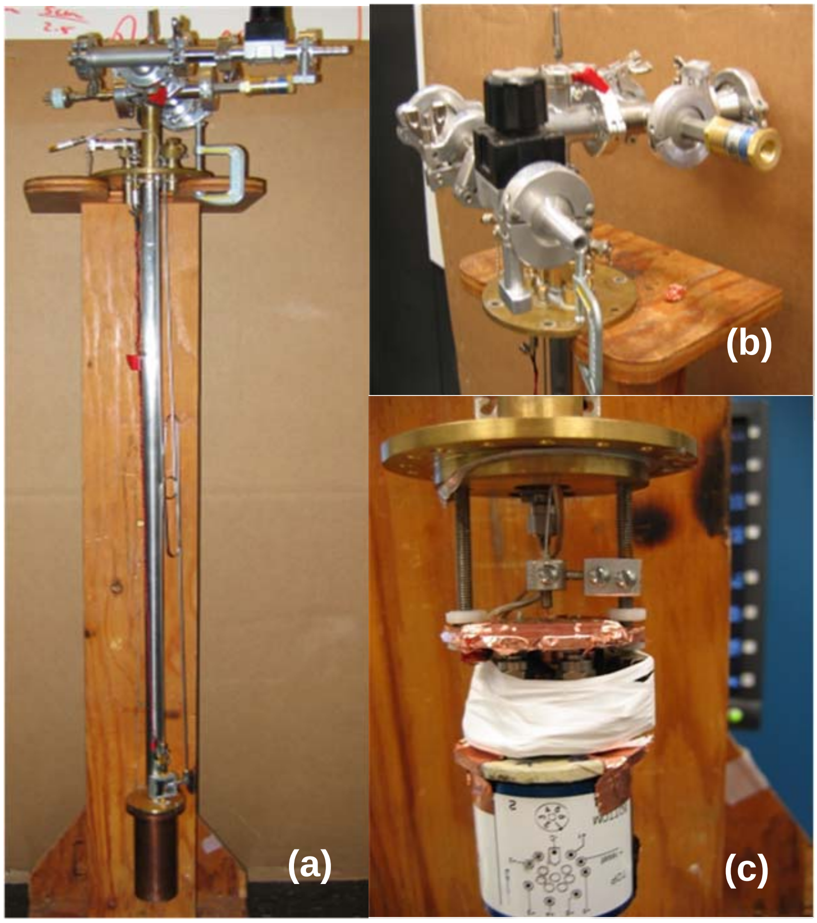

Figure 4.4: Images of the cryogenic insert designed and constructed for this thesis work: (a) the complete insert with vacuum can, ready to be inserted into the dewar; (b) close-up image of the top side of the insert, which shows the KF-vacuum components; (c) close-up image of the internal components of the vacuum can, with the dc wiring and SMA connections between the RF switch and the sample covered by Teflon tape to prevent any of the wires from contacting the walls of the IVC.

4.2.1 Insert Mechanical Design

The insert constructed for this thesis work was modeled on standard designs for cryogenic inserts used in many low temperature physics laboratories. A schematic of the mechanical core of the insert is seen in Figure 4.5. The main shaft consists of a vacuum-tight, stainless steel tube that doubles as a support shaft that connects between the top flange and the vacuum can. The main shaft is used to evacuate the vacuum can and also contains dc wiring and the optical fiber. The vacuum can along with the lower part of this main shaft is immersed in liquid helium at 4.2 K. The top part of the shaft protrudes through a brass flange, which covers the dewar opening and is at room temperature. Stainless steel is used for the main shaft because it has a much lower thermal conductivity than other metals such as brass, aluminum, or copper and therefore results in less heating of the liquid helium.[40] At room temperature, the main shaft is connected to several KF-series vacuum components, seen in Figure 4.4(b). These include a vacuum port and valve for evacuating the tube and vacuum can, two hermetic dc wiring feedthroughs, a hermetic fiber optic feedthrough, an overpressure valve, and a thermocouple vacuum sensor.

The vacuum can is connected to the low temperature end of the main shaft. It contains the device that is being measured along with the RF switch, thermometer, heater, and the cleaved end of the optical fiber used to illuminate the SNSPD. An image of the internal components of the vacuum can is seen in Figure 4.4(c). The vacuum can is copper or brass[41] so as to maintain a constant temperature across its surface. Electrical wiring for the thermometer, heater, RF switch controls, dc connections for measurements of the film resistivity, as well as the optical fiber are fed through the inside of the main shaft into the vacuum can. Coaxial electrical lines for dc and RF measurements on

[40] The thermal conductivity of stainless steel at 200 K is approximately 15 W/m-K; the thermal conductivity of high purity copper at 200 K is approximately 400 W/m-K, of high purity aluminum at 200 K is approximately 200 W/m-K, and of brass at 200 K is approximately 100 W/m-K (Weisand 2003). For more information and an overview of materials considerations at low temperatures, see also Nicol *et al*. (1953), Zaitlin *et al*. (1972) and Rosenberg *et al*. (1955).

[41] Both vacuum can types were used, with no measurable difference in performance.

devices are fed through the top flange of the insert and are secured along the outside of the main shaft, passing through the liquid helium, where the bias tee, cryogenic amplifier, and other components are located (see section 4.1). A single coaxial line is fed into the vacuum can using a hermetic glass bead RF feedthrough.

In Figure 4.6, a schematic of the vacuum can is shown. The vacuum can, device mounting stage, and electrical wiring were designed to maintain the device at a stable temperature that can be adjusted using a heater and thermometer embedded in the stage. The stage is made of copper so as to maintain a constant temperature throughout and is much more massive than the device in order to provide a strong heat sink. The RF switch is connected to the copper stage using rigid stainless steel SMA barrels, which provide some thermal isolation from the sample stage. The device mounting stage/switch assembly is suspended and thermally isolated from the top flange of the vacuum can using stainless steel posts.[42] The low thermal conductance between the device mounting stage and the helium bath is such that some helium exchange gas is required to fully cool the stage to the bath temperature, since there is some heating of the sample/switch assembly due to heat conduction from room temperature along the cables. The pressure of helium exchange gas can be adjusted to change the thermal conductance, and thereby change the thermal time constant for the temperatures of the liquid helium bath and the sample to equilibrate. In most experiments, a large amount of exchange gas was used so that equilibration was fast, ~10 seconds. Much longer equilibration times may be obtained using much less exchange gas.

[42] The geometry of the components within the vacuum can was chosen to allow the assembly to fit within the liquid helium dewar, which has an inside diameter of approximately 3.5”.

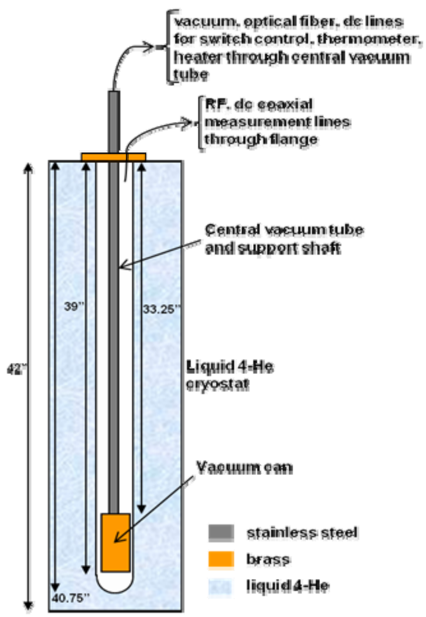

Figure 4.5: Schematic of major components of the cryogenic insert.

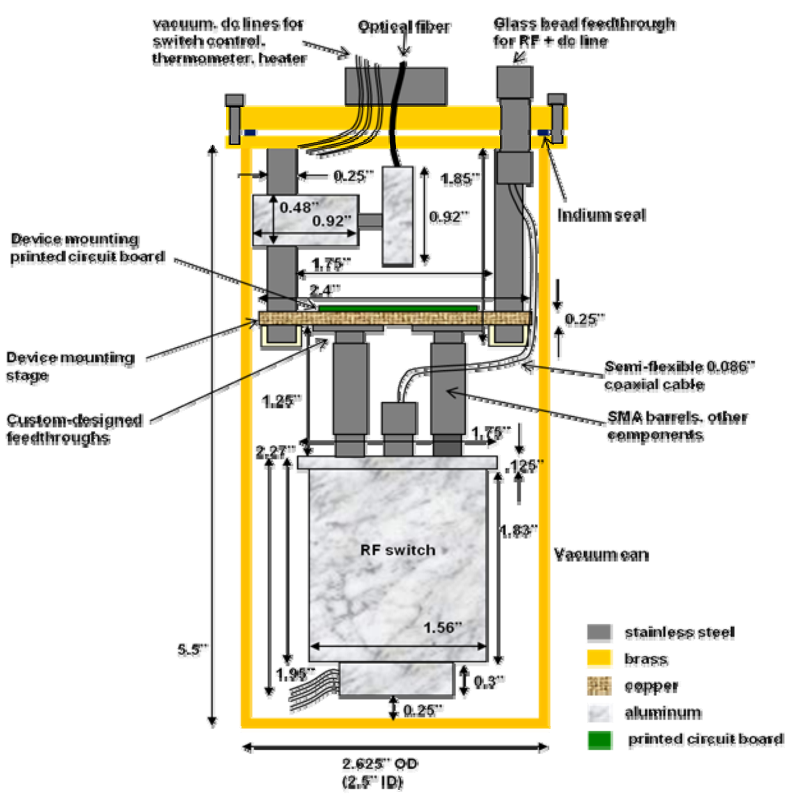

Figure 4.6: Schematic of vacuum can and internal components.

4.2.2 Thermometer, Heater, and RF Switch

The heater consists of a standard 100 Ω through-hole format resistor with a power rating of ¼ W. The thermometer is a Lakeshore Cernox model CX1050-AA-1.6L, factory calibrated from 1.6 K to 325 K. Both components are embedded within the copper device mounting stage so as to be equilibrated with each other and the device. The thermometer requires a four probe resistance measurement, and thus requires 4 dc wires. The heater requires 2 dc wires. The RF switch is operated with 6 magnetic solenoids located within unit. Operation of the RF switch requires 7 dc wires, one for each solenoid plus a common wire. These wires were fed down from room temperature within the main shaft. The wires to the heater were made of copper, so as to ensure nearly all electrical power dissipation was located in the heater resistor. All other wires carry little current, and so were made of manganin, which has a higher resistance and lower thermal conductivity than copper, reducing the heating due to thermal conduction from room temperature. [43] It should be noted that the operation of the Radiall model R573423600 RF switch only requires pulses of current of approximately 10 ms in duration when switching between channel configurations. The switch latches into a given configuration without the need for dc current. The RF switch is designed for room temperature operation, but it was found that by disassembling a new switch, removing the printed circuit boards, and making contact directly with the electrodes of the solenoid by soldering and supporting

[43] Manganin is a copper-manganese metal alloy. The thermal conductivity of manganin at 200 K is approximately 18 W/m-K; at low temperature (~2 K) it is ~0.2 W/m-K (Peroni 1999).

the solder joints by encasing them in Stycast 2651 epoxy, the switch may be reliably operated at cryogenic temperatures.

4.2.3 Sample Mounting Printed Circuit Board and Coaxial-Coplanar Transition

The device being measured is mounted on a printed circuit board (PCB) using GE varnish. Electrical contact to the planar transmission lines patterned in the PCB is made using wire bonding with 0.001" diameter aluminum wire. In general, several bonds are used to connect each device, and the lengths of each bond is kept as short as possible, so as to minimize extra inductance from the narrow wires. The PCB is fixed to the top side of a copper sample mounting stage, with the RF switch underneath (see Figure 4.6) and the coaxial inputs of the switch feeding through the stage (to be explained). The PCB is fabricated from a Rogers RT/duroid two-side coated microwave circuit board blank, which consists of 0.025" thick low-loss dielectric with dielectric constant of 10.2, coated on both sides with 0.5 oz. copper (0.7 mils = 18 μm). A schematic depicting the design of this PCB is shown in Figure 4.7. The bare square in the center is where the device chip is located. GE varnish is used to mechanically and thermally anchor the chip to the bare dielectric surface. On the right side of the figure, four dc contact pads used for four-probe resistance measurements of the Nb and NbN films are seen (see Figure 3.4 and discussion in section 3.6). The four strips radiating outward from the bare square in the center are the center conductors of the conductor-backed coplanar waveguides (CB-CPWs). The width and location of the strips at the center of the PCB is designed to match the width and location of the on-chip contact electrodes (see section 3.6 for a description of the chip layout). The CB-CPW is designed to transition with constant impedance of 50 Ω between the contact electrode patterned on the Nb SNSPD chip and a coaxial input to the RF switch. Since the switch is located underneath this PCB and sample mounting stage, the transition must be made at a right angle, through the PCB and device mounting stage.

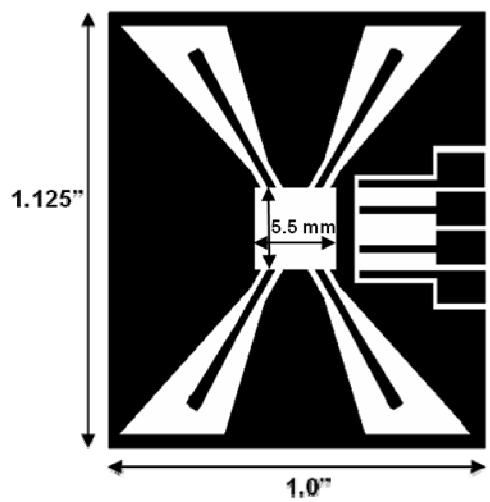

Figure 4.7: Schematic of sample mounting printed circuit board. The dark parts are copper, while the light parts are where the copper top layer has been etched away, exposing the planar dielectric.

The transition through the device mounting stage to the RF switch inputs is at a right-angle and was designed to reduce the mismatch between coplanar and coaxial electromagnetic modes so as to reduce the impedance mismatch. Impedance mismatches can cause reflections in the signal that would distort the shape of the measured voltage pulse that results from detecting a photon. Mismatches may also lead to reflections that cause the device to switch to the normal state at currents below the critical current (see discussion in chapter 5). The reactance associated with these right-angle, coaxial to CB-CPW transitions can be modeled as an equivalent circuit with an inductance and capacitance that results from the non-ideal nature of the transition geometry. The geometry of the transition is seen in Figure 4.8.

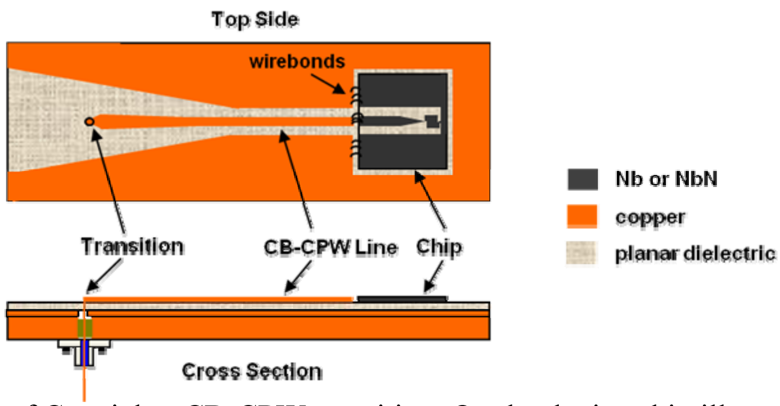

Figure 4.8: Schematic of Coaxial to CB-CPW transition. On the device chip illustrated here, only a single SNSPD device is shown (not to scale).

The origin of the parasitic reactance in a perpendicular transition such as this is well described by Morgan *et al.* (2002). The transition used in this thesis work is based on this work, as well as the work of Safwat *et al*. (2001) and Majewski *et al*. (1981). A significant source of additional inductance is caused by ground currents that must flow around the circumference of the coaxial outer conductor in order to reach the underside of the CB-CPW. Ideally, the transition geometry can be modified to compensate for this parasitic reactance. First, a glass bead is employed to center the coaxial inner conductor and maintain adequate separation when passing through the base plate. Next, a novel compensation technique outlined in Majewski *et al*. (1981) is employed. The aperture in the ground plane of the PCB is made smaller than the diameter of the glass bead, and also offset from the center conductor of the glass bead. This effectively shortens the path ground currents must travel to reach the underside of the CB-CPW, decreasing the inductance. It also concentrates excess capacitance near the side of the aperture where the ground inductance is least, thereby better balancing that area so that an impedance of approximately 50 Ω is maintained. Finally, it must be emphasized that the ground planes for all components of this transition and transmission line structures must be well connected using solder, wirebonds, and copper tape. In Figure 4.9, a plot of the magnitude of the transmission coefficient versus frequency through this transition shows that the impedance is well matched to 50 Ω ($|S_{21}| < 1$ dB) up to frequencies of approximately 8 GHz[44], which is significantly higher than needed for resolving the fast voltage pulses associated with single-photon detection in an SNSPD.

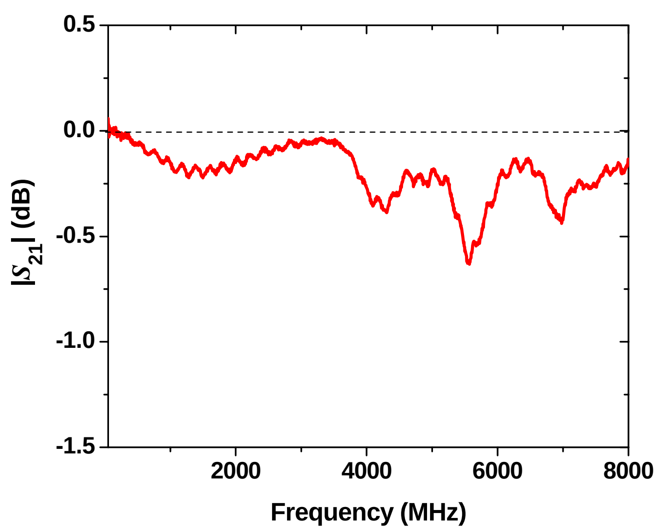

Figure 4.9: Plot of the magnitude of the power transmission coefficient, $|S_{21}|^2$, through a coaxial to CB-CPW transition versus frequency. As can be seen, there is near unity transmission (0 dB) through 8 GHz in this optimized transition structure.

[44] This measurement was done for two nominally identical transitions in series so as to measure the power transmission from coaxial to CB-CPW to coaxial, and then dividing the measured $|S_{21}(f)|$ in dB by 2.

4.2.4 Optical Fiber and Incident Power Characterization

The photon source is a pulsed diode laser, PicoQuant model PDL-800-B driver with 470 nm, 690 nm, and 1550 nm diode heads with an approximately 100 picosecond pulse width and 2.5-20 MHz repetition rate. Much of the jitter measured for single-photon detection and reported and discussed in chapter 5 is due to this finite width of the laser pulse. The average power of the laser is typically ~ 1mW and depends on the diode head used. The laser pulse is attenuated by up to 60 dB using a series of free-space neutral density filters from Ocean Optics. Additional attenuation occurs due to known coupling losses in the fiber optics. The optical fiber that carries laser pulses from room temperature to the detector being measured is an Ocean Optics model ZFQ5426 with an 80 μm core multimode silicon dioxide core coated with aluminum to promote thermalization but without additional cladding. It is contained within the main shaft of the insert, with the vacuum feedthrough located on the room temperature side of the insert, as seen in the upper part of Figure 4.4(b). The cold end of the fiber has been cleaved and left unconnectorized and suspended approximately 3 mm above and centered on the four devices on the chip being measured as seen in Figure 4.4(c) and Figure 4.6. Since the devices are close to each other on a chip (see section 3.6), and the spot size on the chip in this configuration is ~3 mm in diameter for most measurements, the illumination of the devices is uniform with +/- 10% variation in intensity across the spot. To measure the detection efficiency, the amount of power, or number of photons, incident on a device must be accurately determined. This is done by measuring the spot size and power at high optical power densities (low attenuation) using a photodiode power meter to determine the power density per unit area, and then multiplying by the area of a given device. A major source of variation in this measurement is due to interference effects within the fiber, which lead to light and dark regions within the laser spot that differ in intensity by a factor of approximately 2.5. These interference effects are measured and incorporated into the estimate for the standard deviation of the detection efficiency reported in chapter 5. Interference effects add uncertainty of approximately +/- 43% of the value of the detection efficiency and are the dominant source of uncertainty in the values reported in Table 5.2.

Chapter 5

Nb and NbN Single-Photon Detection Performance

5.1 Overview of Detector Performance Characterization

In this chapter, Nb and NbN superconducting nanowire single-photon detectors (SNSPDs) are studied using experiments that determine their detection performance for photons of several wavelengths. As discussed in chapter 3, several hundred chips of SNSPD devices were fabricated and screened for defects using SEM and AFM imaging as well as room temperature dc measurements. This screening procedure resulted in identifying many detectors that were free of lithography defects and therefore potentially fruitful to study. In order to determine which devices were free of internal defects and therefore the most interesting to study, a second stage screening method based on measurements of the dc superconducting properties of individual devices was developed. The procedure for dc screening is discussed in section 5.2. In section 5.3, detection performance for a representative sample of devices is discussed. First, a summary of this sample of devices chosen for study is given. After this, the measured detection performance of these devices, including the detection efficiency, dark counts, and jitter, is discussed. In section 5.4, the performance of the best Nb and NbN detectors is discussed. This shows that NbN SNSPDs are at least equal to and typically exceed the performance of Nb SNSPDs in most measures of detection performance.

