Ultrafast NbN Single-Photon Optical Detectors for Quantum Communications

by

Aaron J. Pearlman

Submitted in partial fulfillment of the Requirements for the Degree Doctor of Philosophy

Supervised by Professor Roman Sobolewski

Department of Electrical and Computer Engineering The College School of Engineering and Applied Sciences

University of Rochester Rochester, New York

2006
To Eleanor and William Pearlman
Aaron Pearlman was born in Madison, Wisconsin on December 15, 1978. He attended Tufts University from 1997 to 2001, and graduated with a Bachelor of Science degree in electrical engineering in 2001. He came to the University of Rochester in the Fall of 2001 and began graduate studies in optoelectronics in the electrical and computer engineering department and received the Master of Science degree in 2003. He pursued his research in superconducting photodetectors and quantum communications under the direction of Professor Roman Sobolewski. Mr. Pearlman is a member of IEEE and Materials Research Society.
List of refereed journals and published conference proceedings

2006


2005


14. "Superconducting nanostructured detectors capable of single photon counting of


2004


2003


Acknowledgements

I would first like to express my deep gratitude to my advisor, Prof. Roman Sobolewski for his support, encouragement, and direction and for all the opportunities he provided me to advance my studies and assist in my development as a researcher.

I would also like to acknowledge past and present members of the Ultrafast Phenomena research group who have helped me during my studies. I have enjoyed studying in this group a great deal on both an academic and personal level. In particular, I would like to thank Allen Cross, Dr. Jin Zhang, Dr. Alexander Verevkin, and Dr. Wojciech Slyszy for their valuable assistance and discussion. In addition, I would like to recognize the kind help from Dr. Arturas Jukna on hotspot photoresponse modeling. In addition, I appreciate the assistance of Dr. Ivan Komissarov, Jennifer Kitaygorsky, and Dong Pan on the SSPD project.

I am also quite grateful to all of our collaborators. In particular, I value the help from Dr. Alexander Korneev on novel SSPD geometries, Dr. Marc Currie on GHz counting rate experiments, and Dr. Val Zwiller on correlation experiments. Also, Dr. Sae Woo Nam, Dr. Aaron J. Miller, Dr. Robert Hadfield, and Dr. Robert Schwall provided with a great deal of assistance for low temperature measurements at NIST in Boulder, Colorado. In addition, I thank Dr. A. James Kerman, Eric Dauler, Dr. Steven Constantine, Dr. William Keicher, Joel Yang, and Dr. Karl Berggren for valuable technical discussions mainly on photoresponse modeling. I thank our collaborators from BBN technologies including Dr. Henry Yeh and Dr. Jonathan Habif for their kind sharing of resources on the DARPA quantum key distribution
network. Of course, I am very thankful to Dr. Gregory Gol’tsman and his research group at Moscow State Pedagogical University for fabricating and supplying SSPDs for this work.

This work was supported by the AFOSR grant FA9560-04-1-0123. Additional financial support was provided by the MIT Lincoln Laboratory and the BBN Technologies.
Abstract

We evaluate the NbN single-photon detector (SSPD) for the purpose of integration into a fiber-based quantum communication system, namely the DARPA quantum key distribution (QKD) network. We first review free-space system measurements to characterize the SSPD in terms of counting rate and timing jitter and then demonstrate its utility in fiber-based systems in two such systems. The first utilizes fiber-coupled SSPDs placed in a cryogen-free refrigerator capable of reaching mK temperatures, and the SSPDs are evaluated in terms of system quantum efficiency (SQE) and dark counts over a broad temperature range. The second system, utilizes fiber-coupled SSPDs assembled on an insert placed in a standard helium dewar with each fiber permanently glued to a device. The SSPDs, evaluated in terms of SQE, dark counts, and timing resolution, show that the system provides relatively high fiber-detector coupling efficiency, good timing resolution, and can integrate easily into the DARPA network.

We also investigate the SSPD’s limitations by analyzing a model which takes into account the SSPD detection mechanism and device inductance to predict its response time. We then optimize the SSPD meander geometry in designing devices with high SQE and counting rate in terms of area, stripe width, fill factor, and thickness using detailed inductance simulations. We will also present a novel low inductance SSPD design and model its photoresponse.

With these designs and measurement results, we will show that the SSPD outperforms its superconducting and semiconducting counterparts for quantum cryptography systems with high clock rates. Thus, the SSPD, with its combination of
high QE, and low timing jitter at telecommunications wavelengths, as well as low dark counts, make it a natural choice for the DARPA network and quantum cryptography systems in general.
Table of contents

1 Introduction to quantum cryptography................................................. 1
   1.1 Basic concepts.............................................................................. 1
   1.2 Hardware..................................................................................... 7

2 Superconducting single-photon detector (SSPD):
   physics and device operation............................................................ 13
   2.1 Superconducting state................................................................. 13
   2.2 Photoresponse............................................................................ 13
   2.3 Hotspot model and phase slip centers ........................................ 16
   2.4 Temperature effects.................................................................... 22
   2.5 Inductance.................................................................................. 29
   2.6 SSPD performance...................................................................... 31
      2.6.1 Single-photon detection......................................................... 31
      2.6.2 Quantum efficiency (QE) and spectral sensitivity................. 32

3 SSPD performance at telecommunication wavelengths:
   free-space system results ................................................................. 35
   3.1 Free-space setup......................................................................... 35
   3.2 Counting rate.............................................................................. 36
   3.3 Timing jitter............................................................................... 40

4 SSPD performance at telecommunication wavelengths:
   fiber system results.......................................................................... 43
   4.1 NIST fiber setup.......................................................................... 43
CONTENTS

4.1.1 QE versus temperature .............................................................. 48
4.1.2 Dark count results ................................................................. 54
4.2 Warsaw fiber setup ................................................................. 56
4.2.1 QE ................................................................. 58
4.2.2 Timing jitter ................................................................. 63

5 SSPD optimization ................................................................. 68
5.1 Motivation – QE and response time ............................................. 68
5.2 SSPD inductance calculations ..................................................... 74
5.3 SSPD high counting rate/high QE design geometry ..................... 77
  5.3.1 Inductance ................................................................. 77
  5.3.2 Device area and fiber-coupling ............................................. 78
  5.3.3 Design specifications ........................................................ 80
5.4 Alternative SSPD geometries ...................................................... 83
5.5 Response time simulation and measurement ................................ 87
5.6 Single-photon detector (SPD) comparison .................................... 96

6 Quantum key distribution system based on SSPDs:

  performance analysis ................................................................. 99
  6.1 DARPA network ................................................................. 99
  6.2 Practical QKD ................................................................. 101
  6.3 Review of recent QKD links .................................................... 103
    6.3.1 APD/SSPD QKD testbed ................................................. 103
    6.3.2 TES QKD link ........................................................... 103
    6.3.3 SSPD QKD link testbed with entangled photon source .......... 104
<table>
<thead>
<tr>
<th>CONTENTS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>6.4 Performance estimates for DARPA QKD link</td>
<td>105</td>
</tr>
<tr>
<td>7 Conclusions and future work</td>
<td>109</td>
</tr>
<tr>
<td>Bibliography</td>
<td>111</td>
</tr>
</tbody>
</table>
List of Tables

4.1-1 SSPDs measured in NIST setup. The data in the table were provided by our Moscow collaborators. 46

4.2.1-1 Parameters of fiber-coupled devices in Warsaw setup. 60

5.6-1 SPD comparison. 98
List of Figures

1.1-1 Depiction of BB84 protocol......................................................... 4

1.1-2 (a) The waveforms resulting from the photons traveling in each arm. (b) A photon at the output of the second coupler is shown. If the interferometers are set up such that they are 180 degrees out of phase, the photons will register in D0 but not D1................................................................. 5

2.2-1 Photoresponse of a superconductor. (a) A photon breaks a Cooper pair to create a highly excited quasiparticles, which interact with electrons and phonons. (b) Quasiparticles recombine to form Cooper pairs and emit phonons. (c) Phonons break additional Cooper pairs. (d) Phonons escape to substrate. .............................. 15

2.2-2 Hotspot profiles showing the concentration of quasiparticles at times 0.8, 2, and 5 in units of thermalization time measured in distance units of thermalization length ................................................................. 18

2.2-3 (a) SSPD meander (b) Hotspot model for single-photon detection................................................................. 19

2.4-1 Normalized concentration of nonequilibrium quasiparticles versus distance in a 4-nm-thick NbN film at several values of $T/T_c$ with $\lambda = 1.55 \mu$m and QE = 1................................................................. 23

2.4-2 Time evolution of the normalized hotspot diameter at several values of $T$, where $T_c = 10$ K. ................................................................. 24
2.4-3 The growth in the diameter versus temperature after photon absorption according to the hotspot model [36] ................................................................. 25

2.4-4 Hotspot diameter versus reduced temperature for a 4 nm-thick NbN film illuminated with a $\lambda = 1550$ nm photon................................................................. 26

2.4-5 Hotspot distribution including initial hotspot of radius ($R_m$), electric field penetration of length ($L_E$), and N-S propagation velocities $v_w$ and $v_l$ across the width and along the length, respectively.................................................. 28

2.6.2-1 Spectral dependence of QE at $T = 4.2$ K. The inset shows the estimated hotspot size at varying photon energies.................................................. 33

2.6.2-2 QE and dark counts as a function of normalized bias. Measurements performed at 4.2 K (closed symbols) and 2.3 K (open symbols). Solid lines represent an exponential fit. [26] ................................................................. 34

3.1-1 Free-space experimental setup................................................................. 36

3.2-1(a) Timing jitter results (b) Histogram technique........................................... 37

3.2-2 Hotspot location in non-uniform stripe causing current density variation in sidewalks................................................................. 38

3.2-3 (a) The hotspot is confined at the edge of the stripe with a higher maximum qp concentration $N_a$. (b) The hotspot is formed at the center, resulting in a lower quasiparticle concentration $N_b$. (c) Time profiles of $N_a$ and $N_b$ onsets of the resistive state $R_a$ and $R_b$, and their resulting time delays $t_{d_a}$ and $t_{d_b}$, for the case (a) and (b), respectively................................................................. 39
3.3-1 (a) SSPD response to 1-GHz and 2-GHz pulses at higher flux. (b) Bandwidth of SSPD system. (c) SSPD response pulse excited with a laser with a repetition rate of 2-GHz at low flux.

4.1-1 (a) Detector holder for fiber coupling in ADR. (b) Wire bonds to SSPD.

4.1-2 Fiber experimental setup utilizing ADR.

4.1-3 SSPD (588/1 84) IV curve.

4.1.1-1 Single-photon power dependence for $T = 4\, \text{K}, \, 2\, \text{K}, \, 1\, \text{K},$ and $60\, \text{mK}$ at $\lambda = 1550\, \text{nm}$.

4.1.1-2 Single-photon power dependence for $T = 2\, \text{K}, \, 1\, \text{K},$ and $100\, \text{mK}$ at $\lambda = 1310\, \text{nm}$.

4.1.1-3 (a) QE versus current for all measured temperatures and wavelengths (b) Highest achievable QE as a function of temperature.

4.1.1-4 Temperature dependence $I_c/I_o$ where $I_o$ is taken at the critical current at $T = 60\, \text{mK}$.

4.1.2-1 Dark counts versus normalized bias current from $T = 0.1\, \text{K}$ to $6.4\, \text{K}$.

4.2-1 (a) SSPD receiver for operation inside a standard helium dewar (b) Details of fiber-coupling arrangement.

4.2.1-1 Single-photon detection dependence for fiber-coupled detectors in Warsaw setup.
4.2.1-2 Coupling efficiency K versus fiber-detector distance and horizontal misalignment……………………………………………………………………… 61

4.2.2-1 Timing jitter histograms for selected fiber-coupled devices in Warsaw setup. A 4 x 4 μm² SSPD jitter histogram measured in free space………………………… 63

4.2.2-2 Correlated counts taken using low time resolution (250 ps) time resolved pulse counter……………………………………………………………………… 65

4.2.2-3 Correlated counts between device #11 and #12 measured in Delft setup……………………………………………………………………………….. 67

5.1-1 Basic time delay setup used for QE vs. counting rate measurements………………………………………………………………………………… 69

5.1-2 (a) Relative detection efficiency versus pulse separation for a 132 nH (open circles) and 56.5 nH (closed circles) SSPD with theoretical fit using hotspot/inductance model. (b) Schematic of hotspot/inductance model. [47]…… 70

5.1-3 Measured photoresponse with fitted response predicted by inductance/hotspot model……………………………………………………………………71

5.1-4 Simulated relative DE versus pulse separation time………………………… 72

5.1-5 Relative detection efficiency (DE) versus reduced current and exponential fit ………………………………………………………………………… 73.

5.1-6 Inductance versus response time, which is defined as the response time necessary to maintain 90% of its nominal QE…………………………………… 74
5.2-1 The dependence of cross-sectional area on inductance for a long (~200 \( \mu \text{m} \) length) wire. .................................................................76

5.3.2-1 Some possible geometries for NbN meanders with 2 GHz counting rates with 4 nm film thickness 25 \( \mu \text{m}^2 \) and 100 \( \mu \text{m}^2 \) area......................................................... 79

5.3.3-1 Meander fill factor dependence of inductance for a 70 nm, 4 nm thickness, 25 \( \mu \text{m}^2 \) area device.................................81

5.3.3-2 Fill factor and stripe width for 17 nH meander SSPD necessary to achieve a 2-GHz counting rate. The calculations were performed using the meander geometry and bridge geometry of equivalent length. The lines are guides to the eye...........82

5.4-1 Parallel device geometries: (a) two parallel meander sections with 17 stripes per section, (b) three parallel meander sections with 17 stripes per section, and (c) ten parallel meander sections with 5 stripes per section............................84

5.4-2 (a) Simulated current in a two-section device geometry. The hotspot section refers to the section of the meander in which the photon is initially absorbed; The parallel section is the section of the meander in which no photon is absorbed.; The transmission line refers to the circuit branch containing the 50 \( \Omega \) transmission line (b) Circuit model for Moscow setup consisting of two parallel meanders each with \( L = 146 \text{ nH} \) ................................................................. 85

5.4-3 Simulated photoresponse comparing a low inductance standard meander device and a higher inductance standard meander device with a device with two parallel sections.............................................................. 87
5.5-1 (a) Voltage versus reduced current data and fitting for several SSPDs. (b) Resulting normalized resistance versus time delay calculated using (a) for device #3 and circuit model (c).................................................................89

5.5-2 Measured photoresponse of time delayed pulses with a 6.3 ns pulse separation and simulated photoresponse in wide-band system.................................90

5.5-3 Measured photoresponse in time delay setup for several values of pulse separation. The left column shows single-shot pulses and the right shows averaged pulses.................................................................92

5.5-4 Measured and simulated photoresponse in the narrow-band system......... 93

5.5-5 Simulated responses for time delay values of 4 ns, 2 ns, and 500 ps in the narrow band setup. The green line shows the minimum relative discrimination level that could be used to count both pulses.........................................94

5.5-6 Maximum time above discrimination on falling edge of pulse versus time delay for narrow- and wide-band systems.................................................96

6.1-1 Phase-coded QKD setup........................................................................100

6.4-1 SPD comparison for key rate versus link distance..............................106

