Superconducting Nanowire Single-Photon Detectors: New Detector Architectures and Integration with Photonic Chips

by

Faraz Najafi

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2015

Author ..... Signature redacted .....
Department of Electrical Engineering and Computer Science
May 20, 2015

Certified by .... Signature redacted .....
Karl K. Berggren
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by ... Signature redacted .....
Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chairman, Department Committee on Graduate Theses

Superconducting Nanowire Single-Photon Detectors: New Detector Architectures and Integration with Photonic Chips

by

Faraz Najafi

Submitted to the Department of Electrical Engineering and Computer Science
on May 20, 2015, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Electrical Engineering

Abstract

Superconducting nanowire single-photon detectors (SNSPDs) are a promising technology for long-distance optical communication and quantum information processing. Recent advances in single-photon generation, storage and detection technologies have spurred interest in integration of these components onto a single microchip, which would act as a low-power non-classical optical processor. In this thesis, I will present a method for the scalable integration of SNSPDs with photonic chips. I will show that, using a micron-scale flip-chip process, waveguide-coupled SNSPDs can be integrated onto a variety of material systems with high yield. This technology enabled the assembly of the first photonic chip with multiple adjacent SNSPDs with average system detection efficiencies beyond 10%. Using this prototype, we will show the first on-chip detection of non-classical light. I will further demonstrate optimizations to the detector design and fabrication processes. These optimizations increased the direct fabrication yield and improved the timing jitter to 24 ps for detectors with high internal efficiency. Furthermore, I will show a novel single-photon detector design that may have the potential to reach photodetection dead times below 1ns.

Thesis Supervisor: Karl K. Berggren
Title: Professor of Electrical Engineering and Computer Science

Acknowledgments

The work presented in this thesis was made possible by the advice, support and contributions of many people. I would like to thank:

Professor Karl K. Berggren for teaching me about nano-fabrication and SNSPDs, and his extensive guidance, advice and support throughout grad school;

Professor Dirk Englund for his guidance throughout the last two years, and for giving me the opportunity the expand my knowledge into the area of photonics;

Professor Rajeev Ram for graciously sharing some of his equipment over the past few years, and for agreeing to be on my Thesis Committee;

Jacob Mower, for being a crucial collaborator and enabling the success of our membrane project, for performing some of the optical simulations in chapters 6 and 7, and for being a good friend;

Adam McCaughan, Dr. Qingyuan Zhao and Francesco Bellei for helping in the lab, and for providing some of the Comsol code for chapter 4;

Andrew Dane and Di Zhu for growing world-class superconducting films, and for their help in the lab;

Dr. Kristen Sunter for performing current crowding simulations, and help in the lab;

Dr. Francesco Marsili, Dr. Yachin Ivry, Nicholas Harris, Dr. Xiaolong Hu, Catherine

Lee and Prashanta Kharel for their help in the lab;

Sara Mouradian and Dr. Tim Schroeder for designing the waveguides in chapter 8;

Jim Daley and Mark Mondol for technical support in the lab;

And finally many past and present members of the Quantum Nanostructures and Nanofabrication Group and the Quantum Photonics Laboratory not mentioned here who have helped me along the way.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The rapid growth of data-driven services and the emergence of smart devices continues to drive the need for processors with increasing computational power and decreasing power consumption. On the other hand, down-scaling electrical logical circuit elements is becoming more challenging. These challenges have sparked interest in photonic circuits as a complementary technology. Photons as carriers of information offer many advantages compared to electrons. They exhibit low scattering and no resistive heating, making them suitable for low-power circuits and novel non-classical approaches to computation. Figure 1-1 illustrates a vision of an optical quantum processor. Single and entangled photons are generated by non-classical sources. A series of tunable optical interferometers act as logic gates. Finally, the output state is detected using an array of single-photon detectors - this stage is equivalent to the sensing of voltage in an electronic processor.

The focus of my work, presented in this thesis, has been the development of processes and technologies for the integration of high-speed single-photon detectors with nanoscale photonic chips. It comprised the new designs of single-photon detectors and the assembly of an optical chip that allowed the first on-chip detection of non-classical light, paving the way towards future optical quantum processors.

Besides their potential to enable chip-scale optical quantum computing, single-photon detectors have gained increasing popularity due to their role in the emerging applications quantum key distribution (QKD) [7], quantum computing and simulation

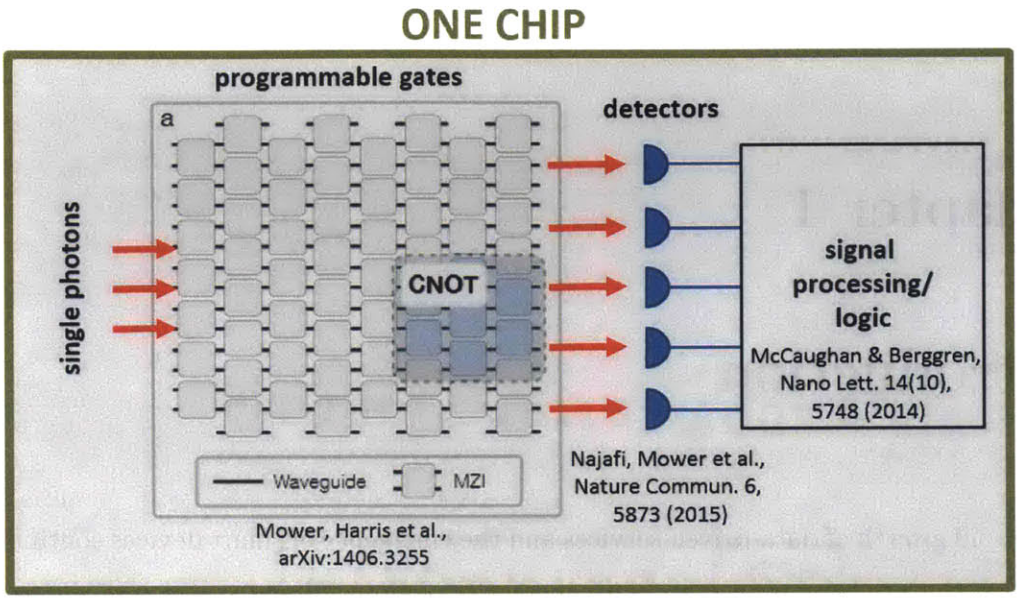

Figure 1-1: The vision of a photonic quantum processor that contains single-photon sources, gates, single-photon detectors and postprocessing electronics on a single chip.

[8, 9, 10] and long-distance optical communication [11, 12].

For the applications mentioned above, several detector performance metrics are of importance: detection efficiency, which limits the spacial sender-to-receiver distance for communications applications and security for QKD applications; dead time, which limits the data rate; and timing jitter, which limits the overall clock rate. Other important metrics are active detector area, which is an issue for free-space-coupled and fiber-coupled detector systems, and scalability, which is important for single-photon-level imaging and chip-scale optical computing. In this chapter, we will review some of these metrics with focus on single-photon detector technologies based on superconducting nanowires. Later in this thesis, we will present how to improve on some of the metrics discussed here, particularly speed and scalability.

1.1 Superconducting Nanowire Single-Photon Detectors (SNSPDs)

Since their invention in 2001 [13] SNSPDs have emerged as a promising technology for long-distance telecommunication [11, 12] and quantum key distribution [7]. SNSPDs with detection efficiencies up to 93% [14] for 1550nm-wavelength photons, dark count rate of 100 counts per second [15], timing jitter down to 18 ps [16] and wideband spectral sensitivity between the UV and mid-IR have been demonstrated [2]. The typical structure of an SNSPD is shown in Figure 1-2. The SNSPD consists of narrow (<150 nm) and thin (<10 nm) superconducting nanowires patterned in a meander structure. Common materials used for SNSPDs are niobium nitride (NbN), niobium titanium nitride (NbTiN) and tungsten silicide (WSi).

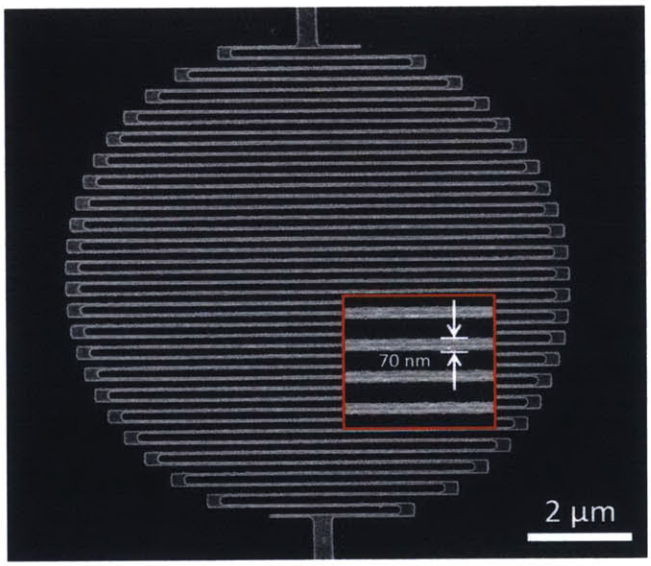

Figure 1-2: Top-down scanning-electron micrograph (SEM) of an SNSPD based on 70-nm-wide nanowires covering an circular active area ~9 µm in diameter. The inset shows a magnified SEM of the nanowire pattern.

1.1.1 The hotspot model

Fig 1-3 illustrates the events that result in single-photon detection in superconducting nanowires according to the hotspot model. When a photon is absorbed by a Cooper pair in the superconducting nanowire, it generates highly-excited quasiparticles (hot electrons [17]). These highly-excited electrons transfer their energy through scattering and break up additional Cooper pairs, generating a localized region filled with secondary quasiparticles. This normal region is called the hotspot [17]. Fig. 1-3(a) illustrates a photon-generated hotspot nucleation event (HSN) in a superconducting nanowire. For NbN, the hotspot is believed to be $\sim$ 30 nm in diameter for 1550-nm-wavelength photons [1]. The hotspot repels the supercurrent to sidewalls around it, increasing the current density in the sidewalls. If the initial bias current in the nanowire is high enough so that the current density in the side walls surpasses the critical current density J_C, the sidewalls will switch to the normal state as well, resulting in a resistive slab across the entire width of the nanowire (Fig. 1-3(b)). The resistive region grows in length due to Joule heating (Fig. 1-3(c)). The current is diverted out of the nanowire into the readout electronics (see next section), allowing for the resistive slab to cool (Fig. 1-3(d)) and to return to the superconducting state. At this point the nanowire can detect the next photon.

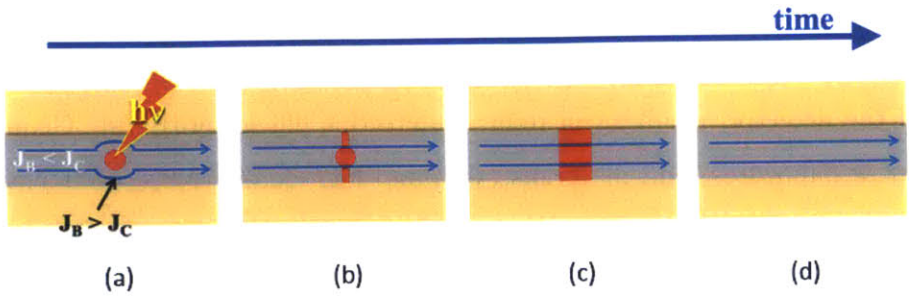

Figure 1-3: Sketch of photodetection process inside a current-biased superconducting nanowire according to the hotspot model. (a) A photon is absorbed in the nanowire, creating a resistive region (hotspot). (b) The sidewalks around the hotspot become resistive as a result of current redistribution in the nanowire. (c) Joule heating results in a growth of the resistive region. (d) The resistive region eventually cools down and returns to the superconducting state.

1.1.2 Detection efficiency and constrictions

The overall detection efficiency of a detector system, referred to as system detection efficiency SDE, can be expressed as

$$\mathrm{SDE} = \eta_{\mathrm{coupling}} \cdot \eta_{\mathrm{absorb}} \cdot \eta_{\mathrm{internal}} \cdot \eta_{\mathrm{trigger}}, \tag{1.1}$$

where η_{coupling} is the coupling efficiency between the optical mode area and the active detector area, η_{absorb} is the optical absorption for light incident onto the detector area, η_{internal} is the internal probability of a photon absorption event resulting in a hotspot nucleation event, and η_{trigger} is the probability of registering the detector signal (photodetection pulse) in the readout electronics. SDE values up to 93% [14] have been demonstrated using fiber-coupled SNSPDs. The product $\eta = \eta_{\mathrm{absorb}} \cdot \eta_{\mathrm{internal}}$ is generally referred to as 'device detection efficiency'. Fig. 3-5 shows the device detection efficiency vs. bias current of SNSPDs based on 30-nm-wide nanowires. In sub-50-nm-wide SNSPDs based on NbN [1] and in WSi SNSPDs [14], it is generally observed that when illuminated with 1550-nm-wavelength light, at low bias currents I_{B} the detection efficiency η increases with I_{B} while at high bias currents the detection efficiency reaches a plateau where η shows a small dependence on I_{B}. A detection efficiency

curve that reaches a plateau is sometimes referred to as saturated. Saturated detectors are a sign of high internal efficiency [18, 14]. Constrictions limit the switching current of the detector, resulting in reduced detection efficiency [1, 19]. Geometries (ultra-narrow nanowires) and materials (WSi and high-quality NbN) that yield saturated detectors are more robust towards constrictions: the purple curve in Fig. 3-5 shows a constricted 30-nm-nanowire SNSPD. Due to a large saturation plateau, the constricted detector can still reach an efficiency value comparable to a less constricted detector (red curve in Fig. 3-5). The inflection point of the η vs. I_B curve is referred to as cutoff current I_{CO}.

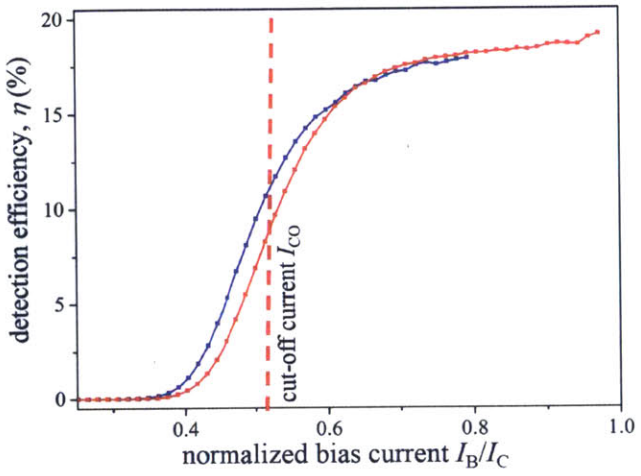

Figure 1-4: Device detection efficiency as a function of the bias current (normalized by the critical current of the less constricted detector) for 30-nm-nanowire-width SNSPDs. The detectors were illuminated with 1550-nm-wavelength light, as outlined in Ref. [1]. The critical current of the less constricted detector (red curve) was used for the calculation of the normalized bias current.