5.2 Device Characterization and Screening at DC

In this section, the dc characterization and screening procedures for both Nb and NbN SNSPDs are discussed. Resistance versus temperature and current versus voltage data are presented for a variety of good devices. In addition, data from several defective devices are presented. The relation between certain features of these data, specifically the critical temperature, the critical current, and poor detector performance, is explained. This provides the basis for the screening procedure used to determine which devices would have the best detection performance.

5.2.1 Resistance Versus Temperature Measurements

One of the hallmarks of superconductivity is that when a superconductor is cooled below its critical temperature, a dc current can flow through it without power dissipation.[45] This zero resistance state is essential for the operation of SNSPDs. Measuring the resistance versus temperature of a device is an important characterization procedure and the first step of the screening process. Devices that do not display a full superconducting transition at temperatures that are at least twice the desired operating temperature are not useful, as such devices would have low critical current and therefore a small signal, and less than ideal detection efficiency (see section 3.1 and section 5.3). A typical resistance versus temperature curve for a Nb SNSPD device with thickness of 7.5 nm, width of 100 nm, and length of 500 μm ($R_{\square}$ = 105 Ω/□) is seen in Figure 5.1. This

[45] The other hallmark of superconductivity is the Meisner effect, which is the expulsion of magnetic flux from the interior of a superconducting material.

device was fabricated according to the procedure outlined in chapter 3, including the Ar^+ ion etching step, which reduced the original device thickness from the directly sputtered value of 14 nm to the final thickness of 7.5 nm.[46] The critical temperature is defined as the temperature where the device resistance, R_d, is equal to one-half the normal state resistance: $R_d = R_n/2$. This is referred to as the midpoint of the transition. The transition width, ΔT_c, is defined as the difference in temperature between when $R_d = 0.9R_n$ and when $R_d = 0.1R_n$.

The critical temperature and transition width are important parameters for the superconductivity of the nanowire. For thin films, the critical temperature depends strongly on the type of material and on its thickness, but for Nb it is found to depend weakly on the resistivity of the film. The critical temperature of Nb devices also depends strongly on the nanowire width when the width is less than 100 nm. The transition width is a measure of how uniform the superconductivity is along the wire. For narrower and thinner wires, ΔT_c is typically larger since local impurities, grain boundaries, and lattice defects, as well as intrinsic temperature-dependent resistance fluctuations such as caused by vortex-antivortex unbinding, are more likely to suppress superconductivity across the wire when the wire is narrow and/or thin. In Figure 5.2, the normalized resistance versus temperature is plotted for three devices with the same width, $w_d = 100$ nm. In red (solid line) is the same device (7.5 nm thick) as plotted in Figure 5.1. In blue (dashes) is a Nb SNSPD with similar thickness (8.5 nm) but with half the resistivity as the device in red. This device was fabricated according to the procedure outlined in chapter 3, but

[46] As will be explained in section 5.3, Ar^+ etching to thin the film is extremely important in enabling high detection efficiency in Nb devices.

excluding the Ar^+ ion etching step; the film was directly sputtered to a thickness of 8.5 nm. As can be seen, resistivity does not strongly influence the critical temperature or transition width. In green (dots) is a Nb device that has lower resistivity than either of the other devices, and is also thicker (14 nm). As is clear from the plots, thinner devices have lower critical temperatures and wider transition widths, fairly independent of the resistivity of the film. These results are typical for Nb SNSPDs.

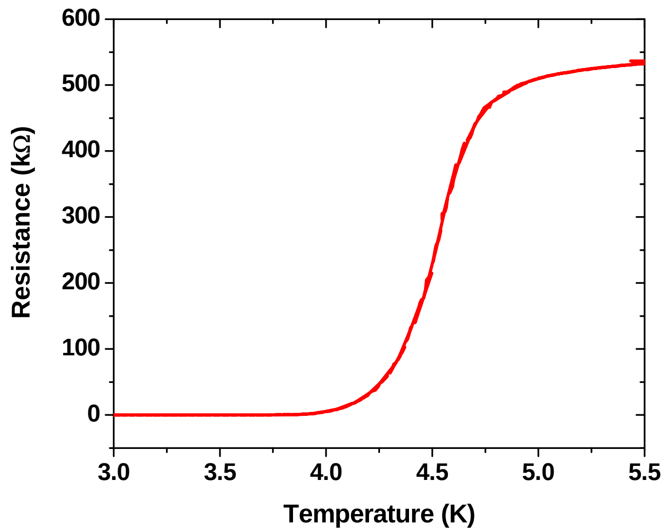

Figure 5.1: Plot of the resistance versus temperature for a 7.5 nm thick, Ar^+ ion etched Nb SNSPD. For this device, $T_c = 4.5$ K (defined at the midpoint of transition) with $\Delta T_c = 0.55$ K (defined as the width from $0.1R_n$ to $0.9R_n$).

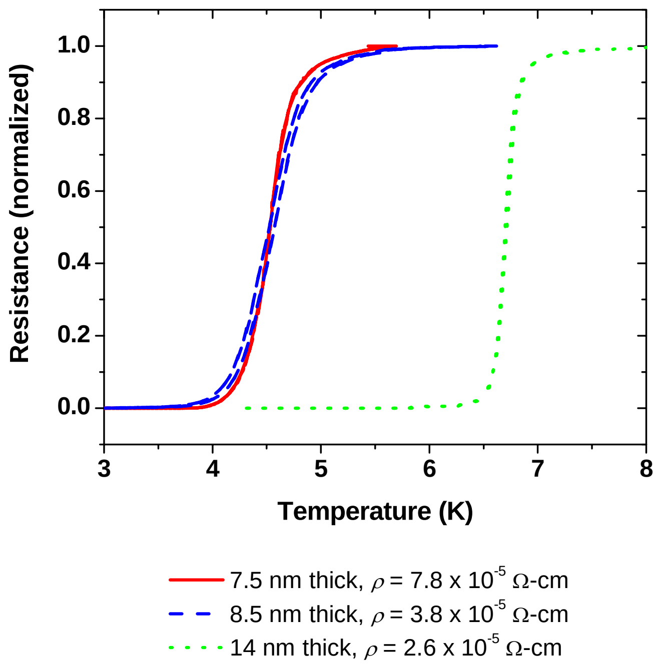

Figure 5.2: Plot of the resistance versus temperature for several Nb SNSPDs of varying thickness and resistivity, with the same width of 100 nm.

The critical temperature also depends on the type of film. In Figure 5.3, resistance versus temperature curves are plotted in red (solid line) for the same Nb SNSPD as in Figure 5.1, in green (dots) for the thick Nb device from Figure 5.2, and in brown (alternating dashes and dots) for a NbN SNSPD with a film thickness of 5 nm and a width of 130 nm. Although the NbN device is significantly thinner than either of the Nb devices, the superconducting state is much stronger in NbN, and so it has a much higher critical temperature. Although this difference in critical temperature between Nb and NbN is well known, it is noted here for pedagogical reasons. Again, these results are typical of many devices tested.

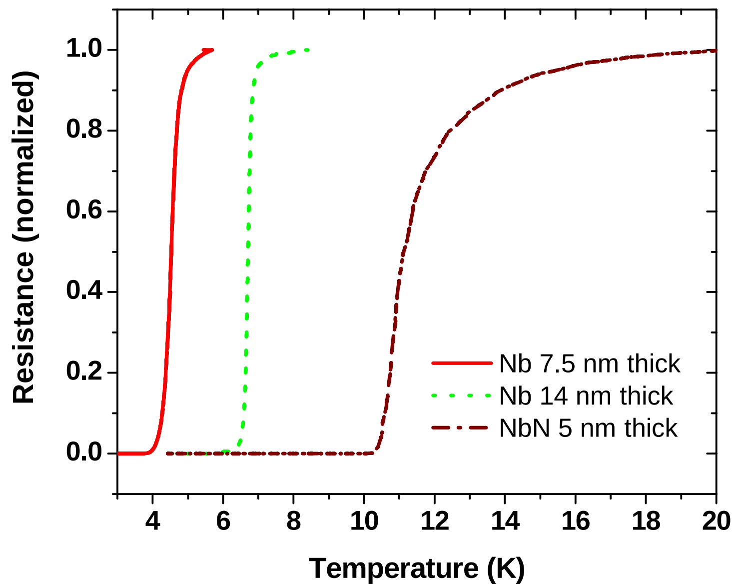

Figure 5.3: Plot of the resistance versus temperature for Nb and NbN SNSPDs.

The superconductivity in nanowires also depends on the width of the nanowire. This is extremely significant, because narrow nanowires are desired in order to reduce the volume of material that the absorbed photon needs to drive into the normal state in order to create a resistive hotspot (see section 2.4). In NbN SNSPDs, nanowires as narrow as 50 nm were shown to display superconductivity well above the typical operating temperature of ~2 K (Yang 2009). These narrow wires can achieve detection efficiency that is as high as wider wires, but with bias currents further below I_c than wider wires (see section 2.4). Because of the reduced volume that needs to be heated by the absorbed photon, narrower nanowires may also be sensitive to longer wavelength photons. In contrast, most Nb devices with widths significantly less than 100 nm did not have high enough critical temperatures, or narrow enough transition widths, to be useful at a typical operating temperature of ~2 K. Figure 5.4 shows resistance versus temperature curves for Nb devices with widths of 100 nm and 70 nm and thicknesses of 14 nm. The device with a width of 100 nm is the same as in Figures 5.2 and 5.3 in green (dots). These results are typical of all Nb devices tested. Devices with a width that was much less than 70 nm did not fully transition to the superconducting state at all, even at temperatures as low as 1.6 K.

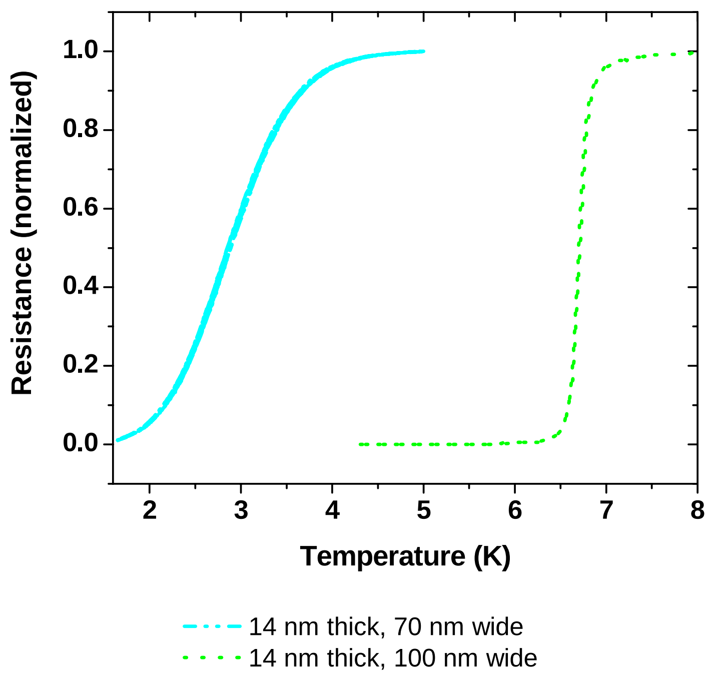

Figure 5.4: Plot of the resistance versus temperature for Nb SNSPDs with the same thickness, but with differing widths of 70 nm and 100 nm.

5.2.2 Current Versus Voltage Measurements

A superconductor with a temperature below the critical temperature will carry a dc current without dissipation if the value of that current is less than the critical current of the superconductor. The critical current of a nanowire depends on both the type of material and on the geometry. Since a dc current flowing without dissipation does not produce a dc voltage, the superconducting state is sometimes referred to as the zero voltage state, although there is only zero voltage for pure dc currents ($\omega = 0$).[47] In an SNSPD with a dc device current, I_d, and a critical current, $I_c(T)$, the voltage across the device will be zero if $I_d < I_c$. For $I_d > I_c$, the device voltage will simply be $V_d = I_d R_n$ if the critical current is uniform, that is, if it is the same at every point along the length of the nanowire. In this ideal case, there is no stable intermediate state where $0 < R_d < R_n$.

A typical plot of dc device current versus dc device voltage in this case is seen in Figure 5.5. All measurements in this section are of devices in the dark (no optical photons incident on the detectors) and with the device connected according to Figure 4.2 with a 50 Ω RF readout line terminated by an impedance matched amplifier powered on. All measurements except those in Figure 5.8 have no additional resistance in parallel to the amplifier, giving a total RF load resistance $R_L = 50\ \Omega$. It is important to ensure that the

[47] For time-dependent currents, there will be an inductive voltage across the device due to the nanowire kinetic inductance, which yields an impedance of $i\omega L_K$ where ω is the angular frequency of the current. The inductive voltage will be important to consider in section 5.3 and chapter 6.

RF line is terminated with an impedance-matched load (amplifier powered on), otherwise RF reflections from intrinsic resistance fluctuations at high current (explained in section 2.7 and chapter 6) may reduce the measured critical current by causing the device to switch prematurely to the normal state. The curve in Figure 5.5 is measured at a temperature of 1.7 K by increasing the current slowly from zero (a) until a voltage is measured. The current at which this voltage is first measured (b) is known as the switching current. With no external noise, and for a uniform nanowire, this should be equal to the Ginzburg-Landau depairing critical current of the device (Tinkham 1996).

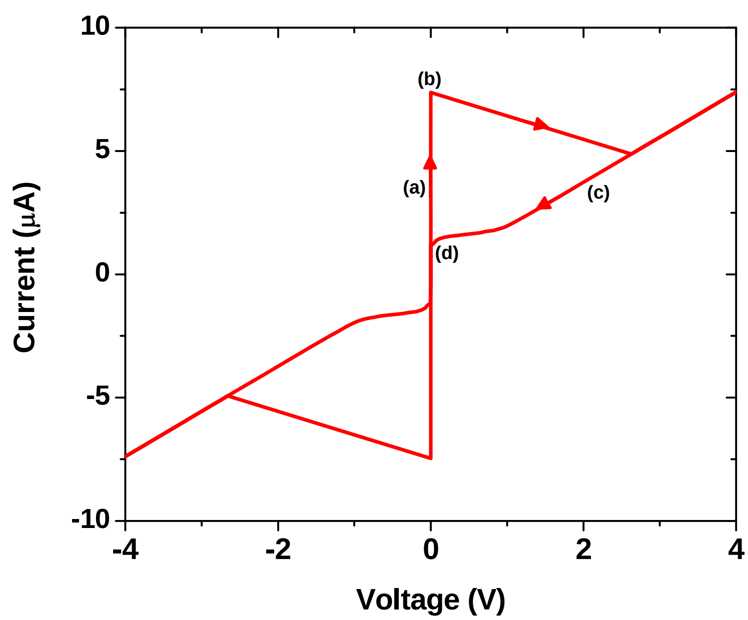

Figure 5.5: Plot of the device current versus device voltage at 1.7 K for a typical Nb SNSPD in a dark environment (no optical photons incident on the device). Here, the slope of the transition from the superconducting state to the normal state is given by the load line of the dc bias circuit. For this measurement, a 1 MΩ dc bias resistor was used and $R_n \approx 550$ kΩ. This device is the same as in Figure 5.1.

In real measurements, the switching current can be ~10% below the actual critical current, depending on the device kinetic inductance and the RF resistance of the readout circuit. This will be discussed in the last paragraph of this section and elaborated further in chapter 6. The slope of the transition from the superconducting state to the normal state, (b) to (c) in Figure 5.5, is given by the load line of the dc bias circuit. For this measurement, a 1 MΩ dc bias resistor was used, so the load line has a slope given by to $-(1\ \mathrm{M}\Omega)^{-1}$ and the device had a normal state resistance of $R_n \approx 550$ kΩ. As the current is reduced from the normal resistive state, the device remains normal even at currents less than the switching current (c). This hysteresis is due to self heating of the device when it is in the resistive state (Skocpol 1974). The device current must be reduced to a value less than the return current, I_r, before superconductivity is restored. This return current is a measure of the cooling capacity of the nanowire: $I_r^2 R_n$ is the maximum amount of Joule heating that the nanowire can dissipate in steady state at the measurement temperature.

As in the case of the resistance versus temperature, the current versus voltage curve depends on the device geometry as well as the type of material. In Figure 5.6, the current versus voltage is plotted for the same Nb and NbN devices as in Figure 5.3. Note that in this graph, the axes are expanded so as to accurately show the switching currents in all three curves. The arrows give the current sweep direction. As can be seen, the thicker Nb device has a switching current that is greater than the thinner Nb device by more than the ratio of the device thicknesses; thus the critical current per area of cross-section is larger in the thicker device, which is consistent with the higher critical temperature (Figure 5.2). In the NbN device, a larger switching current is measured than in the Nb devices. This is consistent with the higher critical temperature of NbN films compared to any of the Nb films (Figure 5.3).

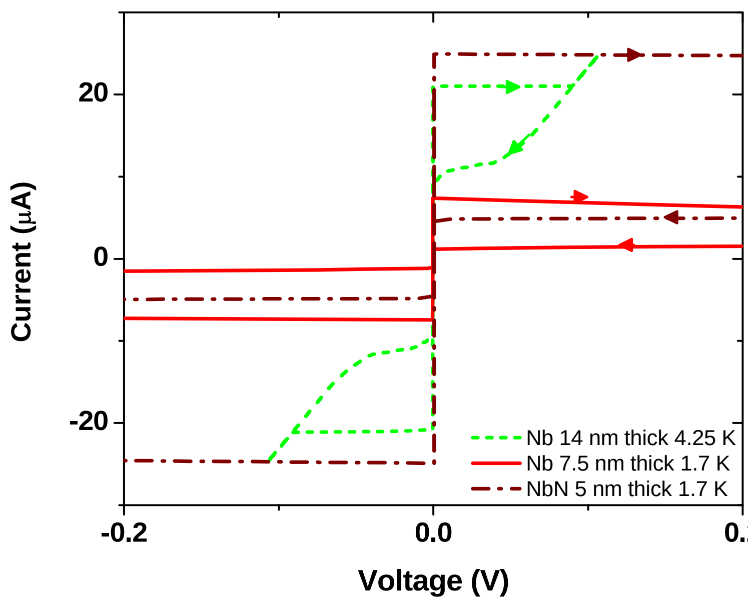

Figure 5.6: Plot of the device current versus device voltage for Nb and NbN SNSPDs in a dark environment (no optical photons incident on the device). These are the same devices as measured in Figure 5.3. The thin Nb and NbN devices are measured at 1.7 K, while the thick Nb device is measured at 4.2 K.

There are many non-ideal cases where the current versus voltage curve does not have the form seen in Figures 5.5 and 5.6. In some cases, part of the device will have a reduced switching current, which results from a non-uniform critical current along the length of the wire. The non-uniform critical current is usually due to one or more sections of the nanowire having a reduced critical current or critical temperature due to impurities, grain boundaries, or lattice defects. This has been studied for NbN SNSPDs in detail by Kerman *et al.* (2007). In that work, these defects are referred to as constrictions, and are shown using measurements of the detection efficiency and kinetic inductance to result from a localized reduction in the superconducting cross section of the nanowire.

In this work, it was determined that if the current versus voltage is measured with a bias resistor that is similar to the normal state resistance of the device (as is done in Figure 5.5), then individual steps can be observed in the current versus voltage curve that correlate to possible constrictions in these non-ideal devices. An example of a severely defected Nb device is seen in Figure 5.7. This device has the same thickness, length, and width as the thick Nb device in Figure 5.6 and not defects in the wire are observable in SEM or AFM imaging. As is clear, a constricted device cannot be current biased near the critical current of most parts of the wire, but only near the critical current of the section wire that is constricted. Thus the detection efficiency of the unconstructed sections of the wire, which is the majority of the wire, is significantly reduced, as we confirmed with measurements of the detection efficiency. In this work, it is found that this type of dc characterization of the current versus voltage with the proper load line is sufficient to determine whether a device is defective, while in Kerman *et al*. (2007), RF and photon detection measurements are required. This thesis work shows that devices with current versus voltage curves like that seen in Figure 5.7 will not have high detection efficiency. This is the second stage of the dc screening process (after determining that the critical temperature is high enough and the transition width is narrow enough).