6.4-2 SPD comparison for maximum key rate versus clock frequency ..........108
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>APD</td>
<td>Avalanche photodiode</td>
</tr>
<tr>
<td>$a$</td>
<td>Number of bits sacrificed for continuous authentication</td>
</tr>
<tr>
<td>$B$</td>
<td>Bit rate (Probability of breaking Cooper pairs)</td>
</tr>
<tr>
<td>BCS</td>
<td>Bardeen-Cooper-Schrieffer (theory of superconductivity)</td>
</tr>
<tr>
<td>$C$</td>
<td>Quasiparticle concentration</td>
</tr>
<tr>
<td>$C_0$</td>
<td>Equilibrium quasiparticle concentration</td>
</tr>
<tr>
<td>$c_e$</td>
<td>Electron specific heat</td>
</tr>
<tr>
<td>$c_p$</td>
<td>Phonon specific heat</td>
</tr>
<tr>
<td>$D$</td>
<td>Diffusivity</td>
</tr>
<tr>
<td>DE</td>
<td>Detection efficiency</td>
</tr>
<tr>
<td>DQE</td>
<td>Device (intrinsic) quantum efficiency</td>
</tr>
<tr>
<td>$d$</td>
<td>Film thickness</td>
</tr>
<tr>
<td>$d_B$</td>
<td>Dark count probability in Bob’s receiver</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Angular mismatch between Alice and Bob (transmitter and receiver)</td>
</tr>
<tr>
<td>$\Delta \tau /2$</td>
<td>Optical pulse width</td>
</tr>
<tr>
<td>$e$</td>
<td>Number of bits in error (or error rate)</td>
</tr>
<tr>
<td>$\eta_B$</td>
<td>Quantum efficiency of Bob’s detector</td>
</tr>
<tr>
<td>$\eta_T$</td>
<td>Transmission efficiency</td>
</tr>
<tr>
<td>$f$</td>
<td>Counting rate</td>
</tr>
<tr>
<td>$f_{\text{clock}}$</td>
<td>Clock rate</td>
</tr>
<tr>
<td>EPR</td>
<td>Einstein-Podolski-Rosen (protocol)</td>
</tr>
<tr>
<td>$F(T)$</td>
<td>Andreev reflection parameter</td>
</tr>
<tr>
<td>$G$</td>
<td>Gain per time slot</td>
</tr>
<tr>
<td>$g_{\text{pa}}$</td>
<td>Privacy amplification security parameter</td>
</tr>
<tr>
<td>$I$</td>
<td>Current</td>
</tr>
<tr>
<td>$I_b$</td>
<td>Bias current</td>
</tr>
<tr>
<td>$I_r$</td>
<td>Superconducting return current</td>
</tr>
<tr>
<td>$I_o$</td>
<td>Number of quasiparticles created by photon absorption, characteristic current</td>
</tr>
<tr>
<td>$j$</td>
<td>Current density</td>
</tr>
<tr>
<td>$K$</td>
<td>Fiber coupling factor</td>
</tr>
<tr>
<td>$L$</td>
<td>Inductance</td>
</tr>
<tr>
<td>LAMH</td>
<td>Langer-Ambegaokar-McCumber-Halperin (theory)</td>
</tr>
<tr>
<td>$L_E$</td>
<td>Electric field penetration length at normal-superconducting boundary</td>
</tr>
<tr>
<td>$L_{eq}$</td>
<td>Equivalent inductance</td>
</tr>
<tr>
<td>$L_k$</td>
<td>Kinetic inductance</td>
</tr>
<tr>
<td>$L_{\text{th}}$</td>
<td>Thermalization length</td>
</tr>
<tr>
<td>$\lambda_L$</td>
<td>London penetration depth</td>
</tr>
<tr>
<td>$n$</td>
<td>Number of photons</td>
</tr>
<tr>
<td>N-S</td>
<td>normal-superconducting (boundary)</td>
</tr>
<tr>
<td>$M$</td>
<td>Quasiparticle multiplication factor</td>
</tr>
<tr>
<td>$m$</td>
<td>Number of bits sent by Alice (transmitter)</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Mean photon number per pulse</td>
</tr>
<tr>
<td>$\mu_0$</td>
<td>Permeability of vacuum</td>
</tr>
<tr>
<td>$\mu_{\text{opt}}$</td>
<td>Optimal mean photon number</td>
</tr>
<tr>
<td>$N$</td>
<td>Number of quasiparticles</td>
</tr>
<tr>
<td>$\text{NbN}$</td>
<td>Niobium nitride</td>
</tr>
<tr>
<td>$N_w$</td>
<td>Number of phonons</td>
</tr>
<tr>
<td>$N_{\text{wT}}$</td>
<td>Thermal concentration of phonons</td>
</tr>
<tr>
<td>$\text{NIR}$</td>
<td>Near infrared</td>
</tr>
<tr>
<td>$\text{MIR}$</td>
<td>Middle infrared</td>
</tr>
<tr>
<td>$n$</td>
<td>Number of bits received by Bob (receiver)</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Frequency</td>
</tr>
<tr>
<td>$\text{PSC}$</td>
<td>Phase slip center</td>
</tr>
<tr>
<td>$\text{PPLN}$</td>
<td>Periodically poled lithium niobate</td>
</tr>
<tr>
<td>$p_{\text{exp}}^{\text{dark}}$</td>
<td>Dark count probability</td>
</tr>
<tr>
<td>$p_{\text{exp}}^{\text{signal}}$</td>
<td>Signal photon detection probability</td>
</tr>
<tr>
<td>$\text{QE}$</td>
<td>Quantum efficiency</td>
</tr>
<tr>
<td>$\text{QKD}$</td>
<td>Quantum key distribution</td>
</tr>
<tr>
<td>$R$</td>
<td>Quasiparticle recombination coefficient, hotspot resistance</td>
</tr>
<tr>
<td>$\text{RT}$</td>
<td>Rothwarf-Taylor (photoresponse model)</td>
</tr>
<tr>
<td>$R_{\text{dk}}$</td>
<td>Dark count rate</td>
</tr>
<tr>
<td>$R_{\text{th}}$</td>
<td>Hotspot radius</td>
</tr>
<tr>
<td>$R_n$</td>
<td>Normal region resistance</td>
</tr>
<tr>
<td>$R_o$</td>
<td>Dark count prefactor</td>
</tr>
<tr>
<td>$R_s$</td>
<td>Source impedance</td>
</tr>
<tr>
<td>$r_c$</td>
<td>Error rate due to misalignment</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Resistivity</td>
</tr>
<tr>
<td>$\text{SSPD}$</td>
<td>Superconducting single-photon detector</td>
</tr>
<tr>
<td>$S$</td>
<td>Effective secrecy rate (key rate)</td>
</tr>
<tr>
<td>$S_m$</td>
<td>Multi-photon pulse probability</td>
</tr>
<tr>
<td>$\text{SQE}$</td>
<td>System quantum efficiency</td>
</tr>
<tr>
<td>$s$</td>
<td>Number of bits sacrificed for privacy amplification</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Cross-sectional area</td>
</tr>
<tr>
<td>$\text{TES}$</td>
<td>Transition edge sensor</td>
</tr>
<tr>
<td>$T_c$</td>
<td>Critical temperature</td>
</tr>
<tr>
<td>$t_d$</td>
<td>Time delay between appearance of supercritical current and registering of voltage pulse</td>
</tr>
<tr>
<td>$t_r$</td>
<td>Superconductivity recovery time</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Detector dead time</td>
</tr>
<tr>
<td>$\tau_3$</td>
<td>Gap relaxation time</td>
</tr>
<tr>
<td>$\tau_{e-p}$</td>
<td>Electron-phonon interaction time</td>
</tr>
<tr>
<td>$\tau_{e-s}$</td>
<td>Phonon-substrate escape time (2-temperature model)</td>
</tr>
<tr>
<td>$\tau_{\text{RT}}$</td>
<td>Phonon-substrate escape time (Rothwarf-Taylor model)</td>
</tr>
<tr>
<td>$\tau_Q$</td>
<td>Charge imbalance relaxation time</td>
</tr>
<tr>
<td>$\tau_{\text{th}}$</td>
<td>Thermalization time</td>
</tr>
<tr>
<td>$V$</td>
<td>Voltage amplitude</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>$V_o$</td>
<td>Voltage amplitude at low counting rate</td>
</tr>
<tr>
<td>$v_l$</td>
<td>Propagation velocity along length</td>
</tr>
<tr>
<td>$v_w$</td>
<td>Propagation velocity along width</td>
</tr>
<tr>
<td>$w$</td>
<td>Stripe width</td>
</tr>
<tr>
<td>$\xi$</td>
<td>Quasiparticle avalanche loss parameter</td>
</tr>
<tr>
<td>$2\Delta$</td>
<td>Superconducting energy gap</td>
</tr>
<tr>
<td>$2T$</td>
<td>Two-temperature (photoresponse model)</td>
</tr>
</tbody>
</table>
Chapter 1 Introduction

The NbN superconducting single-photon detector (SSPD) has recently found applications in CMOS circuit testing [1]. The focus of this thesis is the performance of the SSPD for quantum cryptography. We will present our experimental results and then show that the SSPD outperforms its superconducting and semiconducting counterparts. We will also discuss the SSPD’s inherent limitations and ways to optimize its performance. Other potential applications of SSPDs include single-molecule fluorescence and high-resolution astronomy, linear optical quantum computations, interplanetary communication, and quantum metrology.

1.1 Quantum cryptography: basic concepts and hardware

Cryptography is used to keep a message private by preventing unauthorized parties from accessing it. Various cryptographic methods have been used for centuries in order to transmit classified information [2]. In modern times, this need has expanded both for military and commercial purposes. For instance, the growing amount of information sent over networks has placed a large burden in making sure private information is secure. This demand remains difficult and in reality, much of the information that is exchanged over a typical network is not encrypted. Additionally, the security of the types of encryption most commonly used relies on algorithms that take too much computational power to solve within a practical period of time. Therefore, it is not provably secure. In 1984, the idea of quantum cryptography was proposed, which unlike classical schemes, does not rely on
computational power, but on the laws of quantum mechanics; thus, it is provably secure.

In general, cryptography uses a key to encrypt a message. The key is used in an algorithm called a cipher to process the message into a cryptogram. The message can be restored using a second key. Symmetrical systems require the same key for encryption and decryption, while asymmetrical ones require different keys for encryption and decryption. For instance, in the widely used RSA cipher [3], a public key generated from a private key is used to encrypt the message. Anyone can encrypt the message, but only a party with the private key can decode it directly. The security of the cipher used in this protocol lies in the complexity of factoring large numbers, meaning that an extremely large amount of computing power is necessary to generate the private key from the public one. This method is not provably secure.

There is only one protocol, a symmetrical system called the “one-time pad” that is provably secure using information theory [4]. In this method, a key, which is composed of a random bit string, is added to the message bits (using, e.g., XOR functions) to create the cryptogram. This can only be decoded by subtracting the same key from the cryptogram to retrieve the original message. The problem is that the sender (Alice) and the receiver (Bob) must possess the same key, the key must have the same bit length as the message, and can be used only once. Therefore, a trusted channel is required for the “one-time pad” approach to ultimately be secure. Today, due to the above drawbacks asymmetrical systems are used for the most sensitive messages. Therefore, the security still relies on the limited computational power of the eavesdropper.
The only way to create a completely secure cryptosystem is to utilize a quantum cryptography and, more specifically, a quantum key distribution (QKD) system with a one-time pad, which together do not depend on the limited computational power of the eavesdropper (Eve). Even a quantum computer, which is expected to have dramatically increased computational power due to parallel processing of data, cannot breach the security of a quantum cryptography system [5]. Due to very limited transmission efficiency, the typical quantum cryptography protocol does not use a full one-time pad, but only distributes a key of a certain length using a quantum protocol. This QKD protocol takes advantage of the unique attributes of quantum mechanics by using single photons to transmit the key.

The main attribute of quantum mechanics that makes a single photon a secure medium is that its quantum state cannot be measured without perturbing it. It follows that if two photons are in non-orthogonal states they cannot be distinguished without perturbing at least one of them [6]. Furthermore, the so-called “no-cloning theorem,” states that an unknown quantum state (e.g., our photon state) cannot be copied exactly. In a QKD channel, the eavesdropper can try to measure the photons that are being sent from Alice to Bob, but in the process, will perturb the photon states. This will introduce errors in the transmission process, so Eve’s presence can be detected.

The first protocol for QKD was proposed by Charles Bennett and Gilles Brassard in 1984 and is referred to as the BB84 protocol [7]. In this scheme, the polarization of the photons is used to encode the quantum bits or qubits. Two bases that are not orthogonal are used so that, for instance, vertical and horizontal polarizations could represent a “1” or “0”, respectively, in one basis and orthogonal
and diagonal states would represent the bit values in the other basis. The first step in the protocol is the generation of a random key. This key is encoded by Alice in either of the two bases at random. These qubits are then sent through a “private” (or “quantum”) channel to Bob. Bob tries to guess polarizations imprinted by Alice, by measuring photons in a random choice of basis for each photon. This BB84-type protocol process is depicted in Figure 1.1-1, where both Alice and Bob make a random choice of bases by rotating polarizers. After the transmission, Bob and Alice publicly compare these polarization settings (but not the bit values) and retain only the bit values in cases where their settings are identical. This way they are simultaneously in possession of the same random string of bits— the quantum key.

Another method of sending qubits is by encoding them using phase. In the QKD system that will be analyzed in Chapter 6, the phase of the attenuated laser...
pulse (weak coherent state), which approximates our single photon state, is used to encode each bit using a Mach-Zehnder interferometer [8]. Figure 1.1-2 illustrates the way in which this phase protocol is accomplished. The waveform at Bob’s coupler is shown in Figure 1.1-2(b). For the BB84 protocol, both Alice and Bob adjust the phases by applying voltages to their phase shifters to yield phase shifts of 0 or $\pi$ in one basis or $\pi/2$ or $3\pi/2$ in another. If the phase differences of the waveforms at the central peak are 0 or $\pi$, the bases are compatible, and one detector will respond, while the other does not predictably. Otherwise, the photon will travel randomly to one of the detectors. As a consequence of the Heisenberg uncertainty principle, this scheme would not work in the case a true single-photon source, since we know that there is exactly one photon in each pulse, making the phase indeterminate.

Figure 1.1-2 (a) The waveforms resulting from the photons traveling in each arm. (b) A photon at the output of the second coupler is shown. If the interferometers are set up such
The qubits that are received by Bob make up the raw key. 25% of the bits in the key are incorrect, as they contain errors due to Bob’s measurements, so the useful key is shortened. In order to accomplish this, as we mentioned above, Alice announces the bases she used to measure the photons over a public channel. Then, Bob communicates which photons he measured in the same basis as Alice, and these qubits make up the sifted key, while discarding all of the others. If Eve wants to intercept the qubits, she can measure them the same way that Bob does and then send the new photons with known polarizations to Bob. However, this will introduce an additional 25% of errors into the sifted key and Alice and Bob will easily detect Eve’s presence. Eve can use other strategies that are not quite as easy to detect, but regardless, she always contributes some additional errors in the sifted key. The larger Eve’s “knowledge”, the more “visible” she is to Alice and Bob.

In order to eliminate errors caused by technical imperfections, classical error correction protocols are used. Even though we can detect Eve’s presence, we may still want to transmit the key. In such a situation, we can reduce the correlation between Alice and Bob’s information, so Eve’s knowledge is limited to an arbitrarily low value as long as the errors are kept below a certain threshold. A process called privacy amplification is employed to this end. Once error correction and privacy amplification have been completed and the additional step of continuous authentication, the final key is then obtained. The figure of merit for a QKD system is the effective secrecy capacity $S$, or the effective secrecy rate, or just key rate, which is just $S$ divided by the time it takes for one bit to be sent from Alice to Bob and the system to reset, $\tau$. The expression for $S$ is as follows:
\[ S = \frac{n - e - s - g_{pa} - a}{m}, \]  

where \( m \) is the number of bits sent by Alice, \( n \) is the number of bits received by Bob, \( e \) is the number of bits that are in error, \( s \) is the number of bits sacrificed for privacy amplification, \( g_{pa} \) is the privacy amplification security parameter, and \( a \) is the number of bits sacrificed for continuous authentication. \( g_{pa} \) tells us the number of additional bits which need to be sacrificed during privacy amplification in order to bound Eve’s information to less than \( 2^{-g_{pa}} \ln 2 \). Authentication refers to the verification that Eve is not disguising herself as Bob to obtain the key and requires the sacrificing of additional bits.

The BB84 protocol described above is only one of the available protocols for QKD. Other two-state [9], six-state [10], and Einstein-Podolski-Rosen (EPR) protocols [11] have been proposed. The EPR protocol utilizes an entangled photon source where a pair consisting of two correlated photons is split and one photon is sent to Alice and the other to Bob. Our work will mostly focus on BB84 type protocols, which use a single-photon source and single-photon detector (SPD), the object of our research.