1.1.3 Speed limit and latching

A simplified lumped-element circuit model for an SNSPD is shown in Fig. 1-5(a). The speed of current dynamics in the detector is limited by the kinetic inductance L_K [20] of the superconducting nanowire, which is significantly larger than its geometric inductance. The kinetic inductance limits the current recovery time τ from the load into the detector to $\tau \sim L_K/50\Omega$. The dead time of NbN SNSPDs (sometimes referred to as reset time) can be estimated as $t_D \sim 3 \cdot \tau$ [1]. However, this definition is somewhat outdated. A more accurate definition of the dead time is illustrated in Fig. 1-5(d), which shows the inter-arrival time histogram of two subsequent detector pulses. The dead time t_D can be defined as the time delay at which the inter-arrival time delay histogram reaches 90% of its peak value. While qualitatively useful and intuitive, the simple circuit model is not sufficient to explain the current dynamics in SNSPDs and more complex multi-nanowire detectors. Furthermore, the lumped-element circuit can not explain the speed limit of SNSPDs due to a stable resistive slab, a behavior referred to as 'latching'. A more advanced approximation is the

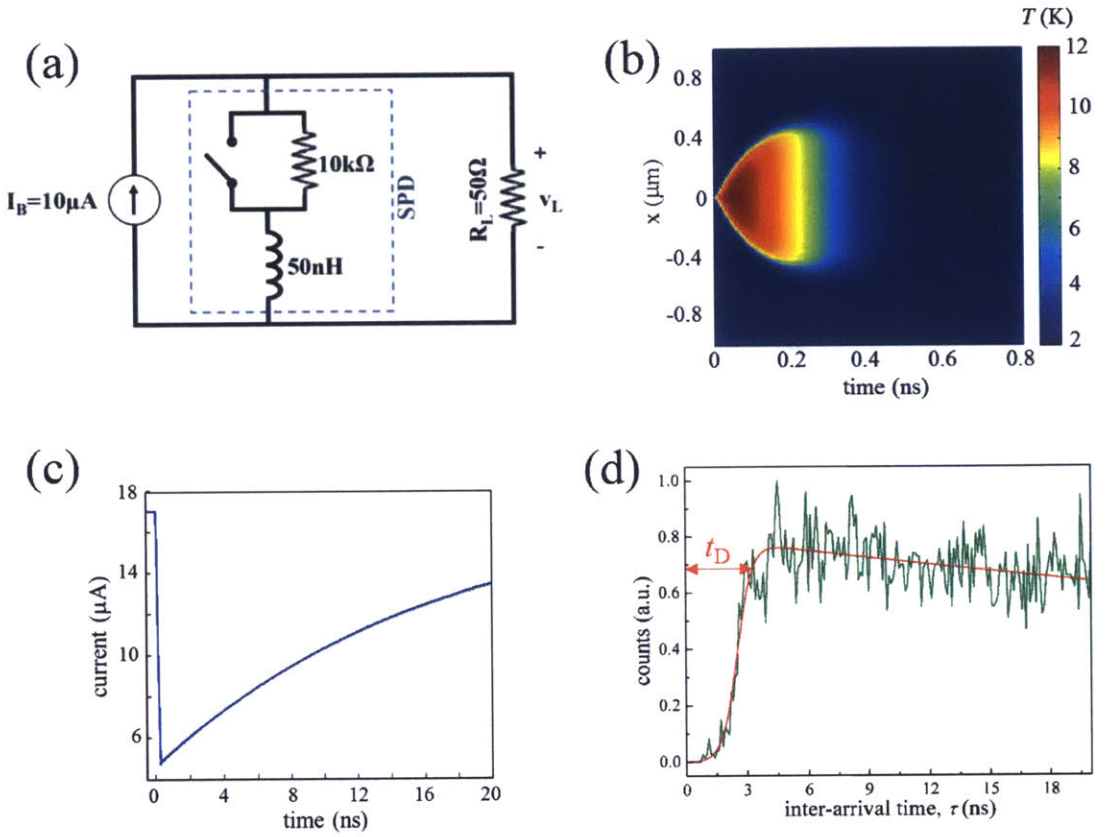

Figure 1-5: (a) Simple circuit model for an SNSPD after photon absorption (open switch). The impedance of the readout electronics is modeled with the load resistor $R_{\mathrm{L}} = 50\Omega$. The inductor in series models the kinetic inductance of the SNSPD. Estimates of typical inductance and resistance values are also listed in the circuit diagram. (b) Calculated time-dependent temperature distribution along an NbN nanowire. After the creation of a resistive slab at t = 0 s, the nanowire temperature increases due to joule heating and reaches a maximum value of 12 K. After ~100 to 200 ps, the resistive region cools back down to the substrate temperature. (c) Simulated time-dependent detector current I after the creation of a resistive slab at t = 0 s. (d) Measured pulse-to-pulse inter-arrival time histogram of a nanowire single-photon detector (4-SNAP, shown in green) and a fit as in Ref. [1] (shown in red). The dead time $t_D \sim 3.3$ ns was extracted from the fit as the inter-arrival time at which the count rate reaches 90% of its maximum value.

electrothermal (ET) model, a macroscopic model that quantitatively describes the formation of a measurable photodetection pulse in the readout circuit, the exponential recovery of the photodetection pulse and latching. The ET model for SNSPDs was first introduced by Yang et al. [21]. It is based on two coupled equations. The first equation is the time-dependent heat equation

$$J^2\rho + \kappa\frac{\partial^2 T}{\partial x^2} - \frac{\alpha}{d}(T - T_{\text{sub}}) = \frac{\partial cT}{\partial t} \tag{1.2}$$

and describes the thermal dynamics of the resistive region governed by Joule heating ($J^2\rho$, where ρ is the NbN resistivity), cooling through diffusion inside the NbN (κ is the thermal conductivity of NbN and c is the specific heat of NbN) and cooling through the substrate (α is the thermal conductivity between NbN and the substrate). The second equation is the differential equation for the circuit shown in Fig. 1-5 [21]. Both equations are coupled through the time-dependent detector resistance $R(t)$. The ET model is now frequently used in the field to design detectors based on superconducting nanowires. Figure 1-5(b) shows the time-dependent local temperature along an NbN nanowire after the formation of a resistive slab following the absorption of a photon. Joule heating results in the growth of the resistive region, which results in the diversion of I_B out of the SNSPD and into the load, as shown in Fig. 1-5(c). The resistive region returns to the superconducting state within a characteristic time on the order of $\sim$ 100 ps - however, due to the kinetic inductance, it takes $\sim$ 1-10 ns for the current through the nanowire to return to its initial value I_B.

1.1.4 Mid-IR detection

One of the advantages of SNSPDs over other single-photon detector technologies is their sensitivity over a broad spectral range. Figure 1-6 shows the device detection efficiency of 30- to 85-nm-wide NbN nanowires as a function of wavelength and bias current for unpolarized light. The detector enters a high-efficiency detection regime for $I_B > I_{CO}$. This regime extends to larger wavelengths for narrower nanowires (up to 5 μm in wavelength for 30-nm-wide NbN nanowires). However, ultranarrow

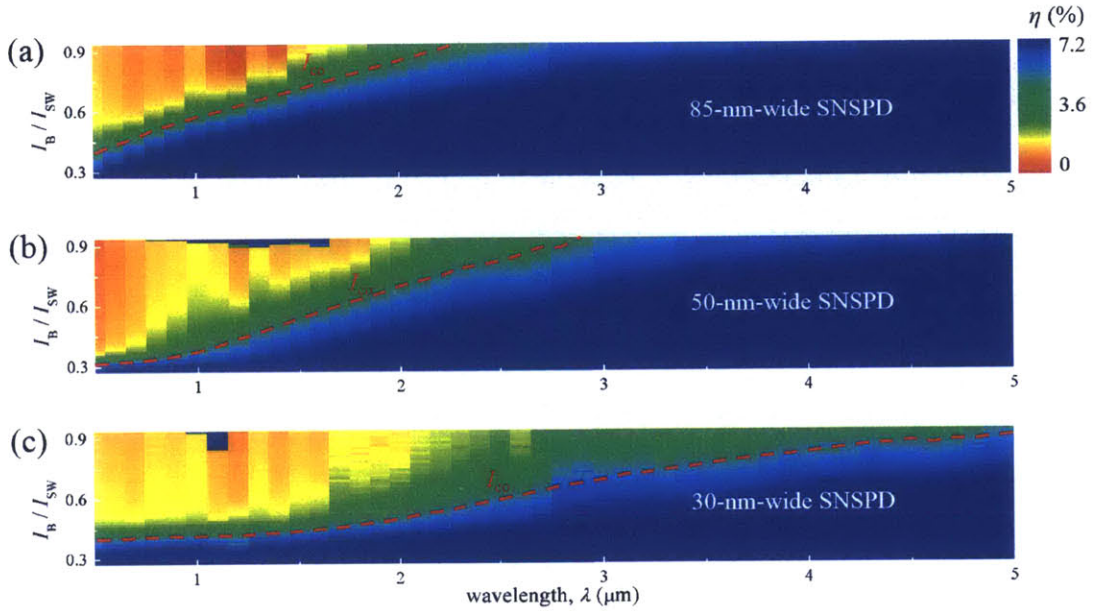

Figure 1-6: 2D thermal map of device detection efficiency η as a function of wavelength λ and normalized bias current (bias current I_B divided by switching current I_{SW}) for SNSPDs based on (a) 85-nm-wide ($I_{\mathrm{SW}} = 20.6\mu\mathrm{A}$), (b) 50-nm-wide ($I_{\mathrm{SW}} = 9.3\ \mu\mathrm{A}$), and (c) 30-nm-wide NbN ($I_{\mathrm{SW}} = 7.4\mu\mathrm{A}$) nanowires. The cutoff current I_{CO}, which increases with increasing wavelength, is marked with red dashed lines. Adapted from Ref. [2]

nanowires are more challenging to fabricate over large areas with high yield. Amorphous WSi has a number of properties that make it a desirable superconducting material for fabrication of large-area of SNSPDs [14, 22, 23]. The reduced carrier density and larger hotspot size in WSi allows the nanowires to be wider than NbN-based nanowires, which considerably improves device yield due to a lower probability of constriction, making WSi a more promising choice for mid-IR single-photon detection. To date, achieving high system detection efficiency at mid-infrared wavelengths remains a challenge due to large mode area of multimode mid-IR fibers and the lack of polarization control of these fibers.

1.1.5 Performance tradeoffs

The different performance metrics (dead time, timing jitter, system detection efficiency, yield) can not be independently optimized due to their inter-dependence, and there are several trade-offs that have to be considered when designing SNSPDs.

System efficiency and speed: In order to achieve high system detection efficiency, it is desirable to maximize the active area of the detector to simplify optical coupling. However, as the kinetic inductance of a nanowire is proportional to its length, the overall dead time of an SNSPD increases with increasing active area. As we will discuss in the next section, multi-nanowire architectures can result in a speed-up compared to conventional SNSPDs.

System efficiency and detector yield: While a larger detector area can enable larger optical coupling efficiency, it also increases the probability of constrictions along the nanowire. Detectors based on amorphous superconducting films (WSi, MoGe, MoSi) have been shown [22, 24, 25] are likely more robust towards constrictions than NbN, as discussed in the previous section.

Signal-to-noise ratio and timing jitter: While SNSPDs based on narrow NbN nanowires and amorphous superconductors offer large saturation plateaus, the smaller detector bias current results in a smaller signal-to-noise ratio (SNR). The limited SNR makes detector readout more challenging [14] and results in increased timing jitter [26]. By optimizing the nanowire width and employing multi-nanowire detector architectures (see next section), this tradeoff can be relaxed.

1.2 Superconducting nanowire avalanche photodetectors

Superconducting nanowire avalanche photodetectors (SNAPs, also referred to as cascade switching superconducting single photon detectors) [27] are based on a parallel-nanowire architecture that allows single-photon counting with higher signal to noise ratio (SNR, up to a factor of ~ 4 higher [1]) than SNSPDs.

1.2.1 The avalanche regime

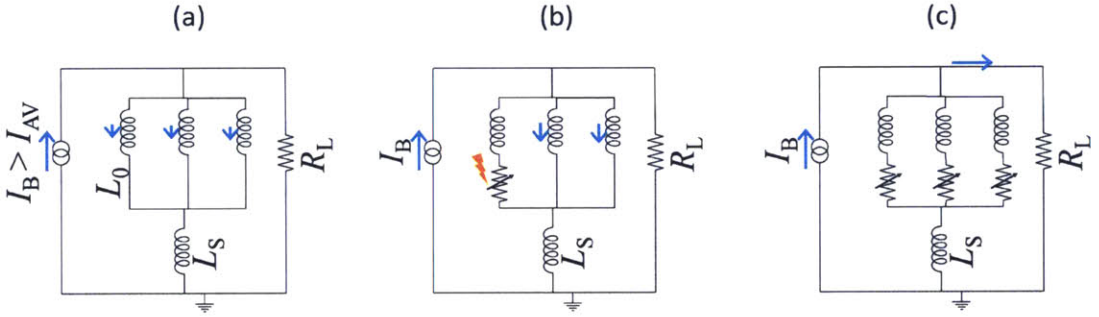

Figure 1-7: Circuit model of a 3-SNAP biased above avalanche current I_{AV}. (a) All three sections are biased at $I_B/3$. (b) The absorption of a photon drives one of the SNAP sections (initiating section) into the normal state. (c) The current redistribution drives the remaining sections (secondary sections) into the normal state (avalanche), resulting in a current redistribution into the load and a measurable voltage pulse across R_L.