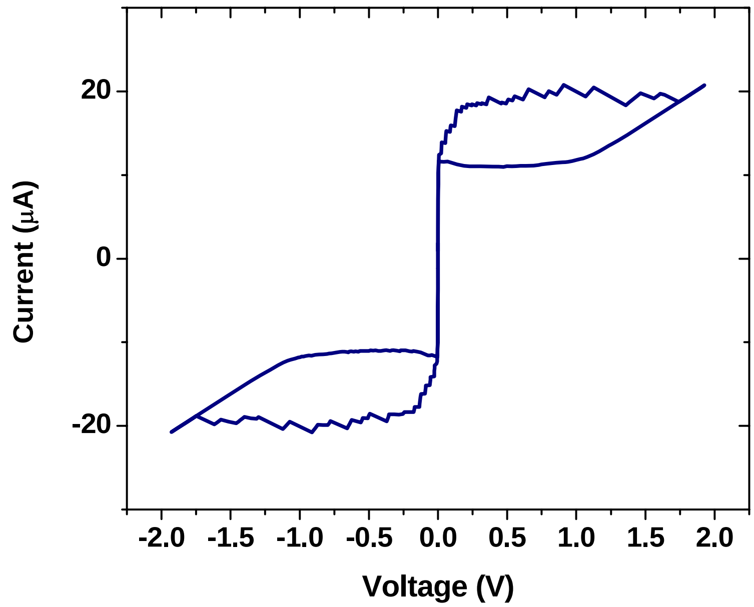

Figure 5.7: Plot of the device current versus device voltage at 4.2 K for a defective Nb SNSPD with width of 100 nm and thickness of 14 nm using a dc bias resistor of 100 kΩ. This device is the same width and thickness as the thick Nb device plotted in Figure 5.6 and should thus have the same switching current and overall shape.

In general, the switching current is not exactly equal to the critical current. The relation between the switching current and the critical current depends on the readout circuit. It was found that the dc bias resistance does not affect the measured switching current when the measurement is done using the circuit detailed in Figure 4.2. However, the RF load that is presented to the device does affect the measured switching current. For finite temperatures, it is only when a dc-coupled, RF-bandwidth, low resistance shunt is added (via the RF switch discussed in section 4.1) that the switching current approaches the critical current at that temperature. In Figure 5.8, a plot of the dependence of the switching current on the total RF load resistance, R_L, is shown. This is obtained from the current versus voltage curves measured using various values of R_L. Here, R_L is given by the amplifier input impedance (50 Ω) in parallel with one or more 50 Ω chip resistors connected via the remote-controlled RF switch (see section 4.1). The dashed lines indicate the estimated value of the critical current for each device. The critical current is obtained by measuring the current versus voltage using the smallest value of the load resistance. This results in the largest measured switching current, which is taken to be equivalent to the critical current when, for the smallest and next smallest value of the load resistance, the switching current rises only by a small amount. As can be seen, the switching current continues to rise slightly even when the load resistance is reduced below that of the amplifier alone. Thus, it is not only RF reflections that cause the device to switch prematurely to the normal state.

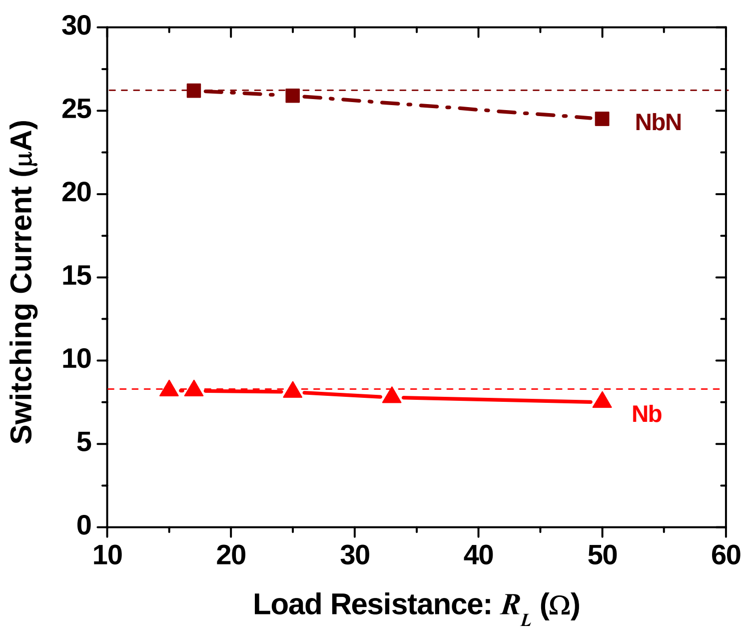

Figure 5.8: Plot of the switching current versus the load resistance for Nb and NbN devices, obtained from current versus voltage measurements at 1.7 K. The dashed lines indicate the estimated value of the critical current for each device. The Nb device is the same as in Figure 5.1 and 5.5, and the NbN device is the same as in Figure 5.3 and 5.6.

The reason that a smaller load resistance yields a higher switching current is the same reason a smaller load leads to a higher latching current (see section 2.6 and chapter 6). It is believed that the measurement environment has no stray photons. In such a completely dark environment, the switching current is simply the latching current due to intrinsic dark counts. When the device is biased with a current just below the critical current, intrinsic thermal or quantum fluctuations lead to the formation of spurious resistive hotspots, which are referred to as dark counts. Since these dark counts only occur near the critical current (see section 2.7), a switching current much below the critical current is never measured. Latching is discussed in detail in chapter 6.

Although the switching current depends on the load resistance, for small loads, the switching current approaches appears to saturate at a value that depends only on temperature. This value of current is believed to be the intrinsic critical current. According to the Ginzburg-Landau theory, the temperature dependence of the intrinsic critical current is particularly strong at temperatures close to the critical temperature; in that regime, the Ginzburg-Landau predicts: $I_c(T) = I_c(0)(1\text{-}T/T_c)^{3/2}$ (Tinkham 1996). Using values of the load resistance that allow for measuring a switching current that is believed to be very nearly equal to the critical current, the critical current versus temperature curves for two Nb SNSPDs and a NbN SNSPD (the same devices as measured in Figure 5.6) were obtained. Plots of $I_c(T)$ versus T for several devices are shown in Figure 5.9. The solid lines are fits using the Ginzburg-Landau expression for the temperature dependence of the critical current; each curve is forced to intersect with the measured data at or near $T = 0.75T_c$. As can be seen, the above Ginzburg-Landau expression for $I_c(T)$ is at best only a rough approximation for $T > T_c/2$.

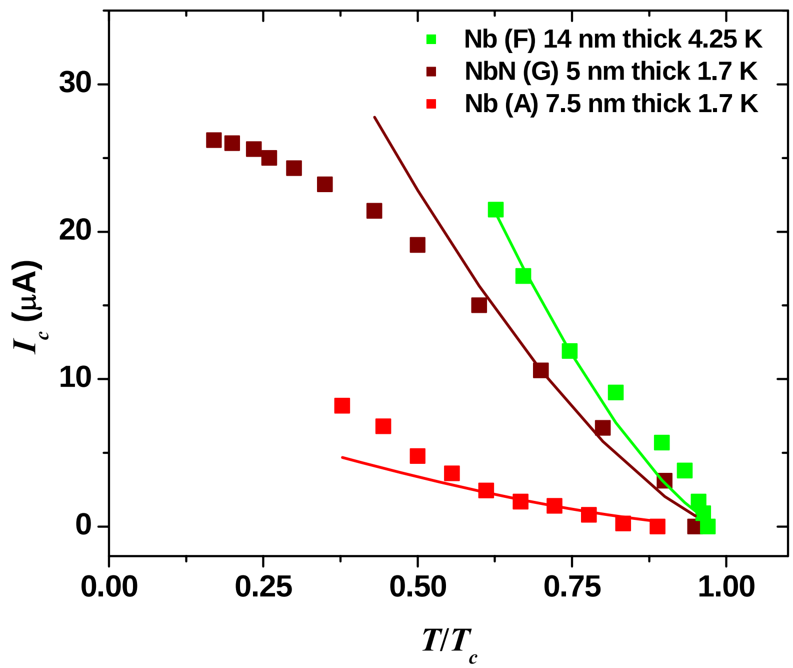

Figure 5.9: Plot of the critical current versus temperature for Nb and NbN devices, obtained from current versus voltage measurements at temperatures from 1.7 K to 10 K. Measurements (symbols) are of the same three devices as in Figure 5.6. The solid lines are fits using the Ginzburg Landau expression for the temperature dependence of the critical current; each curve is forced to intersect with the measured data at or near $T = 0.75T_c$.

5.3 Detection Performance

Using the dc characterization and screening procedure outlined in the previous section, devices with optimum values of the critical temperature, transition width, and critical current were chosen for careful study. A selection of typical devices is now analyzed in terms of the detection performance. A summary of the devices to be discussed in this section is seen in Table 5.1. This selection does not include all devices tested for detection performance. Rather, it includes a representative sample of devices exhibiting the full range of parameters studied. Only the devices listed in Table 5.1 are discussed in the remainder of this thesis. In Table 5.2 is a summary of the detection and reset performance of all devices from Table 5.1. In this section, measurements of the detection performance is discussed. The reset time is discussed in detail in chapter 6.

Device	**Length (l_d)**	**Width (w_d)**	**Thickness (d_d)**	**Resistivity (ρ)**	**Sheet Res. ($R_\square$)**	**T_c**	**I_c at 4.2 K**	**I_c at 1.7 K**	**Ar^+ Ion Etched?**
A (Nb)	500 μm	100 nm	7.5 nm	7.8×10^{-5} Ω-cm	105 Ω/□	4.5 K	0 μA	8.2 μA	Yes
B (Nb)	80 μm	100 nm	7.5 nm	8.2×10^{-5} Ω-cm	110 Ω/□	3.9 K	0 μA	6.2 μA	Yes
C (Nb)	10 μm	100 nm	7.5 nm	8.2×10^{-5} Ω-cm	110 Ω/□	4.0 K	0 μA	6.4 μA	Yes
D (Nb)	300 μm	100 nm	8.5 nm	3.8×10^{-5} Ω-cm	45 Ω/□	4.5 K	0 μA	7.4 μA	No
E (Nb)	20 μm	100 nm	14 nm	2.6×10^{-5} Ω-cm	18 Ω/□	6.7 K	21.0 μA	-	No
F (Nb)	500 μm	100 nm	14 nm	2.6×10^{-4} Ω-cm	18 Ω/□	6.6 K	22 μA	-	No
G (NbN)	105 μm	130 nm	5 nm	4.4×10^{-4} Ω-cm	875 Ω/□	10 K	20.2 μA	26.2 μA	No
H (NbN)	5 μm	130 nm	5 nm	4.4×10^{-4} Ω-cm	875 Ω/□	10 K	21.4 μA	25.4 μA	No

Table 5.1: A summary of the parameters of the devices that were examined in detail in this thesis. The normal state resistance is taken at a temperature approximately 5 degrees above the critical temperature.

Device	Area (A_d)	Inductance at 1.7 K ($L_K + L_M$)	Detection Efficiency at 470 nm	Detection Efficiency at 1550 nm	Reset Time	Dark Counts	Jitter ($\Delta\tau$)
A (Nb)	100 μm^2	235 nH	6 % +/- 2%	0.3 % +/- 0.2%	28.2 ns	~0.1	<110 ps
B (Nb)	16 μm^2	60 nH	5 % +/- 1.7%	-	10.6 ns	~0.1	<110 ps
C (Nb)	2 μm^2	15 nH	5 % +/- 1.7%	-	6 ns (est.)	~0.1	<110 ps
D (Nb)	60 μm^2	47 nH (est.)	≈10^{-2} %	-	8.3 ns	-	-
E (Nb)	4 μm^2	7 nH	~10^{-5} %	-	5 ns (est.)	-	-
F (Nb)	100 μm^2	25 nH	~10^{-5} %	-	7 ns	-	-
G (NbN)	25 μm^2	120 nH	4.8 % +/- 1.5%	4.5% +/- 2.9%	14.4 ns	~0.1	-
H (NbN)	1.3 μm^2	16 nH	4 % +/- 1.3%	3.8%+/- 2.6%	2.8 ns	~0.1	-

Table 5.2: A summary of the detection performance of the devices examined in detail in this thesis. The inductance was measured at 1.7 K with no dc current flowing in the device, and includes a contribution from the magnetic inductance of the leads (primarily from wire bonds).[48] The detection efficiency, jitter, and dark counts were measured with a current bias $I_b = 0.95I_c(T_o)$ at an operating temperature $T_o = 1.7$ K except for measurements of devices (E) and (F), which were done at an operating temperature $T_o = 4.2$ K. The reset time is equal to $3\tau_r = 3(L_K+L_M)/R_{L,max}$ where $R_{L,max}$ is the largest resistance that does not result in latching when $I_b = 0.95I_c(T_o)$ (see section 2.6 and chapter 6). The magnetic inductance, L_M, ranged between 5 and 10 nH, depending on the device and the length of the wire bonds used to connect to it.[49]

5.3.1 Detection Efficiency

The single-photon detection efficiency was measured for both Nb and NbN SNSPDs for photon wavelengths of 470 nm, 690 nm, and 1550 nm at several values of the bias current and at several operating temperatures. A typical single-photon voltage pulse is seen in Figure 5.10 for device (A) operating at 1.7 K with a bias current of 5.0 μA and tested with 470 nm photons. This pulse shows the ideal self-resetting case of operation (see section 2.6 and chapter 6) where the device returns to the normal state after detecting a photon without reducing the bias current. The amplitude of the voltage pulse

[48] Note that in the measurements of L_K discussed in chapter 6, the lead inductance has been subtracted out, or in some cases was eliminated by reducing the length of wire bonds and combining more in parallel.

[49] In the measurements of the kinetic inductance in section 6.2 of this thesis, this magnetic component of the inductance was mostly eliminated. However, the magnetic contribution was present when the latching and detection efficiency of the devices was investigated, so it is included here and in all calculations and measurements in this thesis except those in section 6.2.

is given by $(I_b - I_r) \times R_L \approx I_b \times R_L$. Since I_r is small, the approximate equality usually holds. The rise time of the pulse scales approximately as $L_K/(\max(R_d(t)))$, since $R_d(t) >> R_L$ for almost all times while the hotspot is resistive and the maximum resistance is developed quickly (see section 2.6). The fall time of the pulse is given by $\tau_r = L_K/R_L$; τ_r is referred to as the current return time (section 2.6). The asymmetry in the rise and fall times of the pulse indicates that the maximum device resistance (proportional to the length of the nanowire that is driven normal) is usually much greater than the load; here, $R_L = 50\ \Omega$. The reset time is defined as $3\tau_r$, which is the time needed for 95% of the bias current that is shunted into the amplifier to return to the device. The reason for this definition will become clear in chapter 6, where the reset of SNSPDs is studied in more detail.

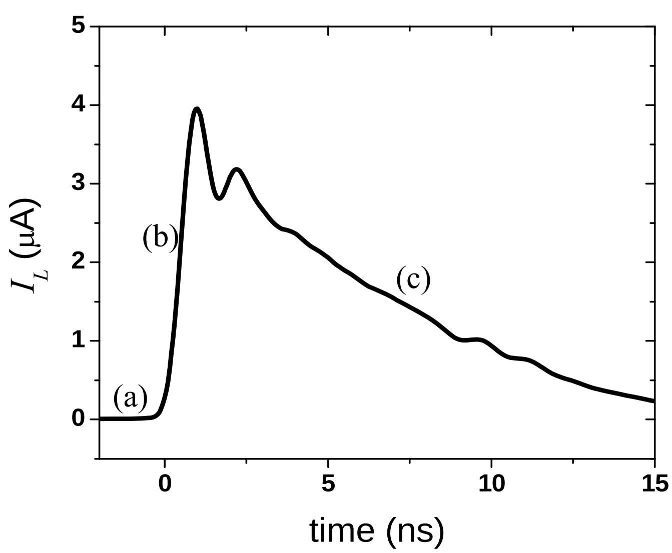

Figure 5.10: Single-photon current pulse measured for device (A) with $R_L = 50\ \Omega$, $I_b = 5.0\ \mu A$, and $T_o = 1.7$ K. The plot shows the self-resetting case (see section 2.6); three regimes are visible: (a) the device is in equilibrium with $R_d = 0$; (b) a photon has been absorbed, the hotspot is growing and the current is transferring into the load in a transfer time ~ $L_K/(R_L+R_d)$; (c) the hotspot resistance has returned to zero, and the current is returning to the device with a time constant $\tau_r = L_K/R_L$. The ripple near the pulse peak is due to small RF reflections in the readout line.

To show that the observed voltage pulses are due to single-photon-initiated events, the Poissonian statistics of the laser photon number distribution are utilized. For a pulse of laser light, the number of photons contained in each pulse is Poisson distributed about an average photon number, n_{avg}. The average continuous power of a laser is proportional to n_{avg}. In the limit that $n_{avg} << 1$, most pulses have zero photons, and occasionally some pulses have one; very few pulses have two photons. In this limit the probability of detecting a single-photon is proportional to n_{avg}. Thus, the detection rate should scale linearly with the average laser power for power levels where $n_{avg} << 1$. Since only photons that are absorbed are counted, even values of $n_{avg} \sim 1$ resulted in a linear dependence of the detection rate on laser power. Since the laser used had a high repetition rate (20 MHz), the average power incident on the detector could be as high as ~5 pW in these measurements. The very high timing resolution of SNSPDs allows for detection and full reset within the time between incident pulses, 1/(20 MHz) = 50 ns, allowing this high pulse rate laser to be used.

The detection rate was measured by counting only the number of voltage pulses that were synchronized (+/- 100 ps) with the laser pulse trigger signal. A schematic of the counting method is seen in Figure 5.11. This method of synchronous counting avoids counting dark counts (discussed later in this section) and therefore leads to a more accurate determination of the detection efficiency. In Figure 5.12, a plot of the detected fraction (detection rate divided by the laser repetition rate) versus laser power is shown for device (A), a Nb SNSPD, for photons with a wavelength of 470 nm and with $I_b = 0.95 I_c(1.7\text{K})$ and $R_L = 17\ \Omega$. The laser power was calculated in terms of the average number of photons per pulse, based on measurements of the average laser power at higher average power than used for these measurements, and based on an accurate characterization of the optical attenuators. This is described in section 4.2.4. This type of plot was obtained for each device studied. In this way, it was shown that all data discussed in this thesis (and reported in table 5.2) were for single-photon detection. In is useful to note that the detection rate still scaled linearly with the laser power even when the bias current was significantly less than the critical current. Also, for devices with very low single-photon detection efficiency at high bias currents (devices (D, E, F) from table 5.2), the count rate scaled linearly with the laser power. When the laser power is increased substantially, such that $n_{ave} >> 1$, the count rate is no longer linear with the laser power but saturates at unity. This is shown in Figure 5.13 for device (A) for photons with a wavelength of 470 nm and with $I_b = 0.95I_c(1.7\text{K})$ and $R_L = 17\ \Omega$. Note here that at low average photon number per pulse, the slope of the curve plotted in Figure 5.13 is the same as the slope of the curve plotted in Figure 5.12.

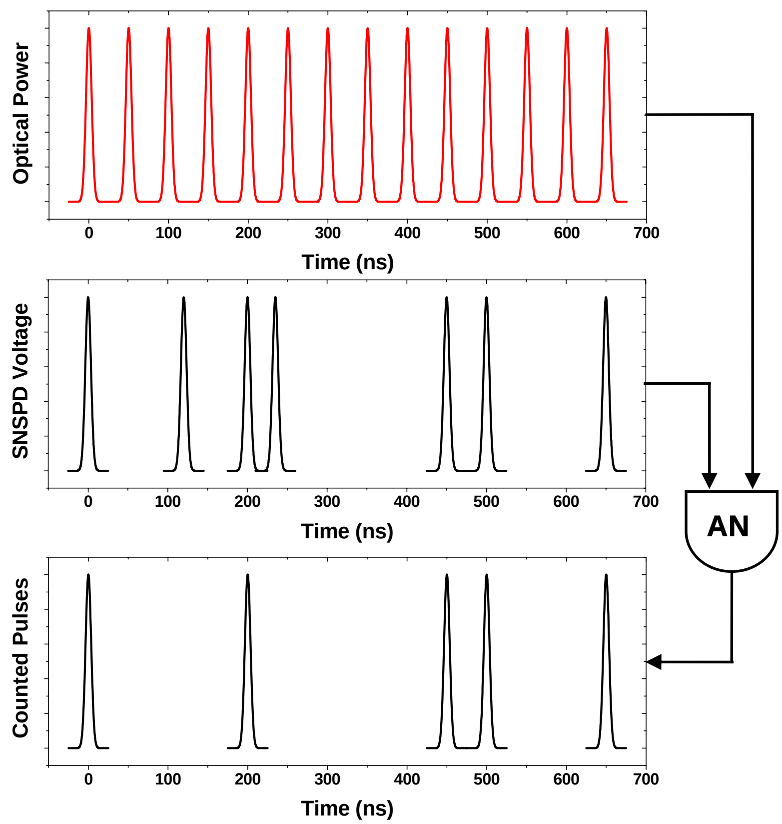

Figure 5.11: The photon count rate (bottom) is measured by counting the number of voltage pulses from the device per unit time (middle) that are synchronized with the laser trigger pulse (top). The laser trigger is synchronized with the optical pulse that excites the detector. In these measurements, the average number of photons per optical pulse is much less than one. Some voltage pulses from the device are not synchronized with the laser pulse (shown in the center plot), but are the result of dark counts. This method of synchronous counting avoids counting dark counts.