1.2 Hardware

Source

For the most part, we will restrict ourselves to highly attenuated laser sources operating at fluxes below one photon per pulse as “single-photon sources”. Since a
laser source obeys a Poisson distribution, a non-zero probability exists of more than one photon per pulse transmitting. If a double-photon event occurs, Eve could split the pulse and measure one photon, while leaving the other to travel to Bob without perturbing it. In such a case, she will not introduce errors in Bob’s key and will remain undetectable. However, the probability of a double-photon event can be made arbitrarily small to prevent Eve from using this beamsplitting attack successfully. Actually, a maximum mean photon number exists for which a secure key can still be distilled with given losses and errors [12]. Thus, Eve’s interference can be ignored, since Alice and Bob can control the number of double-photon events and thereby transmit the key securely. The problem may, however, result in the data rate becoming impractically slow. From this standpoint, a laser with a high repetition rate is highly desirable in order to maintain a reasonably high data rate. Current lasers can produce femtosecond pulses at GHz rates. One option is the Gigajet laser, which is a Ti:Al$_2$O$_3$ laser with a 3 GHz repetition rate with $\leq 30$ fs- wide pulses. The laser used in our free-space setup is the Ultrafast Optical Clock fiber-based mode-locked laser from Pritel that generates 1.6 ps-wide pulses at a repetition rate ranging from 1 GHz to 10 GHz at 1.55 $\mu$m [13].

In addition to analyzing the current DARPA network, which currently uses a highly attenuated laser source, we will consider the future designs of the network, which will employ entangled photon sources [14]. These sources will utilize a highly attenuated femtosecond laser in conjunction with periodically-poled lithium niobate (PPLN) crystals [15] for efficient generation of down-converted, Poisson-distributed photon pairs. The further development in these sources and photon number-resolving
detectors can lead to security gains in quantum cryptography. The reason is related to beamsplitting attacks, which as we discussed, can compromise the security of a QKD system. This security loophole can be avoided because Bob can be warned by Alice when a multi-photon is recorded (by using a photon-number resolving detector), so the qubits associated with these events can then be ignored by Bob [16,17]. This will lead to an extension of the possible length of a point-to-point link and higher key rates.

For completeness, we mention that much progress has recently been made in achieving true single-photon sources at telecom wavelengths. For instance, semiconductor quantum dots have been recently fabricated, which show strong evidence of single-photon emission at 1300 nm [18,19]. Even though these would seem to be the best option for single-photon sources, they are not truly “on-demand” sources of photons and typically suffer from low coupling efficiencies from the quantum dot to a single-mode fiber. Clearly, improvements in on-demand single-photon sources will lead to higher key rates and transmission distances.

**Quantum channels**

**Single-mode fiber**

There are two types of quantum channels typically implemented: single-mode fiber and free-space propagation. As mentioned above, the losses in the channel directly affect the effective speed (data rate) at which the QKD system operates. Single-mode fibers have losses typically ~0.2 dB/km at 1550 nm wavelength due to absorption, scattering, and external factors such as micro-bending. Dispersion is also
an important concern in single-mode fibers, because it also can affect the bit rate. When a pulse broadens too much, consecutive pulses will not be distinguishable in time. If $\Delta \tau_{1/2}$ is the pulse width at Bob, then the bit rate, $B = 0.5/\Delta \tau_{1/2}$.

The type of dispersion that is most likely to affect the bit rate is polarization mode dispersion, which is the effect of having two different group velocities for two different orthogonal polarizations. This effect, which arises due to imperfections in the fabrication of the fiber, splits the optical pulse into fast and slow modes that decouple over time. A typical amount of depolarization dispersion is $0.1 \text{ ps/(km)}^{1/2}$, which yields a dispersion delay of $0.7 \text{ ps}$ for a 50-km long fiber and is not going to limit the system using current technology. Chromatic dispersion originates from the index of refraction dependence of the wavelength. For instance, assuming a fiber length of 50 km and a laser linewidth of 0.8 nm, the dispersion delay or twice the pulse width would be approximately 160 ps. In order to keep this delay lower than the coherence time of the laser, the repetition rate of the laser cannot exceed 6.25 GHz. If a system operates at this speed, dispersion compensating fiber must be inserted into the line.

**Free-space links**

In a free-space link, light experiences attenuation and scattering in the atmosphere that accounts for most of the loss. In addition, beam spreading leads to geometrical losses. These losses depend greatly on weather conditions in the lower atmosphere. Gilbert simulated various conditions in the atmosphere and calculated the total line attenuation at 1550 nm wavelength as a function of the aperture size of
the receiver optics when transmitting from a low Earth satellite (at 300 km altitude) to an aircraft at 35000 feet and to sea level [20]. For instance, for an aperture of 1 m, the line attenuation is about 5 dB to the aircraft and about 15 dB to sea level from the satellite. Misalignment of Alice and Bob is more of a concern in free-space than through fiber due to moving platforms when, for instance, the communication is between a satellite and Earth. The fractional error rate caused by misalignment can be expressed as $r_c = \sin^2(\delta)$, where $\delta$ is the angular mismatch between Alice and Bob. Estimates of $r_c$ using current technology typically yield values less than 0.01 [13].

**Quantum receivers**

Other problems for a QKD system can manifest themselves in imperfections in the reception of the photons. The effect can be seen by considering Equation 1.1.1. Errors can be generated by dark counts in the detector, which will lower the effective secrecy, $S$. In addition, limited quantum efficiency (QE) affects $S$, since the number of bits, $n$, received by Bob will decrease. The counting rate of the detector and its timing jitter also affect $S$, since the counting rate will affect $n$, and large timing jitter can cause errors. Therefore, we require a detector that works in the telecommunications wavelength of 1.55 $\mu$m with high QE, low dark counts, low jitter, and at a very high counting rate.

Typical detectors used for single-photon counting are avalanche photodiodes, photomultipliers, and transition edge sensors. This thesis, however, focuses on the use of the NbN superconducting single-photon detector (SSPD) for QKD communications. After reviewing the performance characteristics of these detectors,
based on our recent results, we will show that these detectors provide the best performance for quantum communications.
Chapter 2 SSPD: Physics and device operation

2.1 Superconducting state

Many materials, mostly metals, when kept below a certain critical temperature, magnetic field, and current density, exhibit zero dc resistance and exclude magnetic fields. These materials are termed superconductors. Kammerlingh Onnes first observed zero dc resistance in a mercury sample in 1911 [21]. After the initial discovery, many scientists proposed models to account for the behavior of superconductors. The most well-known and the first microscopic formulation was published in 1957 by Bardeen, Cooper, and Schrieffer [22]. The BCS theory is applicable to what we call now “ordinary s-wave superconductors” and explains most of their important properties.

2.2 Photoresponse

A central property of superconductors for our purposes is its photosensitivity. For instance, Testardi used lead films to demonstrate that laser light can destroy superconductivity in a small region of the sample [23]. The BCS theory explains this phenomenon in terms of perturbing the ground state of the superconductor.

The superconducting ground state consists of paired electrons with opposite momenta and spins, which are formed through virtual phonon interaction. A minimum amount of energy that is necessary to break these “Cooper pairs” is also referred to as the energy gap $2\Delta$, since it separates the ground state of Cooper pairs from the continuum of single-electron (quasiparticle) states in the superconductor.
energy spectrum. For instance, when a photon with energy much greater than \(2\Delta\) is absorbed, a Cooper pair is broken and two excited quasiparticles are formed. Since the coherence length of a Cooper pair is large (much larger than the interatomic distance), only one of the quasiparticles can be highly excited (\(>>2\Delta\)), while the other quasiparticle remains just above the gap edge. The magnitude of the energy gap of ordinary superconductors is typically on the order of a few millielectron-volts and exhibits a dependence on temperature. This dependence is very strong near the superconductor critical temperature \(T_c\) and saturates when the temperature drops below one half of \(T_c\). The saturation value can be approximated as

\[
2\Delta(0) = 3.52k_B T_c, \quad (2.2.1)
\]

where \(2\Delta(0)\) is the energy gap at \(T = 0\) and \(k_B\) is the Boltzmann constant.

After the initial photon absorption by an electron in a Cooper pair, it is excited far above the energy gap, and a complex relaxation process then occurs that involves electron-electron (e-e) and electron-phonon (e-ph) interactions. Figure 2.2-1 shows the series of steps that occur in this process. The highly excited quasiparticle initially relaxes very rapidly (on \(\sim 10\) fs time scale) through interaction with other electrons and break additional Cooper pairs during a cascading process that occurs in the electron thermalization time \(\tau_{th}\). At some point, interaction with phonons becomes more energetically favorable, thereby continuing to break Cooper pairs. The phonons that are emitted can be reabsorbed, and break other Cooper pairs, but the same time, they can escape the superconducting film into the substrate. When enough phonons
escape into the substrate, the relaxation process ceases [24]. Typically, this avalanche process produced by one optical photon can generate 100 to 1000 quasiparticles.

This photoresponse can be modeled using the Rothwarf and Taylor (RT) model [25], which is expressed through the differential equations:

\[
\frac{dN}{dt} = I_o + \beta N_{\omega} - RN \quad (2.2.2)
\]

\[
\frac{dN_{\omega}}{dt} = \frac{RN^2}{2} - \beta \frac{N_{\omega}}{2} - (N_{\omega} - N_{\omega R}) \tau^{-1}
\]

where \( N \) is the total number of quasiparticles, \( I_o \) is the number of quasiparticles created by photon absorption, \( R \) is the quasiparticle recombination coefficient, \( \beta \) is the probability of breaking Cooper pairs with \( N_{\omega} \) phonons (with energy greater than \( 2\Delta \)),

![Figure 2.2-1 Photoresponse of a superconductor. (a) A photon breaks a Cooper pair to create a highly excited quasiparticles, which interact with electrons and phonons. (b) Quasiparticles recombine to form Cooper pairs and emit phonons. (c) Phonons break additional Cooper pairs. (d) Phonons escape to substrate.](image-url)
\( \tau_T \) is the characteristic time for phonons to escape to the substrate, and \( N_{\omega T} \) is the thermal concentration of phonons. The RT model is accurate for medium to low perturbations at temperatures far below \( T_c \).

Near \( T_c \), another model commonly used to describe the photoresponse of superconductors is the 2-Temperature (2T) model, which can be shown to be equivalent to the \( RT \) model under weak perturbation. The 2T model describes the photoresponse in terms of the flow of energy between electron and phonon subsystems \([26]\). This model characterizes the relaxation time as the average cooling time given by \( \tau_{e-p} + (c_e/c_p) \tau_{e-s} \), where \( \tau_{e-p} \) is the electron-phonon interaction time and \( \tau_{e-s} \) is the phonon-substrate escape time, and \( c_e \) and \( c_p \) are the electron and phonon specific heat, respectively.

### 2.3 Hotspot model and phase slip centers

As described in the previous section, a photon absorbed by a superconductor generates a large number of quasiparticles. These quasiparticles create a “hotspot” in a superconducting thin film (a local volume of suppressed, or even destroyed superconductivity), which expands due to the diffusion of the quasiparticles after they are initially generated. This diffusion occurs according to the 2-D diffusion equation (for an ultrathin film):

\[
\frac{\partial C}{\partial t} = D \frac{1}{r} \left( \frac{\partial C}{\partial r} + r \frac{\partial^2 C}{\partial r^2} \right) + \frac{C - C_0}{\tau},
\]  

(2.3.1)
where $C(r,t)$ is the quasiparticle concentration, $r$ is the distance from the point where the photon absorption occurred, $1/\tau$ is the rate of quasiparticle decay, which is equivalent to the average cooling time, and $C_o$ is the quasiparticle equilibrium concentration. Figure 2.2-2 shows solutions of Equation (2.3.1) expressed as quasiparticle concentration along the stripe width for several different time periods after photon absorption. The time is given in units of thermalization time $\tau_{th}$, and the length is given in units of thermalization length, $L_{th}=(D \tau_{th})^{1/2}$, where $D$ is the diffusivity.

An SSPD is a nanostructured meander [Figure 2.2-3(a)], typically fabricated in a $10 \mu m \times 10 \mu m$ area with 4 nm-thick, 120 nm-wide NbN stripes and a fill factor of $\sim 60\%$. The fabrication procedure has been described in detail in [27]. In an ultrathin and very narrow superconducting stripe, hotspot formation works in conjunction with another phenomenon, namely, phase slip center (PSC) formation, to produce a measurable response. The device is biased initially below its critical current. Upon absorption of a photon, the hotspot is formed leading to a local resistive region in the stripe. This hotspot forms uniformly across the thickness of the film, $d$, since $d<< L_{th}$ [28,29] The volume of the hotspot has been shown to be proportional to the energy of the photon [30].

As the diffusion of quasiparticles occurs, the supercurrent is expelled from this region into the “sidewalks” of the stripe exceeding the sidewalk critical current, leading to the formation of PSCs. This phenomenon occurs when the zero voltage state can no longer be maintained because of a phase change of $2\pi$ in the order parameter. This causes the collapse of the energy gap and the formation of non-
superconducting areas in the sidewalks [31]. These regions, in addition to the resistive region due to the hotspot, make up a resistive stripe across the entire width of the film. Thus, a measurable voltage response can be registered after this resistance develops. Subsequently, the superconducting state is restored, and the detector is then ready to register the next photon absorption event. This phenomenon is depicted in 2.2-3(b). The relaxation mechanisms will be discussed thoroughly later in the chapter.

Figure 2.2-2 Hotspot profiles showing the concentration of quasiparticles at times 0.8, 2, and 5 in units of thermalization time measured in distance units of thermalization length [24]
Figure 2.2-3 (a) SSPD meander (b) Hotspot model for single-photon detection
PSC formation experiments have been performed by us to validate this model [32]. The amount of time between the appearance of a supercritical current (in our case, in the sidewalks of our stripe) and registering the voltage signal is referred to as the time delay $t_d$. This time delay is due to PSC formation. According to superconducting dynamics, $t_d$ corresponds to the time necessary for the superconducting energy gap to vanish. $t_d$ was first measured by Pals and Wolter in 1979 in aluminum films by applying a supercritical current pulse to a narrow superconducting stripe [33]. Geir and Schön developed a model that accounted for the relationship between $t_d$ versus the applied current to fit the Pals and Wolter data [34]. In our case, the approach used by Jelila et al. [35] is most relevant, since it applies to times shorter than or on the order of the gap relaxation time, $\tau_\Delta$ for $T << T_c$.

The SSPD photoresponse differs, of course, from the experiment of Pals and Wolter and Jelila et al., as we did not apply a supercritical current pulse to the sample. Instead, we applied a subcritical current from a constant voltage source, sent incident photons to form the hotspot, cause the supercurrent redistribution, and, eventually, created supercritical-like excitation in the sidewalks. As a result, the SSPD switching dynamics must depend not only the hotspot formation mechanism, but also on the PSC generation process.

For single-photon detection at 810-nm photon wavelength, $t_d$ was measured to be $65\pm5$ ps for an SSPD of 3.5-nm-thick and 130-nm-wide stripe in a 10 $\mu$m x 10 $\mu$m meander, which is in rough agreement with the model [32]. This time delay serves as direct evidence of PSC formation. In addition, $t_d$ adds to the response time of the detector and the variation in $t_d$ also contributes to timing jitter. Finally, PSC
formation also accounts for the appearance of dark counts, since these quantum fluctuations in resistance can cause unwanted voltage transients. These points will be explored further in later sections.

Another property of PSCs that will become relevant in later sections is the resistance that is generated as a result of their formation. Thus, the total resistance upon absorption of a photon contains contributions from both the original hotspot and normal regions formed as a result of PSC formation. The effect of the electric field penetration at the normal-superconducting (N-S) interface for a distance \( L_E = (D\tau_Q)^{1/2} \), where \( \tau_Q \) is the charge imbalance relaxation time, is considered as well as the effect of Andreev reflections. These reflections of holes occur at the N-S interface and result in the transfer of charge directly to Cooper pairs. Thus, the supercurrent can be directly converted into normal current, so no electric field is generated and no resistance is contributed. The parameter \( F(T) \) takes into account the contribution of Andreev reflections, so that the resistance contributed by PSCs and the hotspot in a SSPD with width \( w \) and thickness \( d \) upon absorption of a photon can be estimated as

\[
R = \frac{2\rho F(T)L_E \left(1 + F(T)\frac{2L_E}{R_m}\right)}{dw \left(1 + F(T)\frac{2L_E}{R_m}\right) - R_m}, \tag{2.3.2}
\]
where \(R_m\) is the hotspot diameter and \(\rho\) is the resistivity of the normal film [28].