Figure 3-3 illustrates the operation of SNAPs. In the equilibrium state (Fig. 3-3(a)) all N sections are biased at I_B/N. The detection event begins with the absorbtion of a photon in one of the SNAP sections, shown in Fig. 3-3(b). This section, referred to as initiating section, becomes resistive, diverting its current to the nanowires that are electrically connected in parallel to it. These sections are called secondary sections. If the initial bias current (I_B) of a SNAP with N parallel sections (called N-SNAP) is

higher than the avalanche threshold current (I_{AV}), the current diverted to the $(N-1)$ secondary sections is sufficient to switch these sections to the normal state (Fig. 3-3(c)). We call this process an avalanche. As a result, a current $\approx N$-times higher than the current through an individual section is diverted to the read out [27]. Hence, in avalanche regime the SNAPs operate as single-photon detectors, i.e. a single hotspot nucleation event (HSN) is sufficient to trigger an avalanche and therefore a measurable detector pulse. Fig. 1-8(a) shows the SEM of a 3-SNAP resist mask based on 30-nm-wide nanowires, resulting in increased SNR, as illustrated in Fig. 1-8(b). Fig. 1-8(c) shows the detection efficiency of an SNSPD and a 2-, 3- and 4-SNAP as a function of normalized bias current $I_{\mathrm{B}}/I_{\mathrm{SW}}$. The last inflection point in the η vs. I_B curve is the avalanche current I_{AV}. The normalized avalanche current $I_{\mathrm{AV}}/I_{\mathrm{SW}}$ increases with increasing N. In practice, we were unable to achieve stable operation of detectors with $N > 4$ in avalanche regime.

1.2.2 The arm-trigger regime

When biased below I_{AV}, the current diverted from the initiating section is not sufficient to trigger an avalanche. For an avalanche to form, additional HSNs have to occur in the secondary sections. Fig. 1-9 shows the detection efficiency vs. bias current-curve for a 3-SNAP [1].

When biased below I_{AV} at least two HSNs are necessary to trigger an avalanche in this 3-SNAP. We refer to this operation condition as arm-trigger regime, illustrated in Fig. 1-10(a-c): the first HSN (Fig. 1-10 (a)) 'arms' the 3-SNAP, activating the remaining two sections. These still-superconducting secondary sections operate as a pseudo-2-SNAP, i.e. they will generate an avalanche if an additional HSN (Fig. 1-10(b)) occurs in one of them. In this regime the SNAP operates as a low-jitter multi-photon gate [6] rather than a single-photon detector. Fig. 1-10(d) shows the simulated current dynamics for a 3-SNAP in arm-trigger regime. The arming HSN suppresses the current I_1 in the initiating section and results in increased current flow $I_{2,3}$ through the secondary sections. A second HSN in one of the secondary sections results in an avalanche, diverting the current from the detector into the load. The

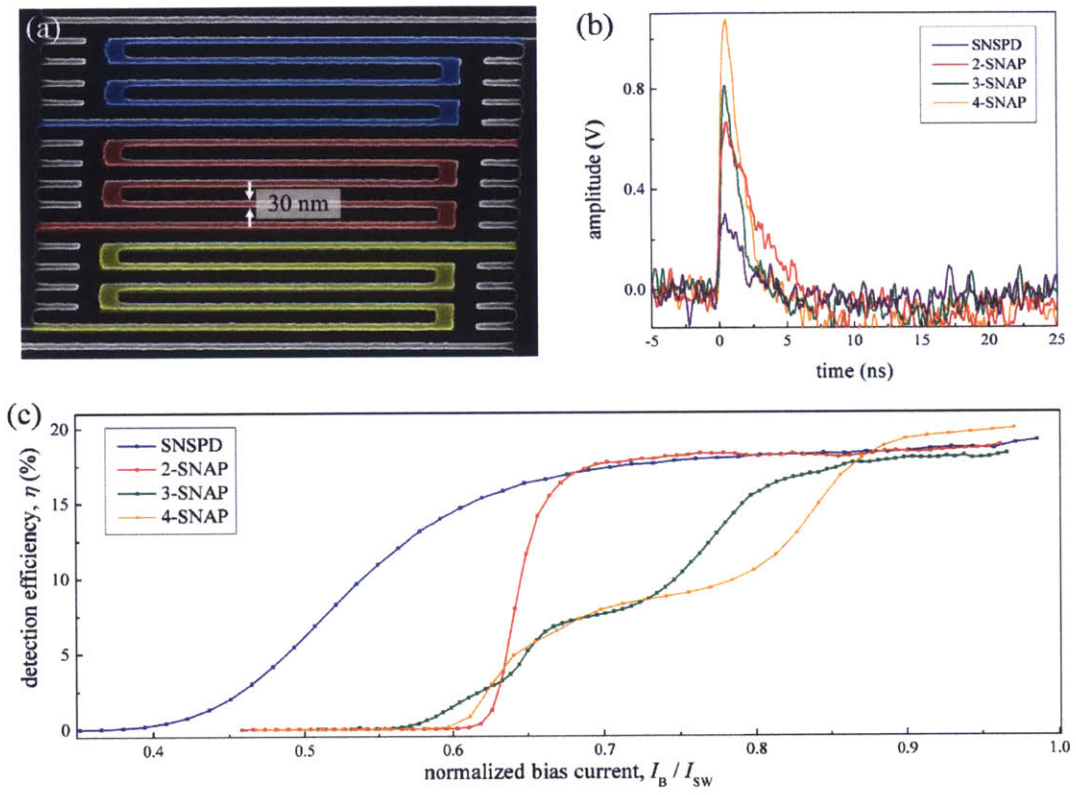

Figure 1-8: (a) False-color SEM of a resist mask representing a 30-nm-nanowire-width 3-SNAP. Each section colored differently. (b) Voltage traces of detector pulses of an SNSPD, a 2-SNAP, a 3-SNAP, and a 4-SNAP based on 20-nm-wide nanowires (in black, red, green, and blue, respectively), showing increasing signal-to-noise ratio as the number N of SNAP sections is increased. (c) Device detection efficiency η at 1550 nm wavelength as a function of normalized bias current (I_B/I_{SW}) for an SNSPD (I_{SW} = 7.2 μA, purple trace), a 2-SNAP (I_{SW}=13.4 μA, red trace), a 3-SNAP (I_{SW}=18.1 μA, green trace), and a 4-SNAP (I_{SW}=28.4 μA, orange trace). All detectors were based on 30-nm-wide nanowires. Adapted from Ref. [1]

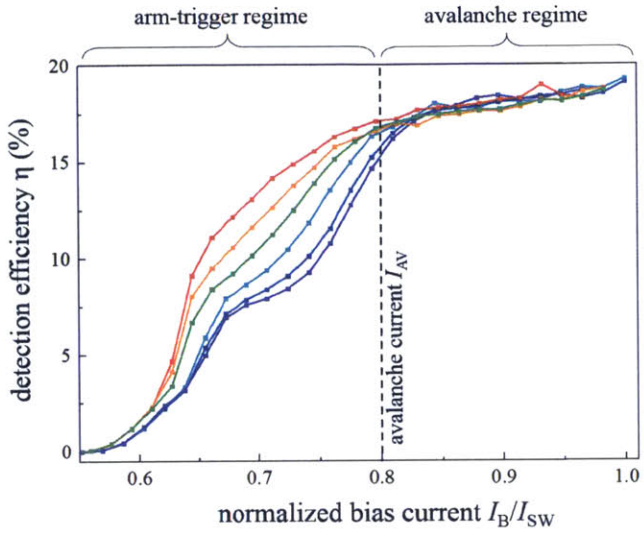

Figure 1-9: Device detection efficiency vs. bias current curves of a 3-SNAP measured at photon fluxes ranging from 0.6 (red curve) to 19 million photons per second (purple curve)[1]. In the avalanche regime, the detection efficiency is independent of the incident photon flux.

diverted current, I_{out}, results in a measurable voltage pulse across R_{L}.

1.2.3 Multi-stage SNAPs

As discussed in the previous sections, in N-SNAPs the leakage current increases with increasing N. As a result a larger current is required to generate an avalanche, and $I_{\text{AV}}/I_{\text{SW}}$ grows with N. Experimentally, the maximum number of SNAP sections in parallel with stable avalanche has been limited to 4 sections ($I_{\text{AV}}/I_{\text{SW}}$ ~0.9 for a 4-SNAP). One approach to circumventing this limit is to implement a multi-stage SNAP structure [3] that allows for a smaller avalanche current $I_{\text{AV}}/I_{\text{SW}}$. Fig. 1-11 shows the structure of a multi-stage 2-SNAP. In this SNAP, a pulse is not generated through a single avalanche but rather through three cascaded avalanches. An HSN results in a first avalanche in a 2-SNAP, marked in blue. The now resistive 2-SNAP in connected in parallel to an identical 2-SNAP and a series inductor L_{S2}, resulting in the re-distribution of the current from the resistive 2-SNAP into the still-superconducting 2-SNAP connected in parallel to it (second avalanche, marked with red arrow). In

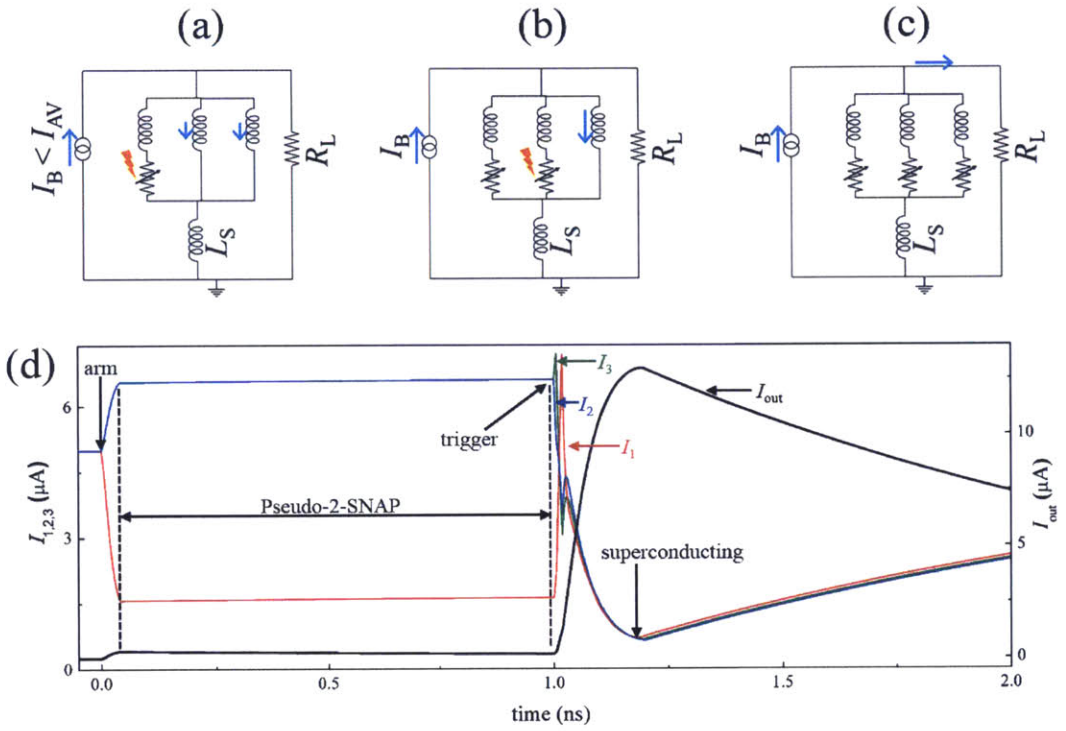

Figure 1-10: (a) Circuit model of a 3-SNAP biased below I_{AV}. An initial HSN drives the initiating section into the normal state. However, the current redistribution is not sufficient to drive the secondary sections into the normal state. (b, c) A second HSN in a secondary section occurs, triggering an avalanche. (d) Simulated time-dependent current through the initiating (arming) section (I_1), the secondary section where the second (trigger) HSN occurs (I_2) and the secondary section that switches to the resistive state (I_3) following current redistribution from the first two sections. The second HSN results in an avalanche and current redistribution from the SNAP into the load (I_{out}, shown in black). The detector was biased below I_{AV}. Adapted from Ref. [2]

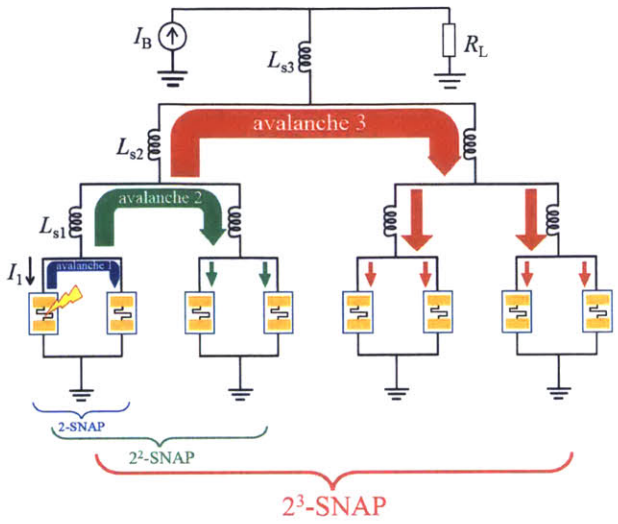

Figure 1-11: Circuit model representing a 3-stage 2-SNAP (2^3-SNAP). The first, second and third avalanches are illustrated with the blue, green and red arrows, respectively. Adapted from Ref. [3]

avalanche regime, the current redistribution drives the second 2-SNAP into the resistive state as well. Subsequent avalanches follow the same scheme. The SNR of the three-stage 2-SNAP, referred to as 2^3-SNAP, is comparable to a traditional 8-SNAP [3], with the advantage of a lower avalanche current: the relative avalanche current of the 2^3-SNAP is comparable to the value of a 2-SNAP, while an 8-SNAP would be extremely challenging to realize due to its high expected relative avalanche current $I_{AV}/I_{SW} > 1 - 1/N$ [1].

1.2.4 Device yield

Physical constrictions and defects limit the bias current of superconducting nanowires. As discussed in section 1.1.2, a limited bias current range of a severely constricted detector can ultimately prevent the detector from reaching the saturation plateau if the constricted critical current I_{cc} is smaller than the cutoff current I_{CO}. As discussed in section 1.2.2, a similar argument applies to SNAPs: a severely constricted detector with $I_{cc} < I_{AV}$ will not operate in the arm-trigger regime. As shown in Fig. 1-12, a saturated SNAP with $I_B \gg I_{AV}$ can operate in a sub-100-ps-jitter regime. It is

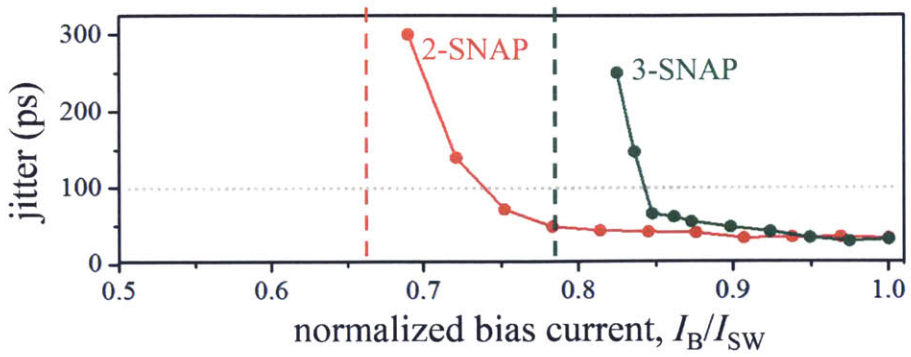

Figure 1-12: Timing jitter of a 2- and 3-SNAP based on 30-nm-wide nanowires as a function of normalized bias current. The vertical dashed lines represent the corresponding avalanche currents I_{AV}. Adapted from Ref. [4]

desirable for SNAPs to operate in this regime where both the efficiency (Fig. 1-8(c)) and the timing jitter (Fig. 1-12) show a low dependence on bias current. For m measured saturated detectors, we can therefore define yield as the ratio n/m, where n is the number of detectors that have a switching current higher than the cutoff current (SNSPDs) or avalanche current (SNAPs). In this high-efficiency regime, SNAPs typically exhibit sub-100-ps timing jitter.