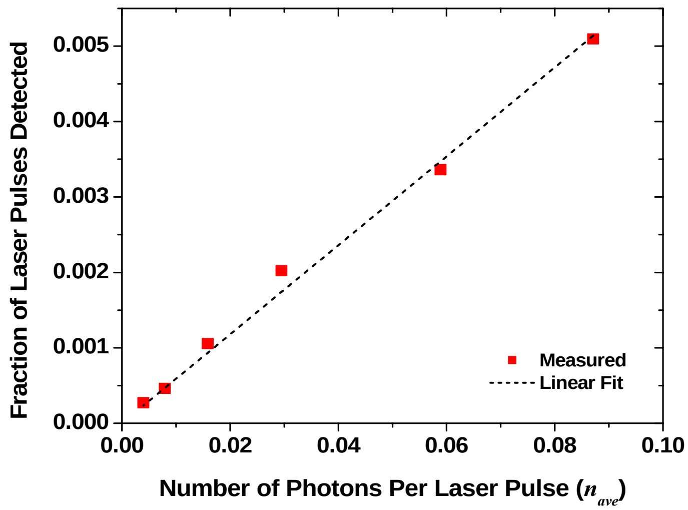

Figure 5.12: Plot of detected fraction (voltage pulse count rate divided by laser repetition rate) versus laser power for low laser power levels where the average number of photons absorbed by the detector from each laser pulse is much less than one. Data from device (A), a Nb SNSPD, for photons with a wavelength of 470 nm and with $I_b = 0.95I_c(1.7\text{K})$ and $R_L = 17\ \Omega$.

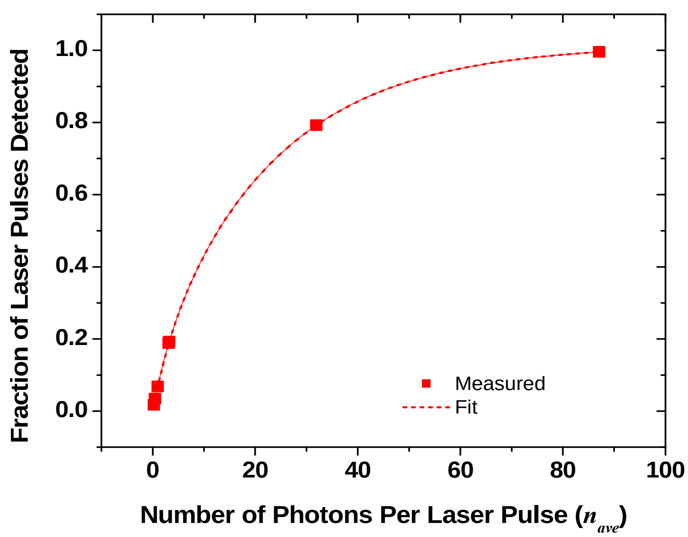

Figure 5.13: Plot of detected fraction (voltage pulse count rate divided by laser repetition rate) versus laser power up to high laser power levels where the average number of photons absorbed by the detector is much greater than one. Data from device (A), a Nb SNSPD, for photons with a wavelength of 470 nm and with $I_b = 0.95I_c(1.7\text{K})$ and $R_L = 17\ \Omega$. Note that the slope for low average number of photons per pulse is the same as the slope of the curve in Figure 5.12.

Once it is shown that measured voltage pulses are initiated by single photons, the detection efficiency can be determined. To determine the detection efficiency from the count rate requires accurately measuring the number of photons that are incident on the detection area. The procedure for doing this is explained in chapter 4. The results of measuring the detection efficiency for 470 nm, 690 nm, and 1550 nm photons are shown in Figure 5.14 for device (A) with $I_b = 0.95I_c(1.7\mathrm{K})$ and $R_L = 17\ \Omega$. As can be seen, the detection efficiency for 690 nm photons is substantially less than the detection efficiency for 470 nm photons; the detection efficiency for 1550 nm photons is less still. This can be understood from the hotspot formation process, described in section 2.6. A higher energy photon is more likely to form a resistive hotspot, because its greater energy increases the local temperature of the superconductor by a greater amount. Plotted results are typical for all Nb devices that were thinned with the Ar^+ ion etching process (devices (A), (B), and (C), as well as others not reported in this thesis).

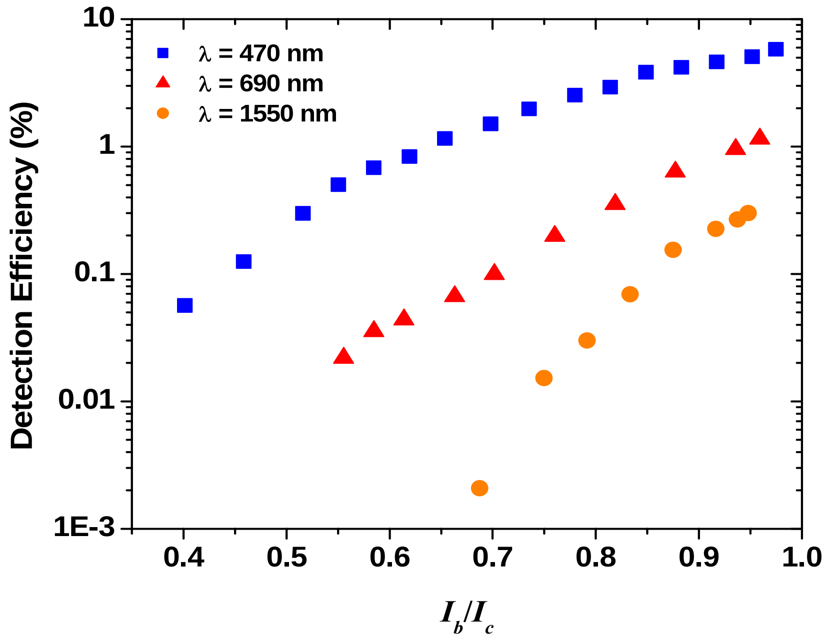

Figure 5.14: Plot of detection efficiency versus bias current, normalized to the critical current at the temperature of operation. Data is from device (A), a Nb SNSPD with $I_c = I_c(T_o = 1.7\text{K})$ and $R_L = 17\ \Omega$.

The detection efficiency of the Nb devices that were not thinned by etching, but rather were directly deposited, was much less than the detection efficiency of the etched devices, even when the final thicknesses of both sets of devices was similar. In Figure 5.15, detection efficiency versus bias current is plotted for device (A) as well as for device (D). Device (A) and device (D) have very similar superconducting properties; the critical temperature, critical current, film thickness, wire width, and material are all nearly equal (see Table 5.1 and Table 5.2). The major difference is in the fabrication procedure used (described in detail in section 3.4) and in the resistivity of the final device. Device (A) was thinned from 14 nm to a final thickness of 7.5 nm using an Ar^+ ion beam. All devices that were thinned in this way had resistivity that was more than twice as large as the resistivity of devices that were directly sputtered to a similar final thickness (see Table 5.1 and Table 5.2). This greater resistivity may be due to increased surface roughness or from increased internal disorder from the Ar^+ ion etching process. This apparent increase in the detection efficiency with device resistivity is not yet understood theoretically. For thicker, directly sputtered films, the detection efficiency was even lower than for thin directly sputtered films (see Table 5.2), typically in the range of 10^{-5}% or less. Thus, both thin and thick directly sputtered Nb SNSPDs had poor detection performance. All results are typical; in all cases, several chips of each type of devices were measured with similar results to those reported here.

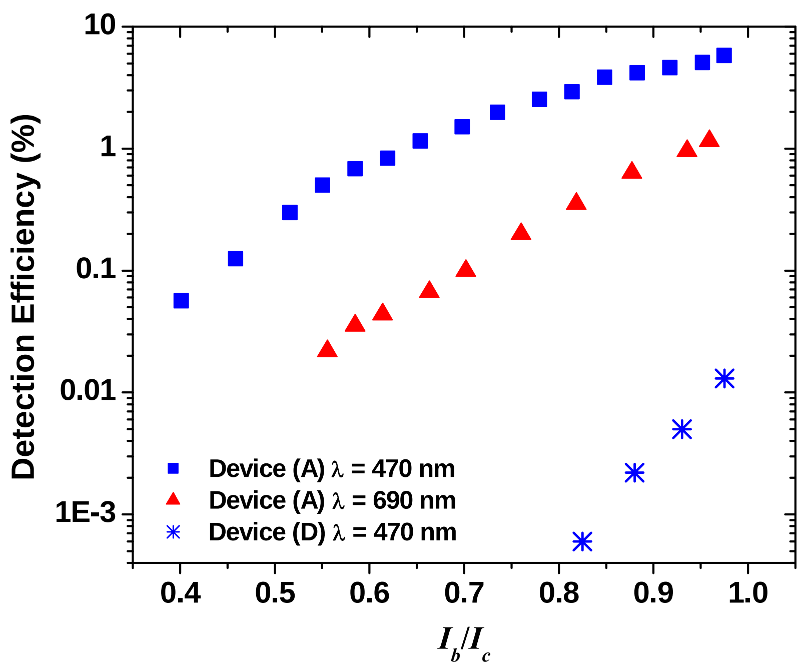

Figure 5.15: Plot of detection efficiency versus bias current for device (A) (I_c = 8.2 μA), which was etched from 14 nm to 7.5 nm using an Ar^+ ion beam, and device (D) (I_c = 7.4 μA), which was directly sputtered to 8.5 nm. Data for taken with $I_c = I_c(T_o = 1.7\text{K})$ and R_L = 17 Ω.

The detection efficiency also depends significantly on temperature. In Figure 5.16, a plot of the detection efficiency for device (A) is shown for two operating temperatures and two wavelengths of incident photons. As is clear, the detection efficiency at optimum current bias points (i.e., near the critical current) is lower at higher temperatures. Although this result has been observed in NbN devices (Korneev 2004), at this time there is no theoretical explanation as to why this is the case in either Nb or NbN devices.

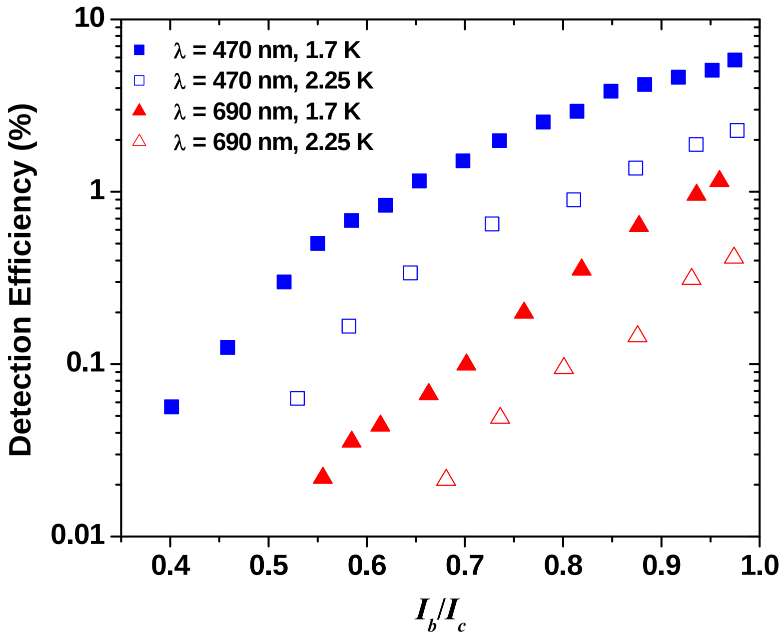

Figure 5.16: Plot of detection efficiency versus bias current, normalized to the critical current at the temperature of operation. Data is from device (A), a Nb SNSPD with a with $I_c = I_c(T_o)$ where T_o = 1.7 K and 2.25 K, and R_L = 17 Ω.

The functional dependence of the detection efficiency of NbN SNSPDs is similar to Nb. In Figure 5.17, the detection efficiency versus current is plotted for device (G) for photon wavelengths of 470 nm, 690 nm, and 1550 nm at operating temperatures of 1.7 K, 4.3 K, and 7.0 K, using R_L = 25 Ω. These results are consistent with other published results for NbN SNSPDs (e.g., Korneev 2004, Kerman 2009). As can be seen, the detection efficiency of all photon wavelengths and at both lower temperatures appears to saturate at a value of approximately 5%.[50] Higher temperatures need higher fractions of the bias current to saturate, however.

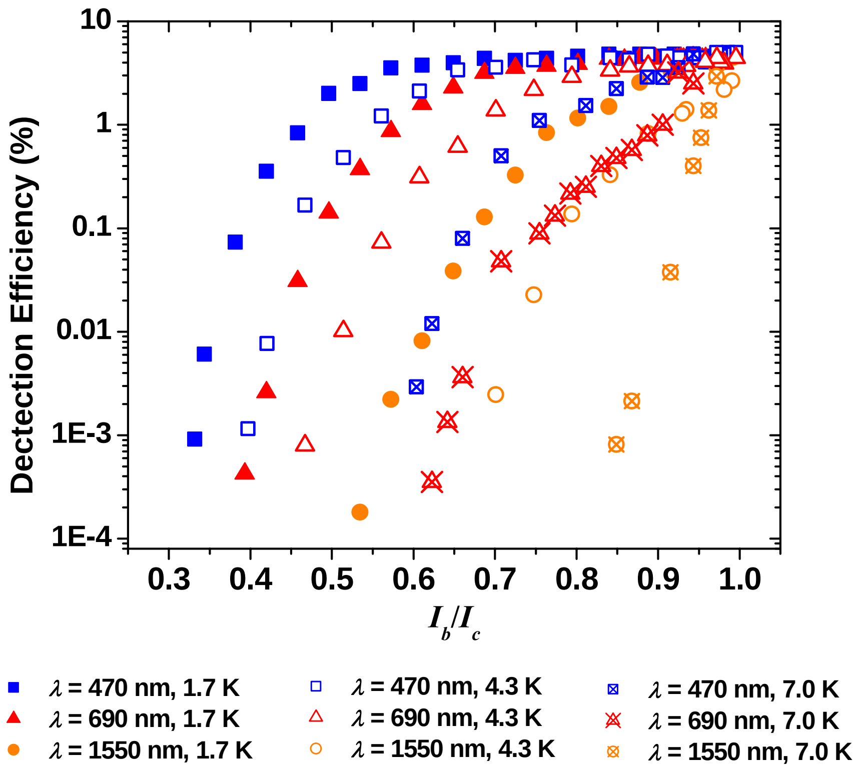

Figure 5.17: Plot of detection efficiency versus bias current, normalized to the critical current at the temperature of operation. Data is from device (G), a NbN SNSPD, for photons with a with $I_c = I_c(T_o)$ where T_o = 1.7 K, 4.3 K, and 7.0 K, and R_L = 25 Ω. Note that the detection efficiency measurement at 7.0 K includes some small error due to a large number of dark counts at that temperature, some of which are counted as synchronized with the laser trigger pulse (see Figure 5.11).

[50] This is true within the uncertainty in the measurements of the detection efficiency at each wavelength, which is listed in Table 5.2.

5.3.2 Dark Counts

Dark counts are spurious voltage pulses that do not result from incident photons but from internal fluctuations within the nanowire. For a complete theoretical description, see section 2.6. The dark count measurements reported in this thesis are not due to the detection of stray photons from external sources such as room light or blackbody radiation, nor are they due to electrical noise. Each of these sources of spurious detection have been almost completely eliminated by careful shielding and filtering. The procedure for this is explained in section 4.2. Voltage pulses created by internal thermal fluctuations at finite temperature, or from quantum mechanical fluctuations, cannot be eliminated by filtering or shielding. In practice, results in the literature suggest that only thermal fluctuations are important in SNSPDs (Kitaygorsky 2005, Engel 2006, Kitaygorsky 2007, Bell 2007, Bartolf 2010). These thermal fluctuations can drive the creation of resistive hotspots through a variety of physical mechanisms, including localized phase slip formation, vortex-antivortex unbinding, and vortex hopping. It is unclear exactly which mechanism(s) is dominant (section 2.6) and it is possible that the observed dark counts are due to more than one of these mechanisms, depending on the temperature and bias current.

Dark count voltage pulses are experimentally indistinguishable from voltage pulses that result from photons. In Figure 5.18, two plots are shown; one (solid) is the result of averaging 1000 photon-induced voltage pulses. The other (dashed) is the result of averaging 1000 dark counts. The pulses are nearly identical. Note that the second, smaller peak on the exponential tail of the pulses is caused by small RF reflections in the readout line due to impedance mismatches. The dark counts in general depend on both temperature and current. At any temperature, there are only appreciable dark count rates (> 1 Hz) at currents greater than ~$0.9I_c$. Both Nb and NbN SNSPDs have significant dark counts when biased near the critical current, and the dark count rate generally increases monotonically with temperature up to the critical temperature.

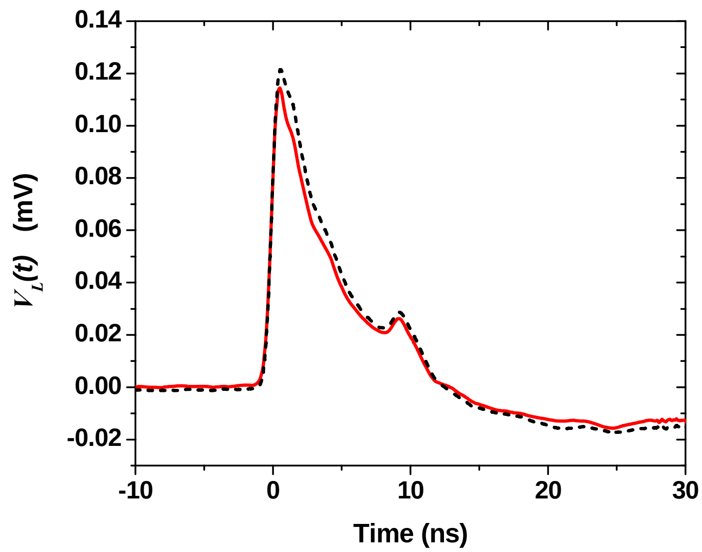

Figure 5.18: Plot of averaged voltage pulses that results from photon detection (red solid) and dark counts (black dashed). Data taken from device (A) at I_b = 3.0 μA, T_o = 2.5 K, and R_L = 50 Ω. Both are averages of 1000 individual voltage pulses. The photon-induced pulse is for photons of 470 nm wavelength. Note that the second, smaller peak on the exponential tail of the pulses is caused by small RF reflections in the readout line due to impedance mismatches. Reflections were reduced in later measurements.

In Figure 5.19, the dark count rate versus bias current is plotted for NbN device (G), for several operating temperatures from 0.25 K up to 8.0 K. In order to have high detection efficiency, the device must be biased with a bias current $I_b \geq 0.95I_c(T_o)$ (see Figure 5.17). The dotted vertical line is drawn at this value of the bias current. As can be seen, the dark count rate at this bias point depends greatly on temperature; for very low temperatures, $T_o < 2.0$ K, the dark count rate at $I_b = 0.95 I_c(T_o)$ is small, < 1 Hz. Thus, low temperatures are desired not only for having the highest detection efficiency (as is clear in Figure 5.17) but also to ensure the lowest dark count rates (as is clear in Figure 5.19). Note that the data at 0.25 K was not measured in the apparatus described in chapter 4, but rather in a Helium-3-based cryogenic system without optics. This is the only data set discussed in this thesis to have been taken from measurements in this apparatus.

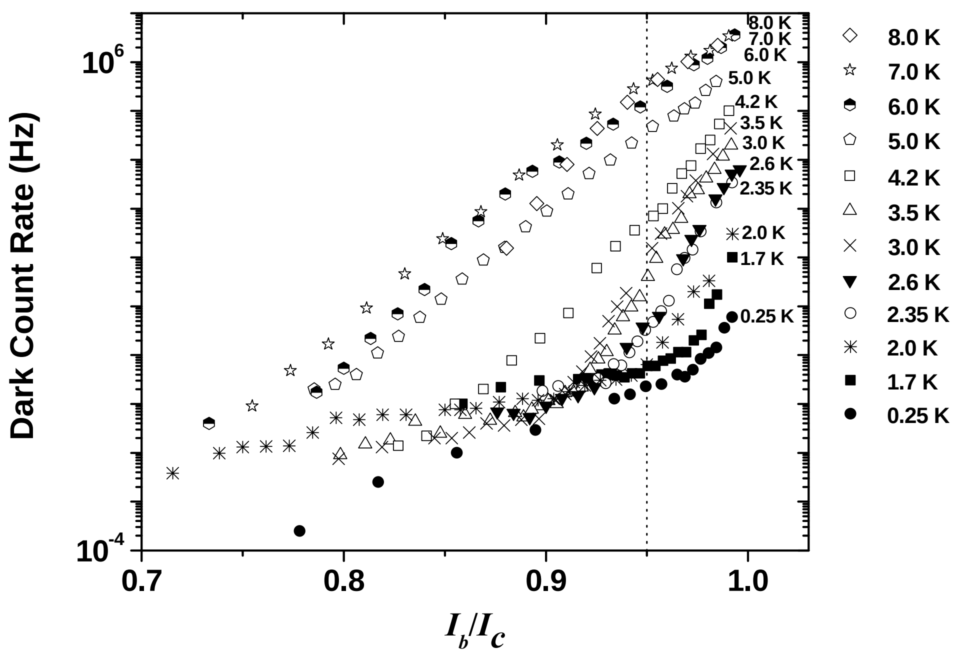

Figure 5.19: Dark count rate for NbN device (G) as a function of bias current and temperature. The dotted vertical line indicates the ideal bias point; for bias currents equal to or greater than this ($I_b \geq 0.95 I_c$), the detection efficiency is best. Note that the data at 0.25 K was not measured in the apparatus described in chapter 4, but rather in a Helium-3-based cryogenic system without optics.