Neglecting the effects of heating from the bias current, we express the current density \(j\) in the sidewalks in terms of this resistance,

\[
j = \frac{RI}{4\rho F(T)L_E}.
\]  

(2.3.3)

### 2.4 Temperature effects

We will later attempt to explain our SSPD measurement results at ultra-low temperature in terms of the hotspot photoresponse model. Here, we review the temperature dependent parameters that may affect SSPD operation. As described earlier, the quasiparticle concentration, \(C\), after the absorption of a photon follows the solution of the two-dimensional diffusion equation [Equation (2.3.1)]:

\[
C(r,t) = \frac{M(t)}{4\pi D \tau} \frac{\exp(-r^2)}{4Dt} \exp\left(-\frac{t}{\tau}\right) + C_o,
\]  

(2.4.1)

where \(M(t)\) is the time dependent multiplication factor, \(D\) is the diffusivity, \(d\) is the film thickness, and \(\tau\) is the average cooling time.
The temperature dependence of $C$ lies in the temperature dependence of $M(t)$ and $\tau$, since they depend on $c_e \propto T$ and $c_p \propto T^3$ and the energy gap $\Delta = 2.15k_B T_c (1 - T/T_c)^2$. Figure 2.4-1 shows the spatial distributions of quasiparticles for a range of temperatures. As the temperature decreases, the hotspot will decrease in area. We also characterize the time evolution of the hotspot using the expression for the hotspot diameter $R_m$ derived from the diffusion model [36]

$$R_m = \sqrt[3]{\frac{\Delta}{\Delta + \Delta^2}} \left( \frac{1}{2D\tau_\eta dN_\eta \Delta} \right) \ln \left( 1 + \frac{\pi^2 \zeta}{\Delta} \frac{h\nu}{D\tau_\eta dN_\eta \Delta} \right),$$

(2.4.2)

Figure 2.4-1 Normalized concentration of nonequilibrium quasiparticles versus distance in a 4-nm-thick NbN film at several values of $T/T_c$ with $\lambda = 1.55 \, \mu m$ and QE = 1.
where $\varsigma$ is a parameter that takes into account losses during the avalanche process and $\nu$ is the frequency of the incident photon. An exponential increase and decay of the hotspot were assumed in the model. The resulting hotspot evolution is depicted in Figure 2.4-2 for several values of temperature assuming that for our NbN stripes, $T_c \approx 10$ K. As the temperature decreases, the hotspot growth and its relaxation time becomes longer. The hotspot growth time is also shown in Figure 2.4-3 and we note that it varies by $<10$ ps in the range shown. Figure 2.4-4 shows that the maximum hotspot size decreases with decreasing temperature down to only $\sim 0.6I_c$ according to

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure2.4-2}
\caption{Time evolution of the normalized hotspot diameter at several values of $T$, where $T_c = 10$ K.}
\end{figure}
the diffusion model. Below this value, the maximum hotspot size does not decrease further. The latter comes from the $\Delta(T)$ dependence, which for BCS superconductors saturates for $T < 0.5T_c$. The values for hotspot diameter are lower than those reported experimentally [37], because self-heating effects were neglected. Thus, the diffusion of quasiparticles does not change appreciably in the temperature range where we normally operate the SSPD ($T < 0.5T_c$).

Figure 2.4-3 The growth in the diameter versus temperature after photon absorption according to the hotspot model [36]
The main phenomenon responsible for the superconducting recovery is not the shrinking of the hotspot as quasiparticles escape into the substrate, but is mainly associated with the recovery of superconductivity from self-heating. Evidence of this effect is presented by Hadfield et al. [38], which shows that a SSPD does not recover its superconductivity until the current drops well below the critical current ($I_r < \ll I_c$), which will also be evident in our results presented in later sections. Skocpol et al. [39] showed that for superconducting microbridges, as the temperature decreases, $I_r$ is reduced further leading to a slower relaxation. If the relaxation were dominated by phonon escape into the substrate, once the hotspot diameter decreased to just below the critical current, superconductivity would be immediately restored. Since this is not the case, the effect of lowering the temperature and in so doing, slowing down the
hotspot dynamics will most likely not lead to a degradation in detector response time. Therefore, the self-heating effects will be included in our photoresponse model in our discussion of detector response time.

The self-heating effects can also be important in finding the effective hotspot size. After quasiparticles diffuse outward, the bias current heats the outer hotspot regions, thus aiding the expansion of the hotspot region. We account for the temperature dependence of this expansion by considering the propagation of the N-S boundary defined by the intersection of the normal hotspot and superconducting surrounding. After hotspot formation, the current density is increased in the sidewalks as quasiparticles diffuse outward. In the regions of current flow farthest from the edges of the stripe, however, the current density is unchanged from the original biasing current close to $I_c$. The propagation velocity of the N-S boundary will depend on the current density in a linear manner according to the models of Whetstone and Roos [40] and Cherry and Gittleman [41], assuming the local $T_c$ is unaffected by the change in current density. This faster expansion towards the stripe edges will, in turn, result in the formation of an elliptical hotspot as opposed to a circular one predicted using the diffusion model alone. The complete picture of hotspot formation is shown schematically in Figure 2.4-5.

In addition, experimental results presented in Ref. 40 show that as the operating temperature decreases, the ratio between the propagation velocity and current density also decreases, but saturates at low temperatures. Thus, the ultra-low temperature behavior of the detector will not be affected directly by a change in propagation velocity.
One factor omitted from consideration thus far is the temperature dependence of the critical current \( I_c \propto [1 - (T/T_c)^2]^{3/2} \), according to BCS theory [42]. \( I_c \) thus changes appreciably until it reaches its saturation value at low temperatures. This consideration is unique to the detector operation, since maintaining a constant \( I/I_c \) across all operating temperatures is important for the purpose of analyzing its performance. This factor will later be used in explaining our results to relate the propagation velocity to the operating temperature of the SSPD.

The effective hotspot size defined by the combination of quasiparticles emanating from the point of photon absorption and phonons at the N-S interface could affect the probability of PSC formation, thus affecting the device QE. We will show that the QE does indeed increase with decreasing temperature and this improvement saturates below 2 K, which we will relate to the temperature dependence of \( I_c \).
2.5 Inductance

Inductance is an inherent limitation in superconducting electronics and has been analyzed rigorously for superconducting circuits to determine their speed limitations [43,44]. Likewise, the high inductance of our superconducting meander can affect its speed as well. Therefore, we must consider it along with the hotspot dynamics to determine the detector’s limitations with respect to counting rate.

In general, the total inductance of a microstrip contains two contributions— the magnetic (or geometrical) inductance and kinetic inductance. The kinetic inductance stems from the movement of charge carriers with some kinetic energy, while the magnetic inductance is associated with the energy stored in the magnetic field. The kinetic inductance is important for nanostructures and can dominate the magnetic inductance for a superconducting stripe with a small cross-sectional area. If the current distribution is uniform in this thin wire, the total inductance is given by the following simple expression for kinetic inductance alone:

$$L_k = \mu_0 \lambda_L^2(t) \frac{l}{wd},$$

(2.5.1)

where \(l\) is the length, \(\lambda_L\) is the London penetration depth, \(t\) is the reduced temperature, \(d\) is the thickness, and \(w\) is the width. Shadowitz derived rigorously the expression for the inductance of whiskers, \(w < \lambda_L\) and \(d < \lambda_L\) and arrived at the same expression for the inductance [45]. In contrast, Baker et al. showed experimental evidence that for thin films with \(w < \lambda_L\), the current may not be completely uniform [46].
Recently, Hadfield et al. have presented evidence of a phase locking phenomenon in a meander-type SSPD by illuminating it with rf radiation and observing that when the SSPD is driven to an intermediate state, relaxation oscillations occur, which are “locked” by the rf source. The voltage pulses generated were fit using a model that took into account the inductance of the SSPD [38]. Their results show a high inductance (∼500 nH) in a typical 10 µm x 10 µm meander device. Furthermore, they used Equation (2.5.1) to calculate the London penetration depth of the SSPD as $\lambda_L = 560$ nm, which as they pointed out, is larger than the bulk NbN value of 200 nm. The inductance was verified by Kerman et al., while at the same time finding that the response time of the SSPD is limited by this inductance [47]. This will be discussed in greater detail in the SSPD optimization section. We will use a simulation program to calculate the inductance of various meander geometries.

This simulation program used, FastHenry, version 3.0 [48], was originally formulated for calculating the inductance of complex conductor geometries [49] and later adapted to account for superconductors as well by including the contribution of the kinetic inductance along with the usual magnetic inductance. This approach uses the two-fluid model. It divides the superconducting geometry into filaments, each with a uniform current distribution and calculates its impedance using the penetration depth input by the user. The inductance is then easily extracted.

We can resolve the correct value for the penetration depth of our SSPD by first calculating the inductances for SSPD geometries used by Kerman et al. using a range of $\lambda_L$ values. We then compare these values to the measured inductance values.
for the real devices used by Kerman. The details of these calculations will be described in later sections. The result of the calculation using $\lambda_L \approx 560$ nm gives an inductance of 44.5 nH, which is equal to the measured value for the 3 $\mu$m x 3.3 $\mu$m meander, with 100 nm–wide, 4 nm–thick stripes, and a 50% fill factor device measured by Kerman et al.. Note that using the same value of $\lambda_L$ in Equation (2.5.1) gives 50.5 nH, which is quite close to the result of the simulation. However, since we will explore a wide range of SSPD geometries, we will continue to employ the simulation program instead of only using Equation (2.5.1) to test the equation’s applicability to various geometries. We will use these calculated inductance values to predict the detector response time of various SSPD geometries.

2.6 SSPD performance

2.6.1 Single-photon detection

In order to show that the SSPD is capable of counting single photons, we have performed statistical studies using the fact that photons arrive from a pulsed laser source, which has a Poisson distribution for the number of photons emitted per pulse, $n$ for a mean photon number per pulse $\mu$.

$$p(n) = \frac{e^{-\mu} \mu^n}{n!}$$

(2.6.1.1)

If the flux is very weak $\mu \ll 1$, the probability of detecting one photon ($n = 1$) is reduced to $p(1) \approx \mu$. Thus, single-photon detection can be shown if the linear
relationship between the detection probability and the average number of photons incident on the detector holds. When the condition $\mu \ll 1$ does not hold, the SSPD still detects in a linear manner, since according to the hotspot model, this will lead to multiple hotspots that will register as one detection event. Thus, using a first order approximation, the detection probability is $1 - p(0) \approx \mu$, which shows that when this relationship holds, the detector is operating in the single-photon regime. For instance, if the SSPD could only detect two or more photons, the probability of detection would be $1 - p(1) - p(0) \approx \mu^2$. Thus, in results presented in subsequent sections, we will verify single-photon detection by observing the proportional relationship between $p(n)$ and $\mu$.

2.6.2 Quantum efficiency (QE) and spectral sensitivity

As mentioned earlier, the hotspot volume is proportional to the photon energy. For an ultrathin stripe of constant thickness (e.g., 4 nm), the spatial distribution of PSC formation is assumed to be uniform in agreement with Ref. 31 and the hotspot area scales with the photon energy. The relationship between QE and wavelength is shown along with the estimated hotspot size for different wavelengths and photon energies in Figure 2.6.2-1. If we assume the QE scales identically as the diameter of
the hotspot does with photon energy, we could expect \( \text{QE} \sim 1/\lambda^{5/2} \). For wavelengths in the visible to NIR, a similar dependence is shown in Figure 2.6.2-1 for a 3.5 nm-thick SSPD [50,30]. We also expect QE dependence to saturate for high photon energies as the total hotspot area covers the width of the stripe, which is also shown in the figure.

Since the QE for the SSPD measured in [30] was measured to be 20% throughout the visible range and the NbN absorption coefficient was calculated to be \( \sim 26\% \), we do not expect any more improvement. Hence, any further expansion of the hotspot will not affect the QE beyond the visible wavelength range. The broad spectral sensitivity is relevant for our purposes, since it demonstrates that our SSPDs are capable of single-photon detection in the near infrared (NIR telecommunications).

![Figure 2.6.2-1 Spectral dependence of QE at \( T = 4.2 \, \text{K} \). The inset shows the estimated hotspot size at varying photon energies.](image-url)
and middle infrared (MIR) radiation.

Our most recent data also show that the performance, in terms of QE and dark counts, improves with decreasing temperature [51]. In the experiment, 1.26 µm photons illuminated an SSPD at 4.2 K and then at 2.3 K [30]. The values of QE were compared at a range of bias currents and a substantial increase in QE was reported across the entire range of bias. In addition, the dark counts were measured at 4.2 K and 2.3 K, and they were shown to be about two orders of magnitude less at 2.3 K than at 4.2 K as shown in Figure 2.6.2-2.

The results presented in the following sections will show the performance of the SSPD at telecommunications wavelengths. Moreover, we will present QE data on a SSPD operating well below 2.3 K and will demonstrate that the QE improvement saturates at ∼2 K. However, the dark counts continue to decrease substantially with decreasing temperature below 2 K.

Figure 2.6.2-2 QE and dark counts as a function of normalized bias. Measurements performed at 4.2 K (closed symbols) and 2.3 K (open symbols). Solid lines represent an exponential fit. [26]
Chapter 3 SSPD performance at telecommunications wavelengths: free-space system results

3.1 Free space setup

In order to test the counting rate and response time of the SSPD, we used a high speed laser, the Opticlock fiber mode-locked laser by Pritel [52]. This laser operates at 1550 nm wavelength and produces 1.6 ps-wide pulses with a variable repetition rate up to 10 GHz. Alternatively, the photon energy could be doubled externally.

An older 4 \( \mu \)m x 4 \( \mu \)m-area, 3.5 nm–thick, 200 nm–stripe width, SSPD was used in this measurement as opposed to the current 10 \( \mu \)m x 10 \( \mu \)m-area standard. The SSPD was placed in a helium optical dewar at 4.2 K and biased with a voltage source through a bias-tee, Picosecond Pulse Labs 5541A-104 (Figure 3.1-1). The output signal connected to the other terminal of the bias tee was fed through two identical room temperature amplifiers, Miteq JS3-00101800-24, each with ~20 dB gain, and 0.1-18 GHz bandwidth. The amplified voltage signal was connected to a 5 GHz Lecroy single-shot oscilloscope or to a 50-GHz Tektronix sampling oscilloscope.
3.2 Timing jitter

The timing jitter measurements were performed using the 50-GHz Tektronix sampling oscilloscope with both the NIR and energy-doubled photons from the Pritel laser. A standard histogram technique was used to measure the jitter of the SSPD’s response to the 1-GHz repetition rate pulses for the 778 nm and 1550 nm wavelengths. Figures 3.2-1(a) and (b) show histograms for the jitter. The profile is nearly Gaussian with a FWHM of ~18 ps for both wavelengths. Since the intrinsic jitter of the Pritel laser is less than 70 fs for 1-GHz pulses, we can, therefore, conclude that the 18 ps jitter reflects the jitter in the detector and electronics. This is the lowest value of timing jitter reported for any optical photon counter.
The timing jitter depends critically on the PSC formation process. As mentioned earlier, the variation in time delay, $t_d$, produces timing jitter. The stripe width variations directly affect the measured meander $I_c$, setting the value of the bias current. This is depicted schematically in Figure 3.3-2. Therefore, $t_d$ is going to vary depending on the location of the hotspot along the length of the detector. In the PSC studies mentioned earlier, an effective width of 80 nm was used in calculating $t_d$, which was really the minimum width of the superconducting channel, since its

Figure 3.2-1(a) Timing jitter results (b) Histogram technique
irregularities on each side were as high as 25 nm, according to AFM data [32]. The minimum width of 80 nm set the SSPD $I_c$ value, but the current bias of $0.85I_c$, in fact, corresponded only to the narrowest meander regions. The sections with the actual nominal width of 130 nm did not detect photons, since in their case, for the $0.85I_c$ bias, the current density in the sidewalks did not exceed the critical value. For instance, the minimum width to achieve 1.22 normalized bias is 86 nm for a 30-nm-diameter hotspot. The difference in $t_d$ between the 80-nm and 86-nm stripe widths contributes about 70 ps of jitter, according to Jelila’s formulation, which is much larger than the jitter value. This formulation assumes a constant hotspot size, which is not indeed the case due to the differing N-S propagation velocities caused by the variation in current density in the sidewalks.

Figure 3.2-2 Hotspot location in non-uniform stripe causing current density variation in sidewalks
The SSPD timing jitter seems to be also limited by the hotspot size variations when the photon absorption occurs not at the very center of the stripe. If the hotspot is absorbed close to the edge of the stripe, the symmetry of the hotspot is not sustained and the hotspot is likely to occupy a smaller region across the stripe compared to a hotspot absorbed at its center.