1.3 Outline of this thesis

The previous sections discussed some of the detector architectures based on superconducting nanowires and the different performance metrics. For chip-scale applications, these detectors are to be integrated with photonic chips.

Photonic integrated circuits (PICs) are optical circuits based on waveguides. PICs allow optical experiments to be performed in a compact and scalable manner on a small photonic chip. PICs are being developed for a wide range of applications in quantum information science, including quantum simulation [8, 28], quantum photonic state generation [29, 30, 31], quantum-limited detection [32], and linear optical quantum computing [33, 34]. A central goal has been to integrate single-photon sources and photon-resolving detectors into PICs to reduce optical losses, latency,

and wiring complexity associated with off-chip components. For high data and clock rates multiple detectors with low timing jitter and high efficiency have to be integrated onto the photonic chip.

Two main challenges have to be overcome when integrating SNSPDs with PICs. The first is the incompatibility of the SNSPD fabrication process with a variety of PIC processes, e.g. photonic chips fabricated using a standard CMOS process. The second challenge is the low fabrication yield α of high-performance SNSPDs. The probability of having n high-performance adjacent SNSPDs is then α^n (assuming the yield is due to independent random events), which yields only a $\sim$35% probability for 10 adjacent detectors and 90% yield. In practice, the direct fabrication yield of NbN SNSPDs, as defined in chapter 2, is around 10-30%, significantly smaller than the required $\alpha \gg 90\%$ for scalable integration of SNSPDs. For practical purposes, the developed processes should enable the integration of SNSPDs with CMOS-compatible material systems.

This thesis presents a method for the scalable integration of SNSPDs with photonic chips. Using a micron-scale flip-chip process, waveguide-integrated SNSPDs can be integrated onto a variety of material systems with unity effective yield. Using this technology, we will demonstrate a prototype photonic chip that enabled the first on-chip detection of non-classical light using multiple adjacent SNSPDs with average system detection efficiencies beyond 10%, an improvement by two orders of magnitude compared to previous demonstrations of on-chip systems with multiple SNSPDs. I will begin with a review of the SNSPD fabrication process, followed by improvements to the process and detector design. I will show how these optimizations, besides improving the direct fabrication yield beyond 70%, help improve device performance such as sub-30-ps jitter, sub-1-ns reset time and large active areas.

Chapter 2

Optimizing the detector fabrication process

SNSPDs offer high detection efficiency [14, 35, 15], sub-50-ps timing jitter [36, 6] and sub-50-ns dead time [20] for near-infrared photons. The unique combination of these characteristics has led to rising interest and development towards SNSPD arrays for high-speed free-space telecommunication [12, 11], imaging applications with single-photon sensitivity [37, 38, 39] and on-chip photonic quantum processors [40]. These applications require a large number of high-performance SNSPDs. One of the core challenges in scaling up the number of detectors has been the nanofabrication yield, which can be small for low-jitter SNSPDs due to non-uniformities in the superconducting film and defects introduced during fabrication [19]. In this chapter we present an optimized fabrication process for SNSPDs designed to reduce processing-related defects in the superconducting film and the nanowires. The updated process allowed us to obtain a yield of ~70% for detectors based on 80-nm-wide NbN nanowires, compared to a ~10-30% yield using the old fabrication process.

2.1 Conventional detector fabrication process

In this section we will briefly review the old SNSPD fabrication process as outlined in Refs. [1, 41, 42]. A ~5-nm-thick NbN layer (thickness estimated from deposition

time) was deposited on top of double-polished R-plane sapphire substrates via DC reactive magnetron sputtering [42] at 900°C. The films were covered with ~1-nm-thick native oxide. The first fabrication step was the deposition of electrical contact pads via liftoff, as illustrated in Figure 2-1. The sample was covered with a ~1-μm-thick positive photoresist layer (Shipley S1813, spun at 5.5 krpm followed by a 3-minute bake at 90°C). The resist was exposed with UV light at $2900\,\mu W/cm^2$ and developed in MS352 (a sodium-hydroxide-containing commercial resist developer solution) for 3 min, resulting in exposed areas for electrical contact pads (Fig. 2-1(c)). In order to facilitate electrical probing of and wire-bonding to the detectors a 5-nm-thick titanium layer and a 20-nm-thick gold layer were deposited via e-beam evaporation (Fig. 2-1(d)). The resist liftoff (Fig. 2-1(e)) was performed in n-methyl-2-pyrrolidone (NMP) at 85°C for 5 min. As outlined in Ref. [41], in order to create a resist undercut and simplify liftoff, the exposed photoresist-covered sample was dipped in chlorobenzene for 15 min before development. This lengthy liftoff process often resulted in incomplete gold liftoff or rough pad edges (see Fig. 2-3(a)).

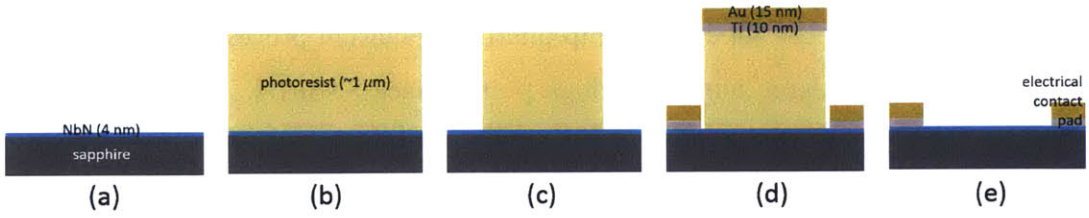

Figure 2-1: Schematic overview of steps in the contact pad fabrication process.

The nanowires were patterned via electron-beam-lithography. The sample was covered with a ~50-nm-thick negative electron beam resist hydrogen-silsesquioxane (HSQ (2%), spun at ~2 krpm, see Fig. 2-2(c)). Before this spinning step, in order to promote adhesion of HSQ to the NbN, a faux development step was performed composed of a 4-min dip in a 25%-solution of tetra-methyl-ammonium hydroxide (TMAH) at room temperature (Fig. 2-2(b)). The HSQ was exposed in a Raith 150 EBL tool with 30 keV acceleration voltage. The nominal exposure dose was in the

range ~1-1.5 mC/cm^2. Within hours after exposure, the HSQ was developed in a 25% TMAH solution at ~24°C for 4 min. The resulting HSQ pattern, illustrated in Fig. 2-2(d), was transferred into the underlying NbN film via a CF_4 reactive ion etch (RIE) at 100 W for 1 min.

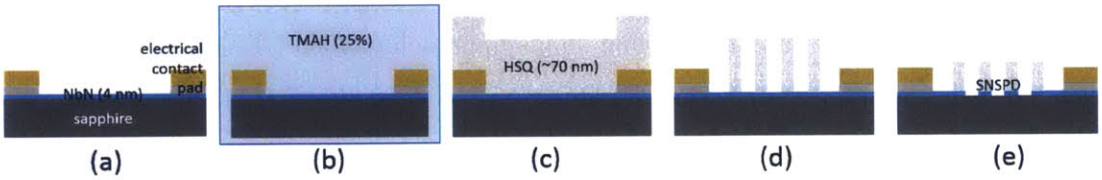

Figure 2-2: Schematic overview of steps in the post-liftoff SNSPD fabrication process.

As discussed in sections 1.1.2 and 1.2.4, constrictions that result in a suppressed critical (switching) current limit the yield of nanowire single-photon detectors. Constrictions can be due to defects introduced during the fabrication of the nanowires, e.g. localized defects in the nanowire pattern, defects in the film, e.g. granularity during the growth process or subsequent oxidation, or design-related limitations on the critical current (discussed in the next chapter). The following sections in this chapter will discuss fabrication-related modifications that helped us improve device yield.

2.2 Optimized fabrication process

Film quality is crucial for detector performance. We used the metrics room-temperature sheet resistance (R_S) and critical temperature (T_C) to characterize our films. High T_C is associated with high sheet current density and possibly low timing jitter, while high R_S could be associated with higher sensitivity due to larger Joule heating [21]. In accordance with previous work [1] we targeted R_S = 450-550 Ω/square and a T_C as close as possible to the bulk T_C . Substrates with an area of 1 cm × 1 cm cut from a 100-mm-diameter wafer, followed by solvent cleaning exposure to an oxygen plasma

(20% O_2 in He) at 100 W for 3 min. We deposited the NbN film via reactive DC magnetron sputtering on top of double-polished silicon substrates covered with ~300-nm-thick chemical-vapor-deposited silicon nitride (SiN_x). During the NbN deposition the substrate holder was heated to ~800°C. In addition to heating the holder, the substrate itself was heated directly from halogen heat lamps inside the deposition system via a hole in the substrate holder. This approach yielded NbN films that, at a given T_C, had comparable or higher R_S than our previous films. Table 2.1 contains a list of R_S and T_C values of films used for data shown in the following chapters of this thesis. Our reference film (film no. 1) had $R_S = 515.5$ Ω/square and $T_C = 10.9$ K. The thickness of this film, estimated by optical transmission measurements, was ~4.3 nm. Compared with our previous films on sapphire in Ref. [1] ($R_S = 480$ Ω/square and $T_C = 10.8$ K) these films had higher R_S and higher T_C.

film ID	old film	1	2	3	4	5	6
R_S (Ω/sq.)	480	516	504	533	586	525	444
T_C (K)	10.8	10.9	10.5	10.5	10.1	10.6	11
used in	Ref. [1]	Fig. 3-11	Fig. 3-5	3-5	3-10	Fig. 2-9(b)	Fig. 2-9(a)

Table 2.1: List of films used for detector data shown in this thesis

We modified our detector fabrication process [1], outlined in the previous section, as follows:

(1) We replaced the single-layer liftoff process, used for the initial fabrication of gold contact pads, with a bilayer liftoff process;

(2) we reduced the exposure time of bare NbN to TMAH, used for adhesion promotion of the detector resist layer;

(3) we reduced the length of the baking steps; and

(4) we performed post-reactive ion etch (RIE) imaging to minimize damage to the nanowires during pattern transfer from the detector resist layer into the NbN.

Figure 2-3 shows the resist mask for two detectors on top of gold pads fabricated with the old (a) and new liftoff processes (b). The Au pads in Fig. 2-3(a) were fabricated using a single layer of photoresist (Shipley S1813) as an evaporation mask for the subsequent gold liftoff. Due to the small size of our samples (1 cm $\times$ 1 cm square), the photoresist thickness on the edges and corners was significantly larger than in the center of the sample (detector region), preventing tight contact between the photomask and the photoresist during exposure. The lack of reliable contact often resulted in rough edges of gold pads after liftoff, as outlined in red in Fig. 2-3(a), and defects at the interface between the detector resist mask and electrical contact pads. These defects, often localized in certain regions on the detector chip, resulted in either severed detector-to-Au-pad connections or large electrical resistance at the interface. As outlined in Ref. [21] and chapter 1, a large resistance in series with the detector can result in latching and suppressed critical current, decreasing the overall detector yield of the chips, which typically included a matrix of $\sim$ 50 - 200 detectors each. In order to address this problem, we transitioned to a bilayer liftoff process. As shown in Fig. 2-3(b), the bilayer process resulted in smooth gold pad edges and reduced the number of detectors that suffered defects at the gold pad-NbN interface.

We exposed a 700-nm-thick PMGI-SF9 layer covered with $\sim$ 1.5 μm-thick S1813 with UV for 13 seconds at 2300 $\mu W/cm^2$ and developed the bilayer for 24 seconds in CD-26 (2% TMAH). Ten nanometers of Ti and 15 nm Au were evaporated and the liftoff was performed in acetone under sonication for 2-4 min followed by a $\sim$1-min-long dip in CD-26 and a 1-min-long DI dip.

We reduced the exposure of our films to TMAH in order to minimize damage to the film. The detectors were patterned via HSQ as a resist. Our previous fabrication process [42] included a 4-minute long TMAH dip as an adhesion promotion step before the HSQ was spun on to the sample. However, we found that this step decreased the T_C by up to 0.3 K while increasing the R_S, as shown in Fig. 2-4, possibly associated with a thinning of the film. Our new process included a TMAH dip of at most 15 seconds, which we found sufficient for promoting adhesion between the HSQ mask and the NbN film.

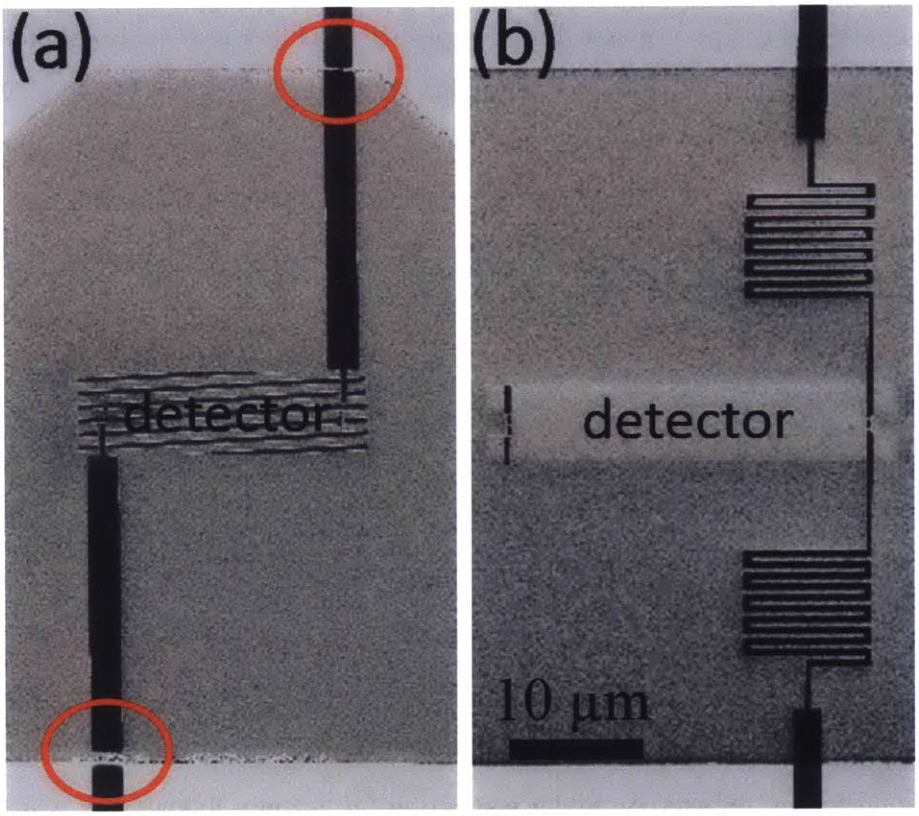

Figure 2-3: (a) Patterned e-beam resist on top of gold pads defined via single-layer liftoff process. The red circles highlight the interface between the detector resist mask and the electrical contact pads (Au pads). (b) Patterned e-beam resist on top of gold pads defined via bilayer liftoff process.

