Cryo-CMOS Readout of SNSPDs

for Diamond-based Quantum Computers

Maoran Li

Delft University of Technology

TU Delft

Cryo-CMOS Readout of SNSPDs

for Diamond-based Quantum Computers

by

Maoran Li

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday August 29th, 2023 at 9:00 AM.

Supervisor:	Dr. Fabio Sebastiano
Co-supervisor:	Luc Enthoven
Project Duration:	September 2022 - September 2023
Faculty:	Faculty of Electrical Engineering, Mathematics & Computer Science, TU Delft
Thesis Committee:	Dr. F. Sebastiano, TU Delft, supervisor
	Dr.ir. C.J.M. (Chris) Verhoeven, TU Delft

Cover:	Photo by Maria Lysenko on Unsplash
Style:	TU Delft Report Style, with modifications by Daan Zwaneveld

TU Delft

Acknowledgement

I am extremely grateful to Dr. Fabio Sebastiano and Luc Enthoven for their support throughout the project for the past year. Weekly conversations and constructive feedback have been invaluable in my academic and personal growth. I would also like to express a special thank to Luc for being an incredible mentor. He is always available to answer my questions and is patient and supportive when it comes to tackling project issues, technical problems, and even personal matters. Being part of CoolGroup is a great experience, everyone is willing to help and the social gatherings are always enjoyable.
Furthermore, I would like to thank my parents and friends for their understanding and encouragement, which has been a great source of comfort during the most difficult times.

Maoran Li
Delft, August 2023

Abstract

Single-photon detection is a critical step in both quantum computing and quantum information technology. For instance, the measurement of qubit states and the establishment of entanglement in nitrogen vacancy (NV) center quantum computing requires the detection of single photons. Superconducting nanowire single-photon detectors (SNSPDs) are very competitive devices due to their performance in high detection efficiency, high count rates, low dark count rates, and low jitter. However, the readout electronics are usually implemented at room temperature by connecting the dilution refrigerator through coaxial cables, while SNSPD need to work in a cryogenic environment that is close to the qubits. In order to have better integration and reduce the complexity of the wiring when the number of qubits increases, it is essential to position the electronics close to, or even integrated with, the SNSPD and qubits.
In this thesis, a cryogenic CMOS readout for SNSPD is designed and taped-out using TSMC 40 nm technology. The SNSPD is designed for color-center quantum computing, which is anticipated to work in the wavelength range of 619-620 nm and 625-750 nm, and at a temperature of 1.8 K. The readout electronics are expected to operate at 4 K. The system is required to have a detection efficiency of more than 90 % and a dark count rate of less than 1 Hz. With the help of SPICE dynamic model, the SNSPD is reproduced in Cadence Spectre for circuit design. Active quenching is implemented in the readout architecture, allowing for an increased readout resistor, which improves the output slew rate and count rates without any latching while still keeping a high bias current for a higher detection efficiency. Under a -40 $^{\circ}C$ simulation, the readout system achieved count rates greater than 20 MHz, an average jitter of 25 ps_{rms}, and a power consumption of 36 μW, while simultaneously expecting to significantly suppress the dark count rates and after pulses.

Contents

Nomenclature

Abbreviations

Abbreviation	Definition
SNSPD	Superconducting Nanowire Single-Photon Detector
NISQ	Noisy Intermediate-Scale Quantum era
CMOS	Complementary Metal-Oxide-Semiconductor
QKD	Quantum Key Distribution
SPAD	Single-Photon Avalanche Diode
NV	Nitrogen Vacancy
PSB	Phonon Side Band
BS	Beam Splitter
ZPL	Zero Phonon Line
TDC	Time-to-Digital Converter
QP	Quasi-Particle
SDE	System Detection Efficiency
FWHM	Full Width at Half Maximum
RMS	Root Mean Square
SPICE	Simulation Program with Integrated Circuit Emphasis
ENBW	Effective Noise BandWidth
DAC	Digital-to-Analog Converter
LSB	Least Significant Bit

1 Introduction

1.1. Quantum Computing and Cryogenic Electronics

Quantum computing has become one of the most exciting technologies due to its potential to accelerate computation exponentially. Traditional computers use bits to store and process information. Each bit is stored as either 0 or 1, and these logics are often controlled by low or high voltages. For a quantum computer, a "qubit" can be in a superposition state of the 0 state ($|0\rangle$) and the 1 state ($|1\rangle$) [1]. Hence, a qubit is in the quantum state of $|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle$, where α_0 and α_1 are complex amplitudes with $\alpha_0^2 + \alpha_1^2 = 1$. If this qubit is measured on a computational basis, the result will be either $|0\rangle$ with a probability of α_0^2 or $|1\rangle$ with a probability of $|\alpha_1|^2 = 1 - \alpha_0^2$. Therefore, the advantage of a qubit being in a superposition of two classical states is that the information stored in qubits increases exponentially with the number of qubits. The n-qubit states can be in the superposition of the 2^n computational basis with the form of $|\psi_n\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle + \cdots + \alpha_{2^n-1} |2^n - 1\rangle$.

In the current NISQ (Noisy Intermediate-Scale Quantum) era, fault tolerance and scalability have become the most significant challenges because qubits are extremely noise sensitive. Any unwanted disturbance can limit the quantum computer from displaying its powerful computing capabilities. DiVincenzo proposed five criteria in 2000 for building a practical quantum computer [2]. Several quantum systems, including Josephson junctions in superconducting circuits [3], silicon electron spins [4], and spins of impurities in solids [5], have met these requirements to some extent. In most cases, qubits need to be cooled to an extremely low temperature (milli-Kelvins) in the dilution refrigerator, and initialization, operation, error correction, and readout are performed by room-temperature control systems through long wires [6]. However, with an increasing number of qubits in the quantum processor, connecting quantum chips from dilution refrigerators to the room-temperature control systems becomes increasingly intricate and costly because of the physical space and heat load of those wires.

One of the solutions for the scalable quantum processor is positioning the electronics close to the qubits, as shown in Figure 1.1, to reduce the number of interconnects required in the dilution refrigerator. CMOS (Complementary Metal-Oxide-Semiconductor) technology is advantageous among the different semiconductor technologies since it can operate at temperatures as low as $30mK$. Furthermore, it has been used to construct complex and power-efficient electronic interfaces for qubits due to its reliability and integration capability [8]. Therefore, cryo-CMOS is promising for the development of large-scale quantum computers.

Photon detection is necessary in both quantum computing and quantum information technologies. Quantum information technology involves using quantum mechanical objects to encode, communicate, manipulate, and measure information [9], and photons are the information carriers when the system utilizes optics. For example, in quantum communication, single photons traveling over a channel prevent eavesdroppers from gaining information in QKD (Quantum Key Distribution) [10]; in photonic quantum computing, qubit information is encoded in the quantum state of the photon by utilizing degrees of freedom such as polarization. This work is based on the color center quantum computing using the negatively charged states of Nitrogen-Vacancy (NV) center, known as NV^-, as the qubit. This type of qubit has a longer coherence time due to the excellent insulation from the external environment

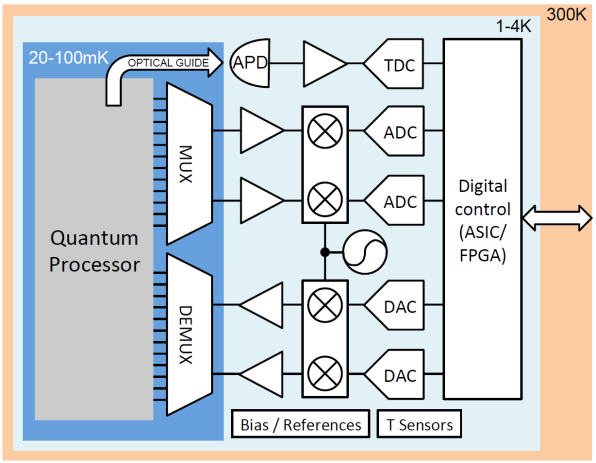


Figure 1.1: A generic system for a scalable qubit array with most of the electronics components operating in cryogenic temperature [7].

provided by the diamond lattice and can be operated even at room temperature, making it advantageous compared to other qubits. The well-defined optical transition makes the detection of emitted photons essential for the state readout and initialization, as well as the establishment of entanglement. Therefore, single-photon detection is required.

1.2. Single-Photon Detection

1.2.1. Single-Photon Detectors

There are various types of single-photon detectors still actively being investigated to push the photon detection limit. Silicon-based SPADs (Single-Photon Avalanche Diodes) absorb photons and generate an electron-hole pair. When the bias voltage is higher than the diode breakdown voltage, the current is self-sustained and will saturate as the charge multiplication (avalanche) process continues [11]. Another class of detectors relies on superconducting material, and can be divided into different types based on different operation mechanisms and topologies: transition edge sensor (TES), superconducting tunnel junction (SJT), microwave kinetic inductance detector (MKID), and superconducting nanowire single-photon detector (SNSPD) [12]. Due to its impressive performance in photon detection efficiency, dark count rate, jitter, and dead time, SNSPDs attract increasing interest [13]. In addition, it is used in this work also because of its advantage of simpler fabrication and the ability to work at cryogenic temperatures together with qubits and cryo-CMOS electronics, which is beneficial for high integration and scalability, compared to SPADs.

1.2.2. SNSPD Readout

The concept of SNSPD reading single-photon information (the detailed detection mechanism will be explained in Section 2.2.2) is to connect it in parallel with a low resistance route. The dashed line box in Figure 1.2 represents the SNSPD, which is modeled by a variable resistor $R_n(t)$ and kinetic inductance L_k (the detailed SNSPD modeling can be seen in Chapter 3). When the SNSPD is biased with a current of $I_b = 20\ \mu$A and connected to a load resistor of $R_L = 50\ \Omega$, the current initially biased in the SNSPD, I_b, will be diverted to the low resistance path ($I_b = I_d + I_L$) since $R_n(t)$ will become much larger than R_L in the detection phase. This results in a maximum voltage output of $V_{out} = I_{L_{max}} \times R_L = I_b \times R_L = 1$ mV after detecting a photon, thus allowing the photon signal to be converted to a voltage signal, which can then be further processed and read out in the following circuits. The voltage produced by the readout is restricted to the product of the load resistance and the bias current. Increasing the load resistance can result in a considerable gain in pulse amplitude, which is advantageous for the timing jitter. However, the maximum load resistance is restricted by the latching problem and the impedance

matching when using 50 Ω cables. Additionally, as the number of qubits increases, the required SNSPD increases, making wiring complicated. Consequently, a compact readout system that requires cryogenic electronics to be placed close to the qubits and SNSPD is necessary.

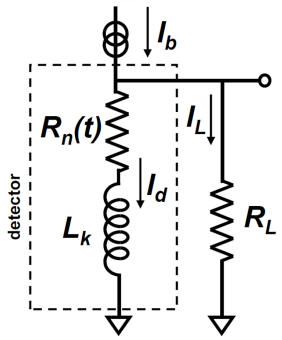


Figure 1.2: Typical SNSPD readout using a low-resistance path R_L [14].

1.3. Thesis Objectives

This work aims to design the first cryo-CMOS SNSPD readout circuit that is capable of operating at 4 K using TSMC 40 nm technology, which can then work closely with SNSPD and be utilized for the development of a diamond-based quantum computer. The main objectives are

- Gain an insight into the fundamentals of how NV center qubits and SNSPDs operate. Examine the performance of the SNSPD system and formulating the requirements for the readout circuits.
- Examine the influence of SNSPD and circuit parameters (readout resistance R_L, bias current I_{bias}, switching current I_{sw}, kinetic inductance $L_k(t)$, and hot spot resistance $R_n(t)$) on the SNSPD's performance in terms of count rates, dark count rates, afterpulses, dead time, latching, and jitter.
- Investigate how the readout circuit (including circuit architecture and electrical parameters) impacts the performance of the SNSPD.
- Design a cryo-CMOS readout circuit meeting the performance required for the application.

The SNSPD single-photon readout system is designed to meet the specifications [15] listed in Table 1.1.

1.4. Thesis Outline

This thesis is structured as follows: Chapter 2 provides an overview of the fundamentals of diamond-based vacancy center quantum computing and the necessity of SNSPD and single-photon detection in this particular quantum system. It also covers the SNSPD single-photon detection mechanism and the associated non-idealities for assessing the SNSPD's performance. Chapter 3 focuses on modeling the SNSPD in SPICE and comparing the Spectre simulation performance with the device effective performance. Chapter 4 outlines the step-by-step process of designing the readout circuit exploiting the active quenching technique. This includes presenting the simulation results of the impact from different SNSPD parameters including R_L, C_{pad}, I_{bias}, L_k, and $R_n(t)$ for determining the system-level architecture, designing the transistor-level implementation of the analog and digital subcircuits, and simulating the performance. Then, the measurement plan will be briefly introduced, followed by the final chapter, which includes the conclusion and potential improvements.

Table 1.1: SNSPD Readout System Specification

Specification	**Required**	**Preference**
Operating Wavelength	619-620 nm (ZPL) + 625-750 nm (PSB)	
Operating Temperature	SNSPD: 1.8 K; Readout electronics: 4 K	
System Detection Efficiency	>80%	>90%
Dark Count Rate	<5 Hz	<1 Hz
Dead Time	<100 ns	<50 ns
Count Rate	>10 MHz	>20 MHz
Jitter	<100 ps_{FWHM} ($\lesssim 40ps_{rms}$)	
Latching	<1 μs	No latching
SNSPD bias Current	<20 μA for NbTiN	with ± 1 μA tolerance
Supply Voltage	1.1 V & 2.5 V	
Power Consumption	<100 μW	As low as possible
Area (for readout electronics)	<1 mm^2	

2

SNSPDs for Diamond-based Color-Center Quantum Computing

This chapter introduces diamond-based defect centers, the principles of entangling two remote qubits, the necessity of single-photon detection, SNSPD, and its readout system, and the system specifications. Defect centers can be created with different vacancies, such as SnV (tin-vacancy), SiV (silicon-vacancy) and NV (nitrogen-vacancy) [16]. This thesis focuses on NV centers as the target application.

2.1. Nitrogen-Vacancy Center Qubits

Diamond-based color centers are especially promising for realizing a large-scale distributed quantum computer due to their long-lived spin quantum states, well-defined optical transitions, and robust and efficient optical interfaces fabricated from diamond crystals [17].

2.1.1. Structure and Energy Level

The NV center is a defect in the diamond lattice consisting of a nearest-neighbor pair of a nitrogen atom which substitutes for the carbon atom and a lattice vacancy, as shown in Figure 2.1a. It has spin degrees of freedom associated with its bound electrons and nearby nuclear spins. There are multiple types of spins in the defect diamond, and this work focuses on the electronic spin. The nitrogen-vacancy center has two charge states, NV^0 and NV^-. The latter, negatively charged NV^0 state [18] is used as a qubit. This extra electron is located at the vacancy site and forms a net spin $S = 1$ pair with one of the vacancy electrons. The corresponding simplified Hamiltonian of the NV^- orbital ground state can be written as [19]:

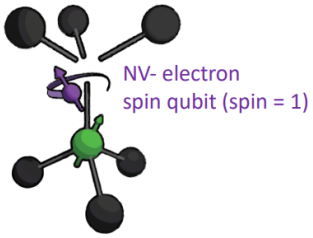
$$H = \Delta S_z^2 + \gamma_e \vec{B} \cdot \vec{S} \tag{2.1}$$

where $\Delta \sim 2.87$ GHz is the zero-field splitting which represents the energy difference between $m_s = \pm 1$ and $m_s = 0$, m_s is the spin energy sublevels; γ_e is the gyromagnetic ratio so that a magnetic field $\vec{B}$ parallel to the NV axis $\vec{S}$ can split the $m_s = \pm 1$ states through the Zeeman effect. Consequently, it produces a well-defined leveled system, and the lowest two levels ($m_s = -1$ & $m_s = 0$) can thus be used as a valid qubit as in Figure 2.1b shows.

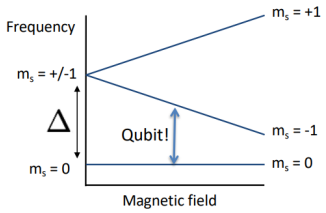
2.1.2. Need for Single Photon Detection

Single-shot Readout

The NV center qubit can be measured with a single-shot readout process. When exposed to resonant excitation, only the electron spin in the state of interest (for example, if we expect the state to be $|0\rangle$, then a resonant excitation for $|0\rangle \rightarrow |1\rangle$ will be applied) will produce fluorescence light and emit photons, as shown in Figure 2.2, which makes the detection of the qubit states relatively simple. This readout technique has exceptionally high fidelity, making it one of the benefits of this diamond-based qubit, and can also be used in qubit initialization.



(a) Simplified physical structure of an NV center.



(b) Energy level structure of the NV center orbital ground state triplet with Zeeman splitting.

Figure 2.1: NV center qubit structure and its energy level [19]. (a) Green: ^{14}N nitrogen atom with the nuclear spin $I = 1$; Purple: one of a vacancy electron paired with an extra electron; Black: ^{12}C carbon atoms with isotopic impurity ^{13}C nuclei $I = \frac{1}{2}$ in the diamond. (b) $|m_s = 0\rangle = |\uparrow\rangle$, $|m_s = -1\rangle = |\downarrow\rangle$, which will be used as a qubit.

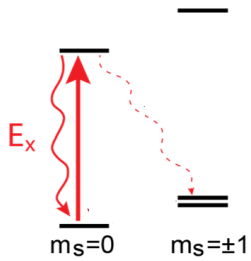


Figure 2.2: Readout scheme of the qubit [20]. Straight line: applied readout laser for $|0\rangle$; wavy line: corresponding fluorescence photon; dashed wavy line: fluorescence to unwanted states.

For the NV center qubit, such a readout scheme emits PSB (Phonon Side-Band) or ZPL (Zero Phonon Line) photons, which are then sent to the SNSPD, resulting in an electrical signal that is further processed by readout circuits. The PSB photons are used for state detection, whereas the ZPL photons are employed to create the entanglement that will be discussed in the next section.

Heralded Entanglement Generation

At cryogenic temperature, the phonon processes are suppressed, revealing the spectrally sharp optical transitions of the NV center [21]. This advantageous optical interface and narrow optical transition spectrums allow the NV centers to generate the entanglement of two distant electron spins. A detailed approach to the heralded entanglement [22], as depicted in Figure 2.3, involves the following steps:

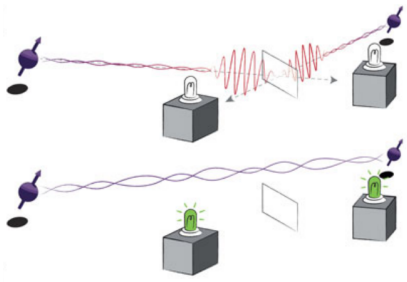


Figure 2.3: Schematic for generating entanglement between distant spins [21]. Direct the locally emitted photons to the BS (Beam Splitter), and the following single photon detection events entangle the two distant qubits.

1. Initialize both qubits to the equally superposed state $\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$.
2. Apply a pulse resonant with the optical transition of $|\uparrow\rangle \rightarrow |e\rangle$. Then spontaneous emission locally entangles the qubit with the photon number, leaving each qubit in the state $\frac{1}{\sqrt{2}}(|1\uparrow\rangle + |0\downarrow\rangle)$. And the

corresponding joint state would be $\frac{1}{2}(|1_A1_B\rangle\,|\uparrow_A\uparrow_B\rangle + |0_A0_B\rangle\,|\downarrow_A\downarrow_B\rangle + |1_A0_B\rangle\,|\uparrow_A\downarrow_B\rangle + |0_A1_B\rangle\,|\downarrow_A\uparrow_B\rangle)$. A and B represent the two NV centers.

3. Then the photons will be diverted to the BS (Beam Splitter) of:

$$\begin{cases} |0_A1_B\rangle \xrightarrow{BS} (i\,|0_C\rangle + |0_D\rangle)(|1_C\rangle + i\,|1_D\rangle) = i\,|01\rangle_C - |0_C1_D\rangle + |1_C0_D\rangle + i\,|01\rangle_D \\ |1_A0_B\rangle \xrightarrow{BS} (i\,|1_C\rangle + |1_D\rangle)(|0_C\rangle + i\,|0_D\rangle) = i\,|01\rangle_C - |1_C0_D\rangle + |0_C1_D\rangle + i\,|01\rangle_D \end{cases} \tag{2.2}$$

and interfere with each other coherently:

$$\begin{cases} |\Psi^-\rangle = |0_A1_B\rangle - |1_A0_B\rangle \xrightarrow{BS} -|0_C1_D\rangle + |1_C0_D\rangle \\ |\Psi^+\rangle = |0_A1_B\rangle + |1_A0_B\rangle \xrightarrow{BS} i\,|0_C1_C\rangle + i\,|0_D1_D\rangle \end{cases} \tag{2.3}$$

where C and D are the output port of the BS. If the photons emitted by the two NV centers are indistinguishable, that is, they are ZPL (Zero Phonon Line), then the detection of a single photon "click" on either output SNSPD port will project the qubits onto the maximally entangled state $|\psi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_A\downarrow_B\rangle \pm e^{-i\phi}\,|\downarrow_A\uparrow_B\rangle\right)$.