Figures 3.2-3(a) and (b) qualitatively illustrate the quasiparticle concentration distributions and hotspot profiles when the photon is absorbed at the edge of the stripe and near the center, respectively. We note that the “edge” hotspot results in higher maximum quasiparticle concentration, as compared to the hotspot absorption at the center. Figure 3.2-3(c) demonstrates that the higher concentration actually leads to the lower resistance and to the sharp increase of the equivalent $t_d$. Differences in $t_d$ values between the different photon absorption acts directly affect the SSPD jitter [53].

**Figure 3.2-3** (a) The hotspot is confined at the edge of the stripe with a higher maximum qp concentration $N_a$. (b) The hotspot is formed at the center, resulting in a lower quasiparticle concentration $N_b$. (c) Time profiles of $N_a$ and $N_b$ onsets of the resistive state $R_a$ and $R_b$, and their resulting time delays $t_{d,a}$ and $t_{d,b}$, for the case (a) and (b), respectively.
3.3 Counting rate

For the first measurements, the Pritel laser was attenuated to a flux of 100 to 1000 photons per pulse in order to allow the SSPD to count essentially all pulses in the incident optical train. Otherwise, at a very low level of flux, e.g. ~1 photon per pulse, the limited QE of the detector would prevent us from registering consecutive pulses. Figure 3.3-1(a) shows that the SSPD can detect photons at 1-GHz and 2-GHz repetition rates as, in both traces, consecutive pulses are resolved clearly. Notice that although the pulses are resolved, their amplitudes decrease considerably as the photon counting repetition rate $f$ increases as is shown in Figure 3.3-1(b), where we present our experimental data and the pulse amplitude $V$ versus $f$ fit:

$$V = \frac{V_0}{\sqrt{1 + (2\pi\tau f)^2}} \quad (3.3.1)$$

$V_0$ is the pulse amplitude at a very low counting rate and $\tau$ is the characteristic time. According to the fit, $\tau = 134$ ps, which corresponds to the response limit of our system. The 3 dB drop off is shown at about 2 GHz in Figure 3.3-1(b). Additionally, in Figure 3.3-1(c), we present a response pulse taken at a very low flux (~1 photon per pulse) excited by the 2-GHz laser. In contrast to when consecutive narrow pulses are registered, we observe a much wider single pulse.

The intrinsic counting rate of the SSPD can be estimated using previous results from electro-optic sampling experiments [54] of NbN films and our results
from PSC formation experiments. Electro-optic sampling experiments previously done show that the electron-phonon interaction time is $\tau_{e-p} \approx 11.6$ ps and the time of phonon escape into the substrate is $\tau_{e-s} = 21$ ps, which yield an average cooling time of 36 ps. The total response time combines this average cooling time with the time $t_d$ for PSC formation as follows: $(36^2 + 70^2)^{1/2} \approx 78$ ps. Thus, the corresponding maximum counting rate is estimated as 13 GHz. This result is, of course, the intrinsic upper bound of the SSPD performance. The counting speed of a real device depends on the particular device geometry, and the inductance is going to play a significant role in limiting its counting rate. We will later show that the geometry of the SSPD measured here provides an inductance that roughly corresponds to the measured bandwidth of 2 GHz, instead of the 13 GHz projected by the hotspot model alone. The bandwidth of our electronics may also limit the measured counting rate of the detector.
Figure 3.3-1 (a) SSPD response to 1-GHz and 2-GHz pulses at higher flux. (b) Bandwidth of SSPD system. (c) SSPD response pulse excited with a laser with a repetition rate of 2-GHz at low flux.
The DARPA network is a fiber-based setup, operated at 1.55 µm telecom wavelength. In order to implement SSPDs in this network, we must integrate them through a fiber-based system. We, therefore, developed two fiber-based systems, one with collaborators at the Institute of Electron Technology in Warsaw, Poland, and the other with researchers at NIST in Boulder, Colorado. The NIST system is based on a cryogen-free refrigerator capable of reaching mK temperatures. The detectors are placed in a removable housing and coupled via single-mode fiber. In the Warsaw design, the fiber is either single-mode or multimode and is permanently glued to the device. The devices (pairs of two) are placed inside a helium transport dewar \((T = 4.2 \text{ K})\), with both their fiber inputs at room temperature. Thus, the low temperature operation of the SSPD is completely hidden from the user.

### 4.1 NIST setup

In order to measure the QE and dark counts at ultra-low temperatures, we employed, together with scientists from NIST, Boulder, CO, the use of a cryogen-free adiabatic demagnetization refrigerator (ADR) [55]. This allowed us to measure the SSPD properties at temperatures as low as 60 mK and switch easily between any temperature of interest. In this system, photons were sent through a single-mode fiber, which was coupled to the SSPD inside the ADR. In order to electrically
connect the SSPD and couple light from the fiber to it, modifications of existing NIST holders (Figure 4.1-1(a)) were made so that the SSPD could be connected to a microstrip by wire bonds as shown in Figure 4.1-1(b). The single-mode fiber was aligned using a microscope equipped with an infrared camera.

The sources used were two laser diodes with wavelengths of 1310 nm and 1550 nm, which were driven by a signal generator that produced pulses of \( \sim 1 \) ns width. A bias source similar to the one used in the free-space setup was used, but a 50 \( \Omega \) room-temperature resistor was placed in parallel with the device to prevent it from dissipating too much heat and raising the temperature of the ADR. In addition, similar broadband amplifiers were used to generate a signal observable using our 8-GHz oscilloscope. A digital counter, Agilent 34401A, was used to count the pulses produced from the photons after amplification with either a Miteq JS3-00101800-24-5A or JS2-0100200-10-10A amplifier and an Agilent high bandwidth amplifier.
Figure 4.1-1(a) Detector holder for fiber coupling in ADR.
(b) Wire bonds to SSPD
The SSPDs measured are listed in Table 4.1-1. All devices are 4 nm-thick, 100 nm-wide, 10 μm x 10 μm, NbN meander SSPDs. Since 588/1 84 has the highest QE, this SSPD was measured most completely; thus, the data for this SSPD is presented.
The IV curve at 4 K for this device is shown in Figure 4.1-3. As evident from the curve, $I_c$ measured at NIST was 14 $\mu$A. The measured resistance across the device and cabling in the superconducting state was 5.7 $\Omega$ and is subtracted in the IV curve shown. This curve shows that as the current increases above $I_c$, the entire meander transitions into the normal state, which we note is not the case when a single photon is absorbed.
4.1.1 Quantum efficiency and temperature

Figure 4.1.1-1 shows the single-photon power dependence for temperatures of 4 K, 2 K, 1 K, and 60 mK for several values of bias current at $\lambda = 1550$ nm. As shown, all curves follow the single-photon (linear) dependence in a very wide range of photon fluxes. At high fluxes, the detector saturates, since all photons are detected, while at the lowest fluxes, the dark counts tend to dominate. The decrease in dark counts with decreasing temperatures is quite apparent. The dark counts even disappear altogether at 60 mK across all measured bias values. The same behavior is shown for a wavelength of 1310 nm in Figure 4.1.1-2 for 2 K, 1 K, and 200 mK. Again, strong evidence of single-photon detection is presented along with decreasing dark counts with decreasing temperature.
Figure 4.1.1-1 Single-photon power dependence for $T = 4\, K$, 2 K, 1 K, 60 mK at $\lambda = 1550\, \text{nm}$
Figure 4.1.1-2 Single-photon power dependence for $T = 2\, \text{K}$, $1\, \text{K}$, and $100\, \text{mK}$ at $\lambda = 1310\, \text{nm}$. 
We can easily calculate the QE from the data shown in Figures 4.1.1-1 and 4.1.1-2. For this calculation, we typically take the detection probability at a flux of ~1 photon per pulse. This was not done, however, in the instances where dark counts dominate at these low fluxes, so we instead use the detection probabilities at higher fluxes where the single-photon regime is shown clearly. In Figure 4.1.1-3(a), the QE is shown for all values of bias where a clear single-photon detection regime is evident for $\lambda = 1550$ nm and $\lambda = 1310$ nm. In Figure 4.1.1-3(b), the highest value of QE is taken at all measured temperatures for $\lambda = 1550$ nm and $\lambda = 1310$ nm. From this figure, we can see that at both wavelengths, the largest improvement in QE occurs when the operating temperature is decreased from 4 K to 2 K. For the 1550 nm data, below 2 K, the QE remains almost constant, but for the $\lambda = 1310$ nm data, some improvement down to 200 mK is still noticeable. In addition, all QE values for $\lambda = 1310$ nm are larger than those taken at $\lambda = 1550$ nm as expected according to the hotspot model and in rough agreement with previous results (see Figure 2.4) showing a factor of four increase in QE as the wavelength is shifted from 1550 nm to 1310 nm [30].
Figure 4.1.1-3(a) QE versus current for all measured temperatures and wavelengths. (b) Highest achievable QE as a function of temperature.
The actual values of QE measured at NIST did not agree with the Moscow values (see Table 4.1-1). In Table 4.1-1, we report that, at \( \lambda = 1550 \) nm wavelength, the QE of the device 588/1 84 was measured as 4% at 4 K and 10% at 2 K, while from Figure 4.1.1-3, we see that the highest QE measured was 1.1% at 2 K and 0.43% at 4 K. The reason is that in our fiber-based experiments, some photons did not reach the detector due to imperfect fiber coupling. We did not account for this loss in our data; the QE we measured at NIST was a system QE (SQE) rather than strictly the device QE (DQE) measured at Moscow. We note however, that the temperature related QE improvement, the ratio of QE at 2 K to that at 4.2 K is almost identical in both measurements, namely, in our case it was 2.56 while for the Moscow data it was 2.5. The NIST measurements confirm the improvement in QE with decreasing temperature down to 2 K as was previously measured in Moscow while also showing that this increase saturates at lower temperatures.

![Figure 4.1.1-4 Temperature dependence \( I_c/I_o \) where \( I_o \) is taken at the critical current at \( T = 60 \) mK](image.png)
This behavior can be understood in terms of the $I_c$ dependence with temperature. In Chapter 3, we assumed that the QE scales linearly with hotspot diameter. Likewise, we assume here that QE also scales linearly with the propagation velocity of the N-S boundary, and as stated earlier, the propagation velocity increases linearly with increasing current bias. We note that in operating the SSPD, we choose a constant reduced bias $I/I_c$. Using the dependence of $I_c$ on temperature (see Section 2.4), we can show that the QE will follow this temperature dependence. Figure 4.1.1-4 shows the $I_c$ dependence on temperature for two SSPDs measured in the NIST setup, which show that $I_c$ only saturates at $T < 2 \text{K}$. Thus, the increase in QE at low temperatures and its eventual saturation can be satisfactorily explained in terms of the temperature dependence of $I_c$.

### 4.1.2 Dark count results

As shown previously in Figure 2.6.2-2, as well as in Figures 4.1.1-1 and 4.1.1-2, the dark counts drastically decrease with temperature. In order to explore this relationship more completely, we measured dark counts over a larger temperature range extending well below 2 K. Figure 4.1.2-1 shows dark-count rates versus bias over the broad temperature range of 0.1 K to 6.4 K. The dark count decrease is verified further. In fact, at 0.1 K, the dark counts are $< 0.1 \text{ counts/s}$ at $0.95I_c$, while at 4.6 K, $\sim 10^5 \text{ counts/s}$ were measured at this bias. As expected, according to our previous data, the dark count dependence the normalized bias current is exponential.
As was mentioned, PSC formation is responsible for the appearance of quantum fluctuations, which produce dark counts. In a separate work on SSPD dark counts [56], the results were presented in terms of the Langer-Ambegaokar-McCumber-Halperin (LAMH) theory for PSC formation in 1-D structures [57]. Using the model and our experimental dependence, $R_{dk}$ is expressed as $R_{dk} = R_0 \exp[I/I_o]$, where $R_0$ is termed the prefactor and $I_o$ is the characteristic current. The model predicts that $I_o$ should have a proportional dependence on $T$, which is verified in the data for $T$ approaching $T_c$. In addition, $R_0$ shows some interesting properties that are indicative of a transition from the quantum phase slip center regime that is normally applicable for $T<<T_c$ to the thermal phase slip center regime, which is applicable for $T \equiv T_c$.

![Figure 4.1.2-1 Dark counts versus normalized bias current from $T = 0.1$ K to 6.4 K](image-url)
The importance of the dark count rate for detectors in the DARPA network for the purpose of secure quantum communication will be analyzed in detail in Chapter 6.

### 4.2 Warsaw setup

The second type of SSPD fiber-coupling system has been developed jointly with scientists in Warsaw and tested in our laboratory. The system consists of a double detector SSPD setup with permanently attached fibers and output wires that can be easily integrated into the DARPA network. The design is shown in Figures 4.2-1(a) and (b), which shows a dip stick that is constructed to fit in a standard helium dewar. The two detectors are placed at the bottom flange. The photons are delivered to each detector via single-mode or multimode fiber, and the electrical signals are delivered via SMA cables to the amplifiers, which are placed outside the transport dewar. Figure 4.2-1(c) shows a detailed view of the fiber coupling arrangement. Micromechanical photoresist rings are positioned over the SSPD and were fabricated using a photolithography process guided by alignment marks made on the structure [58]. The accuracy of the ring location is about 1 \( \mu m \). These rings were then used to place the fiber precisely over the detector. This type of practical QKD receiver typically operated continuously for over two months in our laboratory.
Figure 4.2-1 (a) SSPD receiver schematic (b) Details of apparatus (b) Details of fiber-coupling arrangement
Ten efficient receivers (five pairs), eight of which were coupled via single-mode fibers and two utilizing multimode fibers were tested. As in the NIST fiber coupling setup, the main figure of merit was the SQE for each receiver. We also measured dark counts and performed time resolution measurements. We used highly attenuated, 41-ps-wide pulses from a semiconductor laser diode operating at a 1540-nm wavelength with a repetition rate of 1 MHz for our SQE measurements. The optical power delivered to the detectors was controlled with a digital optical attenuator. The output signals from the amplifiers were connected to either a 50 GHz sampling oscilloscope for time-resolved studies or to a signal counter to perform statistical analysis of photon counts or to measure dark counts.

### 4.2.1 QE results

The typical dependence of the detection probability versus incident power expressed as the number of photons per optical pulse incident on the NbN detector is presented in Figure 4.2.1-1. As in previous results, we observe saturation of detection probability at high photon flux levels and a linear dependence at low photon fluxes consistent with the single-photon dependence. The dark count measurements with the fiber input blocked (0 photons per pulse) resulted in 2 (device #6) to 90 (device #3) counts per second for a bias current $I_b \approx 0.95I_c$. 
Taking the value measured for $I_b \approx 0.95 I_c$, we determined the SQE of our receivers and listed the values for the best and weakest devices in Table 4.2.1-1. We also present in this table the corresponding DQE values intrinsic to the detector structure and measured by our Moscow collaborators directly after SSPD fabrication in a free-space setup. Finally, Table 4.2.1-1 contains the calculation of the efficiency of fiber coupling $K$, which is just the ratio SQE/DQE. We note that device #11 exhibits the best SQE = 0.331 and $K = 0.33$. This value of $K$ is higher than the value of $K = 0.10$ obtained in the NIST setup.

![Figure 4.2.1-1 Single-photon detection dependence for fiber-coupled detectors in Warsaw setup.](image-url)
The following analysis of the fiber-coupling performance was done more thoroughly for this Warsaw setup than for the NIST setup because of the large number of devices. Also, determining the factors that affect the quality of the fiber coupling arrangement will lead to further improvement of this technology. We analyze the fiber-coupling performance by calculating $K$, assuming a Gaussian mode-light profile emitted from the fiber. We obtain the fraction $K$ of power incident on the square ($10 \times 10 \, \mu m^2$) detector with side length $a$

$$K = \frac{1}{\pi w_0^2} \left[ \frac{2}{1 + \left( \frac{\lambda z}{\pi w_0^2} \right)^2} \right]^{a/2} \int_{-a/2}^{a/2} \int_{-a/2}^{a/2} \exp \left\{ -2 \left( x^2 + y^2 \right) \right\} \, dx \, dy, (4.2.1.1)$$

where $w_0$ is the beam radius and $\lambda$ is the optical wavelength.