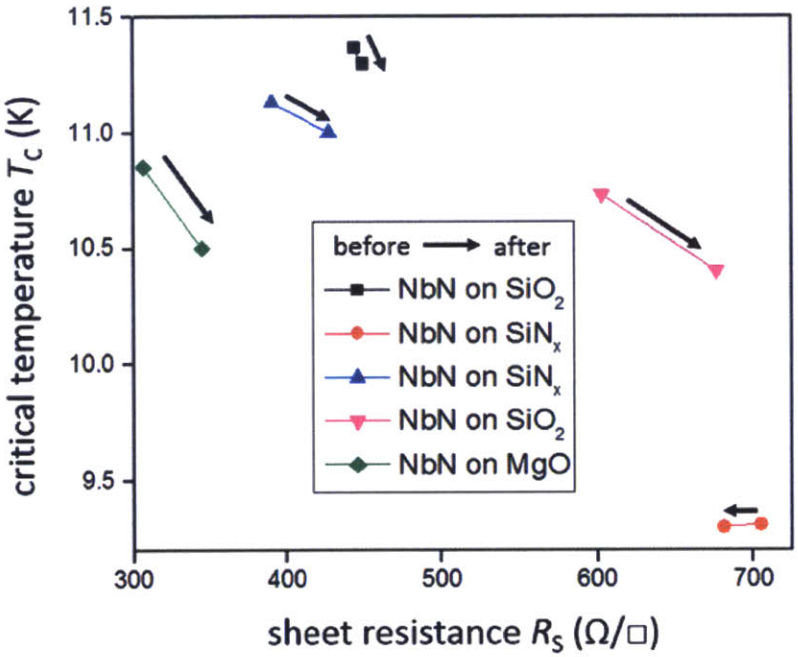

Figure 2-4: Room-temperature sheet resistance vs. critical temperature of NbN samples on a variety of substrates before (bottom of arrow) and after a 4 minute-long dip in TMAH (tip of arrow).

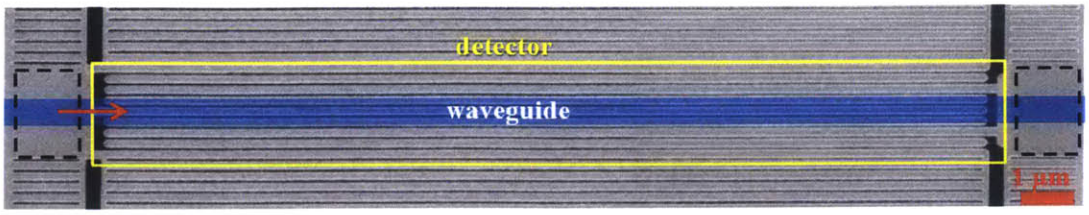

Figure 2-5: SEM of a detector based on ~80-nm-wide superconducting nanowires. The yellow line encloses the active area of the detector. The contact area between waveguide (to be added in a later fabrication step) and detector is highlighted in blue. The red arrow indicates the travelling direction of light coupled into the waveguide. The detector and waveguide are integrated in a subsequent step by releasing the detector on a mambrane and placing it on top of the waveguide.

In order to achieve high uniformity in the mean nanowire width (crucial to the detector performance [19]), we had to ensure a uniform exposure dose. Typically, to correct for the proximity effect and to achieve uniform dose in the side regions of the detector, dummy structures are exposed outside the active region of the device [1]. However, we had to avoid solid proximity effect correction (PEC) features outside the side regions enclosed by dashed lines in Figure 2-5. This limitation arose from the requirement that these detectors had to be integrated with waveguides, sketched in blue in Fig. 2-5. Proximity effect features overlapping with the waveguide, while not part of the active area of the detector, would absorb some of the incoming light before it could reach the hairpin-shaped detector, and hence reduce the system detection efficiency of the waveguide-SNSPD. Figure 2-6 shows a detector fabricated without any proximity effect correction features in the side regions, resulting in narrower nanowires in the side regions.

We compensated for the dose inequality in the side regions by exposing sacrificial PEC regions that did not result in solid HSQ features after development. These features consisted of 2-nm-wide lines (single-pass line exposure) arranged in a 20-nm-pitch. Figure 2-7 shows that these features compensated for the lower dose on the edges of the detector and resulted in a uniform mean nanowire width. While we have applied the sacrificial proximity effect correction method to SNSPDs, it could be applied to other cases where solid PEC features have to be avoided in specific regions.

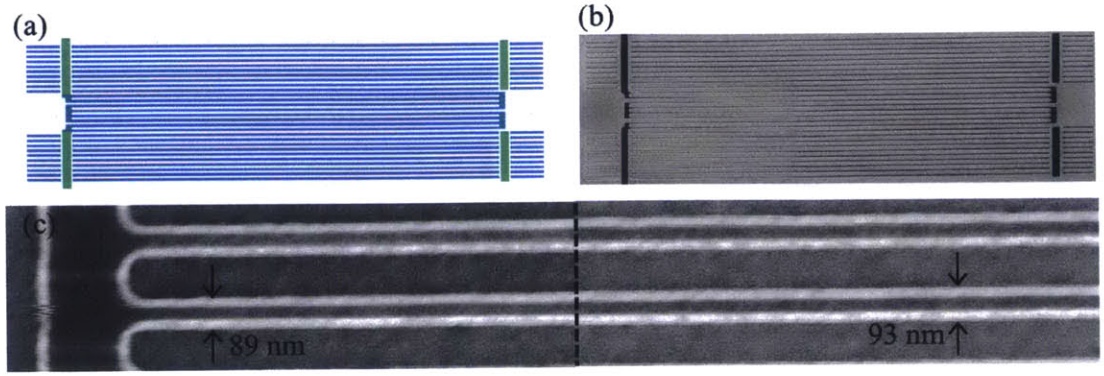

Figure 2-6: Detector fabricated without proximity effect correction features in the side regions. (a) Illustration of the detector pattern exposed via e-beam lithography. (b) SEM of resist structure (HSQ) resulting from exposure of the pattern shown in (a). (c) SEM of nanowires at the edge (left) and in the center (right) of the structure shown in (b).

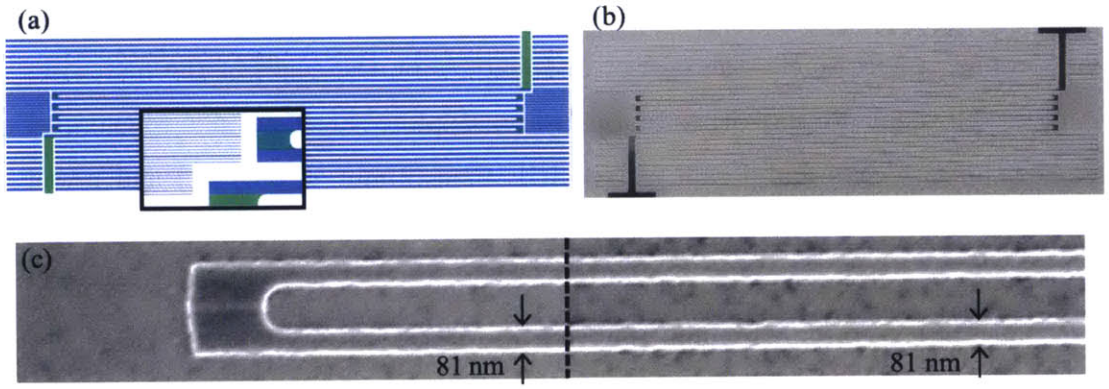

Figure 2-7: Detector fabricated using sacrificial proximity effect correction features in the side regions. (a) Illustration of the detector pattern exposed via e-beam lithography. The inset shows the proximity effect correction features consisting of 2-nm-wide lines in a 20-nm-pitch. (b) SEM of resist structure (HSQ) resulting from exposure of the pattern shown in (a). (c) SEM of nanowires at the edge (left) and in the center (right) of the structure shown in (b).

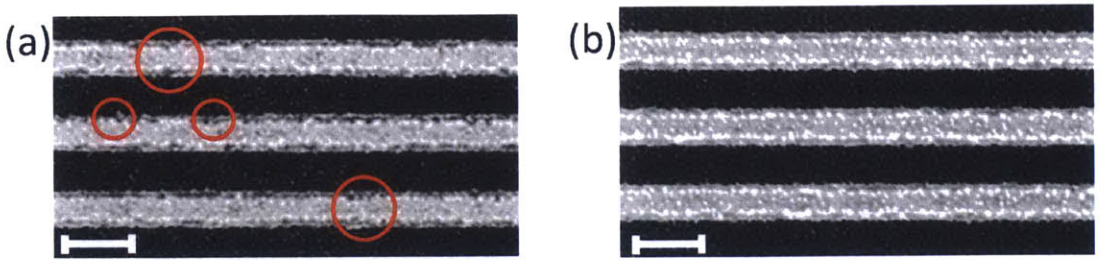

Figure 2-8: (a) Top-down SEM of slightly over-etched NbN nanowires on SiN_x. (b) SEM of NbN nanowires on SiN_x with optimized etch time. The equivalent length of the white scale bar is 200 nm.

After electron-beam exposure and development, the detector resist pattern was transferred into the underlying NbN via RIE (CF_4 at 50 W). In contrast to sapphire, which we used as substrate in previous work [1, 43], our current SiN_x-on-Si substrate did not exhibit significant charging during scanning-electron microscopy (SEM) and enabled high-contrast post-RIE imaging of NbN nanowires, which allowed us to optimize our RIE process. Figure 2-8(a) shows a detector without optimized RIE: while the HSQ mask was exposed with the correct dose (obtained from recent dose matrix), the detectors appeared over-etched. The red circles mark some regions with severe constrictions. Interestingly, the over-etched features on the edges of the nanowires would not be visible for nanowires fabricated on sapphire substrates due to charging and low contrast. We found that a fixed RIE time (2 min 40 seconds) and power (at 50 W) did not always yield the same effective etch rate. Over time the plasma voltage changes, presumably due to contamination in the etch chamber. This resulted in a change in etch rate. We found that over-etch can be avoided by decreasing the etch time ($\sim$ 2 min 10 seconds), imaging the detectors, and etching further in 10- to 15-second steps if needed. Using this modified etch process we obtained better nanowire uniformity, shown in Fig. 2-8(b).

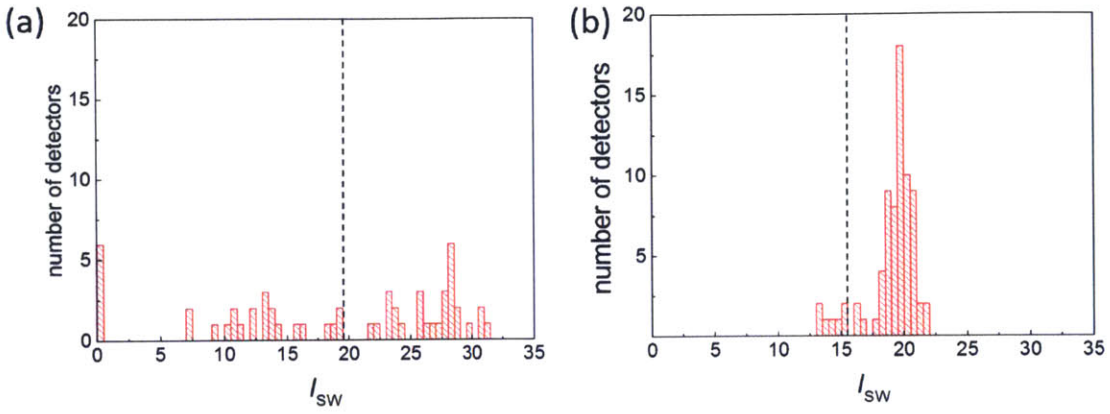

Figure 2-9: (a) Switching current histogram of detectors on a sample prepared with the old RIE process. The dashed line represents the avalanche current. (b) Switching current histogram of detectors on a sample prepared with optimized etch process. The dashed line represents the avalanche current.

2.3 Discussion and yield comparison

The modified fabrication process resulted in higher detector yield. Figure 2-9 shows the switching current histogram of two samples for which SEMs of exemplary detectors were shown in Fig. 2-8. Fig. 2-9(a) shows that the switching current distribution of the sample without optimized etch process (Fig. 2-8(a)) is fairly broad for detectors with the same design. If we define the detector yield as the ratio of the detectors for which $I_{SW} > I_{AV}$, we obtain a yield of ~10-30% for samples fabricated with the old process. In contrast, samples fabricated with the new process (Fig. 2-8(b)) showed a significantly improved detector yield of ~70 %, as shown for an exemplary sample in Fig. 2-9(b).

2.4 Summary and outlook

This chapter outlined several improvements to the standard SNSPD fabrication process that resulted in increased detector yield and a simplified process. It is conceivable that the yield could be further increased by introducing additional improvements to the film growth and fabrication processes. In the next chapter we will show that

the modifications presented in this chapter, combined with several detector design improvements, resulted in detectors with higher efficiency and lower timing jitter.

Chapter 3

Series-SNAPs

Practical SNSPD systems [35, 14] require detectors with large active areas (several μm in diameter). However, large-active-area SNSPDs come with several tradeoffs:
(1) Lower yield due to increased chance of defects along the nanowire;
(2) Larger dead time due to increased nanowire inductance; and
(3) Larger timing jitter due to increased rise time of the photoresponse pulse.
(1) can be addressed with improved fabrication processes (see previous chapter) and new material systems [22]. (2) and (3) can be potentially addressed with SNAPs. However, it is challenging to address these issues with the conventional SNAP design. In this chapter we will present a modified SNAP design, referred to as Series-SNAP, that can be scaled to large active areas without sacrificing saturated detection efficiency, few-ns-scale dead time, ~30 ps timing jitter (without the need for cryogenic amplifiers) and large avalanche bias current range.