2.2. SNSPD

2.2.1. Device Configuration

Superconducting nanowire single-photon detectors consist on an electrically and optically accessible, very narrow ($\simeq$ 100 nm) and ultrathin (4−5 nm) superconducting wire strip [23]. As can be seen in Figure 2.4 (left), one of the SNSPD configurations involves a meandering nanowire that is directly coupled to an optical fiber or through fiber-coupled small gradient index focusing lenses, which are similar in size to the fiber core diameter. However, this type of coupling requires a long and dense meander to compensate for the coupling loss and the low detection efficiency. An alternative configuration (on the

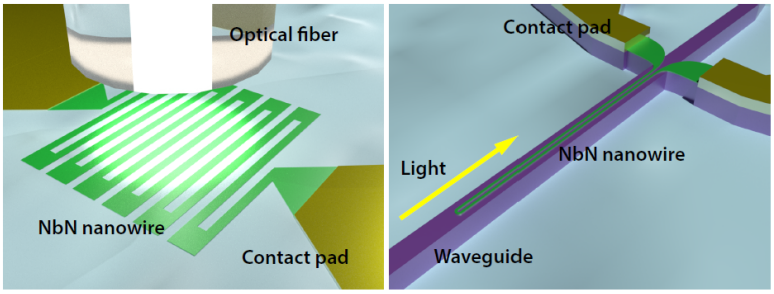


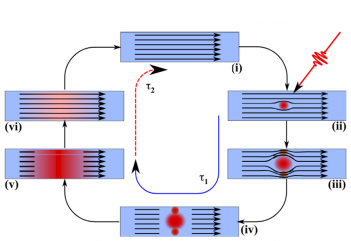
Figure 2.4: SNSPD Configuration [23]. Left: fiber coupled; right: waveguide integrated

right) would be the integrated waveguide : a single nanowire segment is placed on top of a photonic waveguide, so that photons are absorbed in the direction of their propagation. This allows for a drastic improvement in the light interaction length, thus increasing the absorption efficiency. Furthermore, the corresponding shorter wire reduces both geometric and kinetic inductance (A.1), which in turn reduces the response time of the SNSPD, allowing higher count rates.

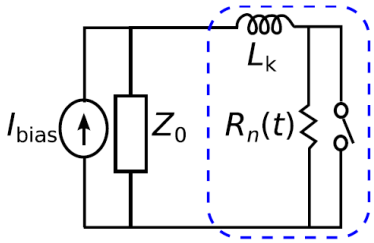
2.2.2. Detection Mechanism

At present, there is no comprehensive model that can accurately explain and forecast all experimental results for the SNSPD detection process to a microscopic extent [23]. This work uses the hotspot theory [24], [25] to explain the formation of the non-superconducting area in SNSPD after a photon is absorbed. The SNSPD is initially kept at a temperature below its superconducting critical temperature T_c and biased with a current I_{bias} that is below its critical current I_c (Figure 2.5a(i)). When a photon hits the

SNSPD, the energy it absorbs is transferred from the photon to the electrons in the nanowire. This leads to an avalanche of quasiparticles (QPs) (i.e., broken Cooper pairs) due to the electron-electron interaction, forming a localized region with a high concentration of excited electrons, known as a hot spot (Figure 2.5a(ii)). The amount of energy transferred and the magnitude of the superconducting energy gap will determine the degree to which the superconducting order parameter in a given area is reduced, resulting in the formation of a "resistive" region. Then this hot spot forces the supercurrent to flow around the resistance area, which leads to an increase in current density beyond the critical current density, thus forming a resistive barrier across the width of the nanowire (Figure 2.5a(iii,iv)). The flow of current through this resistive barrier produces more heat than is dissipated to the substrate, causing the hot spot to grow along the nanowire (Figure 2.5a(v)). Since the hot spot typically has a resistance of kΩ, when connected to a small resistor, $Z_0 = 50\ \Omega$, as illustrated in Figure 2.5b, the current in the SNSPD will be diverted to the low resistance path and results in a voltage across Z_0 that can be used as a measure of the photon-hit event. Finally, as more current is diverted to the readout circuit, the heat generated in the SNSPD decreases, causing the hot spot to shrink (Figure 2.5a(vi)) until it eventually vanishes, allowing SNSPD to return to its original superconducting state and await the next photon (Figure 2.5a(i)).



(a) SNSPD detection cycle.



(b) SNSPD readout circuit schematic. Components in the blue dashed box represents the SNSPD electrical model; I_{bias} is the current used to bias the SNSPD; Z_0 is the readout resistor.

Figure 2.5: SNSPD detection mechanism [24].

2.2.3. Performance and Nonidealities

There are several parameters to judge the performance of an SNSPD and its readout circuit as shown in Figure 2.6:

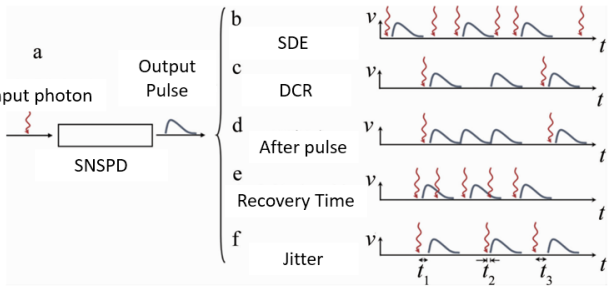


Figure 2.6: SNSPD readout signal specifications and non-idealities, reproduced from [26].

- **SDE (System Detection Efficiency)** is the probablity the SNSPD will trigger when a photon is incident. It can be described by:

$$\eta_{SDE} = \eta_{coupling} \times \eta_{absorption} \times \eta_{intrinsic} \tag{2.4}$$

where $\eta_{coupling}$ represents the probability of an incident photon being coupled to the SNSPD photosensitive surface; $\eta_{absorption}$ is the probability that the nanowire absorbs the photons incident on the photosensitive surface of SNSPD; and $\eta_{intrinsic}$ represents the probability that a photon absorbed by the nanowire generates a voltage pulse. These parameters are mostly related to material and structure. However, intrinsic efficiency $\eta_{intrinsic}$ depends on the magnitude of the bias current I_{bias}, so it is possible to improve the intrinsic efficiency with the help of the design of the external biasing or readout circuit. The basic principle would be to make the trade-off between increasing I_{bias} and leaving the appropriate headroom for its superconducting-to-resistive-state switching current I_{SW}. The current level that determines whether SNSPD will be activated is denoted by I_{SW}. If the total current in the SNSPD due to the incident photon is greater than the switching current, $I_{SNSPD,bias} + I_{photon} > I_{SW}$, then the SNSPD will be activated.

- A **dark count** is an SNSPD triggering without a photon being incident as illustrated in Figure 2.6c. The primary physical causes of dark counts in SNSPDs are vortex-antivortex depairing [27] and blackbody radiation that is transmitted to SNSPDs through optical paths [28] (which can be reduced by optical filtering). The dark count rate generally increases as the bias current increases.
- **After pulses**: after detecting a photon, the SNSPD outputs two or more (N) voltage pulses instead of a single voltage pulse. The second to Nth voltage pulses are called afterpulses. It is noteworthy that afterpulses can also lead to dark counts, although the exact mechanism is not yet fully understood [29]. Therefore, it is clear that a lower I_{bias} (which is a trade-off to achieve a high SDE) can help reduce afterpulses and, consequently, the dark counts rate. The Active Quenching (AQ) technique [29], which will be discussed in the later section, can be used to significantly reduce afterpulses by directly cutting the I_{bias}, thus also reducing dark count rates.
- **Recovery time** is the time required for the SNSPD to recover after a detection event. The response recovery time of SNSPD is defined as the time required for the SDE to recover to 90% of the SDE at low photon flux after completing a photon detection event. As shown in Figure 2.6e, if two or more photons arrive at the SNSPD in quick succession, they cannot be detected effectively. The recovery time of an SNSPD is affected by two factors: the electrical time constant $\tau_e = \frac{L_k}{R_L}$ (which is mainly determined by the kinetic inductance L_k of the device and can be reduced by increasing the load resistance R_L) and the thermal time constant τ_{th} (see Appendix A.3).
- **Latching** is another important phenomenon related to the recovery time that can greatly affect SNSPD performance. Essentially, this phenomenon is caused by rapid negative electrothermal feedback, which is usually used to reset the device [30]. In other words, latching is likely to occur when the ratio of the electrical to the thermal time constant $\frac{\tau_e}{\tau_{th}}$ is too small: When the electrical time constant of the SNSPD ($\tau_e = \frac{L_k}{R_L}$) is less than the thermal time constant τ_{th}, the current returns too fast. This causes the heat produced by the returned current in the SNSPD to equal the heat dissipated to the substrate, resulting in the resistive region being clamped in a stable state. Therefore, the SNSPD now has a stable resistive region, which means that the current diverted to the readout circuit is unable to fully flow back to the nanowire. As a result, the nanowires are unable to continuously detect photons, thus preventing the SNSPD from operating continuously.
- **Jitter** indicates the uncertainty with which the SNSPD measure when a photon is incident. The input signal of an SNSPD is a photon, and the output signal is a voltage pulse. There is a delay between the output and the input, as illustrated in Figure 2.6f. This delay is not fixed, which causes jitter. The detailed expression of the jitter can be found in [31].
  In this work, we are mainly interested in optimizing the electronic jitter $\sigma_{electronic}$. The jitter is given as

$$\sigma_{tot}^2 \approx \sigma_{electronic}^2 + \sigma_{others}^2 \tag{2.5}$$

  where σ_{others} includes the intrinsic jitter of SNSPD σ_{int}, the optical jitter σ_{opt}, etc. The jitter requirement indicated in Table 1.1 is of $j_{jitter} < 100$ ps in terms of the FWHM (Full Width at Half Maximum). As the jitter of most SNSPDs follows the Gaussian distribution [32], the root mean square (rms) jitter can be approximated by [23] $j_{FWHM} \approx 2\sqrt{2ln(2)}\sigma_{j,rms} \Rightarrow \sigma_{j,rms} \approx 42.46$ ps.

Assuming that the electronics jitter dominates the system jitter contribution, the jitter for the readout circuits in this work can be derived by

$$\begin{aligned} \sigma_{j,electronic,rms} &\approx \sqrt{\sigma_{j,rms}^2 - \sigma_{j,SNSPD,rms}^2} \\ &= \sqrt{42.46p^2 - 15p^2} \approx 40ps \end{aligned} \tag{2.6}$$

where $\sigma_{j,SNSPD,rms}$ is taken from the typical value of an SNSPD. To allow some headroom, in this work the rms jitter value is reasonable to set as $\sigma_{electronic} \leq 35$ ps.

- The typical value of the specifications that an commercial SNSPD can achieve is shown in the Table 2.1

Table 2.1: Typical SNSPD (Table top detector provided by Pixel Photonics) Sepcifications [33]

Specification	Typical Value
Operating Wavelength	Visible to IR spectrum
Operating Temperature	2.5 ~ 3 K
System Detection Efficiency	≥ 70%
Dark Count Rate	<100 Hz
Count Rate	≤ 60 MHz
Jitter	<50 ps

2.2.4. SNSPD Readout Technique Comparison

Passive Quenching

The traditional method of reading the signal produced by an SNSPD is known as passive quenching. This was previously introduced in Section 2.2.2. As a reminder, when a sufficient amount of current is diverted to the low-resistance readout path, the SNSPD is quenched due to the heat produced by the current being less than the heat dissipated to the substrate. The passive quenching method for SNSPD readout has a trade-off between latching, signal level, the gain requirement on the following amplifier, and eventually, the jitter performance. As is demonstrated in Figure 2.7, a small readout resistor R_L can

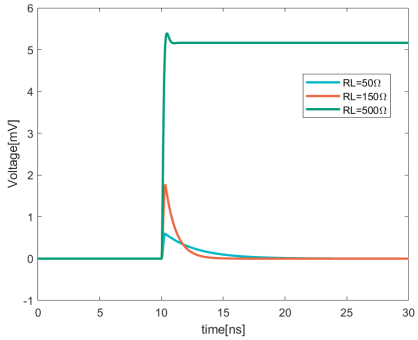


Figure 2.7: Impact of the load resistor R_L in passive quenching. As resistance increases, the slew rate on the rising edge increases, and the recovery time decreases. However, if R_L is too large, the system will latch (green curve).

reduce the risk of latching. However, the signal amplitude of the SNSPD is restricted by $I_{bias} \times R_L$, which requires stricter requirements on the gain of the subsequent amplifier. This results in a decrease in SNR due to the lower signal level and extra noise from the following circuits. Furthermore, the maximum count rate decreases due to the larger $\tau_e = \frac{L_k}{R_L}$, which is inversely proportional to the load resistance. Another problem comes from the impedance matching since the output signal will be transimitted

through the 50 Ω cables to the room temperature readout electronics. Therefore, changing R_L will also distort the output signal, as illustrated in Figure 2.8.

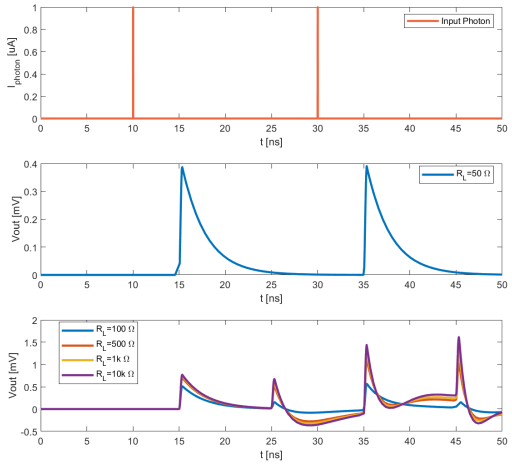
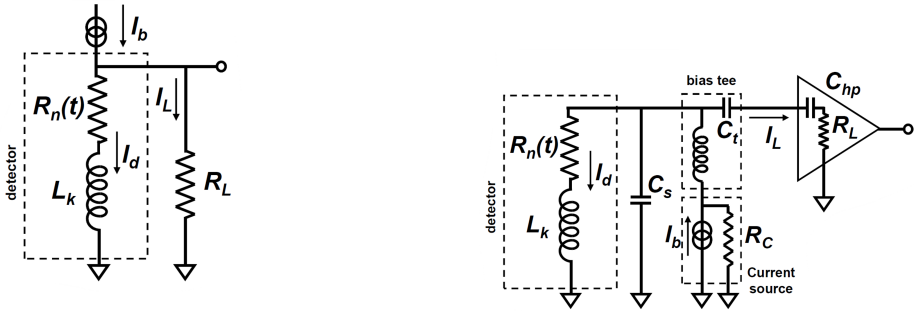


Figure 2.8: Impact of different load resistors R_L when using 50 Ω cables in tradition room temperature readout. Signals are distorted as can be seen from the bottom plots.

The volatge signal across the load resistor which is in parallel with the SNSPD will be further processed by the following readout circuits such as an amplifier. The amplifier can be either DC or AC coupled, as demonstrated in Figure 2.9. However, in practice, the load seen by the detector is not just a resistive



(a) Typical SNSPD readout using a low-resistance path R_L to divert the bias current I_b. The amplifier will be DC coupled.

(b) AC coupled amplifier with a bia tee. C_s indicates the parasitic capacitance; C_{hp} is the input capacitance of the amplifier.

Figure 2.9: SNSPD readout circuit schematic [14]. The dashed box models the SNSPD where $R_n(t)$ is the time-dependent hot spot resistance; L_k is the kinectic inductance of nanowire.

element that is independent of the frequency as demonstrated in Figure 2.9a. This results in a nonlinear feedback that links the average count rate to the detection efficiency. In the cases of extremely high count rates, this feedback can even cause the detector to enter the latched state [14]. The count rate of both schemes will be restricted by the 50 Ω load resistor, as the recovery time $\tau_{falling} \sim \frac{L_k(t)}{R_L}$ is larger. Furthermore, the AC coupled amplifier will limit the system count rate from charging and discharging the capacitor.

Figure 2.10 shows the improved SNSPD readout electronics. The gate electrode of a commercial high-electron-mobility transistor (HEMT) has a wideband high impedance, which is shunted in parallel with the DC readout resistor R_L. The HEMT is used in the typical common source setup, and its high gain and broad bandwidth effectively separate the SNSPD from the readout electronics. To provide extra gain, a second cascaded HEMT, common-source stage is used. The isolation provided by the preamplifier allows R_L to be different from the usual 50 Ω for impedance matching. However, the

improvement in this readout resistance is still limited, as the SNSPD will eventually enter the latched state as R_L increases. Another readout shceme is the capacitor grounded readout [34] to reduce the

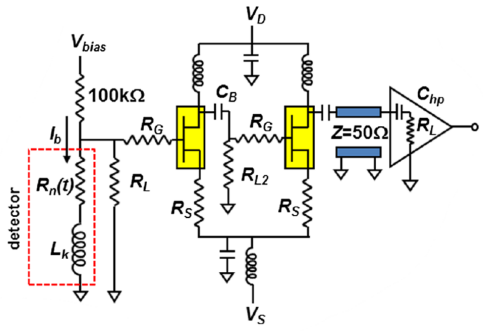
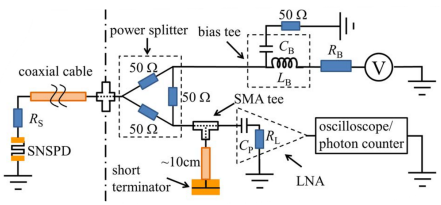
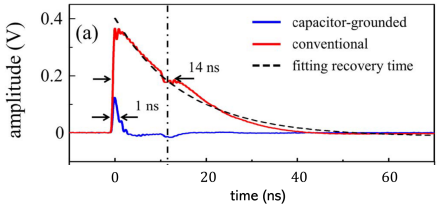


Figure 2.10: Improved SNSPD DC readout with a preamplifier to removes the AC-coupled feedback [14].

current reset time and remove the coupling effect that limits count rates, as illustrated in Figure 2.11a. A resistor R_s connected in series with the detector is utilized to reduce the recovery time and thus enhance the maximum count rate. The input of the low noise amplifier (LNA) is paralleled with a short terminated coaxial cable through an bias tee which will equally split the SNSPD output into two pulses. One goes to the LNA directly while the other one is inverted in amplitude at the short termiator of the coaxial cable, and then arrives at the LNA with a time delay of ~ 1 ns. Therefore, the overlapped input pulses of the LNA are cancelled out, resulting in much narrorwer ones, which charge on the LNA input capacitor with less energy. Also, the rising edges of the pulses are still maintained where the recovery time is decreased, as shown in Figure 2.11b.



(a) Schematic of the capacitor grounded readout



(b) Waveforms acquired by an oscilloscope

Figure 2.11: Capacitor grounded radout sheme [34].

However, in each of this cases the output voltage is still limited to the load resistance and the bias current. And such load resistance has its maximum values restricted by the latching. And it is not preferable to solve these issues by decreasing the bias current because it will affect the detection efficiency. Therefore the readout schemes discussed above will limit the system performance from having high count rates and timing jitter which is due to the associated relative low signal to noise ratio.

Active Quenching

In order to have a large readout resistor during the signal readout phase, and one that is small enough to avoid latching yet still large enough to quickly reset the SNSPD during the resetting phase, the active quenching technique [29] is utilized. As shown in Figure 2.12, a comparator is located at the SNSPD output. If the voltage of the signal v_{DET} exceeds the comparator threshold, the SNSPD bias $i_B(t)$ is deactivated, forcing SNSPD to return to the superconducting state more rapidly. The comparator then rebiases the SNSPD after a time period of τ, preparing it for the arrival of the next photon [29]. The deactivation of bias current during the quenching phase can significantly reduce afterpulses and

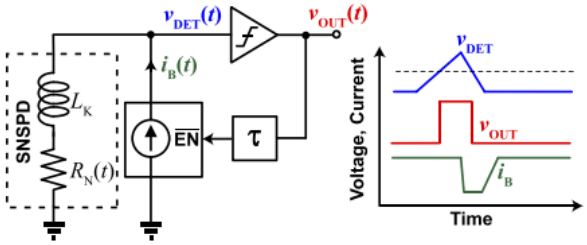


Figure 2.12: Conceptual block diagram of active quenching architecture and key waveforms describing its operation [29].

associated dark counts. Additionally, a larger R_L increases count rates due to a shorter recovery time on the falling edge and improves jitter performance due to a higher slew rate on the rising edge.
Therefore, in this work, a cryo-CMOS readout circuit employing the active quenching (AQ) is designed to make a compact SNSPD readout system that can work at cryogenic tempertature, while the performance of count rate, dark count rate, and jitter of the SNSPD can be simultaneously enhanced.

In conclusion, this chapter first provides a fundamental overview of NV center quantum computing, including the state measurement and entanglement establishment, which necessitate single-photon detection. Subsequently, the SNSPD detection mechanism was explored, followed by the presentation of the performance and imperfections of SNSPD. Finally, various state-of-the-art readout techniques were discussed. The following chapters will start introducing the circuit design process with the SNSPD SPICE model in Chapter 3.

3
SNSPD Model

It is essential to construct a model that accurately portrays the SNSPD photon detection dynamics for the purpose of designing the readout circuit in the subsequent stages. This chapter first explains the model used to simulate the evolution of the SNSPD hotspot. Then, this SNSPD SPICE model will be implemented and simulated in Cadence Spectre, and its performance will be evaluated by comparing it with the laboratory data.

3.1. SNSPD SPICE Dynamic Model

SPICE does not include superconducting nanowires in its component list, making it difficult to model complicated superconducting nanowire detector structures quickly and easily. Karl K Berggren et al [35] present three circuit models to explain the expansion and annihilation process of hotspots:

- The **curve-fitting model** replicates the current-voltage (i-v) relationship of the nanowire, as well as the transition into and out of the superconducting state. However, this model does not account for the physical description of the nanowire dynamics and is therefore too ideal for electrical circuit design consideration in this work. As demonstrated in Figure 3.1, ① is the superconducting state, ② is the hotspot state, and ③ is the state in which the entire nanowire becomes resistive. This curve-fitting model does not take into account the electrothermal dynamics of the hotspot region, which means that it cannot accurately interact with the external readout circuit. Thus, it is not suitable for the purpose of readout circuit design in this work.