The results from these calculations, based on Equation 4.2.1.1 for the detector–fiber misalignment, are presented in Figure 4.2.1-2 with $K$ values for the
selected receivers listed in Table 4.2.1-1. The detectors are labeled according to their most likely fiber-detector vertical distance and the horizontal misalignment that corresponds to the measured $K$ values. The figure shows constant $K$ values for an increasing vertical fiber–SSPD distance ($y$ axis). The stronger dependence of horizontal (radial) fiber misalignment ($x$ axis) than vertical fiber–detector separation on $K$ is clearly apparent. However, the detector misalignment giving the measured $K$ values seem unrealistically large since the fiber was precisely positioned with the photoresist ring, so we conclude that vertical displacement and horizontal misalignment are not the only factors affecting $K$.

The resulting low $K$ observed is, thus, most likely due to the distortion (tilt) of the fiber end face with respect to the SSPD plane that was caused by its movement during the cooldown process. Therefore, the 30-$\mu$m-thick positioning ring may not be sufficiently resistant to lateral distortions of the fiber, resulting in lower than expected $K$ values observed in most of the tested devices.
In an attempt to improve the fiber alignment and reduce lateral distortions, we implemented a multimode fiber in our detectors. Because of the large ~50-μm-diameter core of the multimode fiber, the coupling should be significantly less sensitive to the horizontal displacement. The drawback in the case of the multimode fiber is that only a small portion of radiation reaches the detector because of the disparity in the surface areas of the SSPD surface area (100 μm²) and the fiber core (~2000 μm²). In fact, for a Gaussian mode profile of a graded-index fiber, only 1/10 of the power from the fiber reaches the SSPD, even in the case of perfect horizontal alignment. For this reason, we were unable to surpass the SQE values obtained for single-mode receivers with multimode fiber coupling as reflected in the $K$ value listed in Table 4.2.1-1.

The table also shows that some of our detectors exhibited very low SQE and $K$ values. In these cases, the most probable reason was either a crack of the fiber somewhere along its length or a permanent deformation of the fiber’s ending position with respect to the NbN meander. Evidence of this is that the very low $K$ remained constant after a few thermal cycles, and cracks in the fibers were found in mechanical tests of the two weakest devices.
4.2.2 Timing jitter

The results for timing jitter measurements for fiber-coupled dip-stick devices and an older device measured in the free-space setup are shown in Figure 4.2.2-1. The histograms, compiled using the same technique as used in the free-space measurements, show Gaussian profiles with FWHM of 19 ps, 37 ps, and 58 ps, for the free-space device, device #11, and device #3, respectively. The disparity in values for the fiber-coupled devices can be explained by the difference in fiber properties. Device #11 was coupled via a single-mode fiber while device #3’s coupling was via a multimode fiber, leading to a larger jitter caused by modal dispersion. Using values from typical telecom fibers, ~2 m of fiber can account for ≥24 ps of dispersion [59], which is consistent with our results. The very low value of the free-space jitter is most likely because of the use of very short electrical cables and low dispersion electronics.

Figure 4.2.2-1 Timing jitter histograms for selected fiber-coupled devices in Warsaw setup. A 4 x 4 μm² SSPD jitter histogram measured in free space.
In order to show the operation of the two-detector Warsaw configuration in a way which more closely reflects the way the system would be operated in the DARPA network, we measured correlated counts between two of our detectors. 41-ps pulses with 1-MHz laser at ω = 1540 nm were sent to a 50/50 beamsplitter, then through two fibers of unequal length, and finally detected using two low QE devices (SQE = 0.037% and 0.0018%). The output signals were fed to the time-resolved pulse counter with the output from the one device acting as the trigger (start) pulse after which the output from the other (stop) device was counted for the time window specified. The time histogram for the correlated counts for the devices is shown in Figure 4.2.2-2 for a photon flux of ~60,000 photons/pulse. The peaks are clearly resolved every 1 μs corresponding to the period of the laser. The timing resolution corresponding to the width of the peaks is ~1 ns, but this value may not reflect the true timing resolution of the system because of the low 250-ps resolution of our time-resolved pulse counter (Fastcomtec P7887). Thus, these measurements were repeated by our collaborators in the Delft University of Technology, Delft, The Netherlands, using a similar setup and a time-resolved pulse counter with a <50 ps time resolution (Picoquant Time Harp).
In the Delft setup, a laser with 500-fs pulses at \( \lambda = 940 \) and a repetition rate of 82 MHz was used. Figure 4.2.2-3 shows correlation peaks between device #11 and device #12 with FWHM \( \approx 390 \) ps. The start and stop pulses were switched between the devices and a similar correlation FWHM \( \approx 420 \) was recorded. Note the decreased number of counts for the latter arrangement, since the lower QE device is used as the stop pulse. These values represent the timing resolution of the complete two-detector system. The widths of the correlation traces shown in Figure 4.2.2-3 are relatively narrow compared to standard APD quantum correlation systems, but wider compared to a combination APD-SSPD system presented in Ref. 60.

Accounting for the large difference in the widths of the peaks in the Delft correlation setup and the histogram setup still poses some problems. The width of the correlation peak corresponds to twice the timing jitter of the SSPD and electronics,
while the histogram width in the former setup corresponds to the laser trigger jitter (<10 ps) and the SSPD and electronics. Thus, the peaks measured in the correlation setup should have roughly twice the width of the peaks measured in the former setup. This is not the case, since the correlation peaks are much larger than twice the histogram widths in the previous setup. Further investigation into this discrepancy is warranted. Regardless, the correlation results demonstrate the utility of SSPDs in general and the Warsaw setup in particular for quantum cryptography systems.
Figure 4.2.2-3 Correlated counts between device #11 and #12 measured in Delft setup.
Chapter 5 SSPD optimization

5.1 Motivation - QE and counting rate

We have reported counting rates of 2 GHz for NbN meander structures [51]. In addition, we have made QE measurements independently. Recalling from Chapter 3, the change in the device QE was not monitored as the laser repetition rate was changed, for instance, from 1 GHz to 2 GHz. The response time was merely defined according to the maximum measured counting rate. A more appropriate measurement of the response can be defined by considering the change of QE with counting rate. Thus, a simultaneous measurement of QE and counting rate is necessary to find the true response time of the detector.

In more recent measurements, these high counting rates have not been obtainable using newer larger devices (approximately 100 $\mu$m$^2$ area, 4 nm–thick, 120 nm–wide stripes, and 60% fill factor). In this section, we discuss recent work that reveals the reasons for this result, as well as discuss recent experimental results from measuring the QE and response time of a SSPD simultaneously and the model used to explain these results. According to the model, the relationship between QE and response time is related to the device inductance.

Kerman et al. measured the response time and QE simultaneously [47] using a Brown-Twiss type time delay setup [61] shown in Figure 5.1. An attenuated laser pulse is split, and the pulse in one arm is time delayed with respect to the other. The time delay is adjustable so that the SSPD can detect two pulses with a variable time separation. The number of times both pulses were detected is recorded as the time
delay is varied. This type of setup was also used in our earlier experiments of response time where QE was not monitored closely [32].

By recording the number of photons detected while varying the time delay, Kerman et al. showed that as the pulse separation time becomes smaller, the probability of detecting both pulses decreases. For instance, for a device with \( L = 132 \) nH, once the delay is decreased below a value of 8.7 ns, the QE drops to 90% of its highest value. Likewise, for a device with \( L = 56.5 \) nH, a delay of 6.5 ns drops the QE to 90% of its highest QE value. The results for the 132 nH and 56.5 nH device are presented by Kerman et al. in Ref. 47 and reproduced here for convenience in Figure 5.1-2(a).

Figure 5.1-1 Basic time delay setup used for QE vs. counting rate measurements
Kerman et al. modeled the detector response by considering the hotspot model and the effect of device inductance. This simple hotspot/inductance circuit model, similar to the model used by Hadfield et al., is shown in Figure 5.1-2(b). $R_n$ is calculated using theoretical expressions for the resistance of the hotspot combined with that resulting from PSC formation (Equation 2.3.2). The parameters $R_s$ and $I_b$ are the source impedance and bias current, respectively. According to the model, after a photon is absorbed, the current through the device drops with a fall time of $L/(R_n + R_s)$. Once the current drops below the return current $I_r$, defined as the current necessary for superconductivity to be restored, the current begins to rise again with a rise time of $L/R_s$. $I_r$ is determined by the self-heating of the hotspot region as described in Chapter 2. The voltage is generated by the current through the
inductor/resistor combination before superconductivity is restored, and by just the current through the inductor after it is restored.

To verify the above concept, we used this inductance/hotspot model to fit our measured photoresponse pulse generated from a 10-µm x 10-µm meander SSPD with a 60% fill factor, 4-nm–thick film, and 120-nm–wide stripes (device #3), as shown in Figure 5.1-3. For this simulation, \( L = 420 \, \text{nH} \), \( R_s = 50 \, \Omega \), \( R_n = 200 \, \Omega \), and \( I_r = 0.80I_c \). Pspice Schematics, Evaluation Version 9.1 [62], was used to simulate the response, and a separate filtering program was used to account for the 0.05-4 GHz amplifier-limited bandwidth of our system. As shown, the model (solid red line) fits the measured photoresponse very well.

![Figure 5.1-3 Measured photoresponse with fitted response predicted by inductance/hotspot model](image-url)
This model is used not only to predict the photoresponse pulse shape, but as already shown in Figure 5.1-2, can be used to determine the detector response time [relative detection efficiency (DE) versus pulse separation]. The theoretical fits to the data were generated using the instantaneous current through the device predicted by the circuit model together with the experimentally determined QE versus bias current. For instance, if a photon absorption event occurs, the current drops, so that when the next photon absorption event occurs at some time later, the photoresponse is less probable, since the QE has decreased (according to the measured QE versus bias current dependence). In this way, Kerman et al. calculated the dependence of QE versus pulse response time for two different device inductance values, 56.5 nH and 132 nH over a range of pulse separation times.

Several theoretical curves like the ones used in Figure 5.1-2(a) are shown in Figure 5.1-4 for several device inductance values. For these curves, we used our measured QE versus bias current dependence shown in Figure 5.1-5. The QE saturates for pulse separation times $\gg L/R_s$. As shown, the lower the inductance, the steeper the drop in QE with decreasing response time.

Figure 5.1-4 Simulated relative DE versus pulse separation time
We extended the analysis of Kerman et al. in order to predict the response time over a wide range of device inductances and extracted the relationship between the device inductance and response time. We define the response time as the time delay at which the QE drops to 90% of its maximum value for each curve. We note that the values of response time we generated are longer than those predicted by Kerman et al., due to the difference in the measured relative DE versus reduced current. However, we can extract a linear dependence of the inductance versus response time for response times lower than the superconducting recovery time, as shown in Figure 5.1-6.

![Figure 5.1-5 Relative detection efficiency (DE) versus reduced current and exponential fit.](image)
Using the results of Kerman et al. and the linear relationship between inductance and response time extracted, a 17-nH inductance would be necessary in order to achieve a 500-ps response time. Using the result shown in Figure 5.1-6, this inductance would be only $\sim 5$ nH. We have shown, however, that the model can be used to predict the response time for a given SSPD inductance. We will use the result of Kerman et al., since it has been experimentally verified, to predict the inductance for a number of meander geometries in order to optimize the design of future SSPDs for a given counting rate with respect to SQE.

5.2 SSPD Inductance calculations

As described earlier, the inductance calculations were performed using FastHenry, version 3.0, simulation program designed for calculating the inductance of a conductor or superconductor in an arbitrary configuration [48]. To investigate the
extent to which the simple expression for the kinetic inductance of a superconducting wire [Equation (2.5.1)] applies, simulations were performed to compare the inductance of a meander structure to that of a long stripe with an equivalent length with the same width and thickness. For instance, the inductance values calculated using both geometries were approximately equal (<1% difference) for our most common SSPD geometry of 100 \( \mu m^2 \) area, 120 nm–width, 60% fill factor, and 4 nm-thickness and equivalent bridge of length 508 \( \mu m \).

Moreover, we calculated the inductance of a meander structure that more accurately represented an SSPD by making the edges wider where the meander changes direction, which is done in a real device so that the width near the corners do not set the critical current lower than other straight segments. Similarly, for the common SSPD geometry (10 x 10 \( \mu m^2 \) area, 4 nm-thick, 120 nm-wide stripes, 60% fill factor), the calculated inductance only deviated by ~2% from the inductance calculated with the uniform width meander. Thus, we use the value of \( L = 420 \) nH in approximate agreement with all configurations used.

In order to investigate a wider range of dimensions for the inductance calculations, we use the simple bridge geometry and vary the thickness and width. From Equation 2.5.1 and our analysis of a single SSPD meander geometry, we expect \( L \propto 1/\sigma \), so that width and thickness should contribute equally to the inductance. This relationship is verified in Figure 5.2.1, as the all geometries used follow the same curve. Thus, we conclude that the current distribution is approximately uniform. Moreover, varying the number of filaments from one to ~100 per segment
for both high $\sigma$ and low $\sigma$ devices did not produce any appreciable inductance change.

Two additional parameters can be varied to change the device inductance—the fill factor and the temperature. Changing the device fill factor has the effect of changing the meander length. The temperature, $T$, is changed by simply changing the London penetration depth, $\lambda_L$, according to the following relation:

$$\lambda_L(T) = \frac{\lambda_L(0)}{\sqrt{1 - \left(\frac{T}{T_c}\right)^4}}, \quad (5.2.1)$$
where $\lambda_L(0)$ is the London penetration depth at $T = 0$. In this way, we can input $\lambda_L$ into the inductance calculation to find the inductance at any temperature. The value of the London penetration depth saturates to its lowest value for temperatures below $\sim 0.5T/T_c$, so decreasing the temperature below the device’s normal operating temperature (4.2 K) will not change the inductance significantly.

5.3 SSPD high counting rate/high QE design geometry

5.3.1 Inductance

To illustrate the possibility that our older devices were still limited by their inductance, we take the device which we measured at 2-GHz counting rate and calculate its inductance using its meander dimensions—250-nm stripe width, 70% fill factor, 3.5 nm thickness, and a 16 $\mu$m$^2$ area. We also assume a penetration depth of the NbN film $\lambda_L \approx 560$ nm at 4.2 K as before. This yields an inductance of 17.3 nH calculated. This value corresponds to a response time of 510-ps (1.9-GHz counting rate), which is close to our maximum measured counting rate. We note that the QE at this counting rate was not monitored closely so that the QE may not have been 90% of its maximum value. This suggests, however, that the faster response time is related to the lower inductances of the older devices.

We now use the model to design SSPDs that can be fabricated with current technology that operate with a desired response time while optimizing their SQE. We use FastHenry again to calculate possible meander geometries to achieve this inductance-limited response time. We choose 500 ps response time for the device
and investigate the geometry that maximizes the SQE in a fiber-coupled system at $\lambda = 1550$ nm.

### 5.3.2 Device area and fiber coupling

We start by considering the appropriate meander area by comparing 25 $\mu m^2$ and 100 $\mu m^2$ devices. Figure 5.3.2-1 shows a plot of stripe width and corresponding fill factor for 25 $\mu m^2$ and 100 $\mu m^2$ area and 4 nm and 10 nm–thick devices that yield an inductance of 17 nH corresponding to a 500-ps response time. Note that the range of values of stripe width shown in the figure is arbitrary and the appropriate range for consideration will be discussed later. As evident, the fill factor of the SSPD must be quite small for the larger area devices to maintain this inductance value because of the much longer meander.

From these simulation results, we conclude that for a low inductance/high counting rate device of 4 nm thickness, the device must be fabricated with a large stripe width and a low fill factor. Clearly, if we use a smaller area device, we would require a much larger fill factor than if we use a large area device. In deciding between a 25 $\mu m^2$ area with a higher fill factor or a 100 $\mu m^2$ area with a low fill factor, we consider the coupling factor, $K$, of a Gaussian beam onto the detector. For instance, for a single-mode fiber-coupled system, with spot radius, $w_o = 4 \mu m$, wavelength, $\lambda = 1550$ nm, and a fiber-detector distance, $z = 5\mu m$, the coupling factor $K = 97.3\%$ for the 100 $\mu m^2$ area detector and $K = 61.4\%$ for the 25 $\mu m^2$ area detector. For this analysis, we will neglect the reduction in coupling due to misalignment.
Taking into account the fill factor for equivalent inductance of the two devices (shown in Figure 5.3.2-1) along with the coupling factor at a distance of 5 µm, the 100 µm² area device can couple 7.8% of the radiation, and the 25 µm² area device can couple 19.6% of the radiation. From this analysis, we can narrow our investigation to 25-µm²-area devices and now optimize the QE for this fast detector.