The functional form of the curves for dark counts versus current at high current bias can be described by several fluctuation-induced resistance models, but there are no conclusions in the literature as to which theory fits best. Bartolf *et al.* (2010) as well as

Engel *et al*. (2006) and Bell *et al*. (2007) attribute fluctuations to current-assisted, thermally activated vortex motion within the superconducting strip. The models describing the dependence of the dark counts on temperature and current are discussed in section 2.6. In Bartolf *et al*. (2010), the vortex-antivortex unbinding and the vortex hopping models both describe the dark count versus current curve equally well for bias currents near the critical current ($I_b > 0.9I_c$) . Thus, the experimental measurements of Bartolf *et al*. could not distinguish which physical mechanism was responsible for dark counts.[51]

For bias currents further below the critical current ($I_b < \sim 0.9I_c$) , the dark count curve has a shallower slope; this part of the curve is less well explained by theory. Bartolf attributes this shallow tail to the current-activated flow of single vortices in the film. However, measurements of the dark count rate at the level of this tail (count rates < 0.1 Hz) are very sensitive to electronic noise due to imperfect filtering and shielding and even the occasional high energy 4.2 K blackbody photon. Thus, it is difficult to determine conclusively what internal physical mechanism, if any, causes this shallow tail. It is worth noting that dark count rates of less than 0.1 Hz are several orders of magnitude better than other infrared single-photon detectors (see Table 1.1) and are completely negligible in most applications.

At low temperatures, both the detection efficiency and the dark count rate are optimized. The range of bias conditions that maximizes detection efficiency and minimizes the dark count rate is rather narrow, however. If it is desired to keep the dark

[51] The author wishes to acknowledge A. M. Kadin for helpful input on this subject.

counts below 1 Hz, yet achieve the highest detection efficiency at 1550 nm, it is necessary to operate both Nb and NbN SNSPDs at temperatures below 2.0 K, and at bias currents that are $0.95I_c < I_b < 0.97I_c$. The dark count rate decreases somewhat in NbN for temperatures less than 2.0 K, which enables biasing at slightly higher currents. For Nb, the situation is similar.

A useful way to characterize the optimum bias current as a function of temperature is by plotting $0.95I_c$ versus operating temperature, T_o, alongside a plot of the current where the dark count rate is equal to 1 Hz, I_{1Hz} versus T_o. Plots of these for both Nb and NbN are seen in Figure 15.20. It is desired that $0.95I_c > I_{1Hz}$ for the best detection performance (lowest dark counts and highest near-infrared detection efficiency). As can be seen, there is only a small range of temperatures for Nb or NbN SNSPDs where this is true. Thus, the ideal bias range for SNSPDs is very small. Interestingly, the Nb SNSPD has a wider range of temperatures where $0.95I_c > I_{1Hz}$, and furthermore, the difference between the $0.95I_c$ and I_{1Hz} in this optimum temperature range is larger for Nb than for NbN. Thus, Nb has an apparent advantage in terms of dark counts; however, the detection efficiency for Nb is significantly less than for NbN in the near-infrared range (see Table 5.2, and Figures 5.14 and 5.17), and so the Nb detector would only be advantageous in the visible range, where its detection efficiency is similar to a NbN device.

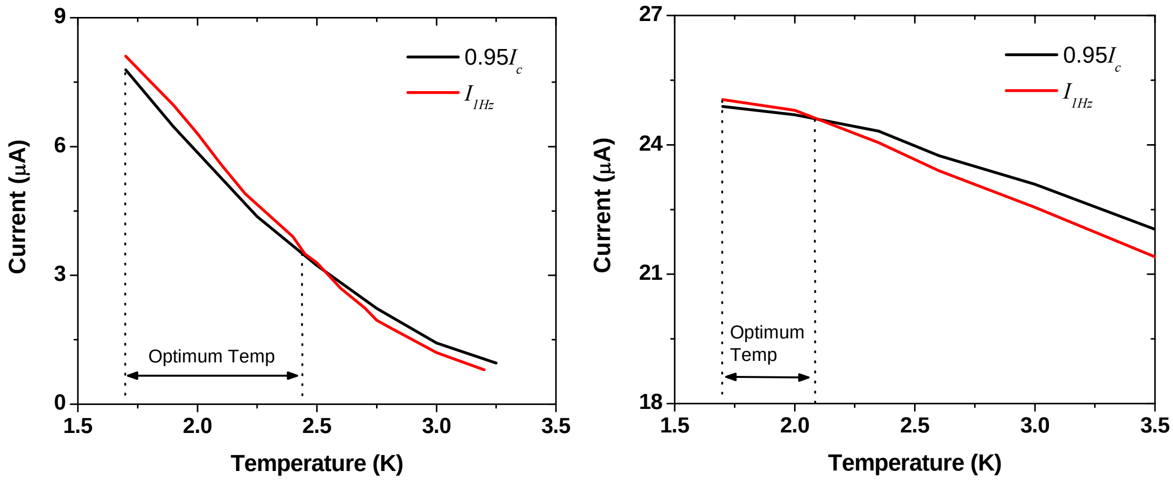

Figure 5.20: Left: a plot of 0.95I_c versus T_o and I_{1Hz} versus T_o for Nb device (A). Right: a plot of 0.95I_c versus T_o and I_{1Hz} versus T_o for NbN device (G). The 1 Hz dark count current, I_{1Hz}, is the current where the dark count rate is equal to 1 Hz. It is desired that 0.95I_c > I_{1Hz} for optimum bias (highest detection efficiency and negligible dark counts).

5.3.3 Jitter

Jitter is uncertainty in the measurement of the time at which a photon is detected. It is measured by determining the full width at half maximum of the distribution of time delays between the leading edge of the photon-induced voltage pulse and the leading edge of the laser electrical trigger pulse, which is synchronized with the optical pulse. A schematic of this measurement is seen in Figure 5.21. The jitter measured in this way is the total system jitter. The system jitter is comprised of the jitter due to the laser, the readout electronics, as well as the intrinsic jitter of the SNSPD: $\Delta\tau_{sys}^2 = \Delta\tau_{laser}^2 + \Delta\tau_{electronics}^2 + \Delta\tau_{SNSPD}^2$, where $\Delta\tau_{sys} \approx 150$ ps. The jitter due to the laser, $\Delta\tau_{laser}$, is dominated by the arrival time uncertainty of a single-photon; this uncertainty arises because the power in the laser pulse is Guassian distributed in time. The pulse widths for the diode lasers used in this thesis were approximately 100 ps (full-width at half-maximum). If the pulse is attenuated such that the average number of photons per pulse, n_{ave}, is less than 1, then a single photon may arrive anywhere within this pulse-width in time, with probability distribution given by the Guassian shape. Thus, $\Delta\tau_{laser} \approx 100$ ps. The jitter due to the readout electronics and noise is dominated mostly by the noise on the leading edge of the voltage pulse, which is from the amplifier. By filtering carefully when measuring the jitter, this can be reduced to much less than the laser jitter; it is estimated that $\Delta\tau_{electronics} \sim 10$ ps. The jitter from the oscilloscope (HP54855A) is extremely small, < 1 ps according to the specifications of the instrument. By subtracting the laser and electronic jitter in quadrature, an upper bound of $\Delta\tau_{SNSPD} \approx 110$ ps (+/- 10%) can be placed on the intrinsic jitter of the best performing Nb SNSPDs (devices (A), (B), and (C)). For NbN, Yang (2009) reports typical values of the jitter of 40 ps(full-width at half-maximum) for NbN devices similar to devices (G) and (H). Thus, it appears that the jitter of Nb SNSPDs is significantly greater than for NbN SNSPDs. The cause of this is uncertain, although it is possible that the slower thermalization of hot quasiparticles in Nb as compared to NbN (see section 2.2) may contribute to more variability in the time to develop a resistive hotspot once a photon is absorbed (see section 2.5). A second source of variability in both Nb and NbN detectors may be the exact location where the photon is absorbed on the nanowire strip (i.e. near the edges or in the center), which may also contribute to variability in the time to develop a resistive hotspot after the photon is absorbed.

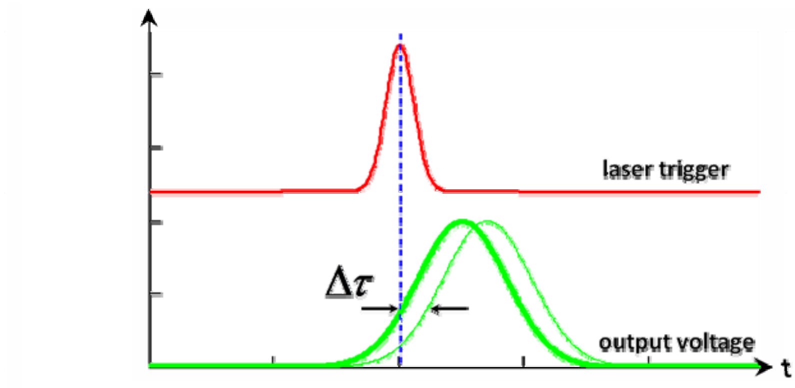

Figure 5.21: Schematic showing how jitter is measured.

5.4 Comparison of Nb and NbN SNSPD Detection Performance

In nearly all performance categories, NbN SNSPDs (devices (G) and (H)) equal or exceed the performance of the best Nb SNSPDs (devices (A), (B), (C)). This is apparent in Tables 5.1 and 5.2. The detection efficiency for photons with a wavelength of 470 nm is similar; the detection of efficiency of Nb for 690 nm and 1550 nm wavelengths is much less than in NbN SNSPDs. The Nb device is somewhat thicker (7.5 nm) than the NbN devices (5 nm), but less resistive. The thicker films may lead to a greater probability of absorbing a photon in Nb, which may account for the similar or slightly greater detection efficiency in Nb SNSPDs at 470 nm. The increased volume that must be driven normal because of the increased thickness, however, is probably the reason that the detection efficiency in Nb decreases so drastically for lower energy photons compared to NbN. Thinner or narrower Nb nanowires are not useful because the reduced critical temperature that results from the thinning makes them impractical to use at easily accessible cryogenic temperatures (~2K). The dark count rate in Nb and NbN SNSPDs is similar when the detectors are at similar operating temperatures and values of the bias current. The range of optimum bias points is slightly greater in Nb. The jitter in Nb SNSPDs is greater than in NbN SNSPDs. Thus, in terms of detection performance, NbN SNSPDs are clearly superior for detecting near-infrared photons, while the performance is similar for detecting higher energy optical photons. In chapter 6, the reset time of Nb and NbN SNSPDs will be explored.

Chapter 6

Reset Dynamics of SNSPDs: Kinetic Inductance and Thermal Relaxation

6.1 Overview

In chapter 5, measurements of the detection performance of Nb and NbN SNSPDs were reported. In the summary contained in section 5.4, the best devices were determined based on detection efficiency, dark count rate, and jitter. It was found that NbN devices equal or exceed the performance of Nb devices for detecting near-infrared photons in nearly all of these performance criteria. The data do suggest that Nb devices offer somewhat lower dark count rates than NbN devices when detecting visible photons, however there are a variety of single-photon detectors available for the visible range. It is unclear how much benefit Nb devices would offer compared to the full range of competition for visible single-photon detectors (see section 1.3 for a comparison to other technologies). In this chapter, the reset dynamics of the SNSPD devices with the best detection efficiency is studied in order to determine the minimum reset time for an SNSPD. A shorter reset time gives a higher single-photon count rate, which is desirable in many applications. The ideal SNSPD would have high detection performance combined with a fast and reliable reset.

In this chapter, the reset of SNSPDs with the best detection efficiency is studied. Specifically, devices (A), (B), (C), (G), and (H) from Tables 5.1 and 5.2 will be discussed. It is found that NbN devices generally have shorter reset times than Nb devices, although only by approximately a factor of two. In sections 2.1 and 5.3, the reset time of a device is defined as $3\tau_r$, where $\tau_r = L_K/R_L$ is the current return time, and where L_K is the kinetic inductance of the device at $I_b = 0$ and R_L is the load impedance (a resistance, since the impedance is predominantly real) of the readout circuit. The reset time, $3\tau_r$, is the time required for 95% of the bias current that has been shunted into R_L to return into the device. The device current must be close to the value of the dc bias current (which itself must be just below I_c) in order to have high detection efficiency. The discussion of detection performance in chapter 5 is predicated upon the SNSPD device self-resetting to the superconducting state after detecting a photon. This self-reset does not require any active quenching circuitry or a gated bias to achieve high count rates. Self-reset occurs automatically when the reset time is longer than the cooling time of the hotspot in the device. The precise criterion for self-reset will now be explored.

There is a minimum value of τ_r for an SNSPD to self-reset. If τ_r is too small, then self-reset will not occur; rather the nanowire will latch into a finite resistance state that is stabilized in time by Joule heating from the bias current. This is not desired, as the nanowire is not sensitive to photons in this state. The minimum reset time of a device reported in table 5.2 is equal to $3L_K/R_{L,max}$, where $R_{L,max}$ is the largest value of the load resistance that allows self-reset for a given value of L_K and other parameters (as discussed in section 2.6). Kinetic inductance is not only important because of the associated L_K/R_L time, but also because it determines the energy dissipated in the nanowire (see section

2.6). Thermal relaxation is the process by which the hot (non-equilibrium) electrons in a hotspot cool back to their initial temperature, T_o, which is essential for the device to fully reset. As discussed in section 2.6, the thermal relaxation time depends on the energy dissipated in the hotspot, E_{dis}, which in turn depends on L_K and the bias current. Thus, the thermal relaxation time determines the minimum value of $\tau_r = L_K/R_L$. Because L_K, τ_r, and the thermal reset time (to be defined in the next section) are interdependent in a nonlinear way, a numerical model (discussed in section 2.6) is necessary to solve for the value of $R_{L,max}$, and therefore the minimum reset time, for a given device.

In section 6.2 of this chapter, measurements of the temperature and current dependence of the kinetic inductance are presented. Measurements are compared to theory developed in section 2.3, which is based on the Ginzburg-Landau theory of superconductivity. In section 6.3, a comparison to the predictions of the model developed in section 2.6 is made using measurements of latching for several values of L_K and R_L in Nb and NbN SNSPDs. In the final section of this chapter, a summary comparing the reset performance of Nb and NbN SNSPDs is given. This is followed by a brief discussion of methods that could be used to decrease the reset time of SNSPDs in light of the theoretical and empirical understanding developed both in chapter 2 and in the present chapter.

6.2 Kinetic Inductance Measurements

In this section, measurements of the temperature and current dependence of the kinetic inductance of Nb and NbN SNSPDs are reported. In Kerman *et al*. (2006), the kinetic inductance in NbN SNSPDs is studied in detail and shown to scale with nanowire length. In this thesis work, kinetic inductance is studied in both Nb and NbN SNSPDs. The results reported here for NbN at low temperatures are consistent with the results of Kerman *et al*. (2006, 2007). Kerman *et al*. do not measure kinetic inductance at elevated temperatures (>4.2 K), however. Measurements reported in this thesis for NbN devices in this range of temperatures contain features not observed at low temperatures (T < 4.2 K). This will be explored in this section.

The current-return process in Nb SNSPDs is conceptually identical to the process in NbN SNSPDs; it occurs on a time scale given by L_K/R_L. The magnitude of L_K for a Nb nanowire, however, is less than for a NbN nanowire of equivalent geometry. The results of this work extend the results of Kerman *et al*. (2006) by comparing kinetic inductance in Nb and NbN devices, and also by studying the current and temperature dependence of L_K in greater detail. Kinetic inductance is an important property to understand accurately in any study of reset in SNSPDs not only because its value determines the current return time, $\tau_r = L_K/R_L$, but because it also determines the energy dissipated during a detection event. Dissipated energy is important because it determines the thermal reset time of a device, as will be clear in section 6.3 (see also section 2.6). The kinetic inductance in nanowires is also interesting to study because it may allow the fabrication of very small lumped element inductors for use in solid state quantum bits (Kerman 2010) and other cryogenic electronics. In detectors as well as in other applications, it is necessary to understand the current and temperature dependence of the kinetic inductance.

Kinetic inductance is also a topic of fundamental interest because it is directly dependent on the superconducting pair density and can therefore be used as a probe of the superconducting order parameter. Although the theory developed in section 2.3 for the current dependence of L_K is universal for any one-dimensional superconductor near its critical temperature, some of the experimental results and other theoretical predictions in the literature for 1-d wires disagree with the predictions of this theory. In Kerman *et al*. (2007), the kinetic inductance of NbN nanowires was measured to have a weak dependence on current (increase by factor of ≈1.2 as $I_b \rightarrow I_c$). This disagrees with the prediction from section 2.3, which predicts a divergence as $I_b \rightarrow I_c$. (See also Yang 2009.)

There are several results in the literature for studies of two-dimensional geometries. For two-dimensional superconductors, the theory of section 2.3 does not apply. However, even in two-dimensional samples, there is disagreement in the literature between theory and experiment, and between devices fabricated from different superconducting materials. In Cho *et al*. (1997), a strong divergence in L_K versus I_b (increase by factor of ~4 as $I_b \rightarrow I_c$) was measured in a two-dimensional strip of Yttrium Barium Copper Oxide (YBCO) very near T_c. In Enpuku *et al*. (1995), the kinetic inductance of two-dimensional YBCO films very near T_c was observed to increase with current, but reached a maximum at $I_b = 0.95 I_c$ which was a factor of 1.5 above the zero current value. For higher currents, Enpuku *et al*. (1995) report a decrease in the value of L_K as $I \rightarrow I_c$. In Anlage *et al*. (1989) and Meservey *et al*. (1969), measurements showed almost no dependence of L_K on current in two-dimensional strips of Nb and Tin, respectively. In Saracila *et al*. (1999), the kinetic inductance in two-dimensional Nb strips was measured to have a weak current dependence (increase by factor of ≈1.25 as $I_b \rightarrow I_c$). In Johnson *et al*. (1997), the kinetic inductance of two dimensional NbN films was measured to decrease with current, but the decrease began at bias currents that were a much smaller fraction of the critical current than observed in Enpuku *et al*. (1995).

In Figure 5.10 (in the previous chapter), a plot of a typical photon-induced voltage pulse is seen for device (A) with R_L = 50 Ω. For this device and load combination, $\tau_r = L_K/R_L$ = 4.7 ns, since L_K = 235 nH (see Table 5.1). The current return time, τ_r, is the time for current to return back into the nanowire from the readout circuit load. For devices with the same resistivity and cross-sectional area, the kinetic inductance is proportional to the length of the nanowire, and thus this return time should also be proportional to the length. In Figure 6.1, plots of typical photon-induced voltage pulses are plotted in terms of the current driven through the load during the detection event, $I_L = V_L/R_L$ (= V_d/R_L), for Nb devices (A), (B), and (C) with R_L = 25 Ω. The current through the load is normalized to the maximum load current for each device so τ_r can be visually compared.[52] An exponential fit is made to each curve, from which τ_r is extracted. For devices where L_K is large, this extracted value of τ_r is approximately equal to the predicted value of $\tau_r = L_K/R_L$ where L_K is the zero-current value of the kinetic inductance. For example, for device (A), $\tau_{r,fit}$ = 10.9 ns where as L_K/R_L = 9.4 ns; for device (B), $\tau_{r,fit}$ = 2.3 ns where as L_K/R_L = 2.4

[52] In this measurement, the peak current through the load was different for each device because a different value of the bias current was used for each. This was necessary to prevent latching in the shorter devices, as will be explained in section 6.3. In this measurement, the bias current values used were: (A) I_b = 7.7 μA, (B) I_b = 5.0 μA, (A) I_b = 3.8 μA

ns. The disparity between predicted and measured values of τ_r for the longer devices is attributed to uncertainty in the fits to the measured pulses, and also to a dependence of L_K on current, which will be examined next. Note, for example, that the exponential decay of the pulse for the longest device (A) is distorted at long times: the pulse undershoots zero current because the bandwidth of the amplifier and filters used in this measurement was such that the roll-on frequency was greater than the lowest frequencies in the spectrum of this pulse. For the short device, however, there is a larger discrepancy between $\tau_{r,fit}$ and L_K/R_L. This is not completely understood, but is thought to be influenced by the finite hotspot cooling timescale (see section 2.6 and section 6.3 on latching and hotspot cooling times). The dip in current in the shorter devices ((B) and (C)) near 10 ns is due to a small reflection caused by an impedance mismatch in the RF readout circuit. This impedance mismatch was later reduced. It should be noted that the value of $\tau_{r,fit}$ was observed to scale with R_L as expected ($\sim 1/R_L$) when R_L = 50, 25, 17 Ω for all but the shortest devices.