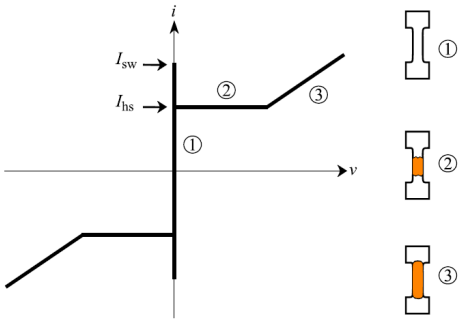


Figure 3.1: The i-v curve used in curve-fitting model [35].

- The **dynamic model** includes the electrothermal dynamics of the nanowire, which is necessary when the dynamics of the nanowire interacts with the dynamics of the external circuit. Therefore, the rise and fall times (i.e. $\tau_{rise} = \frac{L_k(t)}{R_n(t)}$ and $\tau_{fall} = \frac{L_k(t)}{R_L}$), the recovery and dead times, latching, and jitter performance will be more accurately modeled than with the curve-fitting model. Therefore, this model will be sufficient for the purpose of this work.
- The **thermal device model** is the most comprehensive model that takes into account the effects of thermal relaxation in the superconducting state, allowing for a qualitative representation of

the afterpulses and related dark count. However, it is highly sensitive to the selection of model parameters and requires further validation with actual experimental data.

Due to its relatively complete description of electrothermal dynamics and realiability, the dynamic model is used in this work for modeling the SNSPD in SPICE. Figure 3.3 shows a schematic of the dynamic device model that uses nanowire electrothermal physics to model the evolution of the hot-spot resistance over time. The following is the explanation of the components used:

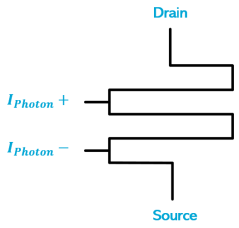


Figure 3.2: SNSPD symbol with terminals

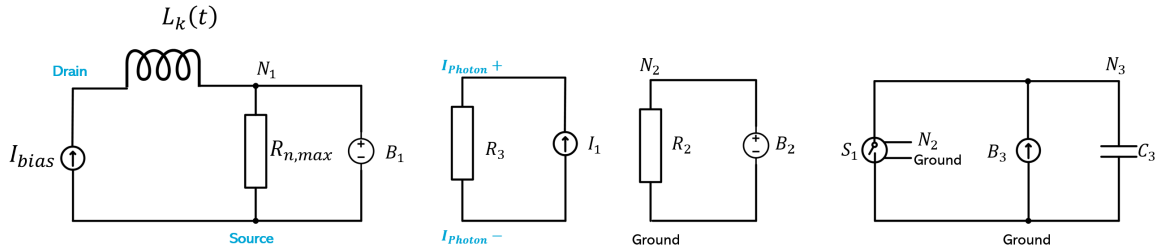


Figure 3.3: SPICE schematic of SNSPD dynamic model with terminals emphasised in blue. The I_{bias} is the bais current source from external circuit which does not belong to the SPICE model.

- The kinetic inductance of SNSPD $\boldsymbol{L_k(t)}$ is current dependent (which is eventually time dependent), and the accurate expression is still unclear. Although the expression can be obtained by fitting the data from [36], it is still not very reliable since the performance is heavily dependent on the property of the SNSPD itself. Also, as shown in Figure 3.4, setting this inductance as a constant value shows little difference. Therefore, in this work, it is enough to take $L_k(t)$ as a constant (i.e., $L_k(t) \approx L_{k0} = 130$ nH).
- $\boldsymbol{R_{n,max}}$ is the maximum resistance of the SNSPD hot spot that can be achieved if the entire nanowire is converted to a resistive state. This is used to restrict the unbounded growth of the hot spot. However, this resistance is usually not reached under normal operating conditions.
- **B1** represents the growth or shrinkage of the resistance of the hot-spot.

$$V(B1) = \frac{V(N3) + |V(N3)|}{2} \times I(B1) \tag{3.1}$$

The value of $V(N3)$ does not physically correspond to the voltage, but is instead numerically equivalent to the resistance of the hotspot (will be explained in the following items). The absolute function of the expression guarantees a real-number value.

- The current source ***I*1** and resistor ***R*3** are employed to mimic the photon hit event. The current generated by the photon, I_{photon}, is produced by I1. R3 is an extremely samll resistor to sense the current that I1 provides. In Cadence Spectre, R3 can be replaced by the current probe device, such as the iprobe. Therefore, the photon-induced current in SNSPD can be represented by

$I_{photon} = I(R3)$. This subcircuit mainly provides an interface terminal $I_{photon+}$ between SNSPD and photon, which can be considered as the input of the SNSPD.
If the current in the SNSPD exceeds the switching current, $I_{SW} - g \times I_{photon}$, which will be set as a judging condition at the behavior source B2, a hotspot will be created. The gain factor, g, is used to guarantee that the photon can activate the SNSPD, but its value can be adjusted based on actual experiments to take into account the effects related to the photon energy.

- The SNSPD state is monitored and judged by **B2** & **R2**. If the current generated by the photon surpasses the switching current, or the voltage across the hot-spot region is higher than the minimum voltage needed to keep the hot-spot, the SNSPD will switch to the resistor-like state, and the value of B2 will be 1. If not, the value of B1 will be 0 and the SNSPD will remain in the superconducting state.
- The subcircuit of hot-spot evolution, consisting of **B3, C1**, and **S1**, is used to calculate the total device resistance by integrating $\frac{\partial R_{HS}}{\partial t}$ as shown in equation (3.2). The operation of switch S1 is depend on the value of $V(B2)$. If S1 is closed, then the current in B3 will be directed to the ground, thus preventing any integration on C3. Otherwise, a capacitor C3 with a value of $\frac{w}{2R_{SH}v_0}$ is employed to integrate the B3 current $f(i) = \frac{\psi \frac{i_D^2}{I_{SW}^2} - 2}{\sqrt{\psi \frac{i_D^2}{I_{SW}^2} - 1}}$ to simulate such current-dependent behavior of the hot-spot dynamics in equation (3.3). Consequently, the voltage across C3 is numerically equal to the instantaneous hot-spot resistance R_{HS} and will be used in (3.1).

$$\begin{cases} \dfrac{\partial R_{HS}}{\partial t} = \dfrac{R_{SH}}{w}\dfrac{\partial l}{\partial t} & (3.2) \\ \dfrac{dl}{dt} = v_{HS} = 2v_0 \dfrac{\psi\left(\frac{i_D}{I_{SW}}\right)^2 - 2}{\sqrt{\psi\left(\frac{i_D}{I_{SW}}\right)^2 - 1}} & (3.3) \end{cases}$$

where i_D is the current flowing in SNSPD kinetic inductance; $v_0 = \frac{\sqrt{h_c \kappa / d}}{c}$ is the characteristic normal domain velocity; h_c, κ, c, T_d and ρ are the parameters related to the SNSPD material; T_S is the bath temperature of the SNSPD substrate; $\psi = \frac{\rho I_{SW}^2}{h_c w^2 d(T_d - T_S)}$ is the Stekly parameter which measures the relative role of Joule self-heating in superconductors [37]; R_{SH} is the sheet resistance of the resistive state in SNSPD; w is the width of the nanowire; l is the length of the hot spot.
If B2 gives an output of 1, the switch S1 will close, meaning that the SNSPD will return to its superconducting state, thus all the current will be directed to the ground, causing $V(N3)$ to be equal to zero, and the resistive region will consequently vanish.

- The two subcircuits with I1 and B2 are not directly connected to the rest of the circuits. The I1 subcircuit offers an SNSPD input terminal that simulates the interface between an incoming photon and an SNSPD, while the B2 subcircuit is an affiliated circuit that monitors the SNSPD's state to give the operational logic to S1 to regulate the development of the hotspot resistance determined on C3.

3.2. SNSPD SPICE Model Performance

Since the circuit design in this work will be carried out in Cadence Spectre, it is essential in this section to compare the performance of the SNSPD dynamic model reproduced in Cadence Spectre against the reference SPICE model provided in [35]. The model has been implemented in Cadence Spectre (the spectreText code can be found in the Appendix B), but some parameters such as h_c, κ, etc., are unknown and would require laboratory measurements to determine. Alternatively, we can fit to measurement data that are also used in [35]. The performance of the reproduced SNSPD model is demonstrated in

Figure 3.4. There are two consecutive photon hit events at $t = 10ns$ and $t = 30ns$, and both Cadence Spectre (used in this work) and LTSpice (used by the [35]) show the expected SNSPD output response. The main difference is due to the performance of the different between Cadence Spectre and LTSpice, and the selection of the kinetic inductance $L_k(t)$ where Cadence simulates with constant kinetic inductance, and LTSpice simulates with the varying L_k. However, this difference is not important for our use case since the slight discrepancy occurs during the falling edge, while this work mainly concentrates on the improvements of the rising edge, where the models match better. Moreover, the behavior of the falling edge will be different after the active quenching is implemented, which will be discussed in detail in Chapter 4. The circuit used for this simulation is the same as the one depicted in Figure 1.2. The dashed

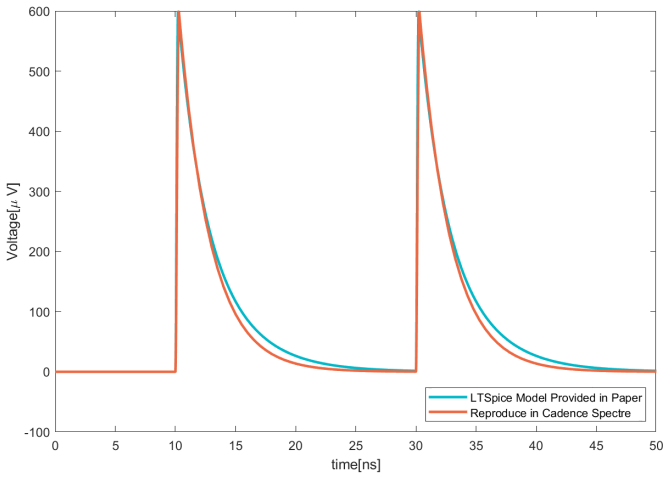


Figure 3.4: Comparison between the model presented in [35] based on laboratory data and the reproduced model simulated in Cadence Spectre.

box representing SNSPD is replaced by the SNSPD model introduced above, which is connected in parallel with a $R_L = 50\ \Omega$ load resistor and biased with $I_{bias} = 15\ \mu$A. The parameters associated with the SNSPD model, which is related to the SNSPD material, can be found in Appendix B.

Comparison with the Laboratory Measurement Data

To further test the performance of the SNSPD model, a comparison between the simulation in Cadence Spectre and the real laboratory measurement data that is provided by *Pixel Photonics* [38] can be found in Figure 3.5. The simulation is still based on the same schematic in Figure 1.2 that the SNSPD is parallel connected to the load resistor. With $R_L = 50\ \Omega$ and the bias current of 10 μA, the SNSPD can achieve the maximum output signal level of around 370 mV, rising edge slew rate of $\sim$ 150 MV/s, and the recovery time of $\sim$ 30 ns. Although the exact material parameters of SNSPD, such as device dimension, thermal conductivity κ, and heat capacity h_c, are still not known, the model can be tuned to closely match the measurement data due to the flexibility of these parameters in the model. By comparing the results, it can be seen that the main discrepancy is at the falling edge, where the time constant of $\tau_{fall} \sim \frac{L_k(t)}{R_L}$ is smaller than in the real device. This trend is similar to the one shown in Figure 3.4. Therefore, it is reasonable to deduce that the difference is mainly due to the absence of the time-varying characteristic of L_k in the model.

3.3. Model Limitations

There are several remarks about this dynamic model:

1. I_{SW} is employed rather than using the depairing critical current I_c to describe the current at which the device can no longer support a supercurrent. This is to stress that in practice, the depairing

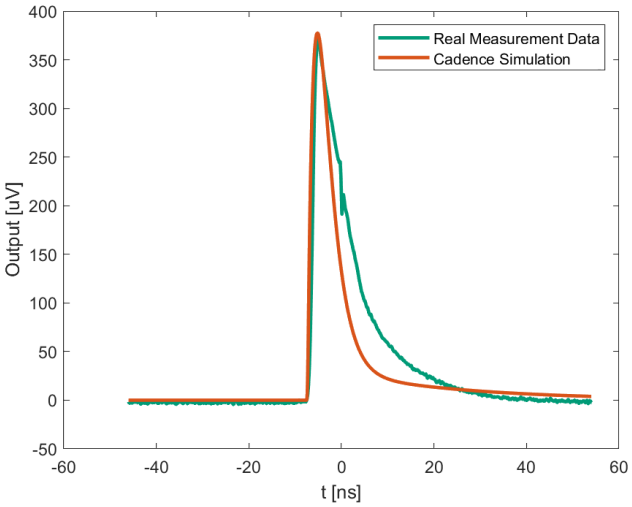


Figure 3.5: The comparison between the SNSPD model build in Cadence Spectre and real measurement data. Orange: simulation in Cadence Spectre; green: laboratory measurement data.

current is not the accurate factor that limits these devices and that vortex crossing, noise, and other elements may also contribute to the suppression of I_c.

2. The SNSPD SPICE model used in this work does not consider the physics of jitter, nor does it take into account the microscopic physics of hot-spot creation and the post-photon-arrival signal propagation delay in the nanowire. Additionally, the current I_{photon} used to simulate the photon-incident event may generate noise which could lead to false or missed counts, but this is not taken into consideration in the model.
3. The temperature dependence of the hot-spot resistance and the thermal parameters, which are not included in this model, can lead to afterpulses. However, the active quenching technique (see Section 2.2.4) should significantly suppress afterpulses.

In conclusion, the SNSPD model employed in Cadence Spectre is comparable to the LTSpice model presented in [35], with the distinction due to the constant L_k being negligible. Furthermore, the model encompasses a range of parameters related to the SNSPD's dimension and material properties, making it relatively simple to adjust to match experimental results. However, the model does not precisely replicate the electromagnetic transmission delay of the signal (for jitter performance), nor does it completely model the SNSPD kinetic inductance and thermoelectric characteristics of the hotspot, which may result in overlooking some nonidealities such as afterpulses. Additionally, future development should also take into account the effect from the photon side, including the statistical model of the photon and the noise in the I_{photon}.

4

SNSPD Cryogenic Readout Circuit Design

This chapter will explain the detailed design process of the SNSPD cryogenic CMOS readout circuit. First, the development of the system architecture is introduced, which takes into account the quenching approach, the pad capacitance C_{pad}, the bondwire inductance L_{wire}, the load resistance (readout resistance) R_L, and the switch resistance R_{SW}. Then both the implementation of analog and digital subcircuits are discussed. An additional focus here is on the transistor parameters choice of the main amplifier and comparator with the derivation on their optimal transconductance. Finally, the performance of the system will be simulated and inspected.

4.1. System Overview

The overview schematic of the readout system is shown in Figure 4.1. The SNSPD chip and the readout chip will be placed near each other and connected through the bond wires $L_{bondwire}$. For the integration of diamond spin qubits with photonic circuits and CMOS electronics for routing, control, and readout of qubits on a large scale, these chips are expected to be 3D integrated in the future [39].
The biasing for SNSPD is indicated by the current source $I_{SNSNPD,bias}$, and the implementation at the transistor level can be seen in Section 4.2.1. When SNSPD is triggered, the voltage across the readout resistor R_L will be amplified by the main amplifier and the count decision will be made by the comparator (see Section 4.2.2). The count signal will be sent to the digital circuit (as seen in Section 4.3) and used to enable the active quenching process. The feedback amplifier (discussed in Section 4.2.6) is used to maintain the correct operating point of the amplifier. Eventually, the SNSPD is quenched by the active quenching switch (AQ SW) and returns to its superconducting state and waits for the next photon. The following sections will discuss the impact of some of the components such as pad capacitance, readout and active quenching switch resistance, etc.

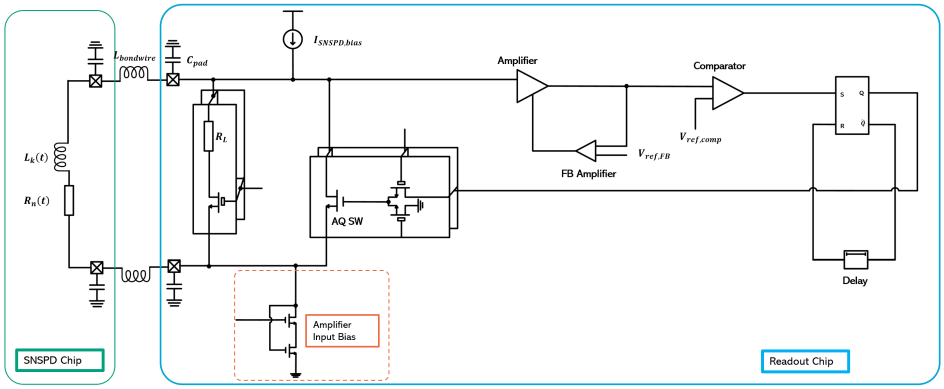


Figure 4.1: Readout system schematic including SNSPD and readout analog and digital subcircuits.

4.1.1. Architecture Decision

When it comes to implementing active quenching in practice, there are a few things to consider, which will be discussed in the following sections:

Impact of Bias Disabling Methods

The proposed scheme to disable the SNSPD bias in [29] is to abruptly break the bias path. An alternative approach (used in this work) is to short the SNSPD and divert the bias current to ground as shown in the Figure 4.2. Instead of breaking the bias path, most of the bias current, I_{Bias}, will be diverted to the

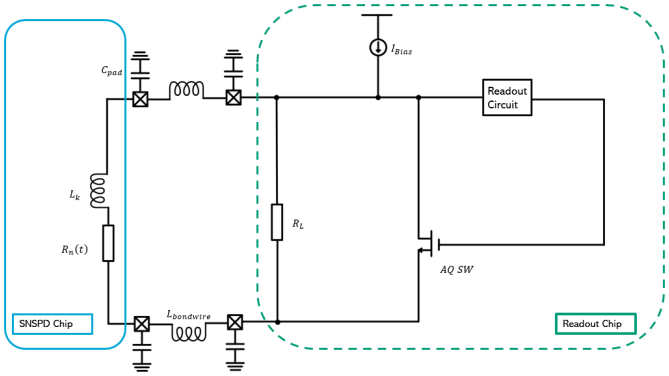


Figure 4.2: Active quenching with short-bias scheme. Blue box: SNSPD chip; green dashed line box: cryogenic readout circuit. The pads and bondwires between these two chips are represented by C_{pad} and $L_{bondwire}$. R_L is the readout load resistor; AQ SW indicates the active quenching switch implemented by the NMOS. When the readout circuit successfully detects a photon signal, the active quenching switch (AQ SW) will be closed, shortening the bias current to the ground.

ground via the active quenching switch (AQ SW) when the readout circuit detects a photon signal. The comparison between the open and short bias scheme is shown in Figure 4.3, the top row displays the SNSPD output voltage for the short-scheme (blue curves) and open-scheme (orange curves). The middle row shows the current in SNSPD, and the bottom row shows the active quenching signal. When the voltage across the readout resistor exceeds the threshold, the comparator gives 1.1 V, and the SNSPD is correspondingly quenched. However, the open-bias scheme has a greater overshoot when the SNSPD is rebiased, which can lead to the voltage exceeding the comparator's threshold, resulting in an incorrect count. This is because when the short scheme is used, $R_n(t)$ will be low at some point and the current will start flowing toward the SNSPD, whereas with the current switch scheme, all current is removed and then is instantaneously brought back to the SNSPD, as indicated in the SNSPD current plot during a time period of 5 ~ 10 ns. Therefore, the short-scheme is preferably used in this work instead.

Impact of Pads' Capacitance and Bondwires' Inductance

As can be seen in Figure 4.3, active quenching is beneficial in aiding SNSPD to return to the superconducting state; however, the oscillation and related overshoot can be a major problem, leading to false counts. Therefore, it is essential to investigate the effect of the capacitance of the SNSPD and the readout chip pad C_{pad} and the inductance of the bondwire $L_{bondwire}$ (indicated in Figure 4.2), which are likely to be the main cause of such oscillations. As demonstrated in Figures 4.4 and 4.5, a larger pad capacitance has a greater impact on the readout response. As C_{pad} increases, the slew rate, which can be estimated by $SR \approx \frac{I_{SNSPD}}{C_{pad}}$, will decrease, resulting in a decrease in SNR and a deterioration in the jitter performance. When C_{pad} approaches 1 pF, the system does not function properly due to the large overshoot in the SNSPD current I_{SNSPD}. This is because when the overshoot is large enough, the current in the SNSPD can be below the latching current, causing the SNSPD to return to 0 Ω instead of being latched. This results in all the current flowing back to the SNSPD, which in turn causes the voltage across the load resistor to be too small to activate the comparator, as shown by the purple curves in Figure 4.4, while the effect of the bondwire is insignificant for the readout performance. In practice, a capacitance of approximately $C_{pad} \sim 200$ fF and an inductance of around $L_{bondwire} \sim 5$ nH are thought to be reasonable

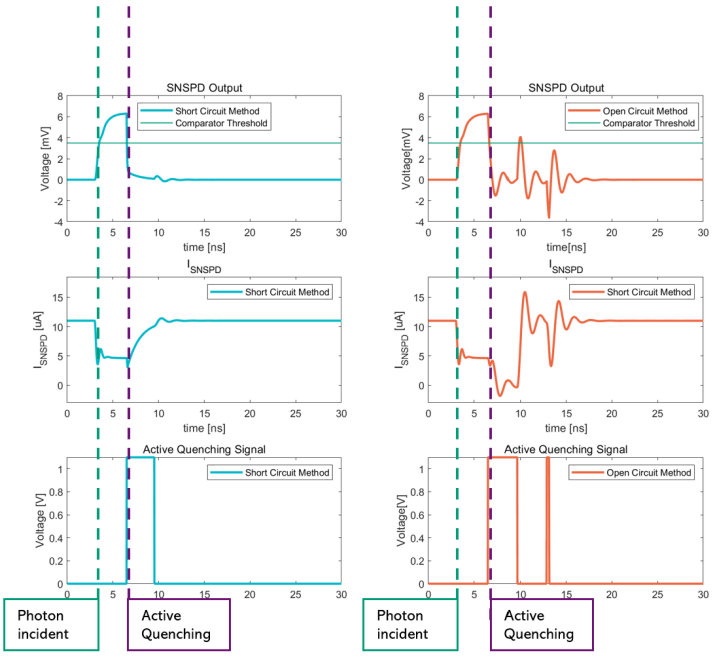


Figure 4.3: Simulation results demonstrate that the open-bias quenching technique is less preferable for deactivating the SNSPD bias. The blue curves are the SNSPD output voltage, the current in the SNSPD, and the active quenching signal (from top to bottom) for the short-scheme and red curves represents the open-scheme; green curve is the comparator threshold. The photon hits $t = 3ns$ and both schemes give a count signal. However, due to the more intense oscillation, the open-scheme is prone to produce other false counts (at $t \approx 13$ ns in the bottom right plot).

approximations, which have also been demonstrated to be safe values based on these simulation results. Consequently, in the following simulations, we will use $C_{pad} = 125$ fF and $L_{bondwire} = 2$ nH.