In calculating the total SQE of the detector for this geometry, we assume the maximum QE of the film to be the absorption coefficient for NbN, of 26% for a 3.5 nm-thick film, which has been approached in recent measurements [37]. The maximum QE for the device suggested above is then 0.26 x 0.196 = 5%. In order to increase this QE, we explore 25-µm² area devices with thicker films.

![Figure 5.3.2-1 Some possible geometries for NbN meanders with 2 GHz counting rates with 4 nm film thickness 25 µm² and 100 µm² area.](image-url)
5.3.3 Design specifications

We now turn to determining the range of values of stripe widths employable for a 25 \( \mu m^2 \)-area SSPD with 500-ps response time for a range of values of film thickness. To this end, we make several assumptions about the SSPD response mechanism. First, the SSPD response was assumed to be generated in a subcritically \((0.95I_c)\) current-biased film by absorbing a 1550 nm single-photon, which creates a hotspot with a diameter experimentally determined as 30 nm in a 3.5 nm–thick film, which corresponds to a constant \(2.5 \times 10^{-24} \text{ m}^3\) hotpot volume at this wavelength [30]. The current in the sidewalks then exceeds the critical current by more then 10\%, and PSCs are created with high probability in the sidewalks. Thus, a resistive barrier is formed across the entire width of the stripe and a voltage response is measured. According to these parameters, the maximum width that will generate a response is 206 nm for a 4-nm–thick film and 75 nm for a 30-nm–thick film. In addition, 70 nm was chosen as the minimum width that can be fabricated using current technology. We, therefore, restrict the range of stripe widths according to their maximum allowable widths, and find the fill factor necessary to maintain an inductance of 17 nH for a 2-GHz counting rate device.

Figure 5.3.3-1 shows the inductance versus fill factor for a 70 nm-width, 4 nm-thick, 5 \( \times \) 5 \( \mu m^2 \) meander and an equivalent bridge (equal length, width, and thickness). The jumps in inductance occur as segments are added to the meander and the bridge length is accordingly changed in the same discrete steps. Thus, no partial segments are used in the geometries, which is consistent with the way such devices would be fabricated. Also notice that the inductance decreases as the fill factor
increases until the jump, because the short segments become even shorter as the fill factor increases. Lastly, the difference between the wide turn meander and the bridge geometry differ by <20%.

We then extract the fill factor corresponding to the 17 nH inductance and repeat for various device dimensions. The results of the simulations for determining the fill factor necessary for a given stripe width and film thickness using the non-uniform meander and bridge geometries in order to obtain a 17 nH-device are shown in Figure 5.3.3-2. As evident, both the meander and bridge geometries give similar values of fill factor for the same thickness and width. Notice that for the 4 nm-thick device, the fill factor becomes extremely small except in for the case of a very wide

![Figure 5.3.3-1 Meander fill factor dependence of inductance for a 70 nm, 4 nm thickness, 25 µm² area device](image)
stripe width. The thicker devices have larger fill factors, and approach 30% for wider stripes. Thus, we choose the 20-nm device thickness, since it will absorb more photons and maintain some flexibility in fabrication with respect to stripe width.

Deciding on the desired fill factor and stripe width for a given thickness is not as straightforward. The reason is that the tradeoff between QE and stripe volume has not been experimentally explored sufficiently. Some older 10-nm–thick devices were measured, but the uniformity of the stripe widths [32] is not comparable with current technology [63] making the comparison difficult. As the stripe width increases, however, the QE is expected to drop, since the hotspot generated from a 1550-nm
photon creates less of a perturbation, and the current density in the sidewalks will not increase as much. This will decrease the probability of occurrence of a PSC, thus decreasing the QE of the device.

We then choose two meander sizes for our 2-GHz counting rate devices: (a) 70-nm-wide stripes, 0.15 fill factor, 20 nm–thick, 25-μm² area and (b) 90-nm–wide-stripes, 0.27 fill factor, 20-nm–thick, 25-μm² area. These choices will both shed light on the relationship between QE and stripe width, while maintaining a high coupling factor.

The maximum SQE is calculated here by taking the product of $K$, fill factor, and $\alpha$, the absorption coefficient. For an SSPD with geometry (b) proposed for 500-ps response time placed in a single-mode fiber setup, the maximum SQE is estimated as $\approx 13\%$. Using the best $K$ obtained in the Warsaw setup and NIST setup, SQE $\approx 4.4\%$ and $1.3\%$ for $T \approx 4$ K, respectively. Taking into account the increase in QE provided by lowering the operating temperature, the SQE would increase to $3.3\%$ at 2 K in the NIST setup. We will compare these values with alternative SPD technologies with respect to the DARPA quantum cryptography system in later sections.

5.4 Alternative geometries

Another approach to designing SSPDs with a desired response time is to maintain the current standard in SSPD fabrication topography, 4 nm–thick, 120 nm-stripe width, 60% fill factor, but maintain a low inductance. This can be done utilizing multiple devices arranged in parallel. For instance, if the inductance of a
single meander device is $L$, then a device consisting of $n$ of these meanders in parallel should have a lower inductance than the parallel meander by a factor of $\sim 1/n$.

The first SSPDs proposed here with this parallel device geometry have three parallel meander configurations: (a) two sections of 17 stripes in each, (b) three sections of 17 stripes in each, and (c) ten sections of five stripes in each as shown in Figure 5.4-1 [64].

FastHenry was again used to compute the inductance values for these simulations. The inductance for the two-section 17-stripe per section device, three-section 17-stripe per section device, and ten-section 5-stripe per section device are 68.2 nH, 46.9 nH, and 4.13 nH, respectively. Note that $L = 141$ nH for one-section of 17 stripes and $L = 41.3$ nH for one section of 5-stripes, which verifies that the parallel meanders combine as inductors in parallel.

![Figure 5.4-1 Parallel device geometries: (a) two parallel meander sections with 17 stripes per section, (b) three parallel meander sections with 17 stripes per section, and (c) ten parallel meander sections with 5 stripes per section.](image)
We will focus here on the two-section device’s response. The model shown in Figure 5.4-2(b) shows the device inductances in parallel in a circuit consistent with the Moscow experimental setup. The model can be easily modified to account for the other device geometries simply by adding more parallel inductive branches for a device with more parallel meander sections.

For the device shown here, once the photon is absorbed in one section, the current would either be diverted to the other parallel section or to the transmission line. A normal region may also develop in a parallel superconducting section depending on the original bias current. The current through each section is shown in Figure 5.4-2(a) assuming that the parallel section stays superconducting after photon absorption in the adjacent section. As shown, the current is slightly increased in the parallel superconducting section once the hotspot is absorbed in this other section. In

Figure 5.4-2(a) Simulated current in a two-section device geometry. The hotspot section refers to the section of the meander in which the photon is initially absorbed; The parallel section is the section of the meander in which no photon is absorbed.; The transmission line refers to the circuit branch containing the 50 $\Omega$ transmission line (b) Circuit model for Moscow setup consisting of two parallel meanders each with $L = 146$ nH.
this case, if $I_c$ of each section is greater than $\sim 12.5 \, \mu A$, the parallel section will not produce a normal region.

On the other hand, if $I_c < 12.5 \, \mu m$ in the parallel section of the device and it becomes normal, a second peak will be generated in the voltage signal. For instance, if $I_c = 10.53 \, \mu A$, and we wish to bias at $0.95I_c$ ($I_b = 10 \, \mu A$), $I_c$ will be attained after a time of $730 \, ps$. Thus, a second pulse will be generated at this time. For a device with a larger number of parallel sections, many peaks in the voltage signal could potentially be generated if they become normal. Thus, a careful choice of biasing current must be made in order to avoid this undesirable effect.

Assuming that only the section where the photon is absorbed becomes normal, we simulate voltage response pulses and compare their response times to demonstrate the effect of device inductance on response time. The resulting voltage pulse with a fall time of $2.5 \, ns$ for a two-section device, each section with $L = 141 \, nH$, is shown in Figure 5.4-3. Also shown in the same figure is a voltage pulse with equal fall time generated in a device consisting of one-section of $L = 68.2 \, nH$, and a pulse with a fall time of $4.5 \, ns$ generated from a one-section device with $L = 141 \, nH$. Thus, the parallel geometry decreases the fall time of the voltage pulse. It turns out that the capacitor branch in the setup does not affect the fall times for these inductor and resistor values. This is reflected in the fact that the fall times are roughly $R/L_{eq}$, where $L_{eq}$ is the equivalent inductance.

The response measured with 100-500 MHz bandwidth amplifiers from the two–section device has a fall time of $3 \, ns$, according to our Moscow collaborators, which is longer than the fall time of $2.5 \, ns$ for the (unfiltered) simulated response
pulse. Using the same model, we simulated responses for the other parallel device geometries. The simulated fall times for the three- and ten-section devices are 1.4 ns and 170 ps, respectively, and, according to the model of Kerman et al., the two-section, three-section, and ten-section devices have estimated response times of 4-ns, 2.6-ns, and 250-ps response time, respectively (while maintaining 90% of their maximum QE).

5.5 Response time simulation and measurement

We are currently developing our own response time measurement setup, which unlike previous setups in our laboratory, keeps an accurate count of
photoresponse pulses for any given time delay. As mentioned previously, a Brown–Twiss–type setup is employed to measure the response time of a SSPD. In our setup, the SSPD response pulse is sent to a fast time-resolved pulse counter with a 250-ps time resolution or to a 6-GHz single-shot oscilloscope. The pulse counter is used to keep track of the QE while recording the time histogram. The traces from the oscilloscope are used in fitting the data with the model.

The model is closely related to the hotspot/inductance model and is shown in Figure 5.5-1(c). The first hotspot it absorbed at \( t = 0 \) and the normal hotspot forms. As before, when the current through the device drops to \( I_r \) at \( t = t_r \), the device recovers its superconductivity and the current begins to increase. At \( t = \tau \), the second photon is absorbed, and a second normal region develops. Finally, this second normal region recovers its superconductivity at \( t = t_r + \tau \). The second normal region, consisting of both the resistance from the hotspot and also PSC formation, may have a different value of resistance than the initial normal region. The reason is that the current through the SSPD decreases after the occurrence of the first hotspot, and from Equations (2.3.2) and (2.3.3), this resistance is current dependent. Figure 5.5-1(a) shows a fitting of voltage versus reduced current for several of our fiber-coupled devices taking \( F(T) \) as a fitting parameter.

Notice that the resistance value for device #3 is lower (\( \sim 25 \, \Omega \) at \( I/I_c \approx 1 \)) than used in the model. Equation (2.3.2) used for this calculation, however, consists of the resistance from the original hotspot (and electric field penetration) which is only \( \sim 1 \) to 5 nm. Self heating is neglected, which stretches the hotspot to the 30-nm diameter (for \( \lambda = 1550 \, \text{nm} \)) value to which we referred earlier. This additional resistance in
conjunction with the resistance from PSC formation will increase the resistance. To extract an accurate value of resistance, we multiply the calculated resistance by the ratio of the total width to the original hotspot diameter yielding \(~200\, \Omega\), which is consistent with the normal resistance previously used.

These devices, including device #3 which is the device used for the simulations presented in this section, show a constant resistance over all bias currents.

Thus, the resistance versus time delay graph follows the shape of the current versus

![Figure 5.5-1](image)

Figure 5.5-1 (a) Voltage versus reduced current data and fitting for several SSPDs. (b) Resulting normalized resistance versus time delay calculated using (a) for device #3 and circuit model (c).
time dependence predicted by the inductance/hotspot model. As shown in Figure 5.5-1(c), the resistance of the second hotspot only changes by a maximum of 1% with respect to the first, so the resistance values for both the initial and time delayed hotspots are assumed to be equal for the simulations presented here. Note that the model shown in Figure 5.5-1(c) does not include the bias-tee and amplifiers. A separate filtering program was used to take into account the effect of these components, whose spectra were measured with a spectrum analyzer.

We analyzed two experimental configurations in order to find the most optimal method for discriminating pulses spaced closely together in time. The first was a more wide-band system which used an 80-kHz to 26-GHz bias-tee and 50-MHz to 4-GHz amplifiers. Figure 5.5-2 shows the response for a pulse separation of 6.4-ns and the simulated response. As evident, the simulated response follows the measured

![Figure 5.5-2 Measured photoresponse of time delayed pulses with a 6.3 ns pulse separation and simulated photoresponse in wide-band system.](image-url)
Chapter 5 SSPD OPTIMIZATION

photoresponse closely. Note that at this value of pulse separation time, the pulses can pulse becomes more difficult to discriminate, as shown in experimental results presented in Figure 5.5-3. As the pulse separation time becomes less than ~3.7 ns, the second pulse is barely visible. According to the model, a second voltage pulse could be generated at any time greater than the rise time of the initial pulse, since this corresponds to the amount of time necessary for superconductivity to be restored in the structure.
Figure 5.5-3 Measured photoresponse in time delay setup for several values of pulse separation. The left column shows single-shot pulses and the right shows averaged pulses.
Without the slow relaxation caused by the inductance, resolving both pulses may be more easily achieved. For this reason, we investigated a narrow band system, where we simply add a 100-MHz high pass filter to the original configuration to remove the slow relaxation. A single pulse using this setup is shown in Figure 5.5-4 along with the simulated result. Experimental results have yet to be obtained using this time delay setup, but since the simulation fits the single response pulse, we can make predictions using the model.

Figure 5.5-5 shows the simulated response for several values of pulse separation time. Two peaks corresponding to the pulse separation appear more visible than in the wide-band system case for times substantially larger than the 470-ps rise time. As the pulse separation is changed to a small value close to the rise time, the two peaks are still visible but do not correspond to the actual pulse separation time. This phenomenon will be quantified shortly for both the wide and narrow-band system.

![Figure 5.5-4 Measured and simulated photoresponse in the narrow-band system](image-url)
Figure 5.5-5 Simulated responses for time delay values of 4 ns, 2 ns, and 500 ps in the narrow band setup. The green line shows the minimum relative discrimination level that could be used to count both pulses.
In order to compare the ability of our digital counter to discriminate pulses in these wide-band and narrow-band system, we use the fact that the pulse counter discriminates on the falling edge of a pulse. We can then compare the amount of time above the minimum discrimination level (shown in Figure 5.5-5) on the falling edge of the pulse for the wide-band and narrow-band configurations. Figure 5.5-6 shows the time available for discrimination versus pulse separation time for both narrow-band and wide-band systems. The pulses generated from the wide-band system have a larger discrimination time than those generated from the narrow-band system. The discrimination times for both systems, however, can be discriminated with available technology [65] for delay times larger than the rise time of the pulse. Figure 5.5-6 also shows the nonlinear behavior of pulses separated by a time approaching the rise time of the pulse. Thus, for times close to the rise time, the pulse separation that can be detected does not correspond to the actual pulse separation time. As mentioned, this rise time corresponds to the amount of time necessary for superconductivity to be restored in the device, so the change in discrimination properties close to the rise time is expected.
5.6 Comparison of SPDs

The SSPD is not the only option for single-photon detection for QC systems. Several current options in single-photon detection technology are compared in Table 5.6-1. For the wavelength of interest, 1550 nm, the InGaAs APDs have a high QE of 14% but have a lower counting rate than SSPDs, higher dark counts, require active quenching, and also frequently suffer from afterpulsing [66]. We note that Si APDs are not usually suitable for this wavelength since the photon energy is not enough to excite electrons into the valence band, so the QE is very low. However, the up-conversion process can be used to convert the telecom wavelength down to the visible for more efficient detection. An up-conversion conversion efficiency $80 \pm 15\%$ has recently been reported and an overall system efficiency of $\sim 33\%$ has recently been demonstrated using a periodically-poled lithium niobate crystal (PPLN) for up-
conversion and a commercial Si APD [67]. The main advantage of APDs is that they can be operated at much higher temperatures and are more readily available than SSPDs.