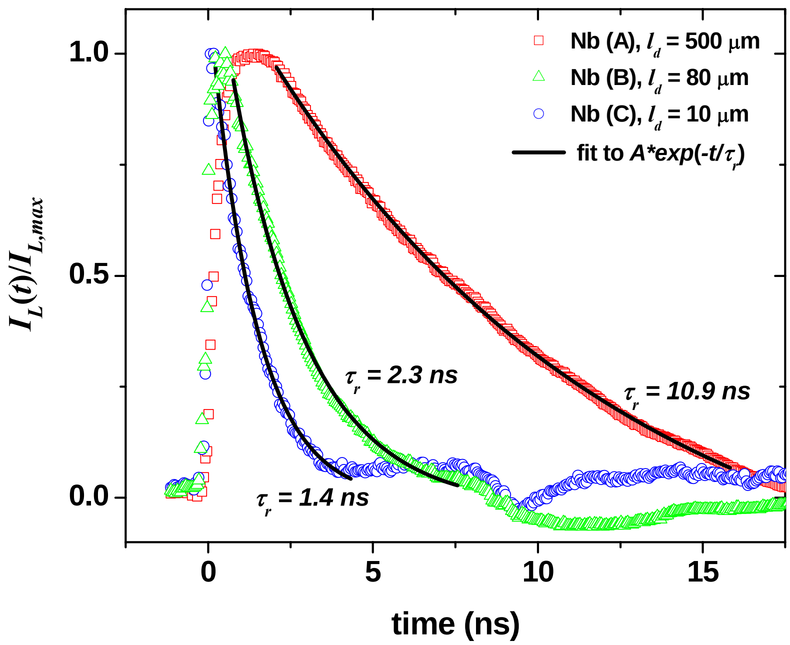

Figure 6.1: A plot of typical photon-induced pulses of current through the load, $I_L = V_L/R_L$ ($= V_d/R_L$), for devices (A), (B), and (C) for a load value of $R_L = 25\ \Omega$. Note that in this plot, the exponential decay of the pulse for the longest device (A) is distorted at long times: the pulse undershoots zero because the bandwidth of the amplifier and filters used in this measurement was such that the roll-on frequency was greater than the lowest frequencies in the spectra of this pulse. The dip in the shorter devices ((B) and (C)) near 10 ns is due to a small high frequency reflection due to an impedance mismatch in the RF readout circuit.

In this thesis work, the kinetic inductance of Nb and NbN SNSPDs was measured directly by incorporating each device into a resonant circuit with a known capacitance. In all measurements, the capacitance was chosen to be significantly greater than any parasitic capacitance in the measurement. The inductance was determined by measuring the resonant frequency of this circuit with a network analyzer. This method is explained in more detail in section 4.1. A series of plots depicting the shift in the resonant frequency as a function of temperature for NbN device (G) is seen in Figure 6.2 (top). In the plots, the reflected power is measured as a function of frequency as the device temperature is increased from 2.5 K to 9.1 K. Figure 6.2 (bottom) shows a fit to the measured curves at 5.0 K and 8.0 K. The fit is obtained by varying L_K, the bias tee capacitor, C_{BT}, and the device resistance, R_d in a simulation performed in Microwave Office. As can be seen, the simple method of determining the kinetic inductance from the minimum in the reflected power versus frequency is quite accurate for all temperatures. For example, at 5.0 K, the value of L_K calculated from the minimum of the reflected power is 113 nH, while the value of L_K obtained from fitting the entire curve in Microwave Office is 112 nH. At 8.0 K, the measured value of L_K is 166 nH while the fit value is 170 nH.

Plots of $L_K(T)$ versus T for both Nb devices (B), (A), and (F), and NbN device (G) are shown in Figure 6.3. The expected temperature dependence of L_K from the Ginzburg-Landau theory is $L_K(T) = L_K(0)\ (1-T/T_c)^{-1}$ (see section 2.3). The solid lines in each plot are fits to the data using the Ginzburg-Landau expression. $L_K(0)$ and T_c were used as fitting parameters; their extracted values are listed in the legend of each plot. As can be seen, except for in Nb device (A), the critical temperature extracted using the Ginzburg-Landau fit to the data for temperatures near T_c is close to the measured value of T_c for both devices. The extracted values of $L_K(0)$ were not consistent with the value expected from measurement, which is not unexpected given that the Ginzburg-Landau prediction is valid only for temperatures near T_c (Tinkham 1996). For device (A), the theory is not particularly accurate even near T_c, since the slope of L_K vs T near T_c decreases; at present, this experimentally observed leveling off cannot be explained theoretically.

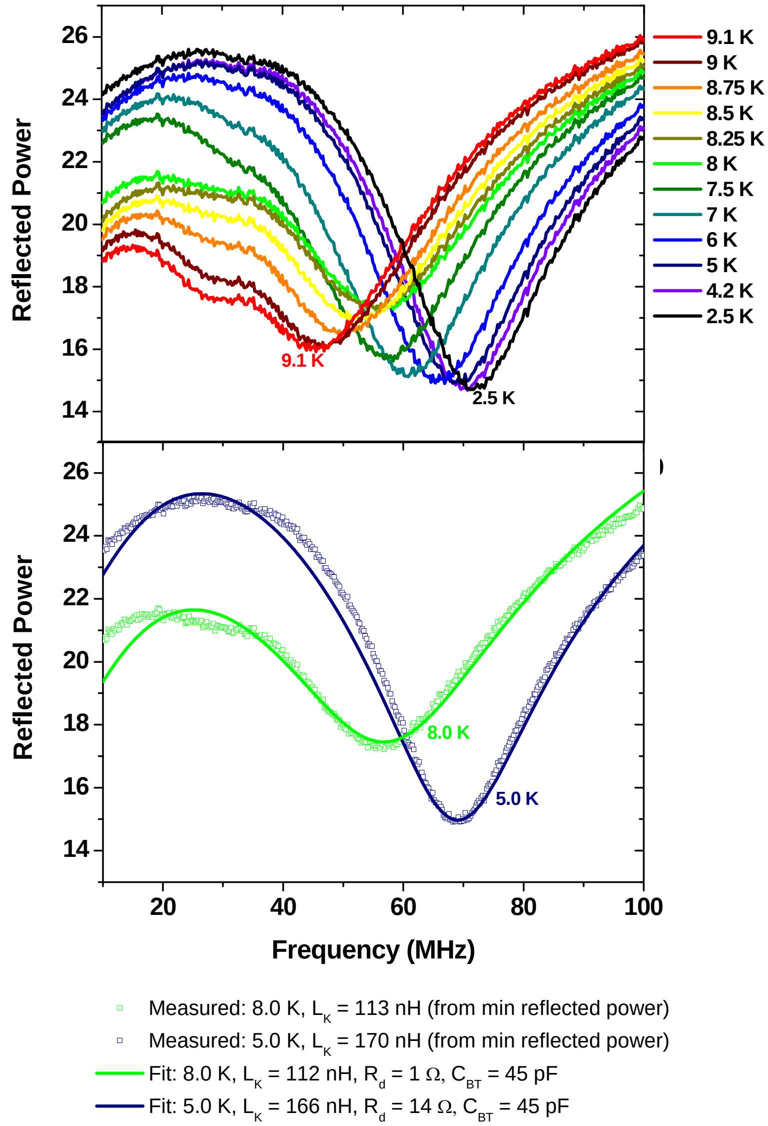

Figure 6.2: (top): A plot of the reflected power versus frequency for NbN device (G) as the device temperature is increased from 2.5 K to 9.1 K, measured at zero dc bias current. The resonant frequency of the circuit, $f_o = (L_K C)^{-1/2}$, defined as the frequency at which reflected power is a minimum, is reduced from 71 MHz to 45 MHz as the temperature is increased from 4.2 K to 9.1 K. (bottom): a theoretical fit (solid) to the measured curves (small squares) at 5.0 K and 8.0 K, fit by varying L_K, the bias tee capacitor, C_{BT}, and the device resistance, R_d.

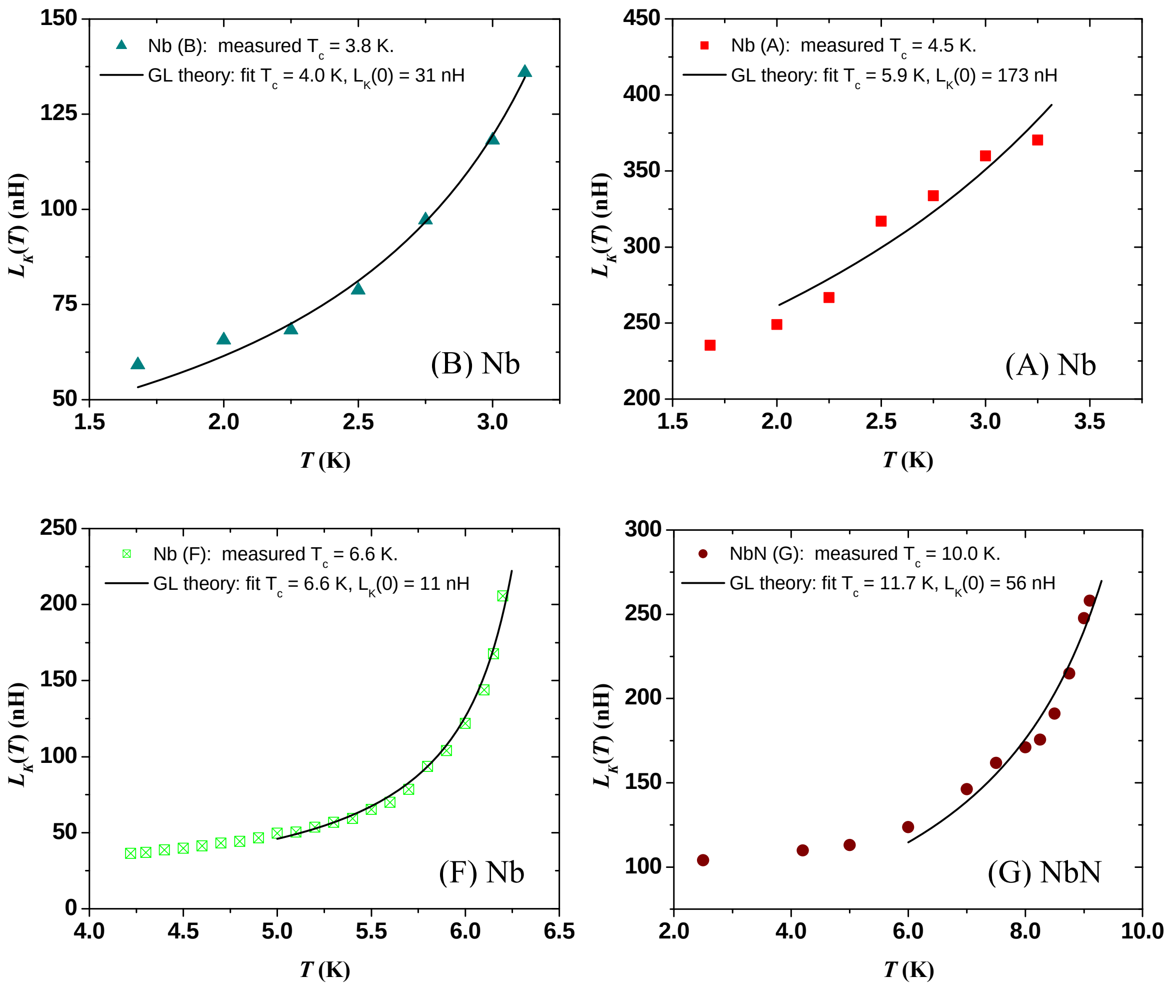

Figure 6.3: Plots of $L_K(T)$ versus T for devices (B) (top left), (A) (top right), (F) (bottom left), and (G) (bottom right). The solid lines are fits to the data using the functional form predicted by the Ginzburg-Landau theory: $L_K(T) = L_K(0)\ (1-T/T_c)^{-1}$. $L_K(0)$ and T_c were used as fitting parameters; their extracted values are listed in the legend of each plot.

A similar investigation can be undertaken for the current-dependence of the kinetic inductance. An expression for the current-dependence of the kinetic inductance based on the Ginzburg-Landau theory is derived and discussed in section 2.3. In Figure 6.4, plots of $L_K(I_b)/L_K(0)$ versus I_b/I_c for NbN device (G) are shown at several temperatures, alongside the theoretical prediction from section 2.3 calculated in two ways (dashed lines).[53] This clearly does not fit the data. First, the theoretical prediction is plotted by setting the Ginzburg-Landau critical current, which is an input to the theory, equal to the measured critical current, I_c (black dashed line). Second, the theoretical prediction is plotted by setting the Ginzburg-Landau critical current equal to $2I_c$ (red dashed line). The factor of 2 is chosen to fit the data. As is clear, the theoretical prediction will fit the data only if the critical current used in the theory is assumed to be much larger than the measured critical current. There are two possible explanations: first, the theory may simply not apply. The Ginzburg-Landau theory is only strictly applicable for temperatures near T_c; all measurements of $L_K(I_b)/L_K(0)$ versus I_b/I_c were done at temperatures significantly less than T_c. Second, it may suggest that the measured critical current in these devices (and the critical current that is important for high detection efficiency) is not the Ginzburg-Landau depairing current but a depinning current for vortex motion, which was predicted by Likharev (1979) to have a value of approximately half of the value of the Ginzburg-Landau depairing current. It is important to note that the experimental measurements of the current dependence of L_K at low temperatures ($T < 4.2$ K) reported here for NbN nanowires are in approximate agreement with measurements of L_K in NbN nanowires reported in Kerman *et al*. (2007) at similar temperatures.

[53] Here, the exact expression (solved numerically) for the current dependence of L_K is used. For analytical approximations at high and low bias currents, see section 2.3.

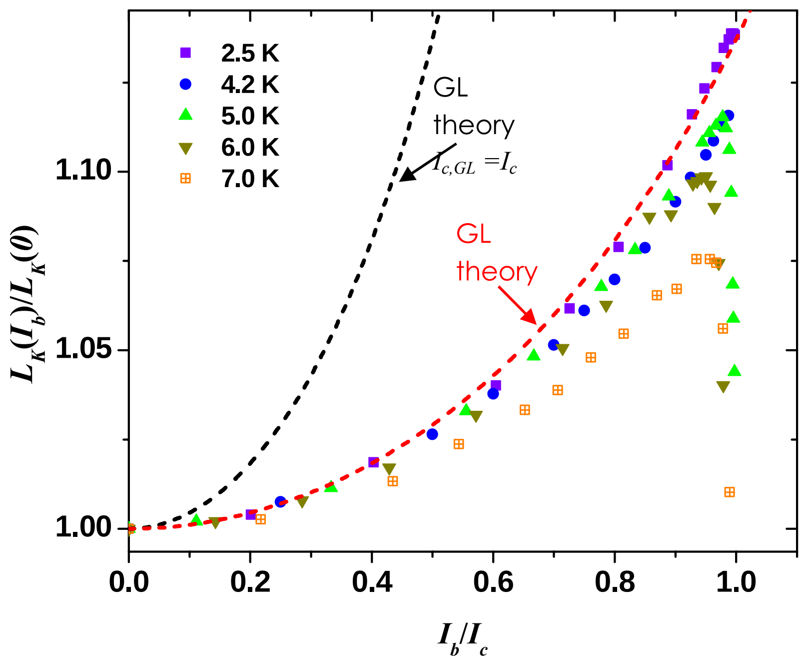

Figure 6.4: Plot of $L_K(I_b)/L_K(0)$ versus I_b/I_c for NbN device (G) at several temperatures, where I_c is the measured value of the critical current at each temperature, as explained in chapter 5. The experimental measurements are plotted alongside the theoretical prediction from section 2.3 and calculated two ways: first, by setting the Ginzburg-Landau critical current equal to the measured critical current, I_c (black dashed line) and setting the Ginzburg-Landau critical current equal to $2I_c$.

In Figure 6.5, it can be seen that for currents very near to I_c, ($I_b > 0.9I_c$) the kinetic inductance at temperatures of 5.0 K, 6.0 K, and 7.0 K decreases substantially as the current is increased. This is effect is not reported in Kerman *et al*. (2006) or Kerman *et al*. (2007), although Kerman does not report measurements at temperatures greater than 4.2 K. A decrease in the kinetic inductance of wider NbN strips at high bias currents was reported by Johnson *et al*. (1997). In that work, it was suggested that intrinsic Josephson junctions within the granular NbN film may lead to a decrease in the inductance with current due to the effect of thermal fluctuations. Preliminary modeling by Johnson was qualitatively consistent with this explanation. It is possible that a similar effect has been observed in this investigation, although there are some significant differences between the devices tested by Johnson and the devices tested in this thesis. Specifically, the devices in this thesis are much narrower, with a higher T_c, and significantly lower resistivity. Furthermore, the decrease of L_K with I_b observed in Johnson *et al*. (1997) occurs over a wider range of bias currents. The kinetic inductance was also observed to decrease with current for two-dimensional YBCO strips in Enpuku *et al*. (1995). In that work, the decrease occurred only for bias currents close to the critical current ($0.9I_c < I_b < I_c$). As is apparent from Figure 6.4, this is also the case for measurements of the NbN devices studied in this thesis work. Enpuku *et al*. (1995) do not propose a specific mechanism for this decrease in kinetic inductance, other than observing that the decrease is associated with an increase in the device resistance. This increase in device resistance is attributed to flux creep in the two-dimensional YBCO film. The values of resistance that correspond to this decrease in kinetic inductance are reported in Enpuku *et al*. (1995) to be $\sim$100 μΩ, which is a very small fraction of the normal state resistance of the strip.

Although a theoretical explanation for the decrease of L_K with I_b for $I_b \rightarrow I_c$ that is consistent with the data measured in this thesis has not yet been developed, the effect appears to be correlated with an increase in the average device resistance as I_b approaches I_c. This is similar to what is reported in Enpuku *et al*. (1995). Furthermore, the decrease of L_K with I_b appears to be an intrinsic effect and not due to an incorrect interpretation of the measurement. The average device resistance increases for I_b near I_c because of dark counts (which are resistance fluctuations; see section 5.3), which yield a non-zero device resistance when averaged at dc. In Figure 6.6 (a)-(c), the average resistance of device (G) obtained from a measurement of the dark count rate, $R_{d,avg}$, is plotted versus current, alongside the kinetic inductance versus current. The value of $R_{d,avg}$ is calculated by multiplying the measured dark count rate (Hz) at a specific value of the bias current by the average resistance of a hotspot in device (G) (taken to be 1 kΩ) and by the approximate duration of the hotspot (taken to be 2 ns). As can be seen, at each temperature, the decrease in kinetic inductance is associated with an increase in the average device resistance. This average device resistance is very small, however. The statistical correlation coefficient (R^2) between the inductance and the resistance for each of these data sets was calculated, yielding -0.913 at 5.0 K, -0.984 at 6.0 K, and -0.874 at 7.0 K respectively, which confirms the observed correlation. Additional studies are necessary to determine the physical origin of this decrease in the kinetic inductance.

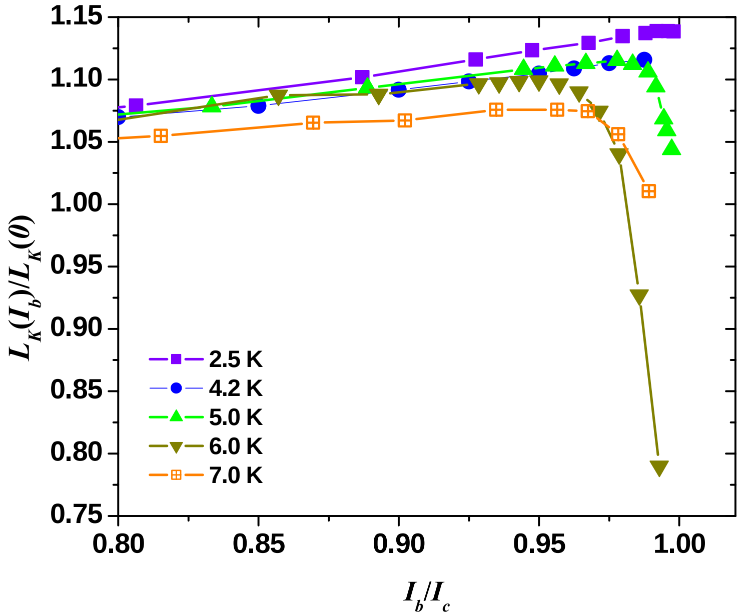

Figure 6.5: Plots of $L_K(I_b)/L_K(0)$ versus I_b/I_c for NbN device (G) at several temperatures, with expanded axes. On this scale, the reduction of the kinetic inductance for bias currents near the critical current is apparent.