Impact of the Load Resistance

By varying the load resistance in the passive quenching scheme, we can determine the resistance at which latching occurs. In our SNSPD model ($L_k(t) = L_k = 130$ nH, $I_{bias} = 15\ \mu$A, $I_{sw} = 18.5\ \mu$A), this value is estimated to be $R_{latch} \approx 200\ \Omega$, as shown in Figure 4.6. Here, I_{SW} denotes the threshold current that when the current in SNSPD is greater than this value, the SNSPD will be triggered. However, the amplifier requirement and the jitter performance of the system are not optimal when using passive quenching with $R_L \approx R_{latch} = 200\ \Omega$. Although active quenching allows for a large load resistance, one should take into account the noise that the resistor can introduce, which can have a detrimental effect on jitter. However, when at cryogenic temperatures, the thermal noise contribution can be significantly reduced. In principle, it is preferable to make the jitter as low as possible, which is beneficial for both the system performance and also the possible design headroom. Jitter can be defined as:

$$\sigma_{jitter} = \frac{\delta V_{noise}}{SR} \tag{4.1}$$

where δV_{noise} is the rms noise amplitude noise of the electronic circuit; $SR = \frac{\partial V}{\partial t}|_{V=V_{decision}}$. This means that the slew rate (SR) should be as high as possible. As shown in Figure 4.7 (passive quenching is used for the purpose of providing a clearer view of how R_L affects the output of the SNSPD), if the comparator threshold voltage is assumed to be $V_{decision} = 3.5mV$, the SR at this point will nearly reach saturation after $R_L \gtrsim 10k\Omega$, which implies that it is possible for choosing R_L as high as possible, but one should be aware of the increasing in noise and potential oscillations it may cause when R_L is relatively large, while the improvements on SR are limited. The possible explanation would be that since R_L is extremely large while R_n experiencing a growth starting from 0, there will be a very small portion of

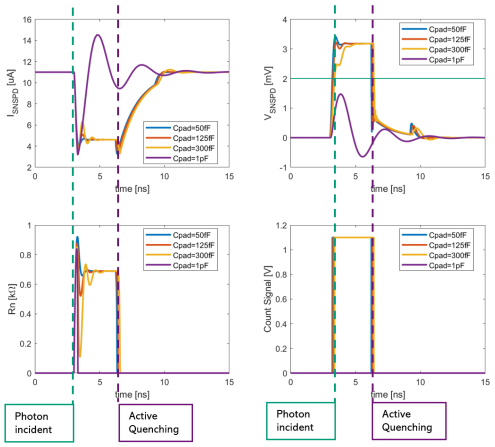


Figure 4.4: Impact of the pad capacitance C_{pad} on the active quenching performance. Upper left: Current in SNSPD; upper right: Output voltage across the readout resistor $R_L = 500\Omega$; bottom left: hotspot resistance of SNSPD; bottom right: count signal.

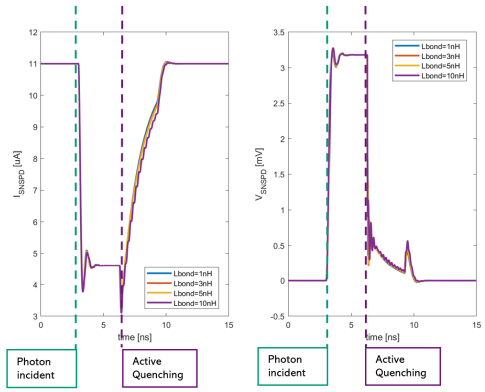


Figure 4.5: Impact of the bond wire inductance $L_{bondwire}$ on the active quenching performance. Left: Current in SNSPD; right: Output voltage across the readout resistor $R_L = 500\Omega$.

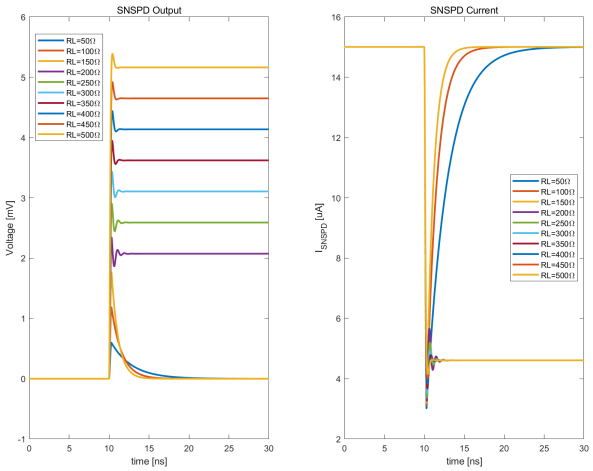


Figure 4.6: By sweeping the load resistance R_L, the system was observed to latch when $R_L \geq 200\ \Omega$.

current that can flow into the readout path. However, such a small current will still generate a 3.5 mV voltage on this large R_L, and the time to achieve this voltage value becomes unfigurable when R_L is large, so that the SR saturates based on $SR = \frac{3.5mV}{t|_{V=3.5mV}}$.

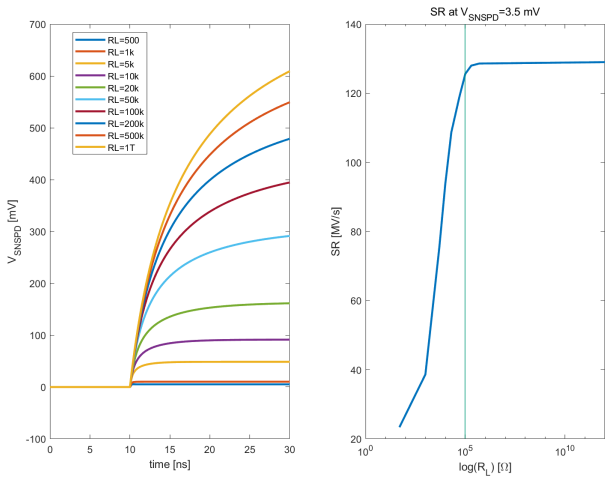


Figure 4.7: Impact of the readout resistor R_L on SNSPD readout performance. Left: Output voltage across different R_L. In each case, the SNSPD latches. Right: Slew rate at $V_{decision}$ = 3.5mV with respect to R_L (in log scale), the green line indicates the value of R_L at which the SR will start saturating.

Impact of the Active Quenching Switch Resistance

The resistance of the active quenching switch R_{SW} is also a significant factor in the readout process. As demonstrated in Figure 4.8, simulations were conducted with three different closed-switch resistances: $R_{SW} = 10\ \Omega, 250\ \Omega, 600\ \Omega$. At $t = 3$ ns, a photon is absorbed by the SNSPD, causing it to enter the normal resistive mode and then the current I_{SNSPD} flows into the readout resistor R_L, resulting in a sharply increasing voltage signal. When the voltage reaches the comparator threshold, the active quenching switch in parallel with the SNSPD and its load resistor will be activated after a delay of 3 ns. This will cause most of the current in the load resistor and SNSPD to be diverted to the switch path and eventually to the ground, resulting in the SNSPD returning to its superconducting state ($R_n(t) \rightarrow 0\ \Omega$). The significance of the closed switch resistance R_{sw} will be evident at this point. Since $R_n(t) \sim 0\ \Omega$, there is a potential difference between the switch path and the SNSPD. The larger R_{sw} is, the greater the potential difference, thus the quicker the current that was just diverted into the switch path will flow back to the SNSPD.

As demonstrated in the first row of Figure 4.8, the transient SNSPD current waveform with different closed switch resistances reveals that during the time period $t \in [6\text{ns}, 9\text{ns}]$, when the current in the switch path returns to the SNSPD due to the potential difference, the speed of this current return in the $R_{SW} = 10\ \Omega$ case is much slower than in the $R_{SW} = 250\ \Omega$ case, leading to a more intense overshoot at the rebiasing operation ($t = 9$ ns, the switch opens). As a result, a false count is generated, which can be found in the corresponding blue curve count signal plot at left bottom corner. This flowback process is quite similar to the resetting stage in the passive quenching system. As R_{SW} increases, the rate of current flowing back from the switch path increases, eventually leading to a situation where the heat generated by the current in the SNSPD is equal to the heat dissipated to the substrate. This causes the SNSPD to remain in a resistive state, which means that the device is latched (as seen in $R_{SW} = 600\ \Omega$ in Figure 4.8), making the active quenching proposed previously ineffective. Therefore, the requirement on R_{SW} should be $R_{sw} \parallel R_L < R_{latch}$ to prevent latching. Additionally, $R_{sw} \parallel R_L$ should not be too small due to the risk of false counts caused by overshoot.

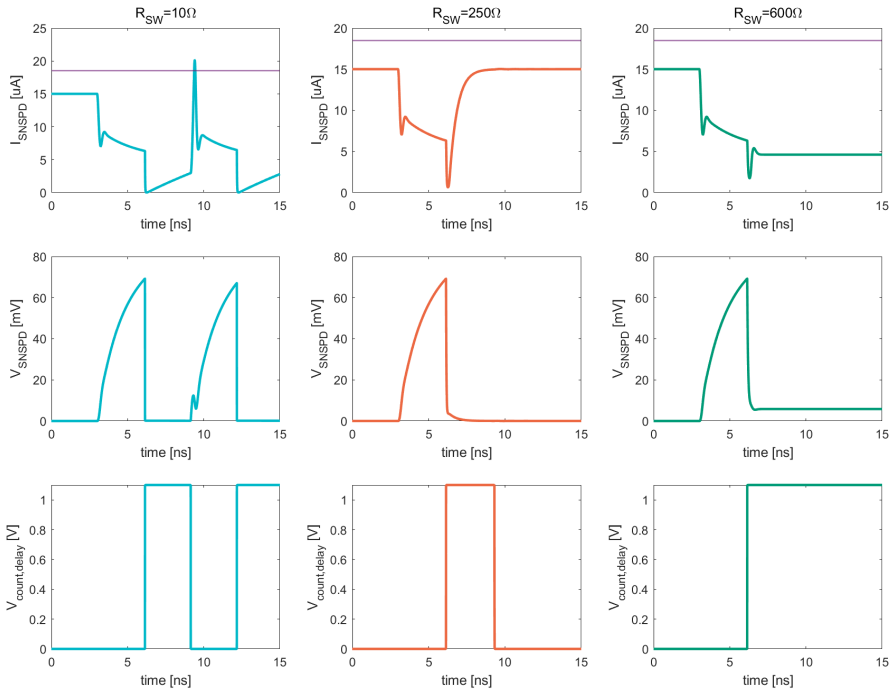


Figure 4.8: Impact of the active quenching switch resistance R_{SW} on the readout performance. Top row: Current in SNSPD I_{SNSPD}; middle row: output voltage across R_L = 10k Ω; bottom row: count signal with delay of t_{delay} = 3 ns. The plots of the same color in different columns represent the different values of R_{SW}. The purple line in the first row indicates the current I_{SW} that SNSPD will switch to the resistive state. Photon hits at t=3 ns and active quenching is enbaled at t=6 ns.

4.2. Analog Sub-circuits

In this section, the design consideration of the analog part of the readout system shown in Figure 4.1 will be discussed.

4.2.1. SNSPD Biasing

To begin with, the SNSPD should be correctly biased. The headroom of the bias current $I_{SNSPD,bias}$ and the switching current I_{sw} is of great importance when biasing the SNSPD, as it is necessary to balance the detection efficiency and the risk of false triggering (higher $I_{SNSPD,bias}$ means higher detection efficiency, but also a higher possibility of false triggering due to noise). However, I_{sw} is mostly dependent on the design and material of the SNSPD itself (see Section 3.1), and the preference value of $I_{SNSPD,bias}$ will usually be provided by the manufacturer. As is demonstrated in Figure 4.9, when the bias current is set between 16 and 18 μA, the system can achieve high detection efficiency while maintaining a relatively low dark count rate. Consequently, it is empirically preferable to set the SNSPD bias current to be:

$$\begin{cases} I_{SNSPD,bias} \in [80\%, 90\%]\, I_{sw} \\ \Delta_{I_{SNSPD,bias}} = \pm 1\mu\text{A} \end{cases} \tag{4.2}$$

In this work, the SNSPD switch current were set to I_{sw} = 18.5 μA and the bias current $I_{SNSPD,bias}$ = 15 μA which is approximately 81.2 % of I_{sw}. These values are flexible and can be adjusted as the paramters in the SNSPD model.

As demonstrated in Figure 4.10, the reference current is mirrored to the middle Mb2 branch and will be mirrored again by Mb3 and programmable M1 to set the current going into the SNSPD, the programmable M1 transistors are connected with the thick oxide 2.5 V switches. The double mirroring here is because the current generator used is of NMOS type, while the SNSPD biasing uses PMOS IDAC

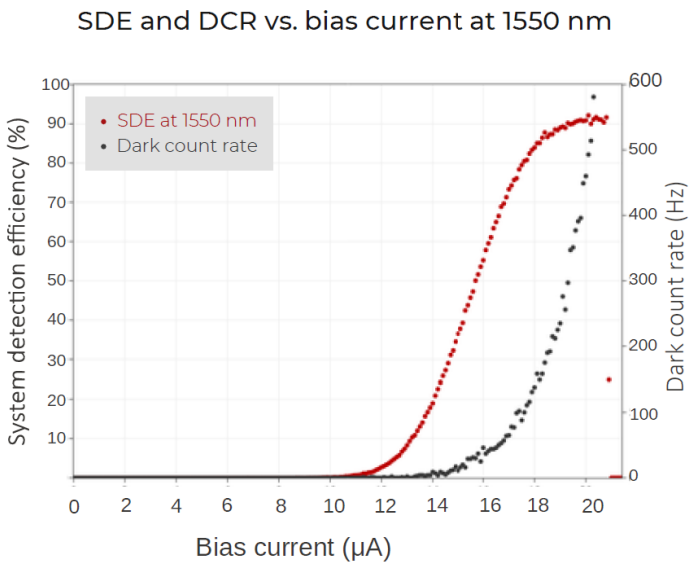


Figure 4.9: System detection effeciency (SDE) and dark count rate (DCR) vs. bias current $I_{bias,SNSPD}$ at 1550 nm [40]

topology, so the additional mirroring is for this conversion. The sources of the switches are connected to the M1 gate and the drains are selectively connected to either the supply or the gate of M4. When operating in a cryogenic environment at 4 K, the threshold voltage of CMOS transistors is anticipated to be 100 mV to 200 mV higher than at room temperature [41]–[43]. This implies that a larger V_{gs} is needed to maintain the same current. However, since the V_{th} of both NMOS and PMOS in the mirror will increase in 4 K, the bias scheme introduced here can still work at low temperature. Therefore, this work uses 2.5 V thick-oxide transistors as switches (except for active quenching switches by using back bias to compensate for the increase in threshold) to guarantee that $V_{gs,sw}$ is sufficiently high to turn the switches completely on irrespective of the V_{gs}.
By making $I_{ref,SNSPD} = 0.5\ \mu$A, the accuracy requirement of the bias current in Equation (4.2) can always be met, as $I_{SNSPD,bias}$ can be adjusted by digitally programming the number of PMOS transistors M1 (the current mirror ratio between Mb3 and M1).

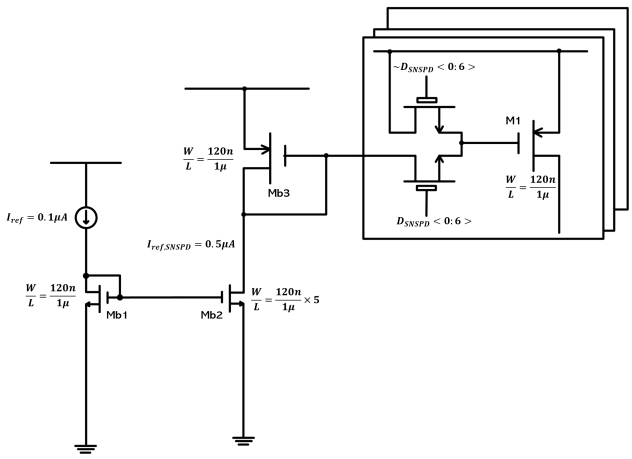


Figure 4.10: SNSPD bias schematic

4.2.2. Main Amplifier and Comparator

In order to reduce power consumption while still meeting the system requirements for count rates and jitter, it is essential to have the optimal implementation of the amplifier and comparator. An amplifier is needed in this work to boost the SNSPD output signal to a higher voltage level with a steeper slope, thus accelerating the detection process. Additionally, the amplifier will reduce the noise contributed from the later stages, which is beneficial for the jitter performance. However, there is no direct correlation between the circuit specifications and those requirements. In this section, we will calculate the optimal transconductance sum of the main amplifier $g_{amp,m1}$ and the comparator $g_{comp,m1}$ to minimize power consumption while meeting the jitter performance requirement at the decision time with $\sigma_{jitter} \leq \frac{v_{n,rms}}{SR_{t=t_{decision}}}$.

Assumptions

1. The topology of the amplifier is the NMOS amplifier with a PMOS active load, as indicated by the blue dashed box in Figure 4.14. The input noise power spectral density of this amplifier is expressed by Equation (4.3). Here, it is assumed that the thermal noise from the put NMOS is dominant.

$$\begin{aligned}\overline{V_{n,amp,in}^2} &= 4kT\gamma\left(\frac{g_{m1,P}}{g_{m1,N}^2}+\frac{1}{g_{m1,N}}\right)+\frac{1}{C_{ox}}\left[\frac{K_P g_{m1,P}^2}{(WL)_P g_{m1,N}^2}+\frac{K_N}{(WL)_N}\right]\frac{1}{f}\\ &\approx \frac{4kT\gamma}{g_{m1,N}}\end{aligned} \tag{4.3}$$

   where k is the Boltzmann constant; T is the temperature in Kelvin; γ is the transistor thermal noise coefficient; g_m is the transconductance of the transistor; W and L are the width and channel length of the transistor; K is the empirical flicker noise coefficient.

2. The topology of the comparator is the differential to single-ended amplifier with a PMOS input pair as shown in Figure 4.18. The corresponding input noise power spectral density is:

$$\begin{aligned}\overline{V_{n,comp,in}^2} &= \frac{[4kT\gamma\,(2g_{m1,2,N}+2g_{m1,2,P}+)]\times R_{out}^2+\frac{\overline{I_{n,s}^2}}{4g_{m,diode}^2}}{A_{comp}^2}\\ &\approx \frac{8kT\gamma}{g_{m1,P}}\end{aligned} \tag{4.4}$$

   where $I_{n,s}$ represents the current noise from M5 in Figure 4.18, and $g_{m,diode}$ is the transconductance of M3.

3. The signals at the amplifier and comparator input are in the form of a ramp function.

$$R(x) := \begin{cases}\alpha x & x \geq 0\\ 0 & x < 0\end{cases} \tag{4.5}$$

4. In the initial nanoseconds after a photon incident event, the SNSPD output has a slew rate of $a = \frac{I_{Bias,SNSPD}}{C_{pad}}$.

5. Both the amplifier and the comparator work as a first-order low-pass filter with some specific gain A, which means that their effective noise bandwidth (ENBW) can be expressed by

$$\begin{aligned}ENBW &= BW_{-3dB}\times\frac{\pi}{2}\\ &= \frac{1}{\tau}\times\frac{\pi}{2}\\ &= \frac{g_m}{A\times C_L}\times\frac{\pi}{2}\end{aligned} \tag{4.6}$$

where τ is the time constant; C_L is the corresponding load capacitor.

6. The comparator output-referred noise should be negligible compared to the noise from the amplifier (at least 10 times smaller).

$$v_{n,amp@comp_{out},rms} \geq 10 \times v_{n,comp@comp_{out},rms} \tag{4.7}$$

7. The comparator's bandwidth is smaller than that of the amplifier, so the amplifier's noise should experience filtering at the comparator's output.This simplification is advantageous for the purpose of computing jitter (see Section 4.2.2). Additionally, it is also beneficial in practice as the amplifier noise can be filtered to improve the signal-to-noise ratio at the comparator output.