Another promising superconducting detector is the transition edge sensor (TES) [68]. The TES is a sensitive microcalorimeter that measures a small change in temperature caused by photon absorption and consists of a 20-nm-thick, 25 x 25 $\mu$m$^2$-area tungsten film placed inside an optical cavity designed to increase absorption at $\lambda = 1550$ nm. This detector has extremely high QE and very low dark counts. In addition, the TES can resolve the number of photons in a laser pulse, which as we mentioned previously, is quite advantageous in some implementations of quantum cryptography. The main drawback of the TES is that counts can be generated at a rate of only 250 kHz and has a timing resolution of 72 ns [69]. Since here we mainly focus on an attenuated laser source as our “single-photon source”, a SSPD will be more suited than a TES due to its higher counting rate. Also, we can increase the nominal QE by decreasing the operating temperature of the SSPD and integrate our devices with resonating structures as was done for the TES [70]. In fact, such devices were recently fabricated and a maximum QE of 47.7% was reported at $\lambda = 1550$ nm and $T = 1.8$ K by the MIT-MIT Lincoln Laboratory group [71].
Table 5.6-1 SPD comparison

<table>
<thead>
<tr>
<th>Detectors</th>
<th>SEMICONDUCTING APDs</th>
<th>SUPERCONDUCTING</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Si</td>
<td>InGaAs</td>
</tr>
<tr>
<td>Temperature (K)</td>
<td>300</td>
<td>200</td>
</tr>
<tr>
<td>Wavelength (µm)</td>
<td>0.25-1.1</td>
<td>1.1-1.8</td>
</tr>
<tr>
<td>Time resolution (ps)</td>
<td>48</td>
<td>~ 300</td>
</tr>
<tr>
<td>Quantum efficiency</td>
<td>25% @ 0.7 µm</td>
<td>14% @ 1.54 µm</td>
</tr>
<tr>
<td>Apertures (µm)</td>
<td>200</td>
<td>30-80</td>
</tr>
<tr>
<td>Dark count rates (cps)</td>
<td>&lt;10^4</td>
<td>&gt;3x10^4</td>
</tr>
<tr>
<td>Data rate (MHz)</td>
<td>&lt;50 MHz</td>
<td>&lt;50 MHz</td>
</tr>
<tr>
<td>Dynamic range (# of photons)</td>
<td>1-3000</td>
<td>NA</td>
</tr>
<tr>
<td>Electrical gating</td>
<td>Yes (0.3 ns)</td>
<td>Yes (0.3 ns)</td>
</tr>
<tr>
<td>Photon number resolution</td>
<td>Limited</td>
<td>No</td>
</tr>
<tr>
<td>Ruggedness</td>
<td>Very high</td>
<td>High</td>
</tr>
<tr>
<td>Availability</td>
<td>Yes</td>
<td>Limited</td>
</tr>
</tbody>
</table>

*Prepared at the NATO Advanced Research Workshop on “Advanced Materials for Radiation Detectors and Sensors: Wide-Gap Semiconductors and Superconductors,” Warsaw, Poland, Sept. 8-10, 2004 by:
Ivan Prochazka, Czech Technical University, Prague, The Czech Republic,
Sae Woo Nam, National Institute for Standards and Technology, Boulder, CO, USA,
Gregory N. Gol’tsman, Moscow State Pedagogical University, Moscow, Russia,
Roman Sobolewski, University of Rochester, Rochester, NY, USA.

** Taken from Ref. [72]
*** Parameters not obtained simultaneously.
Chapter 6 Quantum key distribution system based on SSPDs: performance analysis

6.1 DARPA network

In the introductory sections, the DARPA network was mentioned in which a key is sent from Alice to Bob via single photons. A review of this system is presented by Chip Elliot in [73]. A more detailed diagram of the physical setup of a point-to-point link is shown in Figure 6.1-1. Note that the setup also includes optical components to send a bright 1.3 $\mu$m laser pulse for synchronization between transmitter and receiver.

The figure only shows a single point-to-point link, when, in fact, the DARPA QKD system operates as a network that encompasses Boston University, Harvard University, and BBN Technologies instead of a single point to point link. Operating as a network is inherently more robust, since it allows multiple paths for which information to flow; it has greater resistance to traffic analysis, and extends the physical extent of the network. A single point-to-point link is usually limited to well below 100 km due to the attenuation in the fiber, whereas a network can extend the link indefinitely. Moreover, each point-to-point link is integrated with standard internet protocols for ease of operation. We will now evaluate this network, assuming that the technology could be implemented and compare the outcome with other SPD technologies.
Figure 6.1-1 Phase-coded QKD setup
6.2 Practical QKD

In the introductory remarks on QKD, Equation 1.1.1 was presented to show the factors that must be considered in order to determine the final key rate. We now take a closer look by following results shown by Lütkenhaus [74] in a practical QKD system. The probability that Bob detects a photon is expressed as the combination of the independent probabilities of a detection due to signal photons, \( p_{\text{exp}}^{\text{signal}} \) and dark counts, \( p_{\text{exp}}^{\text{dark}} \).

\[
\begin{align*}
 p_{\text{exp}} &= p_{\text{exp}}^{\text{signal}} + p_{\text{exp}}^{\text{dark}} - p_{\text{exp}}^{\text{signal}} p_{\text{exp}}^{\text{dark}}. 
\end{align*}
\]  

(6.2.1)

\( p_{\text{exp}}^{\text{signal}} \) is dependent on the type of source used, so that for an attenuated laser (weak coherent source) with mean photon number \( \mu \), QE of Bob’s detection apparatus \( \eta_B \) and transmission efficiency \( \eta_T \),

\[
 p_{\text{exp}}^{\text{signal}} = 1 - \exp[-\mu \eta_B \eta_T] .
\]  

(6.2.2)

The error rate, \( e \) sometimes called the quantum bit error rate (QBER) plays a crucial role in this dependence, since error corrections becomes inefficient once \( e \) drops below 11\%. Errors occur when a photon propagates to the incorrect detector because of misalignment or when a dark count is registered, so \( e \) depends on the dark count rate, \( d_B \), the transmission efficiency \( \eta_T \), the QE of Bob’s detector, \( \eta_B \), and the mean photon number.
exp \exp \begin{align} e &= \frac{cp^\text{signal}}{p_{\exp}} + \frac{1}{2} d_B, \\ e &= \frac{p^\text{signal}}{p_{\exp}} + \frac{1}{2} d_B, \\ e &= \frac{p^\text{signal}}{p_{\exp}} + \frac{1}{2} d_B. \end{align}
Before we proceed, we will review some recent results for QKD testbeds and links using current SPD technology.

6.3 Review of QKD links

6.3.1 APD/SSPD QKD link testbed

Recent work performed by the NIST-BBN group has shed light on SPD performance in a QKD system by implementing both an APD and SSPD into the same system [75]. The source in this system produced $\lambda = 1310$ nm photons by generating them from a pulsed generator connected to a laser diode that produced ~1 ns-wide pulses at a 1 MHz-repetition rate. The photons were then sent to either an InGaAs APD or to a SSPD and the photoresponse was recorded with a discriminator and pulse counter. The detector operation was optimized in order to maintain QBER $< 11\%$ for each detector. The corresponding SQE of the APD and SSPD were 23% and 0.5%, respectively. Even though the APD had much greater SQE than the SSPD, its dead time of 10 $\mu$s was much longer than that of the SSPD, assumed in this work to be 10 ns. According to the conclusions in Ref. 75, and as we will later show, the SSPD actually lead to a significant improvement in key rate at high clock frequencies because of its significantly lower dead time.

6.3.2 TES QKD link

The TES has also been integrated into a phase-coded QKD link consisting of 50 km of fiber and evaluated in terms of key rate [76]. $\lambda = 1550$ nm photons were sent from an attenuated pulsed laser source at 1 MHz through telecom fiber. In
addition, a bright $\lambda = 1330$ nm was used for synchronization. Alternatively, electrical synchronization replaced the optical synchronization to yield higher key rates because of diminished stray photon counts. The usual BB84 protocol was used, but with a slight modification made to enable the use of one TES. Note that normally two detectors are used, but one can be used if an additional delay loop is added. Thus, the system in Ref. 76 closely resembled a DARPA QKD link. The highest key rate achieved was $\sim 650$ Hz for a mean photon number $\mu = 0.1$. Operating at this value of $\mu$, their results showed that the link distance could be increased to a maximum distance of 83 km using optical synchronization and to 138 km using electrical synchronization. Increasing the link distance further would lead to a QBER $> 11\%$ and no secure bits could be transmitted.

6.3.3 SSPD QKD link testbed with entangled photon source

Before proceeding to a comparison of all SPDs operating in the DARPA network, we will briefly describe a QKD link which employs an entangled source and utilizes both a SSPD and APD [77]. The correlated photons generated by the entangled source, consisting of a pulsed laser source and PPLN crystals, are split by a beamsplitter and sent to either an APD or SSPD. The coincident counts are recorded between the two detectors. This type of setup is utilized in EPR protocols as discussed in Chapter 2. The coincident counts are recorded as a function of time delay between the two detector arms. In this way, accidental coincident counts were distinguished from ones generated from the same optical pulse. Using results from this measurement, the optimal $\mu$ was calculated as 0.0064 as the value which
produces the maximum ratio of correlated coincident counts to accidental counts. This entangled source’s implementation in the DARPA network is planned in the next phase of this program. This is also a current research direction in our laboratory.

6.4 Performance estimates for DARPA QKD link

We now will give a comparison of SPDs operating in the DARPA network. The results for key rate dependence on link distance are presented in Figure 6.4-1 for SPDs representing the best in current and near-future technology. The optimized SSPD refers to the 20-nm thick, 5 x 5-\(\mu\)m\(^2\)-area device proposed in the SSPD optimization section that was designed to count single-photons at 2 GHz. Furthermore, the key (secrecy) rate is calculated at a temperature of 60 mK, so that the QE is increased and dark counts decreased according to the results obtained in the NIST fiber setup. This is compared to the SSPD that was measured in the NIST setup. The APD and TES parameters used were consistent with the measured parameters in References 75 and 76, respectively. The setup in Ref. 76 was used for modeling the error rate, since the error measurement in Ref. 71 did not account for transmission loss and errors caused by stray light.
In contrast to most analyses done for practical QKD links, the mean photon number $\mu$ was optimized (at each link distance) on order to optimize the key rate. A clock frequency of 3.3 MHz was used, which corresponds to the actual clock frequency at which the DARPA network operates. The dead time of the detectors are taken into account as was done in Ref. 75, so that the key rate $S$ is expressed as

$$S = \frac{G}{\left(\frac{1}{f_{\text{clock}}} + \tau_{\text{dead}}\right)}, \quad \text{(6.4.1)}$$

where $\tau_{\text{dead}}$ is the detector dead time and $f_{\text{clock}}$ is the clock frequency. $\tau_{\text{dead}}$ is set to the gate width used for the TES and APD measurements, and to the calculated response time for the SSPD.
As shown, for this relatively low photon transmission repetition rate, the optimized SSPD outperforms the alternative technologies, except at the longest link distances >95 km. The SSPD in the NIST setup clearly outperforms the APD except over very short distances. The sharp decrease in the key rates of the APD is caused by its high dark count rate ($3 \times 10^4$ Hz). In general, for short links, the key rate is limited by multi-photon events because of the lower errors values. On the other hand, for large distances, the higher error rates dominate the behavior as the link becomes sufficiently long and the error rate becomes too high to distill a secure key.

The large advantage of the SSPD over alternative technologies is most apparent in considering larger $f_{\text{clock}}$. Figure 6.4.2 shows the maximum key rate as the clock rate is varied. As $f_{\text{clock}}$ is increased, $\tau_{\text{dead}}$ limits the key rate. As shown, for $f_{\text{clock}} > 1$ MHz, the optimized SSPD outperforms the TES, and the SSPD in the NIST setup clearly outperforms the APD for $f_{\text{clock}} > 7$ MHz.

Note that the non-optimized SSPD used in the simulation does not represent the best in current technology. For instance, as mentioned previously, Rosfjord et al. reported QE = 47.7% at $\lambda = 1550$ nm and $T = 1.8$ K using an optical cavity to enhance the NbN film absorption [71]. Thus, since the QE of a detector can be increased, the response time becomes the more crucial parameter for QKD, which is the main advantage of the SSPD over competing technologies.

As a final note, Ref. 75 reports the preliminary results from placing the SSPD in the actual DARPA QKD link by replacing one in a pair of APDs in one of the
receivers in the network. Initial tests showed the exchange of secret key material at rates 10-30 bits/s at a high value of $\mu \approx 1$.

Figure 6.4-2 SPD comparison for maximum key rate versus clock frequency
Chapter 7 Conclusions and future work

We analyzed the NbN single-photon detector (SSPD) for the purpose of integration into a fiber-based quantum communications system, namely the DARPA quantum key distribution (QKD) network. After finding that a small area SSPD can count at >2 GHz with 18 ps timing jitter in free-space, we demonstrated its utility in fiber-based systems by performing measurements in two such systems. The first utilized fiber-coupled SSPDs placed in a cryogen-free refrigerator capable of reaching mK temperatures, and the SSPDs were evaluated in terms of system quantum efficiency (SQE) and dark counts over a broad temperature range. We found that the QE decreases with temperature down to ~2 K, after which the QE improvement saturates; the dark count rate, however, decreases with decreasing temperature over the entire temperature range utilized.

The second system consisted of pairs of fiber-coupled SSPDs assembled on inserts, each compatible with a standard helium dewar. For each SSPD, the fiber was permanently attached to the device using a photoresist ring fabricated on the device for alignment purposes. The SSPDs were evaluated in terms of SQE, dark counts, and timing resolution, and we found that the system provides relatively high fiber-detector coupling efficiency, good timing resolution, low dark counts, and can integrate easily into the DARPA network.

We also investigated the SSPD's limitations by analyzing a model which takes into account the SSPD detection mechanism and device inductance to predict its response time. Using the model, we optimized the SSPD meander geometry to design devices with high SQE and counting rate in terms of area, stripe width, fill
factor, and thickness using detailed inductance simulations. We also presented a novel low inductance SSPD designs based on parallel meanders contained in the same SSPD and modeled their photoresponse.

With these designs and experimental results, we showed that the SSPD outperforms its superconducting and semiconducting counterparts for quantum cryptography systems with high clock rates. Thus, the SSPD, with its combination of high QE, and low timing jitter at telecommunications wavelengths, as well as low dark counts, make it a natural choice for the DARPA network and quantum cryptography systems in general.

Several areas of research related to the work presented in this thesis will likely be extended. First, the demonstration of a fiber-based system with high SQE could be accomplished with the further development of the Warsaw fiber-coupling technique and the pre-selection of higher QE devices. The fabrication and testing of low inductance devices using both the optimized geometry and the parallel meander geometry could further validate the hotspot/inductance model using the time delay setup presented and lead to the development of devices with higher counting rates. In addition, the long-term implementation of SSPDs in the DARPA network will be necessary in order to complete the performance analysis. As evident from this thesis, the area of SSPD research has recently become a widespread effort, which draws from a large number of resources at BBN Technologies, NIST, MIT, MIT/Lincoln Laboratory, Moscow State Pedagogical University and others. Thus, we should expect much development in this area in the near future.
Bibliography


14 M. A. Jaspan, J. L. Habif, S.W. Nam, R. Hadfield, “Heralding of entangled telecom
photons with a superconducting single-photon detector.” Submitted to Applied

15 S. Tanzilli, H. De Riedmatten, W. Tittel, H. Zbinden, P. Baldi, M. De Micheli,
D.B. Ostrowsky, N. Gisin, “Highly efficient photon-pair source using periodically


17 G. Brassard, N. Lütkenhaus, T. Mor, B. Sanders, “Limitations on Practical

18 C. Zinoni, B. Alloing, C. Monat, V. Zwiller, L. H. Li, A. Fiore, L. Lunghi, A.
Gerardino, H. de Riedmatten, H. Zbinden, and N. Gisin, “Time-resolved and
antibunching experiments on single quantum dots at 1300 nm.” Appl. Phys. Lett. 88
131102 (2006).


20 G. Gilbert, M. Hamrick, “Practical quantum cryptography: a comprehensive

No. 122 and 124, (1911).

Rev. 108 1175 (1957).

2189 (1971).

24 N. Bluzer, “Temporal relaxation of photoinduced nonequilibrium in niobium.”

25 A. Rothwarf, B.N. Taylor, “Measurement of recombination lifetimes in

26 A. Semenov, G.N. Gol’tsman, R. Sobolewski, “Hot-electron effect in
superconductors and its applications for radiation sensors.” –Topical Review Article,


65 Time Harp 200 manufactured by PicoQuant GmbH Rudower Chaussee 29 (IGZ) 12489 Berlin Germany