3.1 Speed limitations of traditional SNAPs

The N parallel nanowires of an N-SNAP have an N^2-times lower inductance than an SNSPD with the same area. It was initially believed that the SNAPs have the potential to have an N^2-times reduced reset time compared to SNSPDs. However, part of the speed advantage is negated by the choke inductor L_S in series with the nanowires, resulting in a smaller effective speed-up. We found [5] that a sufficiently

large L_S is essential for stable operation in avalanche regime. We can approximate the ratio r of the the leakage current I_{lk} into the load and the current redistributed to all $(N-1)$ secondary sections after an HSN as [44]

$$r = \frac{L_0}{L_S \cdot (N-1)} \tag{3.1}$$

A decrease in L_S/L_0 results in an increase of I_{lk} (see I_{out} in 1-10(d)). The higher I_{lk} results in an increase of I_{AV}, i.e. a larger bias current is necessary to trigger an avalanche. A large I_{AV}/I_{SW} is undesirable because (1) it increases the vulnerability of the detector to noise during operation and (2) it requires the operation of the SNAP at higher dark count rates due to the larger bias current required to operate in avalanche regime. In addition to the undesirable effect of smaller avalanche regime, a small L_S changes the behavior of the SNAP at low bias currents. Fig. 3-1(a) shows the normalized photon count rate (PCR) vs. bias current of a 3-SNAP with $r = 0.28$. The PCR is normalized by the incident photon flux, which corresponds to the device detection efficiency for $I_B > I_{AV}$. Between $I_B \sim 0.7I_{SW}$ and I_{AV} the detector operates in arm-trigger regime. For $I_B < 0.7I_{SW}$ the SNAP emitted many smaller-amplitude pulses in addition to the avalanche pulses, as shown in the lower inset panel (blue trace) in Fig. 3-1(a). We call this operation regime the 'unstable regime'. For sufficiently low trigger levels, the detector pulses in the unstable regime resulted in a spurious, photon-flux-dependent peak in the η vs. I_B curve. We observed that SNAPs with $r > 0.1$ operated in the unstable regime at low bias currents. For some applications a large r might be acceptable, and a question arises regarding the lowest possible L_S value. We found that the reset time of SNAPs in NbN is ultimately limited to ≥ 1 ns [5]. SNAPs with smaller reset times showed afterpulsing, i.e. they generated a series of pulses for each detected photon, as shown in Figs. 3-1(b,c). As a result, in practice the speed limit of SNAPs is similar to SNSPDs ($\sim$ 1-ns dead time for NbN, see Ref. [21]).

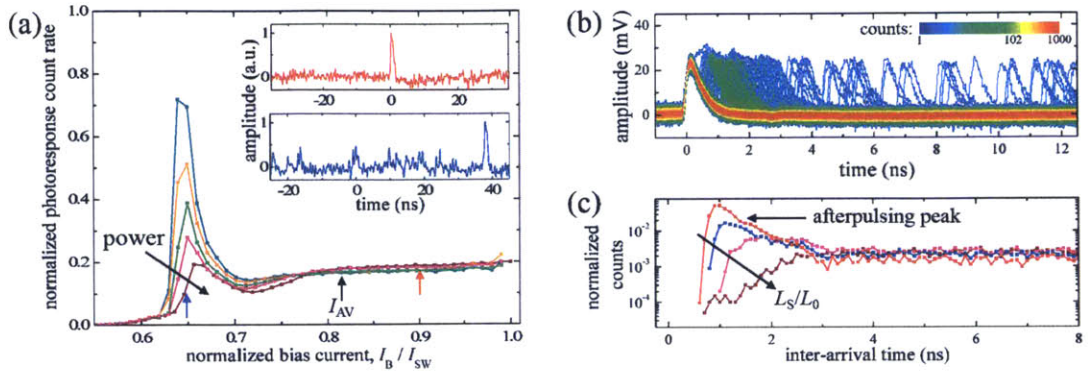

Figure 3-1: (a) Normalized photon count rate (PCR) vs I_B/I_{SW} of a 30-nm-wide 3-SNAP at different incident photon fluxes. The PCR is normalized to the photon flux, corresponding to the device detection efficiency for $I_B > I_{AV}$. I_{AV} (black arrow) was extracted from η vs. I_B curves as in Fig. 1-9. The inset shows detector output voltage traces of the 3-SNAP in the unstable regime (biased at $I_B = 0.65I_{SW}$, bottom panel, blue arrow) and the avalanche regime (biase at $I_B = 0.9I_{SW}$, top panel, red arrow). (b) Persistence map of the detector output voltage traces a 4-SNAP with leakage current parameter $r = 1$ (see text). The detector was biased close to the switching current. (c) Pulse-to-pulse inter-arrival time delay histograms of the detector pulses for 4-SNAPs with r ranging from ~ 0.1 (pink) to 1 (orange). The detectors were biased close to the switching current. Adapted from Ref. [5]

3.2 Scaling SNAPs to large areas

As outlined in the previous section, traditional SNAPs require a large series inductor $L_S > 3L_0$ which scales with the length of a single section and thus the device diameter. The series inductor is designed to minimize the leakage current I_{lk}:

$$\frac{I_{lk}}{I_0 - I_{lk}} = r = \frac{L_0}{L_S(N-1)}, \tag{3.2}$$

where $I_0 = I_B/N$ is the steady-state bias current of a single section. An avalanche will occur if

$$\frac{I_0 - I_{lk}}{N-1} + I_0 \geq \frac{I_{SW}}{N} \tag{3.3}$$

Using (3.2) and (3.3) we obtain for avalanche current I_{AV}

$$\frac{I_{AV}}{I_{SW}} = \frac{1}{1 + \frac{1}{N-1}\frac{1}{1+r}} \tag{3.4}$$

For $r = 0$ (3.4) yields the relation for the case of perfect current redistribution into the secondary sections [27]

$$\frac{I_{AV}}{I_{SW}} = 1 - \frac{1}{N} \tag{3.5}$$

The inductance L_0 of a single section of a traditional SNAP with a square-shaped active area A is

$$L_0 = L_{sq.} \cdot \frac{l}{w} = L_{sq.} \frac{A}{N \cdot w \cdot p}, \tag{3.6}$$

where w is the nanowire width, p is the nanowire pitch and $L_{sq.}$ is the kinetic inductance of the film per square. The overall detector inductance L_{tot} is then

$$L_{tot} = L_{sq.} \frac{A}{N \cdot w \cdot p} [1 + r(N-1)] \tag{3.7}$$

Figure 3-2 shows the calculated minimum dead time $3L_{tot}/50\Omega$ as a function of area A for a 2-, 3- and 4-SNAP. In large-active-area SNAPs L_S negates the relative speed advantage compared to SNSPDs, resulting in small count rates, large rise time and large timing jitter.

The constraint on the series inductor can be relaxed using a design outlined in Fig. 3-3(a), referred to as a series-SNAP. In this configuration, the parallel nanowires are broken down into smaller units. As illustrated in Fig. 3-3, each unit is connected in series with other SNAP units, which act as a series inductor.

Due to the smaller length of the series-SNAP sections of inductance

$$L_0 = L_{sq.} \cdot \frac{\sqrt{A}}{w}, \tag{3.8}$$

the active area of the detector can provide some of the series inductance, and the necessary external L_S can be reduced or, for large areas, entirely eliminated. The overall inductance is then

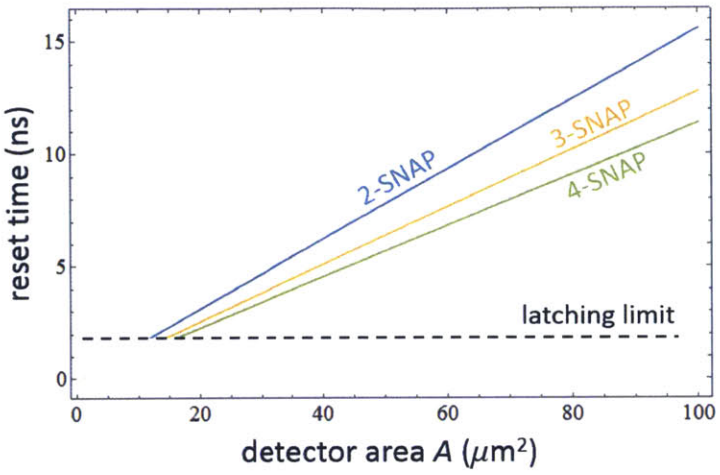

Figure 3-2: Calculated estimated reset time $3L_{tot}/R$ of a traditional 2-, 3- and 4-SNAP as a function of the square-shaped detector area A. The values were calculated for $r = 0.3$.

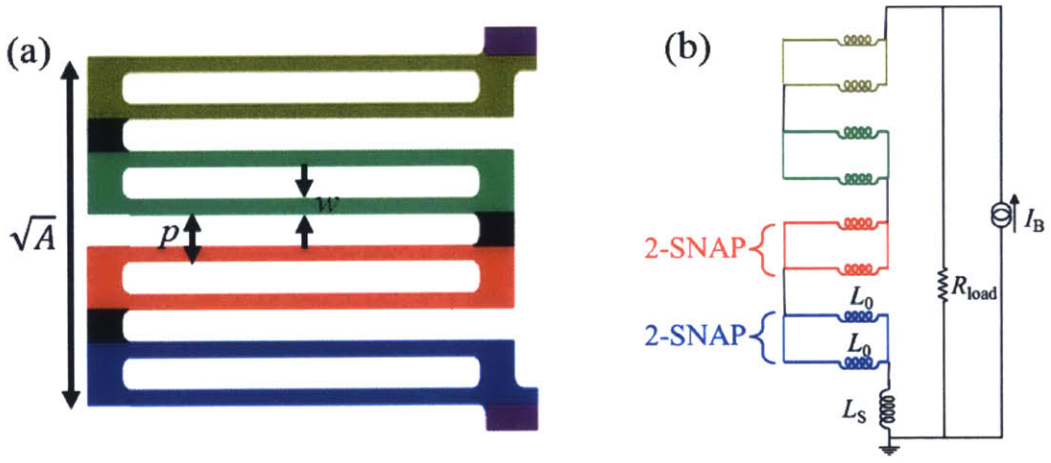

Figure 3-3: (a) Nanowire arrangement sketch for a series-2-SNAP comprising four 2-SNAP units connected in series. (b) Diagram of the equivalent circuit for the series-2-SNAP shown in (a).

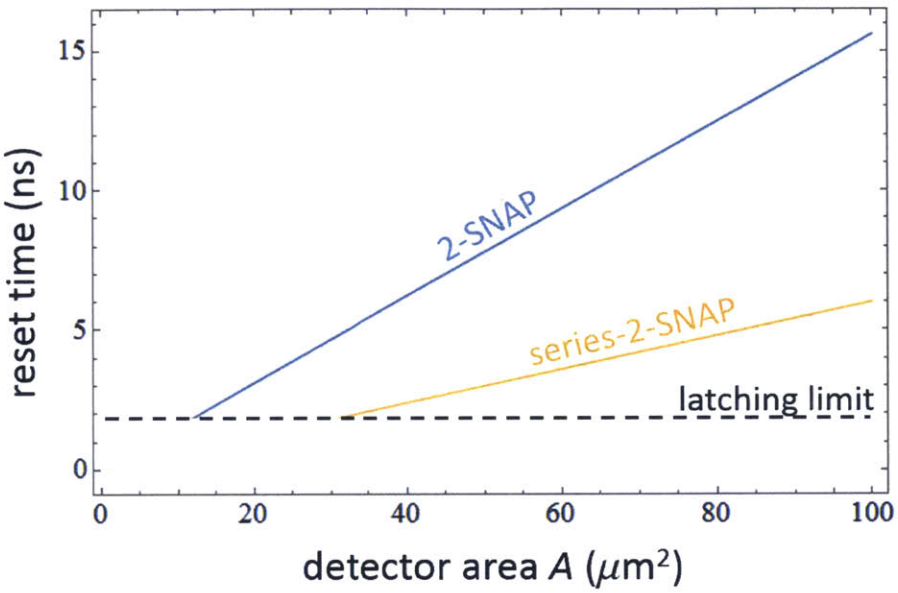

Figure 3-4: Calculated estimated reset time $3L_{tot}/R$ of a traditional 2-SNAP and a series-2-SNAP as a function of the square-shaped detector area A. The values were calculated for $r = 0.3$.

$$L_{tot} = L_{sq.} \frac{A}{N^2 \cdot w \cdot p}, \tag{3.9}$$

resulting in a fast rise time, small dead time and low timing jitter. The significant speed-up for a 2-SNAP is illustrated in Fig. 3-4.

Saturated large-active-area Series-SNAPs based on this design have recently been demonstrated [45].

3.3 Overcoming trade-off between jitter and saturation

Our main objective was to design a detector that showed both saturation and sub-30-ps timing jitter without the need for cryogenic amplifiers. In order to achieve this goal we examined the saturation behavior and timing jitter as a function of nanowire width

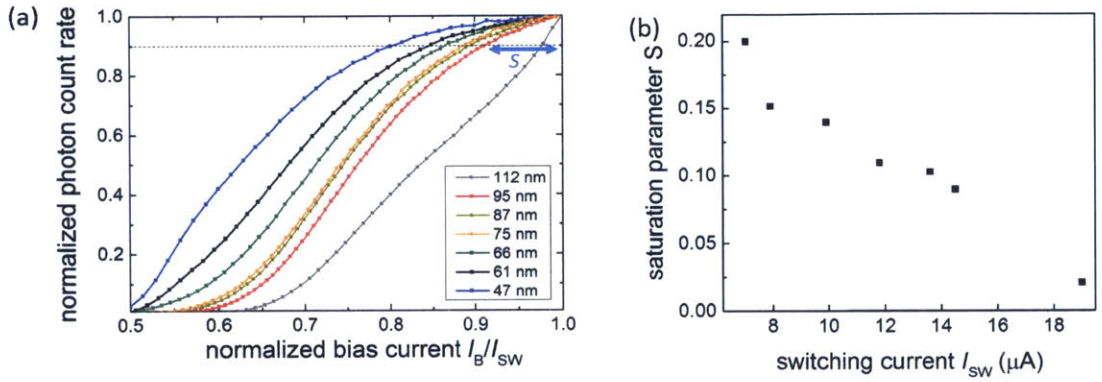

Figure 3-5: (a) Normalized photon count rate vs bias current I_B normalized by switching current I_{SW} for nanowire widths ranging from 47 nm to 112 nm. The red arrow denotes the saturation parameter S for the 95-nm-wide nanowire. (b) Saturation parameter S vs switching current for the detectors shown in (a).

and bias current. Figure 3-5(a) shows the normalized photon count rate (PCR) for 1550-nm-wavelength photons as a function of the detector bias current I_B normalized by the switching current I_{SW}. We quantify saturation by introducing a saturation metric S, defined as

$$S = \frac{I_{SW} - I_B(0.9 \cdot PCR(I_{SW}))}{I_{SW}}, \tag{3.10}$$

where $I_B(0.9\ \mathrm{PCR}(I_{SW}))$ is the bias current at which the PCR has reached 90% of its maximum value. Using this metric, $S = 0$ represents an unsaturated detector while $S = 1$ describes a detector that is saturated over its entire bias current range. Figure 3-5(b) shows S as a function of I_{SW} for the detectors in Fig. 3-5(a). As we had found earlier [1] decreasing width results in a smaller I_B/I_{SW}-dependence of the PCR at high bias currents.