Consequently, when taking into account the assumptions mentioned above, equation (4.7) can be interpreted as:

$$\begin{aligned}
&v_{n,amp@comp_{out},rms} \geq 10 \times v_{n,comp@comp_{out},rms} \\
\Rightarrow &\left(v_{n,amp@comp_{out},rms}\right)^2 \geq \left(10 \times v_{n,comp@comp_{out},rms}\right)^2 \\
\Rightarrow &V^2_{n,amp@amp_{in}} \times A^2_{total} \times ENBW_{f,comp} \geq 10^2 \times V^2_{n,comp@comp_{in}} \times A^2_{comp} \times ENBW_{f,comp} \\
\Rightarrow &\frac{4kT\gamma}{g_{amp,m1}} \times A^2_{total} \times \frac{1}{2\pi\tau_{comp}} \times \frac{\pi}{2} \geq 10^2 \times \frac{8kT\gamma}{g_{comp,m1}} \times A^2_{comp} \times \frac{1}{2\pi\tau_{comp}} \times \frac{\pi}{2}
\end{aligned} \tag{4.8}$$

where $A_{total} = A_{amp} \times A_{comp}$. By referring the equation (4.8) to the input of the amplifier, we obtain the ratio R between the transconductane of amplifier and comparator:

$$\begin{aligned}
&\frac{\frac{kT\gamma A^2_{total}}{g_{amp,m1}\tau_{comp}}}{A^2_{total}} \geq 10^2 \times \frac{\frac{2kT\gamma A^2_{comp}}{g_{comp,m1}\tau_{comp}}}{A^2_{total}} \\
\Rightarrow &R := \frac{g_{comp,m1}}{g_{amp,m1}} \geq \frac{200}{A^2_{amp}}
\end{aligned} \tag{4.9}$$

Ramp Input Response

Based on the assumption 3 and 4, the input of the amplifier and its Laplace form is:

$$\mathcal{V}_{amp,in}(s) = \mathcal{L}\{V_{amp,in}(t)\}(s) = \int_0^\infty ate^{-st}dt = \frac{a}{s^2} \tag{4.10}$$

So the response at amplifier output after a first-order roll-off transfer function with a gain of A_{amp} and time constant τ_{amp} will be:

$$\begin{aligned}
&\mathcal{V}_{amp,out}(s) = \frac{a}{s^2} \times \frac{A_{amp}}{1 + s\tau_{amp}} \\
\Rightarrow &V_{amp_{out}}(t) = \mathcal{L}^{-1}\{\mathcal{V}_{amp,out}(s)\}(t) = aA_{amp}\left[-\tau_{amp}\left(1 - e^{-\frac{t}{\tau_{amp}}}\right) + t\right]
\end{aligned} \tag{4.11}$$

Similarly, the output of the comparator can be determined by inputting equation (4.11) into the comparator and then applying a first-order roll-off low pass filter with a gain of A_{comp} and a time constant of τ_{comp}.

$$\begin{aligned}
&\mathcal{V}_{comp,out}(s) = \mathcal{V}_{amp,out}(s) \times \frac{A_{comp}}{1 + s\tau_{comp}} \\
\Rightarrow &V_{amp_{comp}}(t) = \mathcal{L}^{-1}\{\mathcal{V}_{comp,out}(s)\}(t) = aA_{amp}A_{comp}\left(t - \tau_{amp} - \tau_{comp} + \frac{\tau^2_{amp}e^{-\frac{t}{\tau_{amp}}} - \tau^2_{comp}e^{-\frac{t}{\tau_{comp}}}}{\tau_{amp} - \tau_{comp}}\right)
\end{aligned} \tag{4.12}$$

Slew Rate Calculation

It is tedious to use the equation (4.11) and (4.12) in the later calculation. Thus, the aim in this subsection is to simplify the exponential terms in these expressions into polynomials up to the second or third order.

As can be seen in Figure 4.11, it is reasonable to assume $\frac{t_{rise}}{t_{fall}} \sim \frac{1}{4}$. This means that for the requirement in this work of $f_{count} \geq 20$ MHz $\Rightarrow t_{count} = \frac{1}{f_{count}} \leq 50$ ns, therefore, the time constant of the amplifier and comparator should be $\tau_{amp,comp} \leq t_{rise} \approx 10$ ns. Consequently, by substituting the time constants into

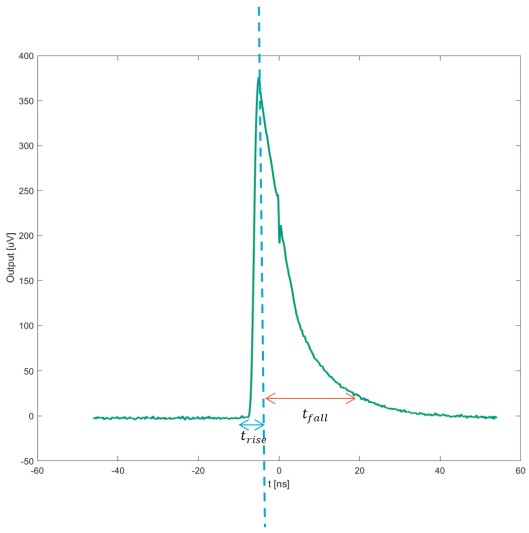


Figure 4.11: Reproduced labotary measurement data of the SNSPD readout signal provided by *Pixel Photonics*. The blue arrow indicates the time period of the rising edge, while the orange arrow is the time period of the falling edge.

equation (4.11) and (4.12), the time domain response can be plotted as shown in Figure 4.12. Figure 4.12 shows that the results of polynomial simplification are accurate enough in the range of t_{amp} and t_{comp}. Therefore, the output response of the amplifier will be simplified to a second-order polynomial, and that of the comparator to a third-order polynomial in the following calculations.

To continue calculating the jitter requirement with $\sigma_{jitter} \leq \frac{v_{n,rms}}{SR|_{t=t_{decision}}}$, this subsection focuses on obtaining the slew rate (SR) at the comparator decision time. After the second-order Taylor series expansion of $e^{-\frac{t}{\tau_{amp}}} \approx 1 - \frac{t}{\tau_{amp}} + \frac{\frac{t^2}{\tau_{amp}^2}}{2!}$, Equation (4.11) can be expressed as:

$$V_{amp_{out}}(t) \approx aA_{amp}\frac{t^2}{2\tau_{amp}} \tag{4.13}$$

and the corresponding slew rate (SR) is:

$$SR_{amp,out} \coloneqq \frac{\partial V_{amp_{out}}(t)}{\partial t} = aA_{amp}\frac{t}{\tau_{amp}} \tag{4.14}$$

Similarly, with the third-order Taylor series simplification of $e^{-\frac{t}{\tau}} \approx 1 - \frac{t}{\tau} + \frac{\frac{t^2}{\tau^2}}{2!} - \frac{\frac{t^3}{\tau^3}}{3!}$, equation (4.12) can be rewritten as

$$V_{comp_{out}}(t) \approx aA_{total}\frac{t^3}{6\tau_{amp}\tau_{comp}} \tag{4.15}$$

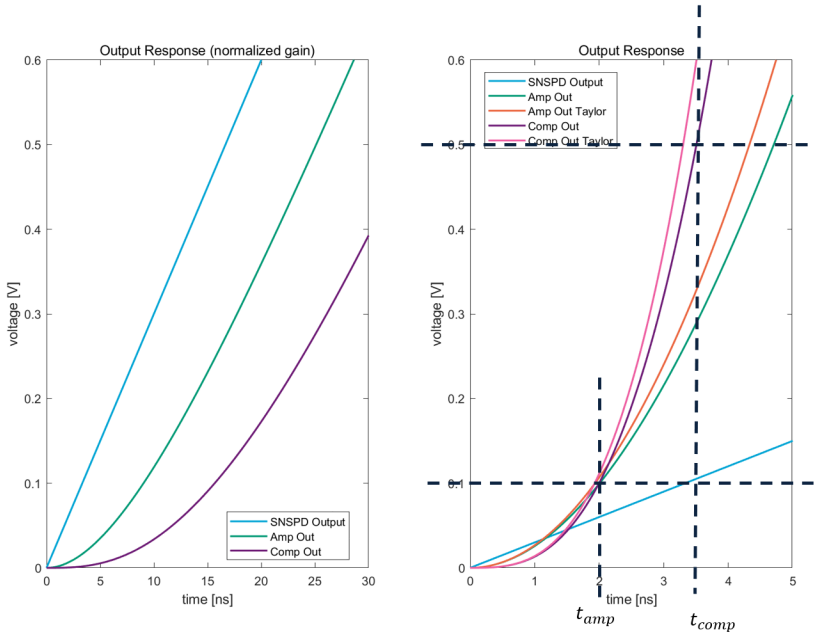


Figure 4.12: Time domain output response for amplifier and comparator in both complete expression and second-order (for equation (4.11)) and third-order (for equation (4.12)) Taylor series simplification. Left: responses normalized by cumulative gain. Right: responses with gains, where t_{amp} indicates the time point when the amplifier reaches its amplification threshold, and t_{comp} indicates the time when the comparator reaches its decision threshold.

where $A_{total} := A_{amp}A_{comp}$. So, the SR at the comparator's output would be:

$$SR_{comp,out}(t) = \frac{\partial V_{comp_{amp}}(t)}{\partial t} = aA_{total}\frac{t^2}{2\tau_{amp}\tau_{comp}} \tag{4.16}$$

Next, the decision time $t_{decision}$ is subsituted to obtained from the equation (4.15):

$$t_{decision} = \left(\frac{6\tau_{amp}\tau_{comp}V_{comp_{decision}}}{aA_{total}}\right)^{1/3} \tag{4.17}$$

into equation (4.16), the comparator output slew rate at the decision time will be:

$$\begin{aligned} SR_{comp,out}|_{t=t_{decision}} &= \frac{aA_{total}}{2}\left(\frac{6V_{comp_{decision}}}{aA_{total}}\right)^{2/3}\left(\frac{1}{\tau_{amp}\tau_{comp}}\right)^{1/3} \\ &= \frac{aA_{total}}{2}\left(\frac{6V_{comp_{decision}}}{aA_{total}}\right)^{2/3}\left(16ENBW_{amp}ENBW_{comp}\right)^{1/3} \end{aligned} \tag{4.18}$$

where $\frac{1}{\tau} = 4ENBW$ is under the assumption of first-order roll-off transfer function.

Optimal Transconductance

After having derived the output referred noise in Equation (4.8), the SR at the decision time in Equation (4.18), and with the assumption 6 to ignore the comparator output noise contribution, jitter requirement

listed in Table 1.1 that $\sigma_{jitter} \leq 40$ ps can be written as:

$$\begin{aligned} &\frac{v_{n,rms}}{SR|_{t=t_{decision}}} \leq \sigma_{jitter} \\ \Rightarrow &\frac{\sqrt{V^2_{n,amp@amp_{in}} A^2_{total} ENBW_{comp}}}{SR_{comp,out}|_{t=t_{decision}}} \leq \sigma_{jitter} \\ \Rightarrow &\frac{\sqrt{\frac{4kT\gamma}{g_{amp,m1}} A^2_{total} ENBW_{comp}}}{\frac{aA_{total}t^2_{decision}}{2\tau_{amp}\tau_{comp}}} \leq \sigma_{jitter} \end{aligned} \tag{4.19}$$

Therefore, the limitation on the amplifier's transconductance can be further obtained by simplifing the Equation (4.19):

$$g_{amp,m1} > \left[R \times \frac{64\left(A_{total}kT\gamma\right)^3}{a^2\sigma^6_{jitter}\left(6V_{out,comp}\right)^4} \times \frac{C^2_{L,amp}}{C_{L,comp}} \times A^3_{amp} \right]^{\frac{1}{4}} \tag{4.20}$$

where R is the ratio between transconductance of comparator and amplifier as calculated in Equation (4.9). From assumption 7, an additional restriction on R can be derived

$$\begin{aligned} &BW_{amp} \geq BW_{comp} \\ \Rightarrow &\frac{g_{amp,m1}}{C_{L,amp}A_{amp}} \geq \frac{g_{comp,m1}}{C_{L,comp}A_{comp}} \\ \Rightarrow &R = \frac{g_{comp,m1}}{g_{amp,m1}} \leq \frac{C_{L,comp}}{C_{L,amp}} \times \frac{A_{total}}{A^2_{amp}} \end{aligned} \tag{4.21}$$

together with Equation (4.9) we have

$$\frac{200}{A^2_{amp}} \leq R \leq \frac{C_{L,comp}}{C_{L,amp}} \times \frac{A_{total}}{A^2_{amp}} \tag{4.22}$$

If $A_{total} := \frac{V_{out,comp}}{V_{in,amp}} \approx \frac{1V}{4mV} = 250$, then based on Equation 4.2.2 we have $\frac{200}{A^2_{amp}} \leq R \leq \frac{C_{L,comp}}{C_{L,amp}} \times \frac{250}{A^2_{amp}}$. Therefore, the load capacitor at the amplifier and comparator output should satisfy $C_{L,comp} \geq \frac{200}{250} C_{L,amp}$. Consequently, by substituting this load capacitance restriction into Equation (4.20), the optimal summation of the amplifier and comparator transconductance G_m can be derived in Euqation (4.23). Here, the sum of transconductance is a useful quality since g_m is proportional to the current, so G_m gives the basic idea of the power consumption of the amplifier and comparator, which is dominant in the readout circuit.

$$\begin{aligned} G_m &= g_{amp,m1} + g_{comp,m1} = g_{amp,m1}(1+R) \\ &\gtrsim \left[\frac{200 \times 64\left(A_{total}kT\gamma\right)^3}{a^2\sigma^6_{jitter}\left(6V_{out,comp}\right)^4} \times C_L \times A_{amp} \right]^{\frac{1}{4}} \left(1 + \frac{200}{A^2_{amp}}\right) \end{aligned} \tag{4.23}$$

G_m with the respect to amplifier's gain A_{amp} is plotted in Figure 4.13. As can be seen in this graph, there is not much difference in G_m for the amplifier gain around the minimal point of $A_{gain,G_{m,min}} \approx 37.4$. The comparator transconductance decreases as the amplifier gain increases, as expected. However, the transconductance of the amplifier increases with the amplifier gain. This is based on the assumption 7 as the bandwidth of the comparator should be smaller than that of the amplifier. Thus, if the amplifier gain is increased, the corresponding transconductance must also increase to widen the bandwidth.
Therefore, one of the possible optimal combination for the amplifier and comparator transconductance,

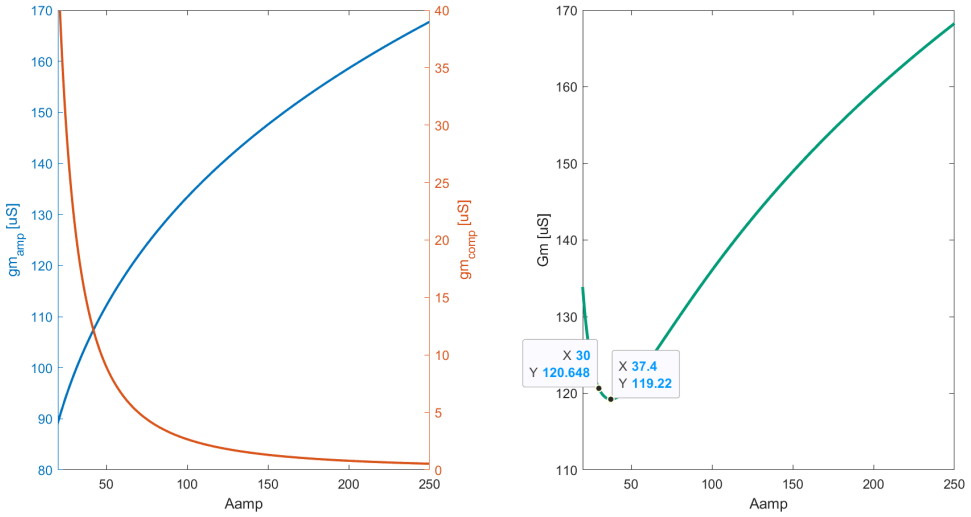


Figure 4.13: Left: transconductance of amplifier and comparator v.s amplifier gain; right: sum transconductance G_m and amplifier gain A_{amp} relation, where $C_{L,amp} = C_{L,comp} = 10$ fF, $A_{total} = 250$, and $\sigma_{jitter} = 30$ ps.

if we choose $A_{amp} = 30 \Rightarrow G_m = 120\ \mu$S, would be:

$$\begin{cases} g_{m1,amp} \approx 105\mu\text{S}, & A_{amp} \approx 30dB \\ g_{m1,comp} \approx 15\mu\text{S}, & A_{comp} \approx 17dB \end{cases} \tag{4.24}$$

4.2.3. Amplifier Implementation

The amplifier is highlighted by the blue dashed box shown in Figure 4.14. The size of the input transistor M1 can be determined by the g_m/i_d methodology as illustrated in Figure 4.15. The gain requirement for the main SNSPD output amplifier is approximately 30 dB based on Equation (4.24), which is not too demanding and can be achieved with the cascoded topology. To reach a gain of 30, the corresponding

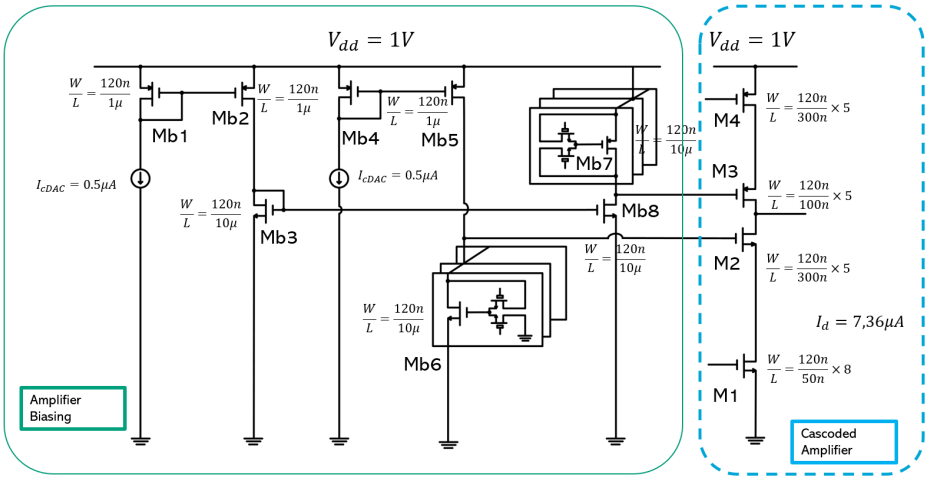


Figure 4.14: Right: NMOS cascoded amplifier with cascoded PMOS active load; left: Programmable biasing network for the amplifier. I_{IDAC} is the output current of the current DAC in Section 4.2.4.

$\frac{i_d}{w} \approx 8$, and the $\frac{g_m}{i_d} \approx 17$ can be determined in Figure 4.15. This means that if $g_m = 105\mu$A, then the required current is calculated as $i_d = \frac{g_m}{\frac{g_m}{i_d}} \approx 6.3\mu$A, and consequently $w = \frac{i_d}{\frac{i_d}{w}} \approx 0.8\mu$m. The results in

Figure 4.15 are simulated with the transistor length of l = 50nm, to allow the gain requirement to be met.

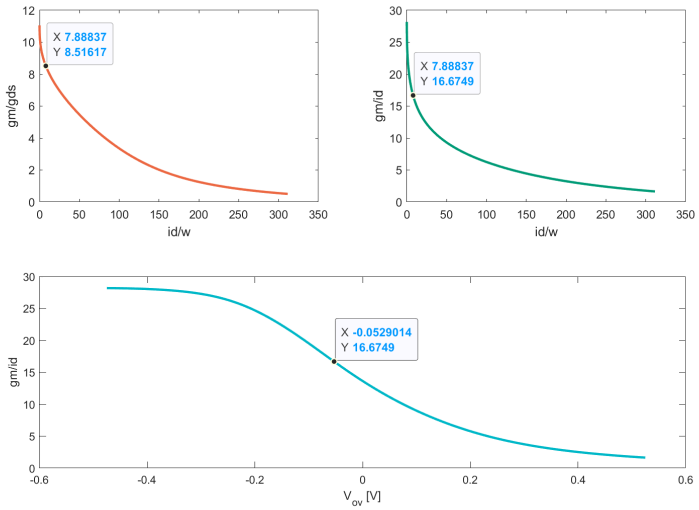


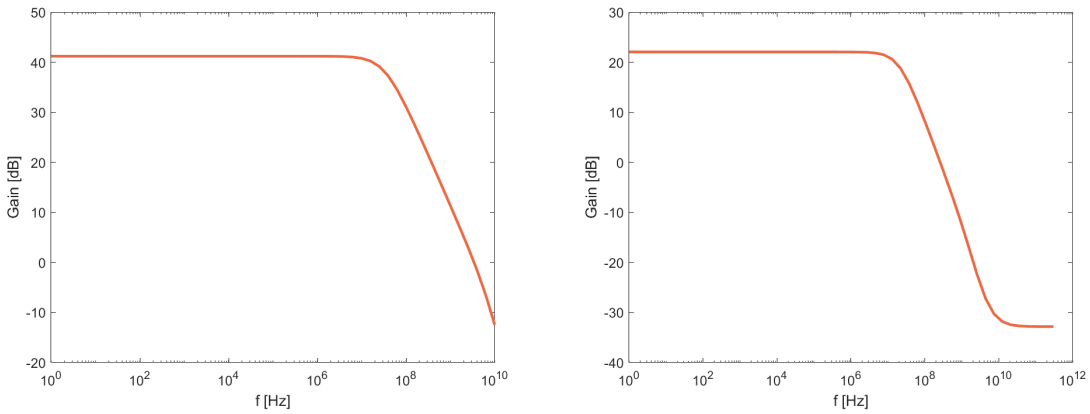
Figure 4.15: g_m/i_d methodology for determining the size of amplifier input transistor M1. Orange: intrinsic gain v.s i_d/w; green: g_m/i_d v.s i_d/w; Blue: g_m/i_d v.s overdirve voltage V_{ov}. w is the width of the transistor.