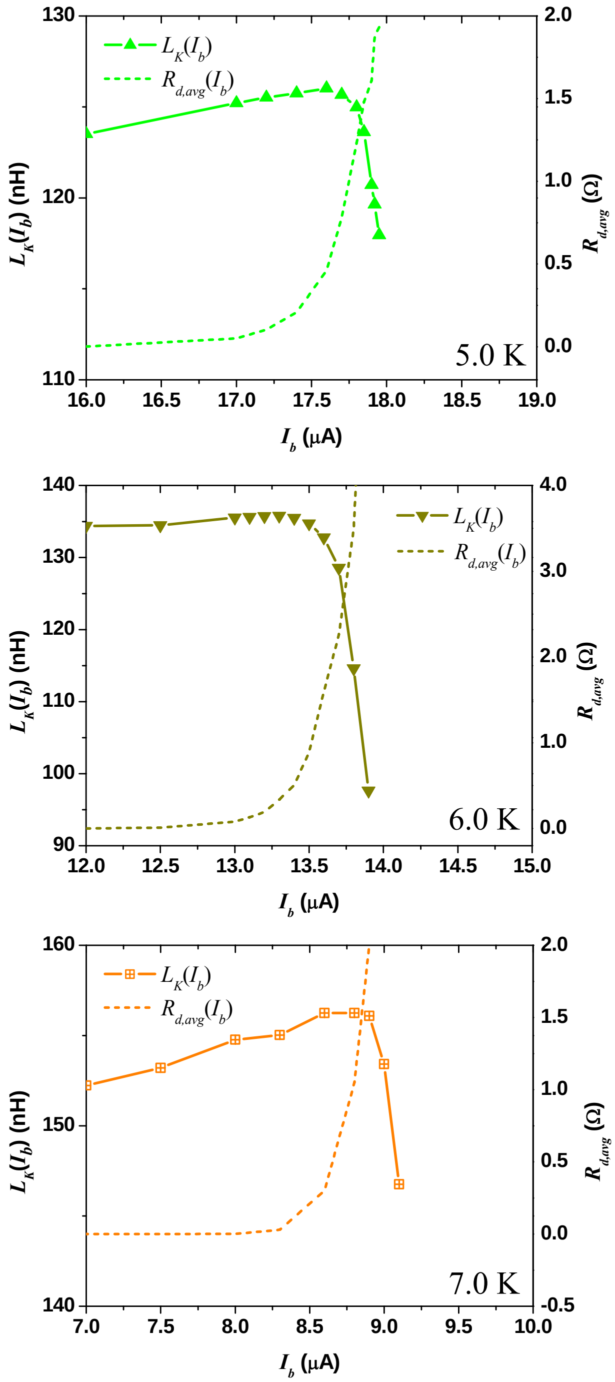

Figure 6.6: Plots of $L_K(I_b)$ and $R_{d,avg}(I_b)$ versus I_b for NbN device (G) at 5.0 K, 6.0 K, and 7.0 K. As can be seen, the decrease in the kinetic inductance is correlated with an increase in the average resistance of the device.

6.3 Thermal relaxation in SNSPDs

An SNSPD operating in the desired fashion will self reset to the superconducting state after detecting a photon. This occurs on a time scale given by the kinetic inductance and the load resistance of the readout circuit. The reset time is defined here as $3\tau_r$ where τ_r is the current return time, equal to L_K/R_L. The reset time can be reduced by reducing L_K or increasing R_L. The kinetic inductance can be reduced by reducing the length of the nanowire, however, this results in decreased active area since the length of the nanowire is proportional to the device area when a meander pattern is used. The load resistance can be increased for a given device, however there is a maximum resistance, $R_{L,max}$ that can be used, which depends on the parameters of the device. For values of $R_L > R_{L,max}$, the device will not self-reset but will rather latch into a finite voltage state after detecting a photon. This sets a minimum reset time for the detector, given by $3\tau_{r,min} = 3L_K/R_{L,max}$.

The minimum current return time, $\tau_{r,min}$, is determined by thermal relaxation. In section 2.6, a detailed discussion and model of this thermal relaxation is presented. In this section, the results of that model will be briefly reviewed. When a resistive hotspot forms, the bias current is shunted out of the device and into the readout circuit load; this reduces the Joule heating, which allows the hotspot to begin to cool. After the hotspot has cooled just enough to return to the zero resistance state, the current begins to return to the device, on a time scale given by τ_r. The electrons in the hotspot are still out of equilibrium, however: the electron temperature is higher than the equilibrium (bath) temperature even though it is less than the critical temperature: $T_c > T_e > T_o$.

As the temperature of the hotspot decreases, the critical current of the hotspot increases toward $I_{co} = I_c(T_o)$ because the hot electrons cool toward T_o. This occurs over a time period given by the average cooling time, $<\tau_c>$, which is determined by the amount of Joule heating from the bias current as well as the intrinsic cooling mechanisms in the film. Thus the electron temperature and the critical current are time dependent: $I_c(T_e(t)) = I_c(t)$. The increase in the device current, $I_d(t)$, must occur on a longer timescale than the increase from zero of the critical current, $I_c(t)$. If the current returns too quickly, then the device current will remain above the value of the critical current at the hotspot temperature. Thus, the hotspot, although it has cooled to below T_c, will remain resistive.

The cooling time, τ_c, is determined by the thermal properties of the films (electron-phonon scattering, phonon-escape, and diffusion) as well as Joule heating. Since τ_c depends on temperature, its value will change in time (section 2.6); thus, τ_c is characterized by $<\tau_c>$, the average value of τ_c over the cooling time period. Using the model of the thermal dynamics of SNSPDs developed for this thesis work, it was shown that the total amount of heat dissipated is given by $E_{dis} = ½L_K(I_b^2 – I_r^2) \approx ½L_KI_b^2$, where the approximation is valid because the return current, I_r, is typically much smaller than the bias current. Aside from reducing the value of R_L, latching can also be avoided by reducing the value of I_b, since this will reduce the heating, which decreases the value of $<\tau_c>$. Reducing the value of R_L also will reduce the magnitude of current returning to the device. However, reducing I_b significantly reduces the detection efficiency, as discussed in section 5.3, and so is an impractical option. Reducing the kinetic inductance will decrease E_{dis}, which will reduce τ_c without effecting the detection efficiency. Thus, reducing L_K is a far more desirable option that reducing R_L.

Since in experiments, it is much more convenient to measure current for a given value of R_L, a latching current is defined, I_{latch}, which is the lowest value of I_b that causes latching in a specific device for a given value of R_L. It is desired that $I_{latch} = I_{co}$. In Figure 6.7, the normalized latching current is plotted as a function of R_L for devices (A), (B), (C), (G), and (H). The curves define the boundary between self-resetting operation ($I_b < I_{latch}$) and latching operation ($I_b > I_{latch}$) for each device. As can be seen, Nb SNSPDs are much more prone to latching for $R_L \approx 50\ \Omega$ than are NbN SNSPDs. Niobium devices have less kinetic inductance than NbN devices for the same geometry (see Table 5.2 and section 6.2), and also a significantly lower critical current (see Table 5.1). Thus, the dissipated energy, E_{dis}, is much less in a Nb SNSPD than in a NbN device if the geometries are similar. This reduces the cooling time in Nb compared to NbN. However, the intrinsic electron-phonon and phonon-escape timescales in Nb are much longer than in NbN (see table 2.1 and section 2.2). This increases the cooling time compared to NbN. In order to understand how this interplay necessitates smaller values of R_L and larger values of τ_r in Nb devices, it is instructive to consider the latching current as a function of the current return time, $\tau_r = L_K/R_L$.

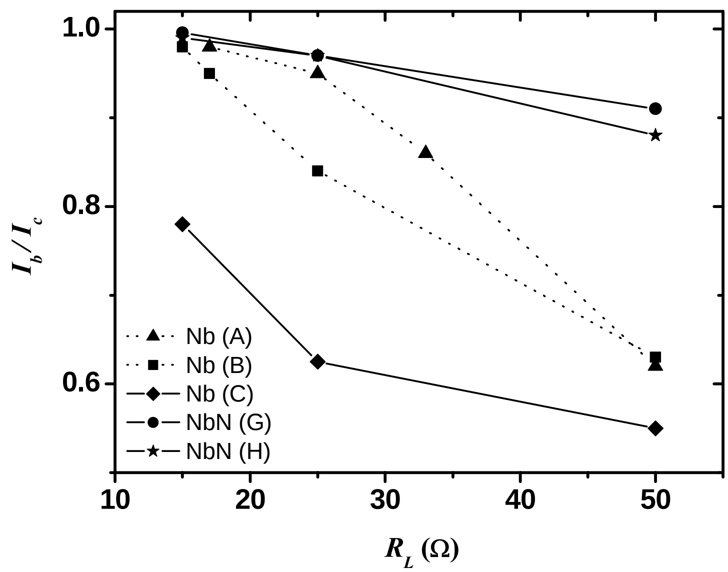

Figure 6.7: Plots of the normalized latching current as a function of R_L for devices (A), (B), (C), (G), and (H).

In Figure 6.8, the normalized latching current predicted by the model discussed in section 2.6 (dashed lines) and measured (solid lines) is plotted as a function of L_K/R_L for devices (A), (B), (C), (G), and (H). For device (A), predictions for two values of the electron diffusion constant, D_e (1.0 cm^2/s and 0.25 cm^2/s), are plotted to show that greater diffusion makes a device less susceptible to latching. For both Nb and NbN, devices with larger kinetic inductance need a larger value of τ_r to achieve the same value of I_{latch}/I_c. For example, in Fig. 2(a), the circled data points for device (A) and (B) are at similar values of I_{latch}/I_c; device (A) has L_K = 235 nH and τ_r = 14 ns while device (B) has L_K = 60 nH and τ_r = 4.0 ns. Reducing L_K is advantageous because it reduces τ_c by reducing E_{diss}. An important caveat is that when L_K is reduced for a Nb SNSPD, in practice the value of R_L must also be reduced. This is because $\tau_r = L_K/R_L$ can only be reduced (approximately) proportional to the amount that $<\tau_c>$ is reduced (see section 2.6). Reducing R_L reduces the peak voltage signal: $V_{L,max} \approx I_b \times R_L$. It is believed that the discrepancies in I_{latch} between simulated and measured data in Figure 6.7 are due to uncertainty in the values of material parameters used in the simulations as well as small RF reflections in the readout circuit, which can reduce the measured value of I_{latch}.

For practical detectors, the latching current should be at least $0.95I_c$. Using this criterion, it can be seen that NbN detectors can reset somewhat more quickly than Nb detectors when the geometries are equivalent. This is apparent by comparing Nb device (B) and NbN device (G) in Figure 6.8. Device (G) has an area of 25 μm^2 with 120 nH of inductance, while device (B) has an area of 16 μm^2 and 60 nH of inductance. Although the NbN device needs a significantly larger value of τ_r than the Nb device to achieve $I_{latch} = I_c$, the devices need similar reset times to yield latching currents near $0.95I_c$ (the NbN device needs only a slightly larger value of τ_r). If the devices were of similar area, and $I_{latch} = 0.95I_c$ is desired, then the NbN device would allow a smaller value of τ_r and thus have a shorter reset time. This is not a large advantage, however; NbN devices do not have a much faster reset time than Nb devices for detectors of useable, but small area (~10 μm^2).

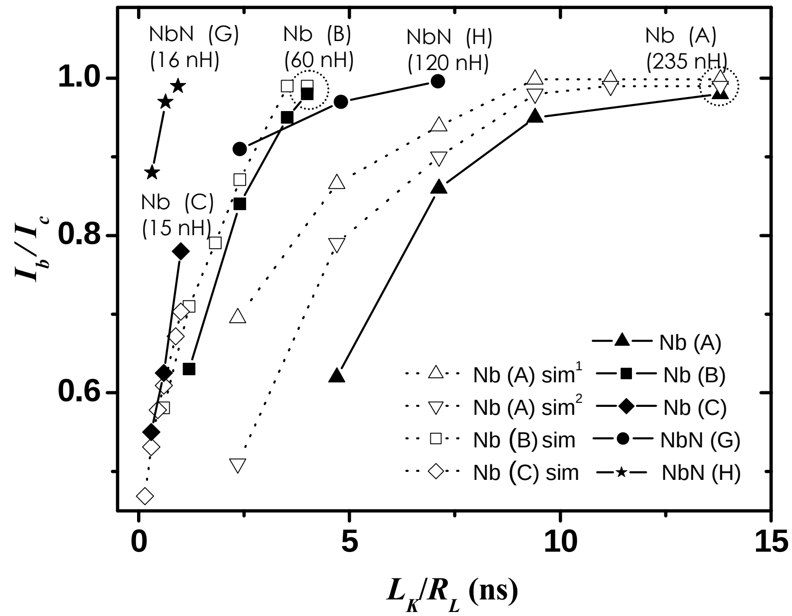

Figure 6.8: Plots of the normalized latching current, I_{latch}/I_c, as a function of L_K/R_L for devices (A), (B), (C), (G), and (H).

In addition to circuit parameters, the time scale of thermal relaxation also depends on material parameters. The cooling time, τ_c, as well as its average, $<\tau_c>$, decreases when $\tau_{e\text{-}ph}$ or the phonon escape time, τ_{es}, are decreased or when D_e is increased. As evident in Figure 6.8, in Nb τ_c is generally larger than in NbN for similar geometry detectors, even though NbN has larger L_K and I_{co}. This causes Nb devices to latch at larger values of τ_r than NbN devices, giving them a longer reset time. The intrinsic reason for this is that NbN has a much stronger electron-phonon interaction than Nb, which more than compensates for lower diffusivity and generally larger values of E_{diss} in typical NbN SNSPDs. The model developed in this thesis suggests that in a Nb SNSPD, electron-phonon relaxation limits the maximum count rate to approximately 150 MHz for a detector with area 10 μm^2. In NbN the maximum count rate for a detector with area 10 μm^2 is approximately 300 MHz (Kerman 2009, Yang 2007, Kerman 2006).

6.4 Summary and Prospects for Reducing the Reset Time

Using the model developed in section 2.6 and data presented in this chapter, it was found that reducing L_K enables higher count rates in SNSPDs as long as R_L is also reduced sufficiently to avoid latching. It is not necessary to reduce the value of R_L by as much as L_K is reduced, thus the net effect of reducing L_K is a shorter reset time. The reason for this is that reducing L_K reduces the stored energy and thus the amount of heating, which reduces the time for the hotspot to cool sufficiently. This allows τ_r to be reduced so that the current can reset more quickly. Although NbN devices should be capable of ~GHz count rates if L_K and R_L are small, the count rates of devices with practical detection area (> 10 μm^2) are limited by dissipation of the energy stored in the kinetic inductance, which increases $<\tau_c>$ substantially. This dissipation depends on I_b and L_K, but is nearly independent of R_L. Thus, the maximum count rate will be reduced as the area is increased, independent of the readout circuit. For extremely large detection areas (>> 100 μm^2), Nb devices may offer a faster count rate than NbN since L_K is smaller. In general, however, NbN devices of area up to 100 μm^2 equal or exceed the count rate of NbN devices, but by only approximately a factor of two. NbN SNSPDs also generally have larger output signals than Nb devices since the critical current (and therefore the bias current) in NbN is generally larger than in a Nb device of the same geometry. Furthermore, NbN SNSPDs can tolerate larger load resistances without latching, which also increases the size of the voltage signal compared to NbN devices, since the peak voltage signal is given by $V_{L,max} \approx I_b \times R_L$. If very small signals and small active areas can be tolerated, short (~10 μm or less) NbN devices may exhibit very short reset times, with potentially as high as GHz count rates. It would be difficult efficiently couple free-space optical or near-infrared photons to such small area devices, however. Lithographically-defined antennas with high coupling efficiency have not yet been developed for the optical and near infrared range.

The ideal material for high detection efficiency, large area, and high count rates would have very small $\tau_{e\text{-}ph}$, large D_e, and small L_K. However, in most known superconducting thin films, a small value of $\tau_{e\text{-}ph}$ is associated with large L_K and small D_e. Interestingly, even if a nanowire is fabricated from a superconducting material with a shorter electron-phonon time than NbN, such as MgB_2, the cooling rate would likely be limited by the phonon escape time, τ_{es}, which probably cannot be reduced significantly from its value for ultra-thin NbN, where $\tau_{es} \sim 40$ ps (Ptitsina 1997). Since it is already the case that $\tau_{e\text{-}ph} \sim \tau_{es}$ in NbN, it is unlikely MgB_2 (Shibata 2008) or any high-T_c superconductor (Sergeev 1994) would be able to achieve a significantly higher count rate than NbN. Other, more novel methods to increase the count rate of SNSPDs can also be envisioned. For example, there may be advantages in new geometries that incorporate parallel nanowires (Ejrnaes 2007, 2009) or layered or alternating width structures to promote diffusion cooling of hot electrons.

Chapter 7

Conclusion and Future Work

7.1 Conclusions

In this thesis, a detailed study has been undertaken of superconducting niobium and niobium nitride nanowires used as single optical and near-infrared photon detectors. In the first part of this study, Nb and NbN SNSPDs were compared based on practical performance parameters including the detection efficiency, count rate, jitter, and dark count rate. In the second part of this study, the details of the thermal reset mechanism and the kinetic inductance of the supercurrent in both Nb and NbN nanowires were explored.

The detection efficiency for photons with a wavelength of 470 nm is similar for both Nb and NbN SNSPDs; however, the detection efficiency of Nb SNSPDs for 690 nm and 1550 nm wavelengths is much less than for NbN SNSPDs. Nb devices must be thicker (typically $\approx$ 7.5 nm) than the NbN devices (typically $\approx$ 5 nm). The increased volume that must be driven normal because of the increased thickness is probably the reason that the detection efficiency in Nb is so much less for near-infrared photons than for NbN SNSPDs. Thinner or narrower Nb nanowires are not useful because the reduced critical temperature that results from the thinning makes them impractical to use at easily accessible cryogenic temperatures (~2K).

In other measures of detection performance, NbN is only marginally superior. The dark count rate in Nb and NbN SNSPDs is similar when the detectors are at similar operating temperatures and values of the bias current. The range of optimum bias points is actually slightly greater in Nb SNSPDs. The jitter in Nb SNSPDs, however, is a factor of 3 greater than in NbN SNSPDs. Finally, Nb has a relatively long thermal relaxation time compared to NbN, which reduces the single-photon count rate in a Nb device with practical area (~10 μm^2) by approximately a factor of 2 compared to a NbN device. Niobium detectors were also found to be more difficult to bias in a regime where the devices self-reset to the zero voltage state after detecting a photon. Related to this issue, because Nb devices have generally a smaller critical current and necessitate a smaller load resistance to read out, the single-photon voltage pulse in Nb is much smaller in amplitude than in typical NbN devices.

The second part of this thesis consisted of a study of the physics of the reset mechanism and of the kinetic inductance of SNSPDs. In this section, thermal relaxation and its relation to the reset time in both Nb and NbN SNSPDs was studied using a combination of experiments and numerical simulations. The minimum reset time after absorbing a photon was found to be set by thermal relaxation. Results from simulations suggest that the electron-phonon scattering time is the dominant limit to the cooling time in niobium nanowires, but that the diffusivity of the wire is also important in determining the total cooling time. Using this theoretical framework, the difference in reset times between Nb and NbN SNSPDs was explained.

The temperature and current dependence of the kinetic inductance of Nb and NbN nanowires was also investigated in detail. This is important to know for both fundamental reasons and because measurements of the kinetic inductance may be used as a diagnostic for determining the uniformity of the critical current along the nanowire. It was found that the temperature dependence near the critical temperature was reasonably well described by the Ginzburg-Landau theory. The current dependence, however, was not well described by the Ginzburg-Landau theory unless the Ginzburg-Landau critical current was assumed to be approximately twice as large as the measured value of the critical current .

7.2 Future Work

Future work targeted toward improving the detection performance of SNSPDs should focus on two areas: new materials and new geometries. SNSPDs fabricated from new materials with shorter electron-phonon interaction times may decrease the reset time of SNSPDs after a photon is detected. Materials with short electron-phonon times and relatively high (> 4.2 K) critical temperatures include magnesium diboride (MgB_2), niobium titanium nitride (NbTiN), and several high-T_c superconducting materials such as YBCO. Early work with MgB_2 (Shibata 2008) and NbTiN (Dorenbos 2008, Miki 2009) suggests that it may be possible to construct SNSPDs with shorter reset time than NbN SNSPDs, however the susceptibility to latching of SNSPDs fabricated from these new materials needs to be fully investigated. In addition, new geometries for SNSPDs that incorporate parallel structures should be further investigated through modeling and measurements. Early work with parallel SNSPDs (Ejrnaes 2009) suggests that this approach may lead to shorter reset times. Parallel nanowires reduce the total kinetic inductance but add more degrees of freedom for the bias current. Since the current dynamics are possibly very different in these detectors due to the many more degrees of freedom, further modeling would be necessary to determine thermal reset dynamics and susceptibility of these parallel SNSPDs to latching.

Bibliography

Anant, V., A. J. Kerman, E. A. Dauler, J. K. Yang, K. M. Rosfjord, and K. K. Berggren, "Optical properties of superconducting nanowire single-photon detectors," Opt. Exp. **16**, 10750 (2008).

Anlage, S. M., H. J. Snortland, and M. R. Beasley, "A current-controlled variable delay superconducting transmission line", IEEE Trans. Appl. Supercond. **25**, 1388 (1989).