Figure 3-6 shows the timing jitter, defined as the full-width-at-half-maximum (FWHM) of the photodetection delay histogram (instrument response function IRF) as a function of nanowire width and bias current. The measurement setup was identical to Ref. [6]. At high I_B/I_{SW} the FWHM jitter seems to be mainly limited by the signal-to-noise ratio [6], although at low bias currents wider nanowires appear to have larger timing jitter.

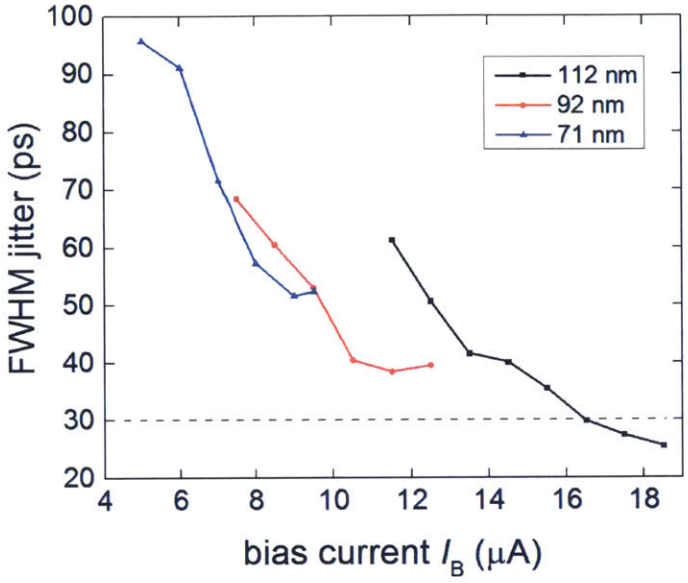

Figure 3-6: Timing jitter vs. bias current for SNSPDs based on 71 , 92- and 110-nm-wide nanowires.

While narrower nanowires resulted in a more saturated PCR vs. I_B curve, indicated by increasing S, they had smaller I_{SW} and hence larger jitter. In order to achieve sub-30-ps timing jitter with the setup used here, a detector would need to have $I_{SW} > 16.5$ μA (see dashed line in Fig. 3-6).

3.4 Current crowding and other design considerations

The detector can be optimized beyond the circuit design and nanowire cross section that were discussed in the previous sections. When current flows around a 180° bend, it tends to concentrate close to the inner boundary of the bend , referred to as current crowding, as illustrated in Fig. 3-7. Berggren and Clem [46] demonstrated that for thin (thickness d less than London length λ) and narrow (less than Pearl length $\Lambda = \lambda^2/d$) nanowires current crowding can significantly reduce to switching current of the nanowire. The suppression of I_{SW} can be quantified with the suppression factor

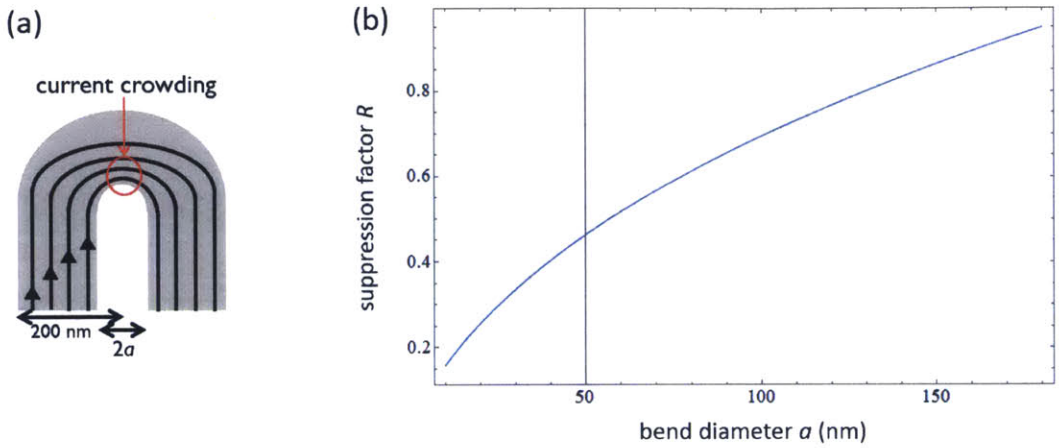

Figure 3-7: (a) Sketch of a nanowire with a 180°-turn of radiaus a. (b) Calculated suppression factor R vs bend radius a for the wire shown in (a).

R

$$R = \frac{I_{SW}}{J_C \cdot w \cdot d} = \frac{I_{SW}}{I_{C0}}, \tag{3.11}$$

where J_C is the critical current density of the thin film and w the width of the nanowire. For a circular turnaround bend of diameter $2a$ and boundary distance of b, as sketched in 3-7(a), the suppression factor is

$$R = \frac{a \cdot ln(b/a)}{b - a} \tag{3.12}$$

Figure 3-7(b) shows R as a function of bend radius for $b = 200nm$. Sharper bends result in large suppression of I_{SW}, while a more gradual change in the inner bend reduces the effect of current crowding on the switching current.

Current crowding around corners is of significance for SNAPs and SNSPDs since a larger I_{SW} increases the bias current range, resulting in larger S and lower timing jitter. Figure 3-8(a) shows the calculated current density distribution in the turnaround region of a series-2-SNAP with strictly circular bends and the corners. In this design, current crowding results in an increase of the current density by up to ~60%. Following Ref. [46], an optimized design for the bends, illustrated in Fig. 3-8(b), can

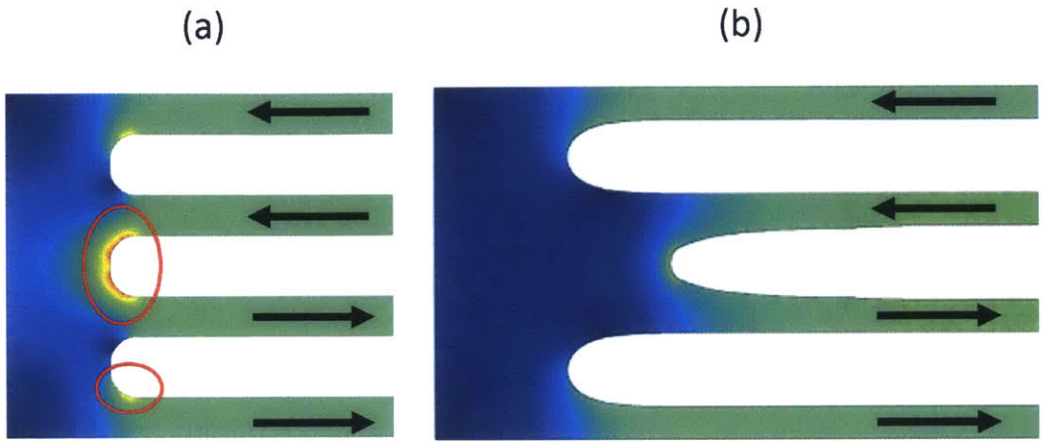

Figure 3-8: Calculated current density around a turn of a series-2-SNAP based on 80-nm-wide nanowires with circular rounded (a) and optimally-rounded (b) bends.

eliminate the crowding effect.

3.5 Experimental results

A 2-SNAP based on 60-nm-wide nanowires appears to fulfill both requirements of high S and low timing jitter. An SEM of an exemplary detector is shown in Fig. 3-9. The detectors consisted of 4 2-SNAP units in series [27, 47] as discussed in Ref. [40], each unit comprising two ~60- to 80-nm-wide nanowires (200 nm pitch) in parallel. We refer to this design as a series-2-SNAP. This detector had a long rectangular shape for subsequent integration with waveguides [40].

Based on the experiments on SNSPDs described in the previous section we designed SNAPs that could meet our requirement of large-bias-current-range saturation and low timing jitter. The detection efficiency was measured at 2.4 K in a cryogenic probe station using a polarized incoherent CW source. The detectors were back-illuminated (through the Si substrate) and the polarization adjusted to achieve maximum efficiency. Fig. 3-10(a) shows the device detection efficiency vs bias current for a series-2-SNAP based on ~60-nm-wide nanowires. Here we will not discuss the unstable regime below the avalanche current I_{AV}, as this regime is discussed in detail

Figure 3-9: SEM of waveguide-detector. (a) The detector is shown with the series inductor consisting of 300-nm-wide nanowires. (b) Magnified SEM of detector region encircled with red lines in (a). The detector consists of four 2-SNAPs in series with each 2-SNAP comprising two ~60- to 80-nm-wide nanowires in parallel.

in Ref. [44]. When biased above I_{AV} the detector showed a characteristic 'saturation plateau' with a detection efficiency value close to the calculated optical absorption of ~14.5 % (see Ref. [26]). The absorption values on our substrate were lower than typical absorption values on sapphire [5] due to the higher refractive index of silicon, which results in significant back-reflection at the substrate-vacuum interface. This is a common problem with silicon that can be solved with several approaches (see Refs. [48, 14, 35, 15]).

When biased well within the saturation regime, at I_B= 15.9 μA, this detector had a timing jitter of ~35 ps, as shown in Fig. 3-10(b). Fig. 3-6 implies that in order to achieve sub-30-ps timing jitter for detectors biased well within the saturation regime higher I_{SW} is required. The revised detectors were based on wider ~80-nm-wide nanowires. The results are shown in Fig. 3-11.

The revised 2-SNAPs enabled a high signal-to-noise ratio (~ 8 - 9, as defined in Ref. [1]), resulting in a timing jitter of 24 ps when biased well within the avalanche regime (I_B ~ 22 μA). Furthermore, this detector showed saturated DE over a bias current range of 3.5 μA, the largest range reported to date. As a result we could bias the detector at > 3 μA below I_{SW}, which reduced the dark count rate by more than one order of magnitude, while maintaining maximum internal efficiency. The optical absorption of this detector was calculated as ~18.4 % (see appendix in Ref. [26]). A single-shot pulse trace of this detector, biased at 22 μA, is shown in Fig. 3-12.

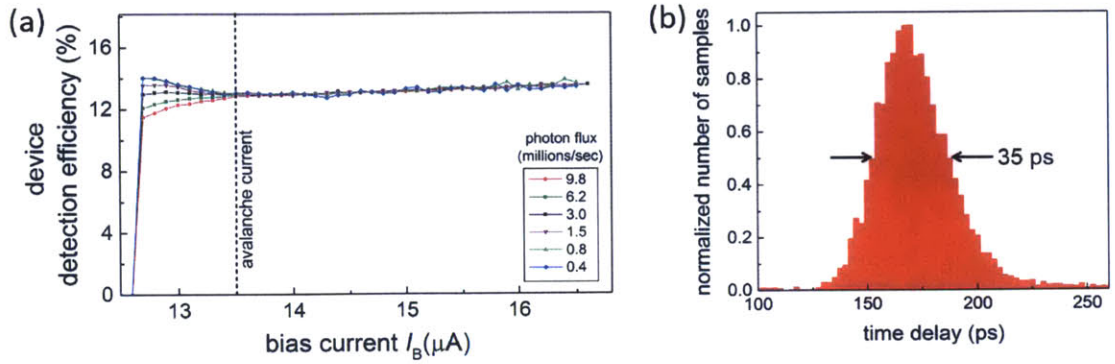

Figure 3-10: (a) Back-illuminated device detection efficiency vs bias current for a series-2-SNAP based on ~60-nm-wide nanowires. The incident photon flux was varied between 0.4 million photons/second (blue) and 9.8 million photons/second (red). The trigger level was set to 310 mV. (b) Instrument response function (IRF) of the same detector as in (a) biased at 15.9 μA. The IRF was measured using the setup described in Ref. [6].

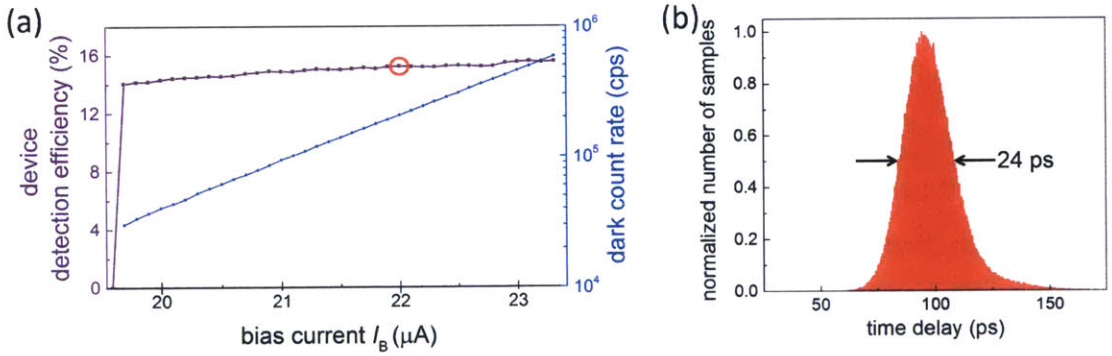

Figure 3-11: (a) Back-illuminated device detection efficiency vs bias current for a series-2-SNAP based on ~80-nm-wide nanowires (see Fig. 3-9). The incident photon flux was 2.3 million photons/sec. The trigger level was set to 380 mV to show only the saturated avalanche regime ($I_B >$ 19.5 μA, see chapter 1). The red circle denotes the operation point at which the timing jitter measurements were performed. (b) IRF of the same detector as in (a) biased at 22 μA. The IRF was measured using the setup described in Ref. [6].