The plot in Figure 4.15 shows that the transistor is in its weak inversion and can be further pushed by increasing the width of the transistor to make g_m/i_d close to its saturation point, meaning that the input transistor operates near its highest transconductance efficiency. The g_m/i_d will be further improved when the circuit is operated at cryogenic temperatures around 4 K due to the increase in the threshold voltage V_{th}. The other transistors are sized in strong inversion to reduce noise contribution and parasitic capacitance, while still maintaining the gain.
The green box indicates the biasing of transistors M2 and M3 of the amplifier. The current DAC introduced in 4.2.4 later generates the bias current I_{IDAC} = 0.5 μA. The Mb6 and Mb7 are designed to be programmeable for biasing M2 and M3, which will be useful in tuning $V_{DS,M2,3}$ to ensure that M2 and M3 will still work in the saturation region when the supply voltage of the amplifier is decreased (the V_{DS} headrooms will consequently decrease) for the purpose of reducing power consumption in the experiments. Here, the supply voltage for the amplifier will be separated from the supply to other components. The bias for M4 comes from the feedback amplifier and is used to fix the operating point of the amplifier during the readout procedure, which will be discussed in Section 4.2.6; The bias for the input transistor M1 is shown in Figure 4.1, which is a diode-connected NMOS with a cascoded NMOS to reduce the impact of V_{DS}. This M1 bias has been designed to be programmable, as illustrated in Figure 4.20. This allows for tunability and the capacity to regulate the amplifier current $I_{D,amp}$ and its bias, which will be beneficial for further experiments and a better understanding of the system.
The gain performance of the amplifier is shown in Figure 4.16a which meets the theoretically calculated requirement that the gain is above 30 dB while the -3 dB bandwidth is around 50MHz based on

$$BW_{-3dB} = \frac{g_{amp,m1}}{2\pi C_{L,amp} A_{amp}} = \frac{100\mu S}{2\pi \times 10fF \times 30} \approx 50MHz \tag{4.25}$$

4.2.4. Current DAC for Anolog Subcircuits Biasing

Instead of using the PMOS IDAC introduced before, the NMOS version is used here, as it is imperative that the analog subcircuits are correctly biased to guarantee the proper functioning of the system. As demonstrated in Figure 4.17, a 7-bit current DAC with a 0.1 μA least significant bit (LSB) is used to



(a) Amplifier gain vs frequency

(b) Comparator gain vs frequency

Figure 4.16: Amplifier and comparator gain

generate the reference current. Then selective switches between the gates of Mb1 and Mb2 are used to programmably adjust the mirroring ratio to produce the bias current. The output current of this NMOS IDAC will be further mirrored through Mb3 to the bias of the other analog circuits, including the main amplifier for SNSPD output amplification, comparator (see Section 4.2.2), and the feedback amplifier (in Section 4.2.6).

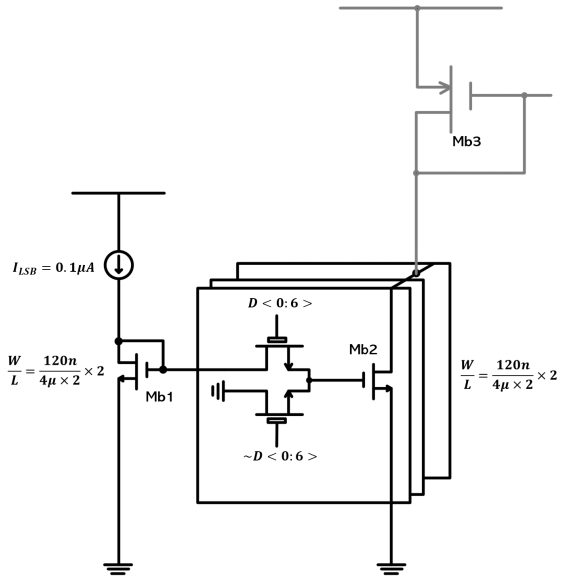


Figure 4.17: Schematic of the current DAC

The programming switches are located at the gate instead of being cascoded with the IDAC NMOS transistor Mb2 to prevent the variation in the mirrored current due to the change in $V_{DS,Mb2}$.

4.2.5. Comparator Implementation

Figure 4.18 shows the used differential-to-single-ended amplifier that generates the count signal which also actives quenching enable signal. The bias current $I_{IDAC} = 0.1\mu$A is generated by the current DAC, which is mirrored through Mb1 to M5 with a ratio of 10. M1 and M2 are used as input pairs. The output of the amplifier is connected to M2 in order to be compared to the reference voltage that is connected to M1. The output is on the same side as M2 to improve response speed. Furthermore, the M2 side

of the circuit (M2 and M4) is designed to be 4× larger than the M1 side (M1 and M3), so the current supported by M5 will be distributed in a 4 : 1 ratio, meaning that the current on the output side is also 4× larger than the other half, which will increase the speed of comparison. The total current needed for the differential structure will not be twice as much as each branch, which can help to reduce power consumption.

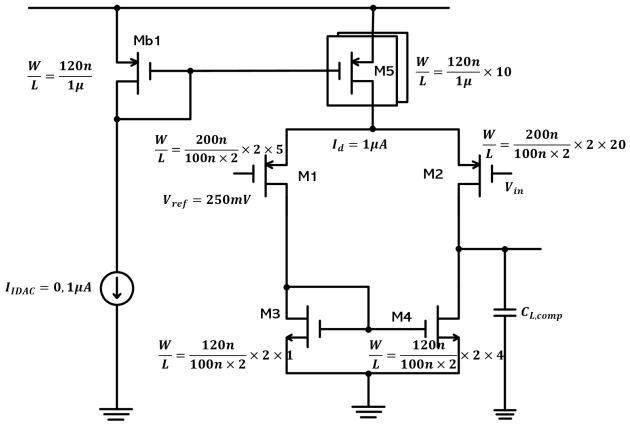


Figure 4.18: Schematic of the differential to single-ended amplifier used as the comparator.

The gain of the comparator is shown in Figure 4.16b, which shows that the DC gain is greater than ∼ 17 dB and the gain bandwidth (GBW) product is close to the theoretical calculation result based on (4.24):

$$GBW_{comp} = \frac{g_{m1,comp}}{2\pi C_{L,comp}} = \frac{15\mu S}{2\pi \times 10fF} \approx 240MHz \tag{4.26}$$

The choice of the comparator threshold is less strict and can be determined by the following: When the output of the SNSPD is $V_{out,SNSPD} = 4$ mV, after amplification, the output of the amplifier should be $V_{out,amp} = 400 - 4 \times 30 = 280$ mV. This requires that

$$V_{out,comp} \begin{cases} \geq 0.5V & V_{in,comp} \leq 0.28V \\ 0V & V_{in,comp} > 0.28V \end{cases} \tag{4.27}$$

By sweeping the reference voltage connected to M1, the different input-output relation can be plotted as shown in Figure 4.19. The blue arrow in x-axis indicates the input range of the comparator (which is also the output swing of the amplifier) in the range of 100 mV ∼ 400 mV. As can be seen in the figure, the reference voltage in the range of $V_{ref} \in [0.3\text{ V}, 0.2\text{ V}]$ meets the requirement specified in Equation (4.27). Here, the reference is set to 250 mV since the transition occurs slightly above the midpoint where the slew rate is at its highest. This also provides some headroom to avoid possible false decisions cuased by offset and noise.

4.2.6. Feedback Amplifier

The amplifier combined with comparator operate in an open loop manner, hence the bias points can shift a lot due to process variations and mismatch. To ensure that the output of the main amplifier is in a range where the comparator can operate effectively, a feedback loop is implemented that regulates the current in the main amplifier. A feedback that can sense the output and fix the amplifier bias is needed as shown in Figure 4.20. The feedback amplifier has a differential-to-single-ended topology similar to that of the comparator described in Section 4.2.5. The current source of this feedback amplifier M_{fb5} is also designed to be programmable to ensure proper function when the V_{DS} headroom is reduced

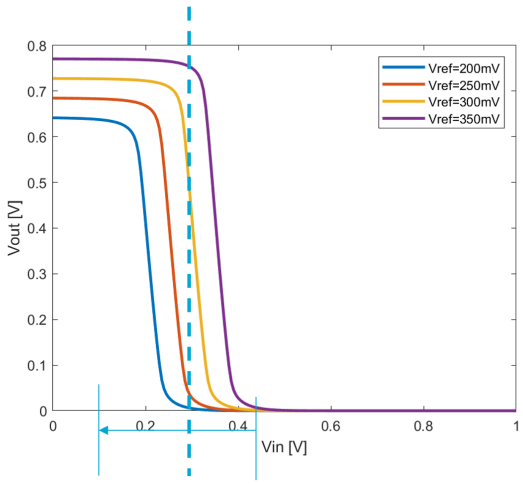


Figure 4.19: Comparator output with respect to the input with different reference voltages.

due to the decrease in supply for power reduction and the increase in V_{th} when working in a cryogenic environment.
The feedback amplifier senses the output of the amplifier through M_{fb1}, and compares it to the reference voltage connected to M_{fb2}. The output is then used as the bias of M4 in the main amplifier to help the amplifier fix and return to its bias point after the amplification. Here, the input M_{fb1} is on the different side of the output since transistor M4 adds another inversion. Ideally, the feedback gain should be

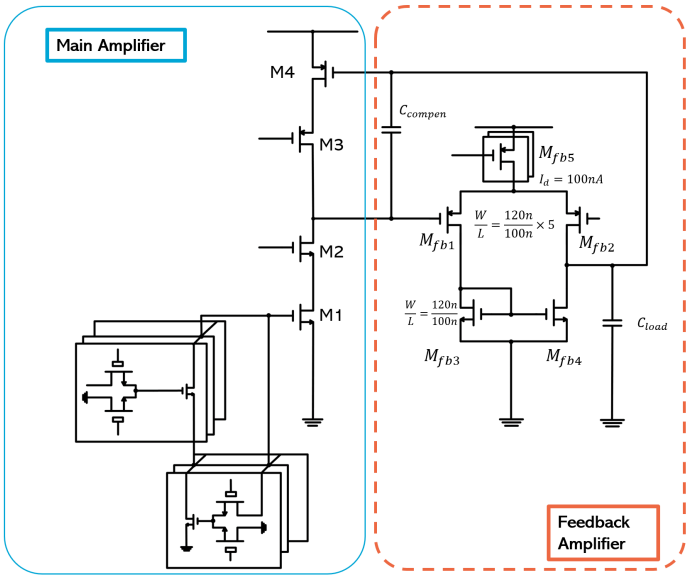


Figure 4.20: Schematic of the main amplfier with a feedback to fix the bias point.

sufficiently large to minimize the error, as indicated in Figure 4.21. The gain of the feedback amplifier can be computed as:

$$\begin{aligned} (\Delta x - y)\, A_{FB} A_{amp,input@pmos} &= y \\ \Rightarrow \frac{A_{FB} A_{amp,input@pmos}}{1 + A_{FB} A_{amp,input@pmos}} &= \frac{y}{\Delta x} \end{aligned} \tag{4.28}$$

Here, it is worth mentioning that since $A_{amp,input@pmos}$ is the gain when M4 is the input transistor and is approximately $A_{amp,input@pmos} \approx \frac{1}{10} A_{amp,input@M1} = \frac{1}{10} \times 30 = 3$ due to $g_{m,M4} \approx \frac{1}{10} g_{m,NMOS}$. In order

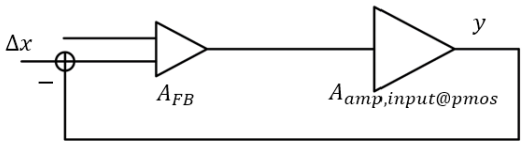


Figure 4.21: Feedback gain v.s frequency

to avoid over-designing and wasting power, the gain can be limited by selecting an appropriate error tolerance. Here, an error of ~ 10 mV is sufficient due to the relatively relaxed requirements for the comparator's threshold. Consequently, the gain of the feedback amplifier should be at least $A_{FB} \geq 20$ dB. Additionally, the bandwidth of the feedback amplifier should be low to ensure that signals at higher frequencies are sufficiently attenuated, so that only DC signals can be sensed and compared. As shown in Figure 4.22, when the load capacitance increases, the bandwidth of the feedback amplifier decreases, resulting in a great attenuation of the high frequency signal.

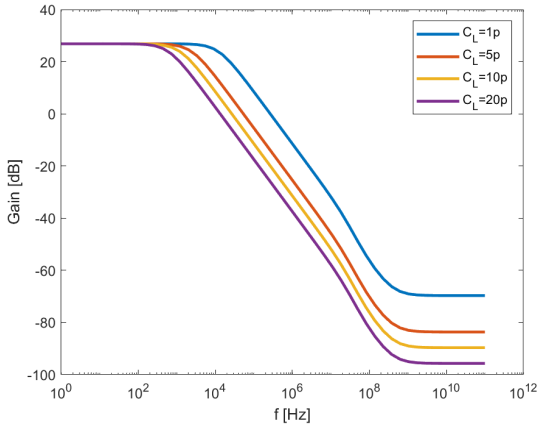
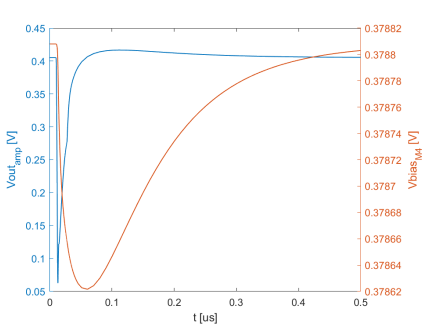
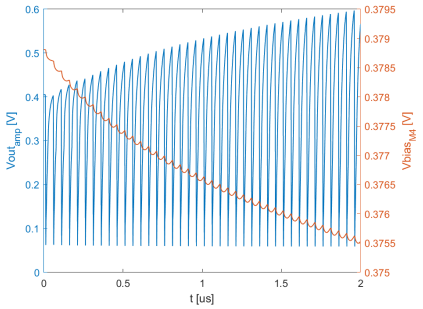


Figure 4.22: Feedback gain v.s frequency

However, using such an huge capacitor is not preferable in layout. Another issue with using slow feedback is that the amplifier will take a long time to recover. In certain extreme testing scenarios, such as when photons continuously trigger the SNSPD at a rate of 20 MHz or higher, the errors generated in each detection period (see Figure 4.23a) cannot be adequately compensated for and will accumulate, leading to a drift in the amplifier operating point. This will eventually reduce performance of the readout system after multiple photon detection events, as shown in Figure 4.23b. Therefore, the load



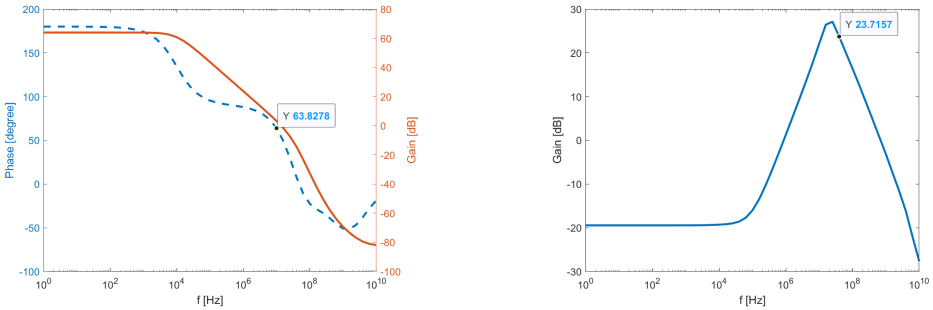
(a) Closer look at the response in a 500 ns period for settling.



(b) The system can not porperly function after several detecting operations.

Figure 4.23: Amplifier output response with the slow feedback.

capacitance at the output of the feedback amplifier should be took off to enhance the speed of the feedback, and a compensated capacitor should be connected between the amplifier output and the gate of transistor M4 for the stability compensation (see Figure 4.24).



(a) Stability of the loop with faster feedback

(b) Amplifier gain with faster feedback

Figure 4.24: Stability and the gain with faster feedback amplifier.

The use of faster feedback can reduce the gain performance of the main amplifier due to the fact that as the bandwidth of the feedback increases, some high-frequency signals will be attenuated, as shown in Figure 4.24b. However, the SNSPD output will continue to increase before it is actively quenched, so some loss in amplifier gain can be tolerated as long as the count signal decision can be made correctly and the jitter performance is met. The transient simulation of the main amplifier with feedback loop can be seen in Figure 4.25.

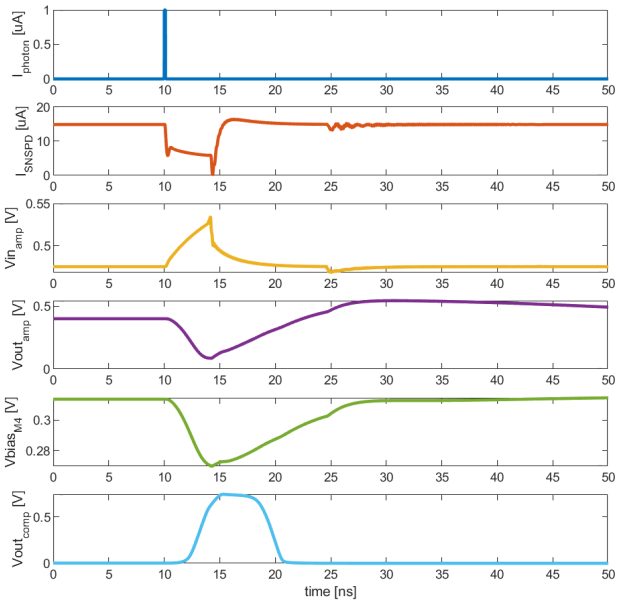


Figure 4.25: Transient simulation of the amplifier output cooperating with the faster feedback within 50 ns period. Plots from top to bottom: incident photon signal I_{photon} that triggers SNSPD; current in SNSPD I_{SNSPD}; input of the amplifier (also is the output of SNSPD) $V_{in,amp}$; output of the amplifier $V_{out,amp}$; output of the feedback amplifier (also is the bias voltage for the PMOS active load M4 of the main amplifier) $V_{out,fb}$; output of the comparator $V_{out,comp}$.

4.2.7. Mismatch and Robustness in Different Corners

This section aims to introduce design consideration in mismatch. Since both the comparator and the feedback amplifier use a differential topology, mismatch in the input pair and current mirror can result

in shifts in operating point. As introduced in Section 4.2.5, when the output of the amplifier reaches 300 mV, the output comparator should be below the midpoint, which means that the comparator will not produce a count signal. Figure 4.19 showed that the reference voltage should be less than 300 mV. Therefore, if the reference is set to 250 mV, then the mismatch in the input pairs of the comparator should follow $3\sigma \leq 300 - 250 = 50$ mV. The accuracy of the comparator will be impacted by the mismatch on the feedback amplifier since the error on the feedback amplifier will eventually appear at the output of the amplifier, which is directly connected to the comparator input. In summary, the width and the length of both the feedback amplifier and comparator should be scaled to meet the requirement of $3\sigma_{comp} = 50mV$.
Here, we first scale (simultaneously increasing the width and length of the transistor while keeping the W/L ratio unchanged) the comparator by 2 to obtain the standard deviation $\sigma_{comp} = 10$ mV on comparator transition voltage when V_{ref} is set to be 250 mV. This means that the headroom for the offset on the amplifier output is $\sigma_{amp,out} \approx \frac{100-6\times10}{6} \approx 6.5$ mV, which can be achieved by scaling the feedback amplifier by 8. The combination of the scaling factors of comparator and feedback amplifier can also be other values. However, due to the the size of comparator is larger than feedback amplifier, the small scaling in comparator is preferable to avoid increasing in the paracitic capacitance, and minimizing the total area also helps to minimize the power dissipation. The Monte Carlo (MC) simulation under 100 runs in TT corners for the amplifier with feedback is shown in Figure 4.26.

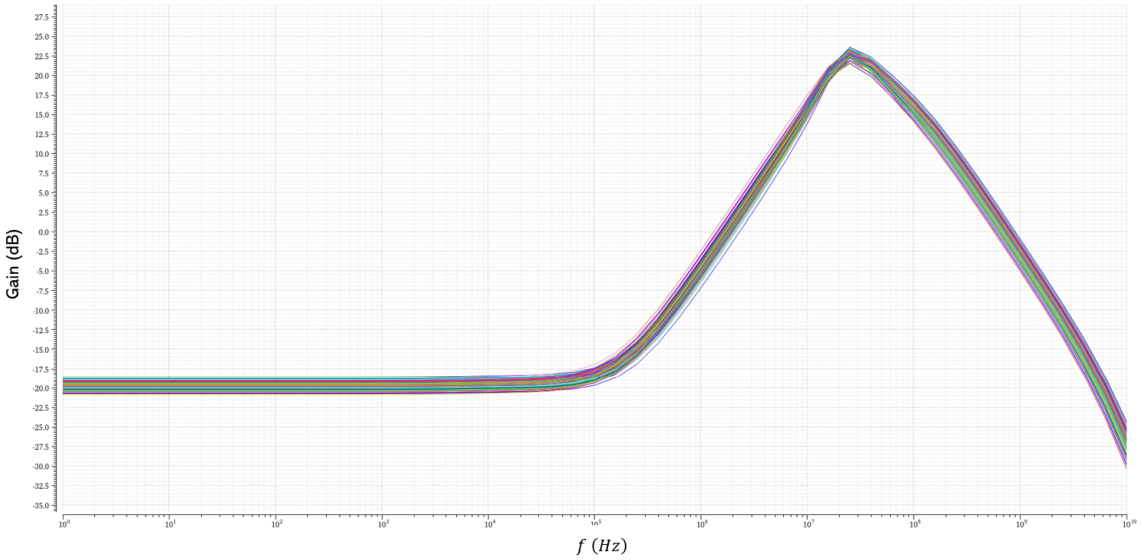


Figure 4.26: Monte Carlo analysis: AC simulation of the amplifier with FB after mismatch compensation.