Annunziata, A. J., "A coaxial to CPW transition," unpublished, available from author (2005).

Annunziata, A. J., A. Frydman, M.O. Reese, L. Frunzio, M. Rooks and D.E. Prober, "Superconducting niobium nanowire single-photon detectors," Proc. SPIE 6372, 63720V (2006).

Annunziata, A. J., D.F. Santavicca, J.D. Chudow, L. Frunzio, M.J. Rooks, A. Frydman and D.E. Prober, "Niobium superconducting nanowire single photon detectors," IEEE Trans. Appl. Supercond. **19**, 327 (2009).

Ashcroft, N. W., D. N. Mermin, *Solid State Physics*, Thomson Learning (1976).

Bartolf, H., A. Engel, A. Schilling, K. Il'in, M. Siegel, H.-W. Huebers, A. Semenov, "Current assisted, thermally activated flux liberation in ultrathin nanopatterned NbN superconducting meander structures," Phys. Rev. B **81**, 024502 (2010).

Bell, M., A. Sergeev, V. Mitin, J. Bird, A. Verevkin, and G. Gol'tsman, "One-dimensional resistive states in quasi-two-dimensional superconductors: experiments and theory," Phys. Rev. B **76**, 094521 (2007).

Blakesley, J. C., P. See, A. J. Shields, B. E. Kardynal, P. Atkinson, I. Farrer, D. A. Ritchie, "Efficient single photon detection by quantum dot resonant tunneling diodes," Phys. Rev. Lett. **94**, 067401 (2005).

Cabrera, B., R. M. Clarke, P. Colling, A. J. Miller, S. Nam, and R. W. Romani, "Detection of single infrared, optical, and ultraviolet photons using superconducting transition edge sensors," Appl. Phys. Lett. **73**, 735 (1998).

Carrol, J. J., S. T. Ceyer, and A. J. Melmed, "Optical constants of (001) niobium in the visible region," J. Opt. Soc. Am. **72**, 668 (1982).

Chi, C. C., M. M. T. Loy, D. C. Cronemeyer, "Transient responses of superconducting lead films measured with picoseconds laser pulses," Phys. Rev. B **23**, 124 (1981).

Cho, S., “Temperature and current dependence of inductance in superconducting meander line,” J. Kor. Phys. Soc. **31**, 337 (1997).

Conradt, R., “Indirekte band-band stossrekombination in Ge,” Z. f. Physik **209**, 445 (1968).

Dixon, A. R., Z. L. Yuan, J. F. Dynes, A. W. Sharpe, A. J. Shields, “Gigahertz decoy quantum key distribution with 1 Mbit/s secure key rate,” Opt. Express **16**, 18790 (2008).

Dorenbos, S. N., E. M. Reiger, U. Perinetti, V. Zwiller, T. Zijlstra and T. M. Klapwijk, “Low noise superconducting single photon detectors on silicon,” Appl. Phys. Lett. **93**, 131101 (2008).

Einstein, A. "On a heuristic viewpoint concerning the production and transformation of light," Annalen der Physik, **17**, 132 (1905).

Ejrnaes, M., R. Cristiano, O. Quaranta, S. Pagano, , A. Gaggero, F. Mattioli, R. Leoni, B. Voronov, and G. Gol'tsman, “A cascade switching superconducting single photon detector,” Appl. Phys. Lett. **91**, 262509 (2007).

Ejrnaes, M., A Casaburi, O. Quaranta, S. Marchetti, A. Gaggero, F. Mattioli, R. Leoni, S. Pagano, and R. Cristiano, “Characterization of parallel superconducting nanowire single photon detectors,” Supercond. Sci. Technol. **22**, 055006 (2009).

Ejrnaes, M., A Casaburi, R. Cristiano, O. Quaranta, S. Marchetti, N. Martucciello, S Pagano, A. Gaggero, F. Mattioli, R. Leoni, P. Cavalier, and J. C. Villegier, “Timing jitter of cascade switch superconducting nanowire single photon detectors,” Appl. Phys. Lett. **95**, 132503 (2009).

Engel, A., A. D. Semenov, H. W. Hubers, K. Il’in, M. Siegel, “Fluctuation effects in superconducting strips,” Physica C **444**, 12 (2006).

Engel, A., A. Semenov, H. W. Hubers, K. Il’in, M. Siegel, “Superconducting single-photon detector for the visible and infrared spectral range,” J. Mod. Opt., **15**, 1459 (2004).

Enpuku, K., H. Moritaka, H. Inokuchi, T. Kisu, M. Takeo, “Current-dependent kinetic inductance of superconducting YBaCuO thin films,” Jpn. J. Appl. Phys. **34**, 675 (1995).

Floet, D. W., E. Miedema, T. M. Klapwijk, J. R. Gao, “Hotspot mixing: A framework for heterodyne mixing in superconducting hot-electron bolometers,” Appl. Phys. Lett. **74**, 433 (1999).

Fowles, G. R., *Introduction to Modern Optics*, Second Edition, Dover Publications (1989).

G’oltsman, G. N., A. Korneev, O. Minaeva, I. Rubtsova, I. Milostnaya, G. Chulkova, B. Voronov, W. Słysz, J. Kitaygorsky, A. Cross, A. Pearlman, and Roman Sobolewski, “Superconducting nanostructured detectors capable of single-photon counting in the THz range,” from a presentation at the 2005 IRMMW-THz Conference.

Gershenzon, E. M., M. E. Gershenzon, G.N. Gol’tsman, A.M. Lyul’kin, A.D. Semenov, and A.V. Sergeev, “Electron-phonon interaction in ultrathin Nb films,” Sov. Phys. JETP **70**, 505 (1990).

Gousev (1), Yu. P., A.D. Semenov, G.N. Gol'tsman, A.V. Sergeev, E. M. Gershenzon, “Electron-phon interaction in disordered NbN films,” Physica B **194**, 1355 (1994).

Gousev (2), Yu. P., G. N. Gol'tsman, A. D. Semenov, E. M. Gershenzon, R. S. Nebosis, M. A. Heusinger, and K. F. Renk, “Broadband ultrafast superconducting NbN detector for electromagnetic radiation,” J. Appl. Phys. **75**, 3695 (1994).

Hadfield, R. H., “Single-photon detectors for optical quantum information applications,” Nature Photonics **3**, 696 (2009)

Haecker, W., O. Groetzinger, and M.H. Pilkuhn, “Infrared photon counting by germanium avalanched diodes,” Appl. Phys. Lett. **19**, 113 (1971).

Hamamatsu website, http://sales.hamamatsu.com/en/products/electron-tube-division/detectors/photomultiplier-tubes/part-r5509-73.php# (19 Dec. 2009).

Iams, H. E. and B. Salzberg, "The secondary emission phototube," Proc. IRE **23**, 55 (1935).

Il'in, K. S., M. Lindgren, M. Currie, A. D. Semenov, G. N. Gol'tsman, R. Sobolewski, S. I. Cherednichenko and E. M. Gershenzon, “Picosecond hot-electron energy relaxation in NbN superconducting photodetectors,” Appl. Phys. Lett. **76**, 2752 (2000).

Irwin, K., S. Nam, B. Cabrera, B. Chugg, G. Park, R. Welty, J. Martinis, “A self-biasing cryogenic particle detector utilizing electrothermal feedback and a SQUID readout,” IEEE Trans. Appl. Supercond. **5**, 2690 (1995).

Jaspan, M. A., J.L. Habif, R.H. Hadfield, and S.W. Nam, “Heralding of telecommunication photon pairs with a superconducting single-photon detector,”Appl. Phys. Lett **89**, 031112 (2006).

Jiles, D., Introduction to Magnetism and Magnetic Materials, CRC Press (1998).
Johnson, M. W., A. M. Kadin, “Anomalous current-dependence of kinetic inductance in ultrathin NbN meander lines,” IEEE Trans. Appl. Supercond. **7**, 3492 (1997).

Kadin, A. M., Duality and fluxonics in superconducting devices,” J. Appl. Phys. 68, 5741 (1990).

Kadin, A. M., M. Leung, A. D. Smith, J. M. Murduck, “Photofluxonic detection: A new mechanism for infrared photodetection in superconducting thin films,” Appl. Phys. Lett. 57, 2847 (1990). (1)

Kadin, A. M., M. Leung, A. D. Smith, “Photon-assisted vortex depairing in two-dimensional superconductors,” Phys. rev. Lett. 65, 3191 (1990). (2)

Kadin, A. M. and M.W. Johnson, "Nonequilibrium photon-induced hotspot: A new mechanism for photodetection in ultrathin metallic films," Appl. Phys. Lett. **69**, 25 (1996).

Kerman, A. J., J. K. W. Yang, R. J. Molnar, E. A. Dauler, and K. K. Berggren, “Electrothermal feedback in superconducting nanowire single-photon detectors,” Phys. Rev. B **79**, 100509 (2009).

Kerman, A. J., “Metastable superconducting qubit,” Phys. Rev. Lett. **104**, 027002 (2010).

Kerman, A. J., E. A. Dauler, J. K. W. Yang, K. M. Rosfjord, V. Anant, K. K. Berggren, G. N. Gol’tsman, B. M. Voronov, “Constriction-limited detection efficiency of superconducting nanowire single-photon detectors,” Appl. Phys. Lett. **90**, 101110 (2007).

Kerman, A. J., E. A. Dauler, W. E. Keicher, J. K. W. Yang, K. K. Berggren, G. Gol’tsman, and B. Voronov, “Kinetic-inductance limited reset time of superconducting nanowire single-photon detectors,” Appl. Phys. Lett. **88**, 111116 (2006).

Kitaygorsky, J., I. Komissarov, A. Jukna, D. Phan, O. Minaeva, N. Kaurova, A. Divochiy, A. Korneev, M. Tarkhov, B. Voronov, I. Milostnaya, G. Gol’tsman, and R. R. Sobolewski, “Dark counts in nanostructured NbN superconducting single-photon detectors and bridges,” IEEE Trans. Appl. Supercond. **17**, 275 (2007).

Kitaygorsky, J., J. Zhang, A. Verevkin, A. Sergeev, A. Korneev, V. Matvienko, P. Kouminov, K. Smirnov, B. Voronov, G. Gol’tsman, and R. R. Sobolewski, “Origin of dark counts in nanostructured NbN single-photon detectors,” IEEE Trans. Appl. Supercond. **15**, 545 (2005).

Kogan, V. G., “Pearl’s vortex near the film edge,” Phys. Rev. B **49**, 15874 (1994).

Kogan, V. G., “Erratum: Pearl’s vortex near the film edge,” Phys. Rev. B 75, 069902 (2007).

Korneev, A., A. Lipatov, O. Okunev, G. Chulkova, K. Smirnov, G. Gol’tsman, J. Zhang, W. Slysz, A. Verevkin and R. Sobolewski, “GHz counting rate NbN single-photon detector for IR diagnostics of VLSI CMOS circuits,” Microelectronic Eng. **69**, 274 (2003).

Korneev, A., P. Kouminov, V. Matvienko, G. Chulkova, K. Smirnov, “Sensitivity and gigahertz counting performance of NbN superconducting single-photon detectors,” Appl. Phys. Lett. **84**, 5338 (2004).

Korneev, A., P. Kouminov, V. Matvienko, G. Chulkova, K. Smirnov, B. Voronov, M. Currie, W. Lo, K. Wilsher, J. Zhang, W. Slysz, A. Pearlman, A. Verevkin, R. Sobolewski, G. N. Gol’tsman, “Sensitivity and gigahertz counting performance of NbN superconducting single-photon detectors,” Appl. Phys. Lett. **84**, 5339 (2004).

Lamb, W. E., M. O. Scully, *The Photoelectric Effect without Photons* (Univ. France Press, Paris, 1969).

Likharev, K. K., “Superconducting weak links,” Rev. Mod. Phys. **51**, 101-159 (1979)

Lita, A. E., A. J. Miller, S. W. Nam, “Counting near-infrared single-photons with 95% efficiency,” Opt. Express **16**, 3032 (2008).

Majewski, M. L. and J. R. Scott, “Modeling and Characterization of Microstrip-to-Coaxial Transitions,” IEEE Trans. Micro. Th. and Techniques **29**, 799 (1981).

McManus, M. K., A. Kash, S. E. Steen, S. Polonsky, J. C. Tsang, D. R. Knebel and W. Huott, “PICA: Backside failure analysis of CMOS circuits using Picosecond Imaging Circuit Analysis, ” Microelectronics Reliability **40**, 1353 (2000).

Meservey, R., P. M. Tedrow, “Measurements of the kinetic inductance of superconducting linear structures,” J. Appl. Phys. **40**, 2028 (1969).

Miki, S., M. Takeda, M. Fujiwara, M. Sasaki, A. Otomo, Z. Wang, “Superconducting NbTiN nanowire single photon detectors with low kinetic inductance,” Appl. Phys. Express **2**, 075002 (2009).

Miller, J.R., I. Novak, C. Tai-yu, "Calculating partial inductance of vias for printed circuit board modeling," Electrical Perf. of Elecronic. Packaging **123** (2002).

Morgan, M. and S. Weinreb, “A Millimeter-wave perpendicular coax-to-microstrip transition,” IEEE Trans. Micro. Th. and Techniques **50**, 817 (2002).

Nahory, R. E., M. A. Pollack, and J. C. DeWinter, “Growth and characterization of liquid - phase epitaxial $In_xGa_{1-x}As$,“ J. Appl. Phys. **46**, 775 (1975).

Nam, S. W. from a presentation available online at: http://qubit.nist.gov/qiset-PDF/Nam.QISET2004.pdf.

Nicol, J., T. P. Tseng, “Thermal conductivity of copper between 0.25 K and 4.2 K,” Phys. Rev. **92**, 1062 (1953).

P. Mirin, D. Rosenberg, "Single-photon detection using a quantum dot optically gated field-effect transistor with high internal quantum efficiency," Appl. Phys. Lett. **89**, 253505 (2006).

Pavesi, L., M. Ceschini, "Stretched-exponential decay of the luminescence in porous silicon," Phys. Rev. B **48**, 17625 (1993).

Peroni, I., E. Gottardi, A. Peruzzi, G. Ponti, G. Ventura, "Thermal conductivity of manganin below 1 K," Nuc. Phys. B **78**, 573 (1999).

Planck, M. "On the law of distribution of energy in the normal spectrum," Annalen der Physik 4, 553 (1901).

Polonsky, S., presentation at the Yale Dept. of Applied Phys. (2005).

Prober, D. E., Teufel, J.D., C.M. Wilson, L. Frunzio, M. Shen, R.J. Schoelkopf, T.R. Stevenson, and E.J. Wollack, "Ultrasensitive Quantum-Limited Far-Infrared STJ Detectors," IEEE Trans. Appl. Supercond. **17** (2), 241 (2007).

Ptitsina, N. G., G. M. Chulkova, K. S. Il'in, A. V. Sergeev, F. S. Pochinkov, E. M. Gershenzon and M. E. Gershenzon, "Electron-phonon interaction in disordered metal films: The resistivity and electron dephasing rate," Phys. Rev. B **56**, 16 (1997).

Quranata, O., private communication (2008).

Reese, M. O., "Superconducting Hot Electron Bolometers for Terahertz Sensing," Ph.D dissertation, Yale University (2006).

Robinson, B. S., A.J. Kerman, E.A. Dauler, R.J. Barron, D.O. Caplan, M.L. Stevens, J.J. Carney, S.A. Hamilton, J.K.W. Yang, and K.K. Berggren, "781 Mbit/s photon-counting optical communications using a superconducting nanowire detector," Opt. Lett. **31**, p. 444 (2006).

Rooks, M., private communication (2010).

Rosa, E. B., "The Self and Mutual Inductances of Linear Conductors", Bulletin of the Bureau of Standards **4**, 301 (1908).

Rosenberg, H. M., "The thermal conductivity of metals at low temperatures," Phil. Trans. Royal Soc. London A **247**, 441 (1955).

Rowe, M. A., E. J. Gansen, M. Greene, R. H. Hadfield, T. E. Harvey, M. Y. Su, S. W. Nam, R. P. Mirin, D. Rosenberg, "Single-photon detection using a quantum dot optically gated field-effect transistor with high internal quantum efficiency," Appl. Phys. Lett. **89,** 253505 (2006).

Safwat, A. M. E., W. Johnson, C. H. Lee, “Novel Design for Coplanar Waveguide to Microstrip Transition,” IEEE Trans. Micro. Th. and Techniques **49**, 607 (2001).

Santavicca, D. F., private communication (2010).

Santavicca, D. F., “Bolometric Response of Superconducting Microbridges and Single-Walled Carbon Nanotubes,” Ph.D dissertation, Yale University (2009).

Santavicca, D. F., B. Reulet, B. S. Karasik, S. V. Pereverzev, D. Olaya, M. . Gershenson, L. Frunzio, D. E. Prober, “Energy resolution of terahertz single-photon-sensitive bolometric detectors,” Appl. Phys. Lett. **96**, 083505 (2010).

Santhanam, P., S. Wind, D. E. Prober, “Localization, superconducting fluctuations, and superconductivity in thin-films and narrow wires of aluminum,” Phys. Rev. B **35**, 3188 (1987).

Saracila, G. F., M. N. Kunchur, “Ballistic acceleration of a supercurrent in a superconductor,” Phys. Rev. Lett. **102**, 077001 (2009).

Semenov, A. D., G. N. Gol’tsman, A. A. Korneev, “Quantum detection by current carrying superconducting film,” Phys. C **351**, 349 (2001).

Semenov, A. D., P. Haas, B. Gunther, H. W. Hubers, K. Il’in, M. Siegel, “Energy resolution of a superconducting nanowire single photon detector,” J. Low Temp. Phys. **151**, 564 (2008).

Semenov, A. D., R. S. Nebosis, Yu. P. Gousev, M. A. Heusinger, and K. F. Renk, “Analysis of the nonequilibrium photoresponse of superconducting films to pulsed radiation by use of a two-temperature model,” Phys. Rev. B **52**, 581 (1995).

Semenov, A., A. Engel, K. Il’in, G. Gol’tsman, M. Siegel, H. W. Hubers, “Ultimate performance of a superconducting quantum detector,” Eur. Phy. J. AP **21**, 171 (2003).

Semenov, A., P. Haas, H. W. Hubers, K. Il’in, M. Siegel, A. Kirste, D. Drung, T. Shurig, A. Engel, “Intrinsic quantum efficiency and electro-thermal model of a superconducting nanowire single-photon detector,” J. Mod. Opt. **56**, 345 (2009).

Sergeev, A., A. Semenov, V. Trifonov, B. Karasik, G. Gol'tsman and E. Gershenzon, “Heat transfer in YBaCuO thin film - sapphire substrate system,” J. Supercond. **7**, 341 (1994).

Shibata, H., T. Maruyama, T. Akazaki, H. Takesue, T. Honjo, Y. Tokura, “Photon detection and fabrication of MgB2 nanowire,” Physica C: Supercond. **468**, 1992 (2008)

Shields, A. J., M. P. O'Sullivan, I. Farrer, D. A. Ritchie, R. A. Hogg, M. L. Leadbeater, C. E. Norman, and M. Pepper, “Detection of single photons using a field-effect transistor gated by a layer of quantum dots,” Appl. Phys. Lett. **76**, 3673 (2000).

Skocpol, W. J., M. R. Beasley, and M. Tinkham, “Self-heating hotspots in superconducting thin-film microbridges,” J. Appl. Phys. **45**, 4054 (1974).

Stevens, M. J., R.H. Hadfield, R.E. Schwall, S.W. Nam, R.P. Mirin, and J.A. Gupta, “Fast lifetime measurements of infrared emitters using a low-jitter superconducting single-photon detector,”Appl. Phys. Lett. **89**, 031109 (2006).

Tinkham, M., *Introduction to Superconductivity*, Second Edition, Mineola, NY: Dover Publications (1996).

Weisand, J. G. “Properties of materials at low temperatures,” presentation from Stanford Linear Accelerator Laboratory, located at: http://extenv.jpl.nasa.gov/presentations/Low_Temperature_ Properties.pdf, and accessed on Jan. 18, 2010, presentation from (2003).

Wilson, C.M., K. Segall, L. Frunzio, L.Li, D.E. Prober, D. Schiminovich, B. Mazin, C. Martin and R. Vasquez, "Optical/UV Single-Photon Imaging Spectrometers Using Superconducting Tunnel Junctions," Nuclear Instr. and Methods in Phys. Research A **444**, 449 (2000).

Yang, J. K. W., A. J. Kerman, E. A. Dauler, V. Anant, K. M. Rosfjord, and K. K. Berggren, “Modeling the electrical and thermal response of superconducting nanowire single-photon detectors,” IEEE Trans. Appl. Supercond. **17**, 581 (2007).

Yang, J. K., “Superconducting Nanowire Single-Photon Detectors and Sub-10-nm Lithography,” Ph.D dissertation, Massachusetts Institute of Technology (2009).

Zaitlin, M. P., A. C. Anderson, “The thermal conductivity of silver and aluminum in the temperature range 2-5 K,” J. Low Temp. Phys. **9**, 467 (1972).