To further test the robustness of the main core of the analog subcircuits, the different dc operating points including the SNSNPD bias current $I_{SNSPD,bias}$, the amplifier current I_{amp}, the transconductance of the amplifier $g_{m,amp}$ and the operating point of the amplifier output $V_{out,amp}$, are simulated in different corners, as illustrated in Figure 4.27.

4.3. Digital Sub-circuits

The digital circuits used in this design include an SR latch, a delay component, and inverters. Inverters after the comparator to buffer the count signal. An SR latch combined with a delay element is used to maintain a count signal for a certain duration in order to facilitate digital signal processing. The delayed signal will also be used as the active quenching enable signal that is connected to the NMOS switch (AQ SW) shown in Figure 4.1, which will give enough time for SNSPD to quench.

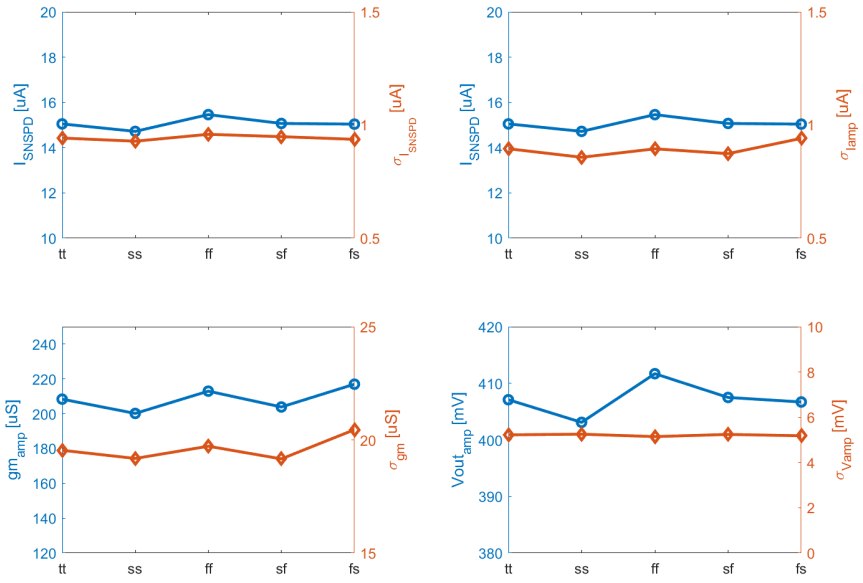
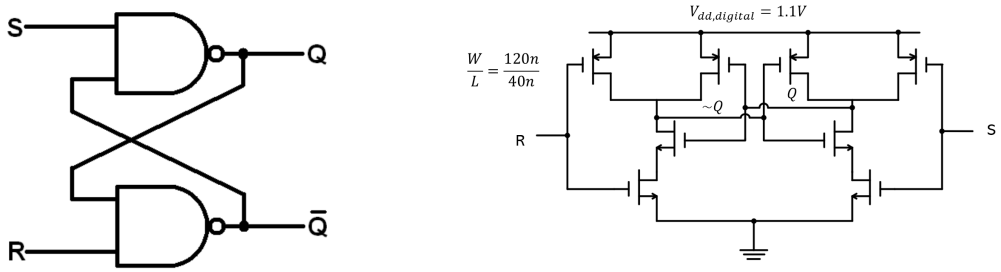


Figure 4.27: Analog subcircuit operating point with the corresponding standard deviation in different corners. Upper left: SNSPD bias current I_{SNSPD}; upper right: current in the main amplifier I_{amp}; bottom left: transconductance of the input transistor of the main amplifier M1 $g_{amp,M1}$; bottom right: dc point of the amplifier output $V_{out,amp}$.

4.3.1. SR Latch

The SR latch used here is the NAND-based SR latch whose schematic is shown in Figure 4.28, and the corresponding truth table shows in Table 4.1. Basically, the output of the comparator is connected to S, and the output Q is used as both the count signal and the active quenching enable signal. Q is sent to the delay element and then connected to R, thus providing the count signal with the corresponding hold time.



(a) Gate-level NAND-based SR latch

(b) Transistor-level NAND-based SR latch

Figure 4.28: NAND-based SR latch

4.3.2. Current-Starve Delay Chain

The current-starve delay chain is used to create a time delay, as illustrated in Figure 4.29. The DAC in Section 4.2.4 supplies the reference current, which is then mirrored on the current-starve route via Mdb2 and Md4, as well as Mdb3 and Md1. The SR latch's output will be connected to the input of the first current-starve path, which will be repeated 15 times in series. The output of the last 15th path will be connected to an inverter to restore the sharp transition edges. The delay is achieved by charging and discharging the parasitic capacitance, mainly from Md2 and Md3, and the delay time can

Table 4.1: NAND-based SR Latch Truth Table

S	R	Q	$\overline{Q}$	Operation
0	0	1	1	Not Allowed
0	1	1	0	Set
1	0	0	1	Reset
1	1	Not Changed	Not Changed	Latch

be controlled by varying the charging current. In this work, the accuracy on controlling of the delay time is not required, since the primary purpose of this component is to provide sufficient time for the active quenching operation. The accuracy requirement can also depend on other following circuits such as TDC or digital processing circuits, which are not taken into account in this work.
As illustrated in Figure 4.30, by changing the input current from 0.4 μA to 1 μA, we can obtain the delay from 2 ns to 6 ns.

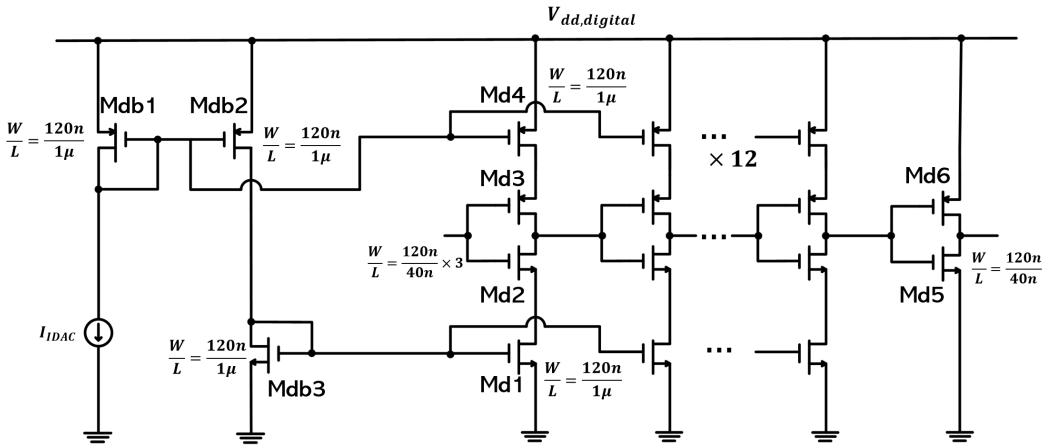


Figure 4.29: Schematic of the current-sarve delay chain.

4.4. System Performance

The performance of the system shown in Figure 4.1, with all components replaced by transistor-level implementation, is demonstrated in Figures 4.32. Additionally, Figure 4.31 is plotted in comparison with Figure 4.23b to show that with the help of the faster feedback, there is no heavy shift of the bias point. The system is now able to operate correctly at a photon count rate of 20 MHz. A more detailed look within a time period of 50 ns is provided in Figure 4.25. In brief, at a rate of 20 MHz, the photon will trigger the SNSPD to convert to a high-resistance state, causing the current biased in the SNSPD to be diverted to the readout resistor and generating a voltage across it. The amplifier then amplifies this small voltage until it exceeds the comparator threshold, producing a count signal that is sent to the digital circuit. After a time delay, the count signal is used as an enable signal to turn the active quenching switch on that shunts the SNSPD, redirecting the SNSPD bias and allowing the SNSPD to return to its superconducting state. The amplifier is also reset to its biasing point with the help of the feedback amplifier. The entire system, including the SNSPD and the readout electronics, is restored within 50 ns and ready for the next photon hit.
The value of the edge-delay jitter $J_{Delay_{photon-Q}}$ between the incident photon signal I_{photon} and the output count signal Q in different corners is shown in Figure 4.33, which demonstrates that the overall system meets the jitter performance requirement listed in Table 1.1. The performance of the readout electronics system is summarized in Table 4.2.

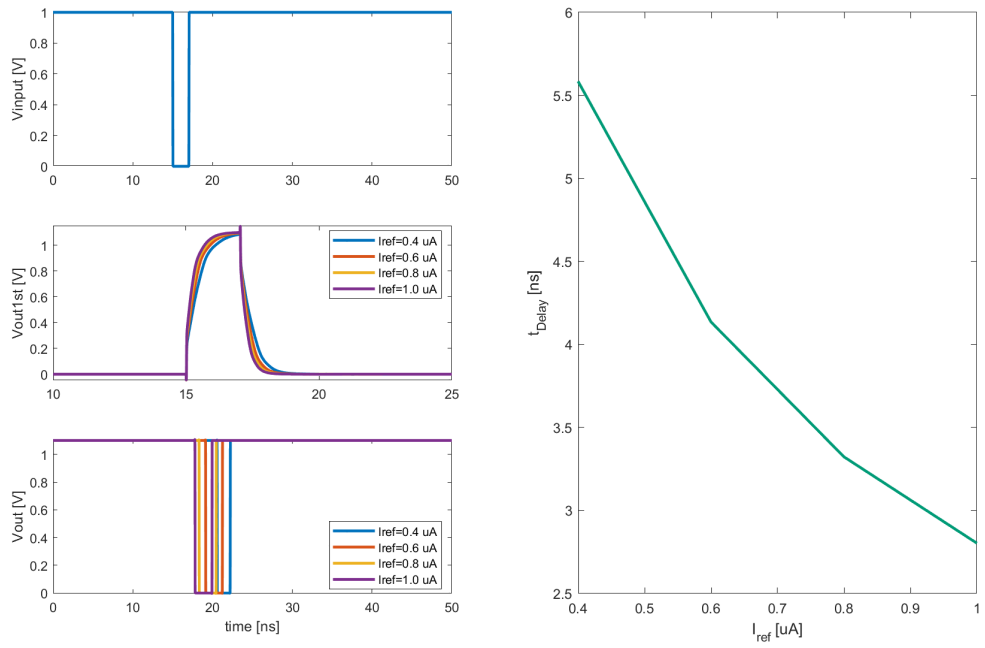


Figure 4.30: Transient simulation on the current-starve delay. Left top: input signal; middle: charging and discharging response at the first current-starve path; bottom: delayed output. Right: delay time v.s reference current.

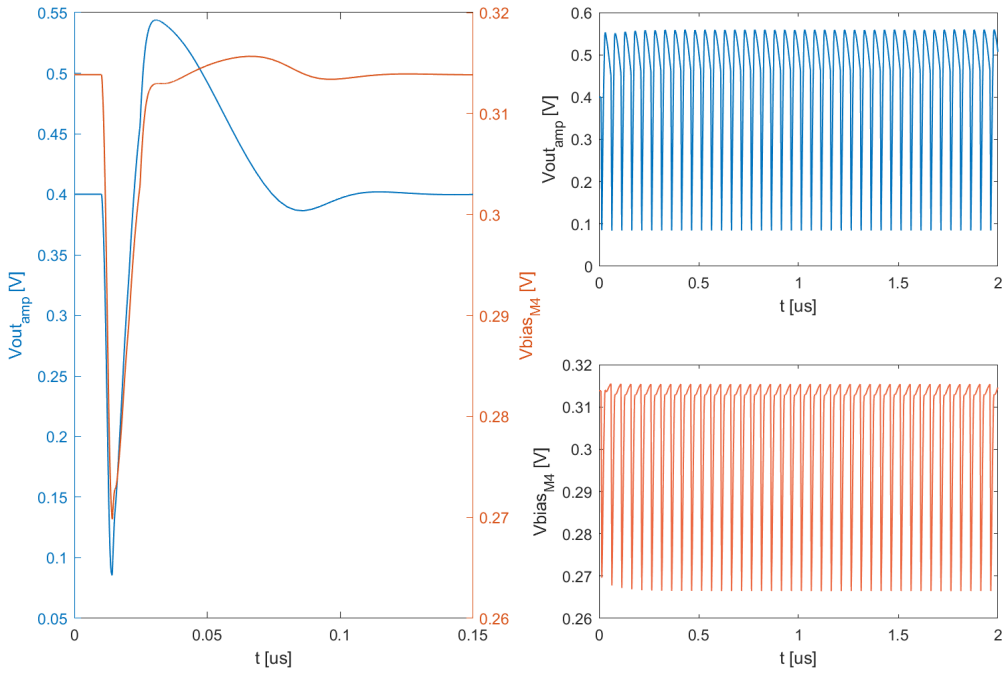


Figure 4.31: Amplifier output response with the faster feedback. Left: closer look at the response in a 150 ns period for settling. Right: the system has no heavy shift on the bias point compared to the Figure 4.23b.

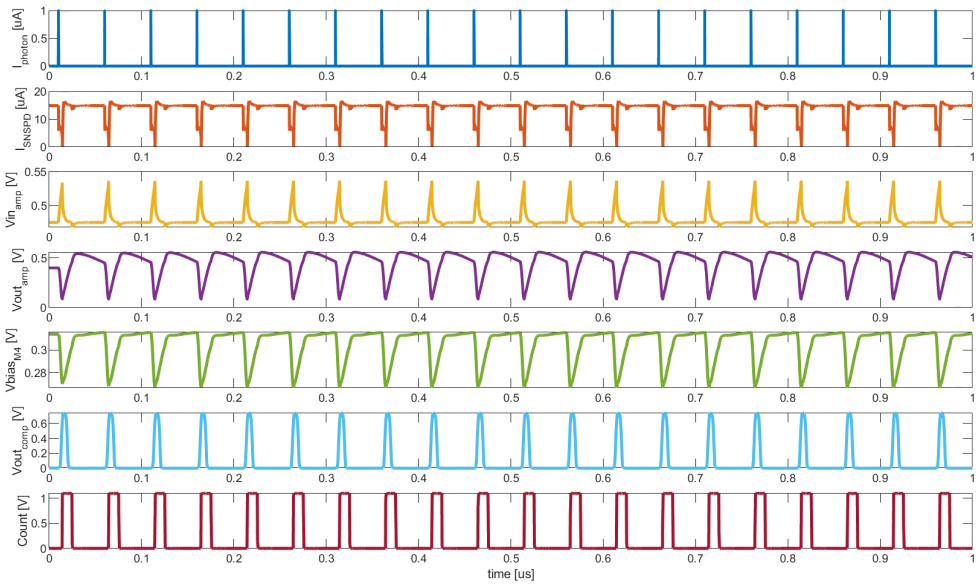


Figure 4.32: Overall system performance. Plots from top to bottom: incident photon signal I_{photon} that triggers SNSPD at 20 MHz rate; current in SNSPD I_{SNSPD}; input of the amplifier (also is the output of SNSPD) $V_{in,amp}$; output of the amplifier $V_{out,amp}$; output of the feedback amplifier (also is the bias voltage for the PMOS active load M4 of the main amplifier) $V_{out,comp}$; output of the comparator $V_{out,comp}$; count signal Q at the count rate of 20 MHz.

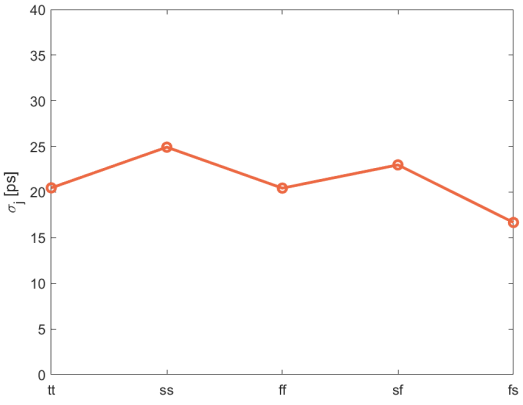


Figure 4.33: $J_{Delay_{photon-Q}}$ in different corners.

Table 4.2: SNSPD Cryo-CMOS Readout System Performance

Specification	Required	Preference	Simulated
Operating Temperature	SNSPD: 1.8 K; Readout electronics: 4 K		
System Detection Efficiency	>80%	>90%	TBD
Dark Count Rate	<5 Hz	<1 Hz	TBD
Dead Time	<100 ns	<50 ns	<50 ns
Count Rate	>10 MHz	>20 MHz	>20 MHz
Jitter	$<100\text{ps}_{\text{FWHM}}$ ($\lesssim 40\ \text{ps}_{\text{rms}}$)		$25\ \text{ps}_{\text{rms}}$
Latching	<1 μs	No latching	No latching
Power Consumption	<100 μW	As low as possible	36.49 μW
Area (for readout electronics)	<1 mm ^{2}		<1 mm ^{2}

4.5. Layout

The readout electronics for SNSPDs presented in this work is the first of its kind to be able to operate at a cryogenic temperature of 4 K using TSMC 40 nm CMOS technology. The full chip dimensions are 0.825×1.075 mm^2 shown in Figure 4.34. The readout circuit has a dimension of 0.404×0.358 mm^2 and includes an on-chip current generator, the shift registers for programming the DACs, a source follower for measuring the analog signal, and a buffer for measuring the digital count signal Q. The shift registers and current DAC occupy the most area, followed by the analog components that include the main amplifier with bias, the feedback amplifier, and the comparator. Digital circuits take the least space. The amplifiers, feedback amplifiers, and comparators in the analog circuits are all of a similar size.

The on-chip current generator produces a reference current for the current DAC. This current can also be supplied from an external source via the "Input Current" pad. The source follower can switch between the amplifier output and the comparator output for the measurement, and send the analog signal to the "Output Analog" pad. $V_{\text{back-biasing}}$ provides the body voltage for back-biasing the active quenching switch, and $V_{\text{ref,FB}}$ and $V_{\text{ref,comp}}$ supply the reference voltage for the feedback amplifier and the comparator respectively, with $V_{\text{ref,comp}}$ also used to calibrate the source follower. "Input±" denotes the two input terminals of the SNSPD that correspond to the two terminals being connected to the C_{pad} in Figure 4.1. The main amplifier is powered by a 1.1 V supply, while the other analog circuits are powered by a different 1.1 V supply. A third 1.1 V supply is used for the source follower to separate the measurement power consumption from the readout circuits, and a fourth 1.1 V supply is used for the digital circuits. Additionally, a 2.5 V supply is used for all the thick-oxide switches in the circuit.

When designing the layout, several considerations were taken into account. Generally, most of the components were designed and arranged in a square configuration to reduce the length of the wiring and decrease the paracitic capacitance. Specifically, the three core analog components (main amplifier, feedback amplifier, and comparator) should be designed and arranged in a way that the complete layout is square-like while keeping the overall W/L unchanged. This is also beneficial in increasing the via contacts and decreasing the wire length to reduce the loss. Additionally, they should be placed close to each other to further reduce the paracitic capacitance caused by the metal wires. Particular attention should also be paid to the design of active quenching switches with back-biasing. This requires the creation of an n+ guard ring within the NW region to form a valid deep n-well (dnw) region. Moreover, the substrate here must be separate from the other substrates to pass the design rule checks (DRC).

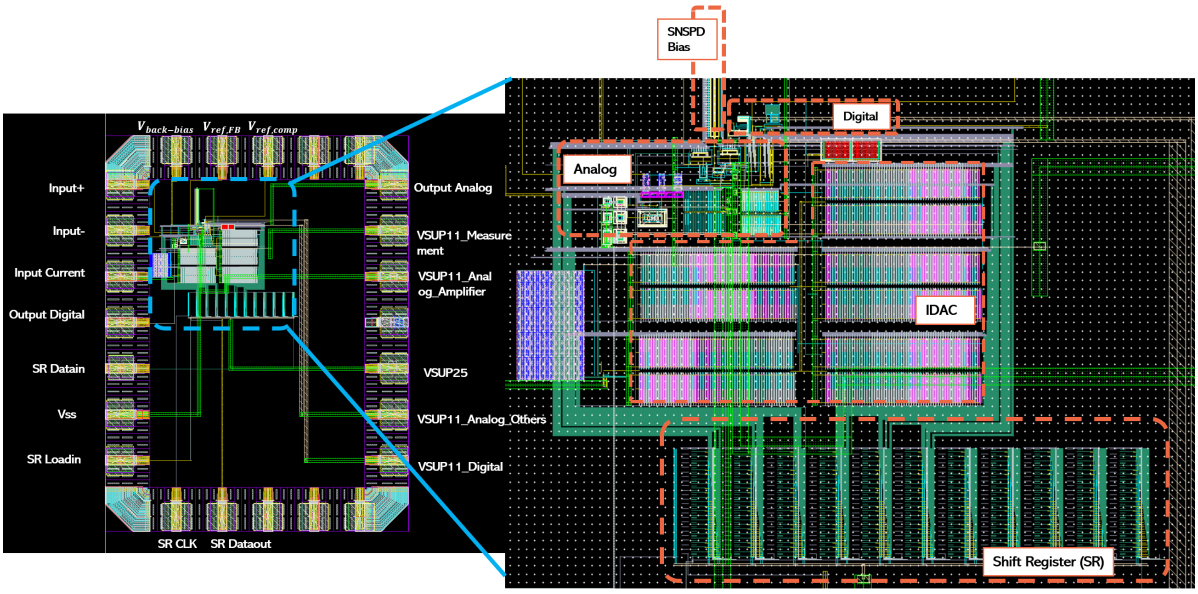


Figure 4.34: Full chip layout overview

4.6. Measurment Plan

The layout in Figure 4.34 is designed to allow multiple experiments to gain a better understanding of the system. The chip is expected to conduct several measurement plans, many of which are due to programmable components that have been incorporated into the design. An example of a jitter measurement setup is shown in Figure 4.35. The input triggers, which mimic the output of the SNSPD, are sent to both the cryogenic readout circuits and the TDC at the output of the readout circuits. This allows us to measure the time relation between the input and the readout output. It is difficult to replicate the SNSPD response, one potential solution is to employ the voltage response at a capacitor with the step-shaped voltage input.

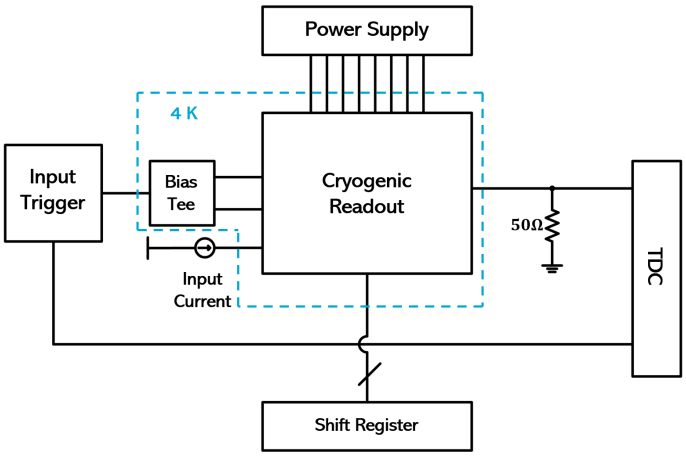


Figure 4.35: Diagram of the jitter characterization setup.

The following are some possible measurement plans:

- Varying the supply of the main amplifier to measure the limitation on the supply and the corresponding power consumption.
- Varying the supply of the other analog parts to measure the limitation on the supply and the

corresponding power consumption.

- By changing the amount of reference current produced by the current DAC, it is possible to control the delay time of the current-starve delay chain to measure the effect of the time duration on the active quenching technique.
- Varying the SNSPD bias current to observe its effect on the amplifier and the corresponding system performance.
- Evaluate the performance of the system with the five different resistances R_L.
- Examine the effect of the impedance of the closed active quenching switch by programming the size of the switches and the various back-biasing voltages.

In conclusion, the measurement plans outlined above are intended to validate the predictions and simulations discussed in Chapters 3 and 4, such as the effect of circuit parameters, including the resistance of the readout resistor R_L and the resistance of the closed active quenching switch R_{sw}. The actual jitter performance should also be compared with the simulation results to gain a better understanding of the sources of the jitter. Furthermore, the limitation of the power supply for potential power reduction can also be a subject of measurement. Additionally, the system is also expected to be connected to the SNSPD to get an initial idea of how the circuit will affect the SNSPD performance, including dark count rates and after pulses, which are not taken into account in the SNSPD model (see Chapter 3) used in this work.

5

Conclusion and Future Improvements

5.1. Main Conclusion

This thesis presented a CMOS readout circuit for SNSPDs that will be used in color-center quantum computers at a cryogenic temperature of 4 K. The proposed design aims to achieve a photon detection system with no latching and afterpulses, extremely low dark count rates, high count rates with little dead time, and low jitter while keeping the power consumption as low as possible. The thesis first investigates the basic working principles of both the NV center qubits and the SNSPD. Then, a SNSPD dynamic model based on SPICE, which takes the hotspot thermoelectric evolution into account, is reproduced in Cadence Spectre. An analisys of the influence on the readout performance from different quenching techniques and several SNSPD and electronic circuit parameters, including the SNSPD bias current, pad capacitance and bondwire inductance, readout resistance, and quenching switch resistance, is implemented. Finally, the readout circuit is designed with parameters determined by theoretical calculations with several necessary assumptions.

This design implements the active quenching technique to increase readout resistance, thereby increasing slew rate and SNR without causing latching problems. The analog readout subcircuit consists of an amplifier and a comparator to achieve the counting of SNSPD signals. Fast feedback is used to fix the bias points of the amplifier and comparator, and transistors are optimized to sufficiently reduce mismatch and minimize parasitic capacitance. Current DACs are used to bias the SNSPD, the amplifier, the comparator, and the feedback amplifier. These DACs are programmable to make the system flexible and able to function properly under different bias conditions and supply voltages. Swithches in programmable components are 2.5 V thick-oxide transistors to guarantee that they can be completely conducted when operating in cryogenic temperatures due to the rise in threshold voltage in cryogenic temperatures. Active quenching is implemented using 1.1 V low-threshold transistors, to ensure that they can be fully conducting at cryogenic temperatures, back-biasing is used to counter-act the threshold voltage increase. The digital subcircuits, consisting of an SR latch and a current-starve delay chain, are used to produce the count signal, which will also be used as the active quenching signal. In addition, different readout resistances and programmable active quenching switches are incorporated to gain a better understanding of their effect on the system's performance.

The design was simulated at -40°C and the SNSPD was biased with $I_{bias,SNSPD}$ = 15 μA, and other parameters of the SNSPD model were similar to those in [35]. The system achieved a count rate of ≥ 20 MHz, an average jitter of 25 ps_{rms} (which is expected to be even lower at cryogenic temperatures), and power consumption of 36.49 μW without any latching being observed. The design was taped out in TSMC 40 nm CMOS technology. To further evaluate the performance, including the dark count rate and afterpulses (which requires an SNSPD device), measurements will be conducted after the chips are received.

5.2. Further Improvements

There are several possible improvements can be considered:

- Chapter 3 discussed the model used in this work. However, a more accurate SNSPD model can be achieved.
  - The kinetic inductance can be made dependent on the current, which can be improved by using laboratory-measured data to fit it.
  - Examine the noise and statistical distribution of the incident photons to model the intrinsic jitter and explore the potential nonidealities that may be caused by this noisy input.
  - Incorporate the advanced model presented in [35] to account for non-idealities associated with SNSPD such as afterpulses and associated dark counts.
- The ratio in the middle branch of the SNSPD biasing depicted in Figure 4.10 does not necessarily have to be 5. A lower ratio can provide more precise and flexibility in biasing of the SNSPD.
- The simulation results suggest that the system power consumption can be reduced while still meeting the jitter requirement, as the jitter caused by thermal noise is expected to be even lower when the system is operated at cryogenic temperatures.
- The assumptions in theoretical calculation, such as the noise distribution between the amplifier and the comparator and bandwidth limitation, can be relaxed to determine if optimal operating points exist.
- Find an optional solution to the biasing point shift caused by slow feedback. And try to avoid the gain loss that this work has with the fast feedback.
- Put variable capacitors in each current-starve delay path to have better control over delay and power consumption.
- Investigate the influence of the initial condition and metastability on the system. In rare cases, the simulation fails, which may be due to an incorrect initial condition, making it difficult to determine whether the failure lies with the software, the SNSPD model, or the circuit itself.
- The layout can be more condensed and compact to further reduce the loss and paracitic capacitance within the analog subcircuits.

References

[1] F. Jazaeri, A. Beckers, A. Tajalli, and J.-M. Sallese, "A review on quantum computing: From qubits to front-end electronics and cryogenic mosfet physics," in *2019 MIXDES-26th International Conference" Mixed Design of Integrated Circuits and Systems"*, IEEE, 2019, pp. 15–25.

[2] D. P. DiVincenzo, "The physical implementation of quantum computation," *Fortschritte der Physik: Progress of Physics*, vol. 48, no. 9-11, pp. 771–783, 2000.

[3] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "A quantum engineer's guide to superconducting qubits," *Applied Physics Reviews*, vol. 6, no. 2, p. 021 318, 2019.

[4] L. Vandersypen, J. Elzerman, R. Schouten, L. Willems van Beveren, R. Hanson, and L. Kouwenhoven, "Real-time detection of single-electron tunneling using a quantum point contact," *Applied Physics Letters*, vol. 85, no. 19, pp. 4394–4396, 2004.

[5] D. D. Awschalom, R. Hanson, J. Wrachtrup, and B. B. Zhou, "Quantum technologies with optically interfaced solid-state spins," *Nature Photonics*, vol. 12, no. 9, pp. 516–527, 2018.

[6] R. Van Meter and C. Horsman, "A blueprint for building a quantum computer," *Communications of the ACM*, vol. 56, no. 10, pp. 84–93, 2013.

[7] E. Charbon, F. Sebastiano, A. Vladimirescu, *et al.*, "Cryo-cmos for quantum computing," in *2016 IEEE International Electron Devices Meeting (IEDM)*, IEEE, 2016, pp. 13–5.

[8] B. Patra, R. M. Incandela, J. P. Van Dijk, *et al.*, "Cryo-cmos circuits and systems for quantum computing applications," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 1, pp. 309–321, 2017.

[9] M. D. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov, "Invited review article: Single-photon sources and detectors," *Review of scientific instruments*, vol. 82, no. 7, p. 071 101, 2011.

[10] H. Inamori, N. Lütkenhaus, and D. Mayers, "Unconditional security of practical quantum key distribution," *The European Physical Journal D*, vol. 41, pp. 599–627, 2007.

[11] F. Ceccarelli, G. Acconcia, A. Gulinatti, M. Ghioni, I. Rech, and R. Osellame, "Recent advances and future perspectives of single-photon avalanche diodes for quantum photonics applications," *Advanced Quantum Technologies*, vol. 4, no. 2, p. 2 000 102, 2021.

[12] L. You, "Superconducting nanowire single-photon detectors for quantum information," *Nanophotonics*, vol. 9, no. 9, pp. 2673–2692, 2020.

[13] R. H. Hadfield, "Single-photon detectors for optical quantum information applications," *Nature photonics*, vol. 3, no. 12, pp. 696–705, 2009.

[14] A. J. Kerman, D. Rosenberg, R. J. Molnar, and E. A. Dauler, "Readout of superconducting nanowire single-photon detectors at high count rates," *Journal of Applied Physics*, vol. 113, no. 14, p. 144 511, 2013.

[15] E. J. van Zwet, J. Morits, F. Sebastiano, *et al.*, "Blueprint for a scalable modular quantum computer using diamond qubits," Tech. Rep., 2021, pp. 12–31.

[16] T. Van Der Sar, T. H. Taminiau, and R. Hanson, "Diamond-based quantum technologies," *Photoniques*, no. 107, pp. 44–48, 2021.

[17] S. Johnson, P. R. Dolan, and J. M. Smith, "Diamond photonics for distributed quantum networks," *Progress in Quantum Electronics*, vol. 55, pp. 129–165, 2017.

[18] C. Schreyvogel, V. Polyakov, R. Wunderlich, J. Meijer, and C. Nebel, "Active charge state control of single nv centres in diamond by in-plane al-schottky junctions," *Scientific reports*, vol. 5, no. 1, pp. 1–12, 2015.

[19] G. De Lange, Z.-H. Wang, D. Riste, V. Dobrovitski, and R. Hanson, "Universal dynamical decoupling of a single solid-state spin from a spin bath," *Science*, vol. 330, no. 6000, pp. 60–63, 2010.

[20] M. H. Abobeih, J. Cramer, M. A. Bakker, *et al.*, "One-second coherence for a single electron spin coupled to a multi-qubit nuclear-spin environment," *Nature communications*, vol. 9, no. 1, p. 2552, 2018.

[21] L. Childress and R. Hanson, "Diamond nv centers for quantum computing and quantum networks," *MRS bulletin*, vol. 38, no. 2, pp. 134–138, 2013.

[22] H. Bernien, B. Hensen, W. Pfaff, *et al.*, "Heralded entanglement between solid-state qubits separated by three metres," *Nature*, vol. 497, no. 7447, pp. 86–90, 2013.

[23] S. Ferrari, "Characterization of nbn waveguide-integrated superconducting nanowire single-photon detectors in the near-infrared range," Ph.D. dissertation, Westfälische Wilhelms-Universität Münster, 2019.

[24] C. M. Natarajan, M. G. Tanner, and R. H. Hadfield, "Superconducting nanowire single-photon detectors: Physics and applications," *Superconductor science and technology*, vol. 25, no. 6, p. 063 001, 2012.

[25] J. K. Yang, A. J. Kerman, E. A. Dauler, V. Anant, K. M. Rosfjord, and K. K. Berggren, "Modeling the electrical and thermal response of superconducting nanowire single-photon detectors," *IEEE transactions on applied superconductivity*, vol. 17, no. 2, pp. 581–585, 2007.

[26] X. Hu, N. Hu, K. Zou, Y. Meng, L. Xu, and Y. Feng, "Twenty-year research and development of snspds: Review and prospects," *Laser Technology*, vol. 46, no. 1, pp. 1–37, 2022.

[27] T. Yamashita, S. Miki, K. Makise, *et al.*, "Origin of intrinsic dark count in superconducting nanowire single-photon detectors," *Applied Physics Letters*, vol. 99, no. 16, 2011.

[28] H. Shibata, K. Fukao, N. Kirigane, S. Karimoto, and H. Yamamoto, "Snspd with ultimate low system dark count rate using various cold filters," *IEEE Transactions on Applied Superconductivity*, vol. 27, no. 4, pp. 1–4, 2016.

[29] P. Ravindran, R. Cheng, H. Tang, and J. C. Bardin, "Active quenching of superconducting nanowire single photon detectors," *Optics express*, vol. 28, no. 3, pp. 4099–4114, 2020.

[30] A. J. Kerman, J. K. Yang, R. J. Molnar, E. A. Dauler, and K. K. Berggren, "Electrothermal feedback in superconducting nanowire single-photon detectors," *Physical review B*, vol. 79, no. 10, p. 100 509, 2009.

[31] F. Garzetti, N. Lusardi, E. Ronconi, *et al.*, "Snspds next generation high-performance timing solutions based on fpga,"

[32] D. F. Santavicca, B. Noble, C. Kilgore, *et al.*, "Jitter characterization of a dual-readout snspd," *IEEE Transactions on Applied Superconductivity*, vol. 29, no. 5, pp. 1–4, 2019.

[33] P. Photonics. "Our table top detector: Focussing on precision." (2023), [Online]. Available: `https://www.pixelphotonics.com/en/products/table-top-detector/` (visited on 08/21/2023).

[34] Q. Zhao, T. Jia, M. Gu, *et al.*, "Counting rate enhancements in superconducting nanowire single-photon detectors with improved readout circuits," *Optics letters*, vol. 39, no. 7, pp. 1869–1872, 2014.

[35] K. K. Berggren, Q.-Y. Zhao, N. Abebe, *et al.*, "A superconducting nanowire can be modeled by using spice," *Superconductor Science and Technology*, vol. 31, no. 5, p. 055 010, 2018.

[36] A. J. Kerman, E. A. Dauler, J. K. Yang, *et al.*, "Constriction-limited detection efficiency of superconducting nanowire single-photon detectors," *Applied Physics Letters*, vol. 90, no. 10, p. 101 110, 2007.

[37] A. V. Gurevich and R. Mints, "Self-heating in normal metals and superconductors," *Reviews of modern physics*, vol. 59, no. 4, p. 941, 1987.

[38] P. Photonics. "Startseite | pixel photonics." (2023), [Online]. Available: `https://www.pixelphotonics.com/` (visited on 08/21/2023).

[39] R. Ishihara, J. Hermias, S. Yu, *et al.*, "3d integration technology for quantum computer based on diamond spin qubits," in *2021 IEEE International Electron Devices Meeting (IEDM)*, IEEE, 2021, pp. 14–5.

[40] S. Quantum. "Single quantum eos datasheet." (2023), [Online]. Available: `https://singlequantum.com/solutions/excellence/` (visited on 08/23/2023).

[41] M. Babaie, E. Charbon, A. Vladimirescu, F. Sebastiano, *et al.*, "Characterization and modeling of mismatch in cryo-cmos," *IEEE Journal of the Electron Devices Society*, vol. 8, pp. 263–273, 2020.

[42] A. Beckers, F. Jazaeri, and C. Enz, "Cryogenic mosfet threshold voltage model," in *ESSDERC 2019-49th European Solid-State Device Research Conference (ESSDERC)*, IEEE, 2019, pp. 94–97.

[43] A. Beckers, F. Jazaeri, A. Grill, S. Narasimhamoorthy, B. Parvais, and C. Enz, "Physical model of low-temperature to cryogenic threshold voltage in mosfets," *IEEE Journal of the Electron Devices Society*, vol. 8, pp. 780–788, 2020.

A

Addtional Basic Concepts

A.1. Kinetic Inductance

One of the essential properties of the SNSPD is its kinetic inductance. Charged carriers have mass and show inertia when an external force is applied to change their motion state. Kinetic inductance represents the inertia of the charge carrier under the influence of an external electric field [23]. For a superconducting film with a Cooper-pair mass of $m = 2m_e$ and charge $q = 2e$, the kinetic inductance is:

$$L_k = \frac{m_e}{2ne^2}\frac{l}{A} = \frac{m_e}{2ne^2}\frac{l}{wd} \tag{A.1}$$

which is proportional to the nanowire length l and inversely proportional to the width w and depth d.

A.2. Thermoelectric Dynamics for Hot-Spot Dynamics

The exact physical process that leads to the formation of a hotspot region in a nanowire after a photon is absorbed is still not completely understooded. However, it is more clear how the resistive region expands, shrinks, and eventually dissipates along the length of the nanowire due to electrothermal interaction, which is accompanied by the dynamic flow of current and the generation of voltage pulses. This thermoelectric model explains the process in which a photon is absorbed by a superconducting nanowire that is correctly biased. The energy from the photon is enough to break the Cooper pair in the superconducting material, increasing the temperature of the wire above the superconducting critical temperature. This causes the wire to transition to the normal, resistive state. The temperature of the hot-spot area increases as a result of Joule heating, and the resistive section quickly expands as a result of the diffusion of non-equilibrium quasi-particles until a portion of the wire becomes resistive. The mathematical representation of this process can be expressed by Equation (A.2), which was solved in [25].

$$J^2\rho + \kappa\frac{\partial^2 T}{\partial x^2} - \frac{\alpha}{d}(T - T_{sub}) = \frac{\partial cT}{\partial t} \tag{A.2}$$

where T is the temperature on each Δx segment of the nanowire; J is the current density through the wire; ρ is the electrical resistivity; κ is the thermal conductivity; α is the thermal boundary conductivity, and c is the specific heat per unit volume. The first term on the left is the heat produced as a result of Joule heating; the second term is the heat that is conducted away by the through the wire; the third term is the heat that is dissipated into the substrate. The term on the right is the rate of change in the local energy density.

A.3. Damping coefficient

The damping coefficient is a useful quantity for understanding the latching phenomenon in SNSPD. The self-resetting behavior of SNSPDs can be seen as a (unstable) negative electrothermal feedback to the wire by the shunt resistance R_L. The normal superconducting (NS) boundary propagates at constant

velocity v_{NS} for fixed device current I_d can be expressed as [30]:

$$v_{NS} = v_0 \frac{\alpha \left(\frac{I_d}{I_c}\right)^2 - 2}{\sqrt{\alpha \left(\frac{I_d}{I_c}\right)^2 - 1}} \approx \frac{1}{\gamma}(I_d^2 - I_{ss}^2) \tag{A.3}$$

where $I_{ss} \equiv \frac{2h(T_c - T_0)}{\rho_n}$. The physical meaning of Equation (A.3) is that the NS boundary is stationary only if the local power density ($\propto I_d^2$) is equal to a fixed value; if it is greater, the hotspot will expand ($v_{NS} > 0$), if less it will shrink ($v_{NS} < 0$).
The damping coefficient ζ can be expressed by [30]:

$$\zeta = \frac{I_0}{4I_{SS}} \sqrt{\frac{\tau_{th}}{\tau_e}} \tag{A.4}$$

where I_0 is the bias current, I_{SS} is the steady state current, $\tau_{th} \equiv \frac{R_L}{2\rho_n v_0}$, ρ_n is the normal resistance of the wire per unit length, $v_0 \equiv \frac{\sqrt{A_{cs}\kappa h}}{c}$. When at the steady state $I_d \to I_{SS}$, $R_n \to R_L(I_0/I_{ss} - 1) \equiv R_{SS}$, the above Equation (A.3) can be re-wrote as:

$$\zeta = \frac{1}{4} \sqrt{\frac{\tau_{th,tot}}{\tau_{e,tot}}} \tag{A.5}$$

with $R_{tot} \equiv R_L + R_{SS}$, $\tau_{e,tot} \equiv \frac{L}{R_{tot}}$, $\tau_{th,tot} \equiv \frac{R_{tot}}{2\rho_n v_0}$. This shows that the stability is determined by a ratio of electrical and thermal time constants.

B

SNSPD Model SpectreText Code

```
// "spectre" description for "SNSPD", "SNSPD_DM", "spectreText"


simulator lang=spectre

subckt SNSPD_DM gate gatereturn drain source N3
L1 (drain N1) inductor l=130n r=1e-100
B1 (N1 source) bsource v=(v(N3)+abs(v(N3)))/2*i("B1")
R1 (N1 source) resistor r=7.6087e5
B2 (N2 0) bsource v=((abs(i("L1"))>(1.85e-5-10*abs(i("R3"))))||(abs(v(N1)-v(source))>4.6098e-6)) ? 1 : 0
R2 (N2 0) resistor r=1
B3 (0 N3) bsource i=(v(N2)==1) ? (32.2118*(i("L1")/1.85e-5)**2-2)/sqrt(((32.2118*(i("L1")/1.85e-5)**2-1)+
abs(32.2118*(i("L1")/1.85e-5)**2-1))/2+0.005) : 0
C1 (N3 0) capacitor c=4.7336e-13
S1 (N3 0) bsource r=(v(N2)>0.5) ? 10e9 : 1e-3
R3 (gate gatereturn) iprobe

ends SNSPD_DM
```

The parameters associated with the SNSPD used in this spectretext can be found in [35]. A summary of these is presented in Table B.1.

Table B.1: SNSPD Model Parameters

Parameters	**Value**
L_k	130 nH
thickness	4 nm
width	100 nm
R_{SH}	400 $\Omega/\square$
T_c	10.5 K
T_{sub}	2 K
J_C	50 GA/m^2
κ	0.108 $\frac{W}{m \cdot K}$
c	4400 $\frac{J}{m^3 \cdot K}$
h_c	50000 $\frac{W}{m^2 \cdot K}$