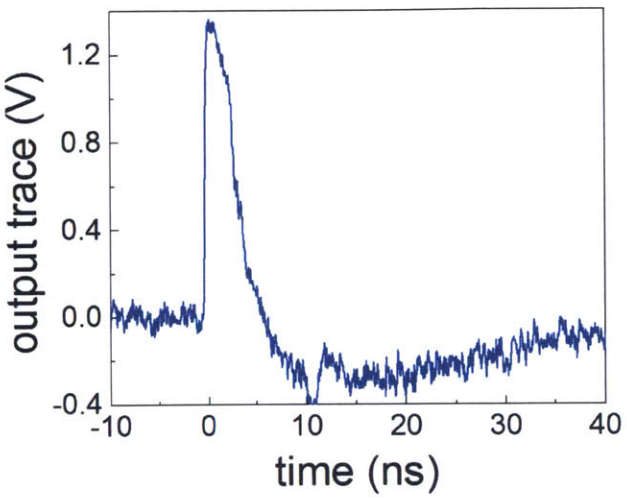

Figure 3-12: Single-shot pulse trace of an exemplary 2-SNAP based on 80-nm-wide nanowires. The detector was biased at $I_B = 22$ μA.

The room-temperature amplifiers had a bandwidth of 20-3000 MHz. When biased in this regime, the detector reached a photon count rate of 17 million counts per second (Mcps) at maximum efficiency.

3.6 Summary and outlook

We studied the area-dependence of the reset time of superconducting nanowire-based single-photon detectors, and introduced a new design, the series-SNAPs, that allowed sub-10-ns reset times and sub-30-ps timing jitter for detectors with active areas on the order of several tens of μm^2. However, the device speed remained ultimately limited by the latching limit. In the next chapter we will discuss a novel approach to designing SNSPDs that has the potential to overcome the latching limit.

Chapter 4

Nano-mesh single-photon detector

Photons absorbed in an SNSPD result in a large resistive region [21]. As outlined in chapters 1 and 3, the cool-down time of this resistive region and the inductance [20] of the detector limit the reset time ($3L/R$) of SNSPDs and SNAPs to ~2 ns for detectors based on niobium nitride (NbN) and to ~40-100 ns in tungsten silicide (WSi). Figure 4-1 shows a sketch of the timeline of photodetection in a superconducting nanowire. After an HSN event the nanowire goes through a hotspot expansion phase, which is assumed to last 100-200 ps [21], and a cooldown phase. During the cooldown phase the superconductivity in the nanowire is restored and current flows back into the nanowire from the load. However, during the initial cooldown phase the local nanowire temperature T is higher than the substrate temperature T_{sub}, and the superconducting gap $\Delta(T) < \Delta(T_{sub})$ remains suppressed. If the current in the nanowire I_B recovers too quickly (within 1-2 ns in NbN) during the cooldown phase so that $I_{\mathrm{B}} > I_{\mathrm{C}}(T)$, the nanowire can switch back to a stable resistive state [49]. The hotspot growth due to joule heating and the suppressed gap during the cooldown hence ultimately limit the reset time of nanowire-based superconducting detectors. Limiting hotspot growth and enhancing the cooldown appear to be promising paths towards reducing the reset time.

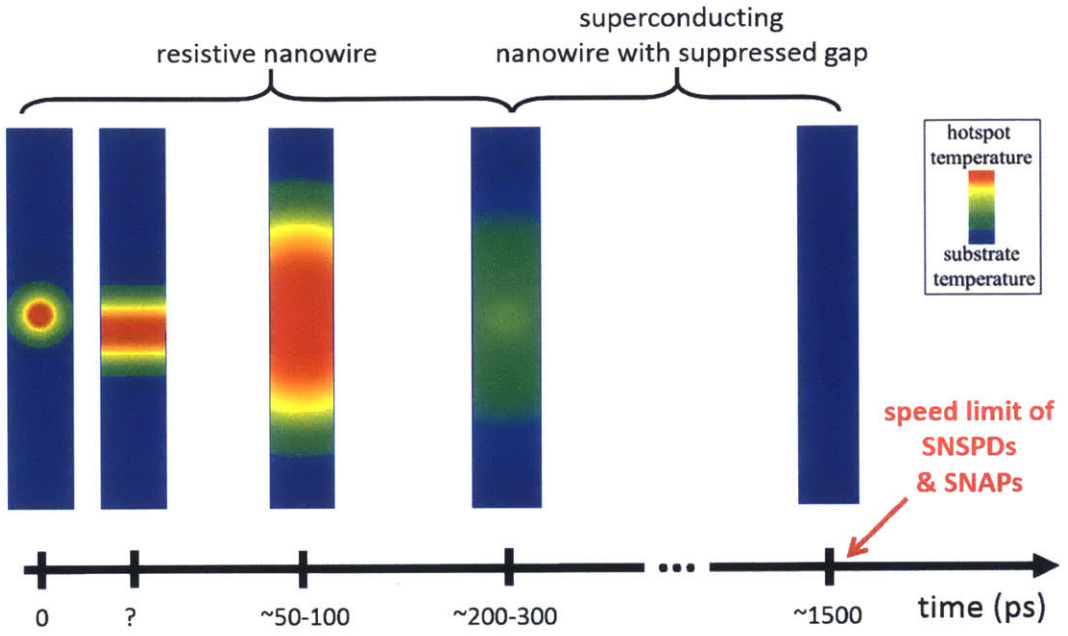

Figure 4-1: Schematic timeline of the temperature profile of a superconducting nanowire after the formation of a hotspot at t=0.

4.1 The cooling pad concept

In 1974 Tinkham and Skockpol [50] reported the effect of large leads on the expansion of resistive regions in short superconducting micro-bridges. Due to diffusive cooling, the leads significantly limited the growth of the resisitive region. The cooling pad concept is based on the same principle. Figure 4-2(a) shows a sketch of the temperature profile of a resistive region that has grown due to joule heating. Using the similar arguments to Ref. [50], the growth of the hotspot could be limited using 'cooling pads', which comprise wider regions attached to the nanowire, as illustrated in Fig. 4-2(b).

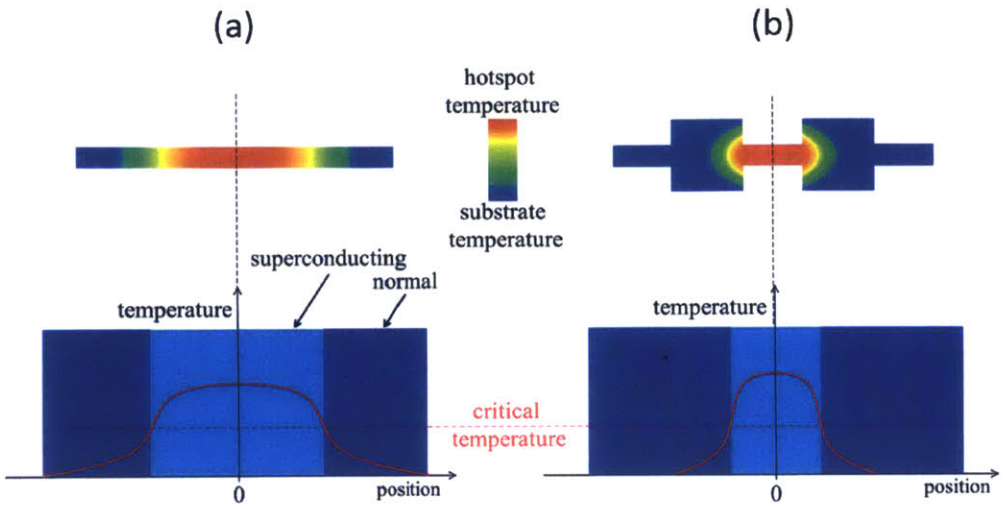

Figure 4-2: (a) Schematic temeprature distribution of a growing cross-sectional resistive region in a superconducting nanowire. (b) Schematic temeprature distribution of a resistive region in a superconducting nanowire with cooling pads. The cooling pads limit hotspot growth the growth of the resistive region.

4.2 Device design and modeling

As outlined in chapter 3, for large-area detectors the kinetic inductance L_{tot} limits the device speed in addition to the hotspot growth. A SNAP implementation is a promising path to reducing the overall device inductance. Figure 4-3(b) shows the sketch of a series-4-SNAP with short nanowire sections. Incidentally, the wider connecting sections (green) could potentially serve as cooling pads. However, these wider sections are not photo-sensitive and would introduce 'blind spots' into the detector active area. Figure 4-3(b) shows the design of a Nanomesh single-photon detector (NMSPD) that combines the requirements for low inductance and a diffusive cooling region while minimizing blind spots. Furthermore it avoids sharp corners, which would result in significant current crowding (see chapter 3). The NMSPD acts as a series-4-SNAP. In order to ensure that the entire active area of the detector

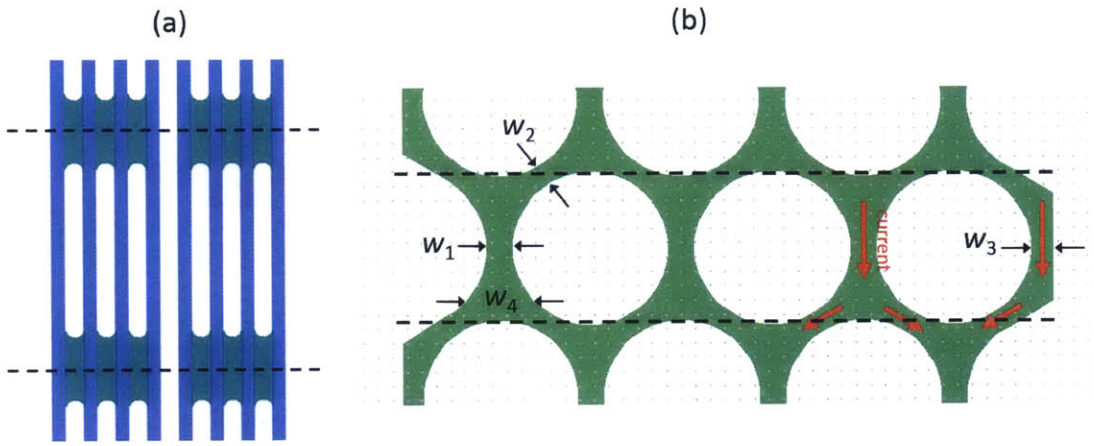

Figure 4-3: (a) Series-4-SNAP with short sections separated by dashed lines. (b) Nanomesh single-photon-detector based on ~30-nm-wide photosensitive sections.

remains optimally-biased (within the saturation region, see chapter 1), a thin film ($R_S = 500 - 600\Omega$/sq.) and narrow gap dimensions ($\leq$50nm) are required. The dimensions of the NMSPD are: w_1=28nm, $w_2 = w_3$=22nm and w_4 =75nm.

The next step was to simulate the behavior of the proposed design. Two properties had to be confirmed:

(1) Whether the wider regions provide sufficient cooling to limit hotspot growth,

(2) Whether an HSN in one section resulted in an avalanche, and

(3) Whether the avalanche is limited to a single SNAP region (section within the dashed lines in Fig. 4-3(b)).

The one-dimensional electrothermal model [21] is not sufficient to model the hotspot evolution in an MNSPD. Therefore a two-dimensional electrothermal COMSOL model was used to simulate the spatial distribution and temporal evolution of the temperature T and resistivity ρ across the NMSPD. This model was based on code developed for simulations in Ref. [51]. Figure 4-4(a) shows the temperature profile of an initiating section. The initial HSN is triggered by a small constriction (wedge) in the initiating section. Figure 4-4(b) shows the resistivity profile for the NMSPD in Fig. 4-3(b) at t =20ps after the initial HSN. While it fulfills the first two

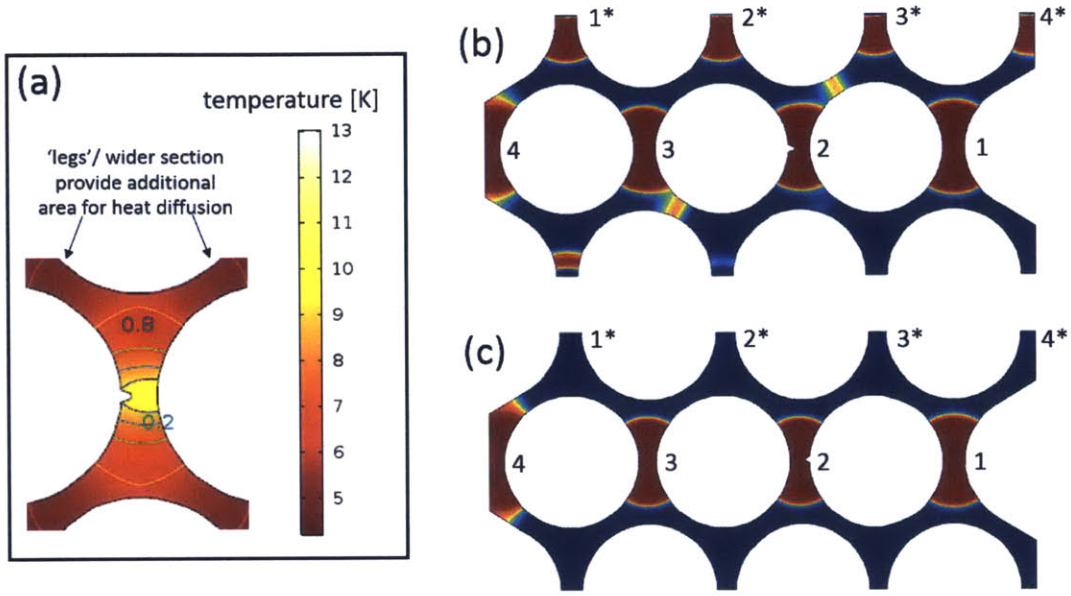

Figure 4-4: Simulated temperature profile of an NMSPD unit cell (a), a 4-NMSPD with fixed unit cell width of w=28nm for sections 1-3 (b) and of a 4-NMSPD with varying unit cell widths (w=32nm for section 4, w=28nm for section 3, w=24nm for section 2 and w=20nm for section 1).

requirements listed above, it does not fulfill the third requirement since the resistive region spreads across multiple SNAP regions, resulting in larger heating of the NbN film. In order to avoid an expansion of the avalanche to sections 1*-4*, it is crucial to ensure a more uniform distribution of inductances for every path. Figure 4-3(c) shows a modified design where the width of sections 1-3 is not the same. The width gradually increases from w_4 =20 nm to w_1=32 nm, resulting in a confined avalanche region.

The direct comparison of a 4-NMSPD and a nanowire of the same equivalent width $w = w_1 + w_2 + w_3 + w_4$, illustrated in Fig. 4-5, reveals the significant advantage of the NMSPD design. The resistive region in the wide nanowire is ~8-times longer than the 4-NMSPD, resulting in larger Joule heating $\propto R \cdot I_B^2$. Due to the complex effects of the readout circuit [52], the current recovery is less straight-forward to simulate, and the modeling-based design modifications were mainly focused on reducing the

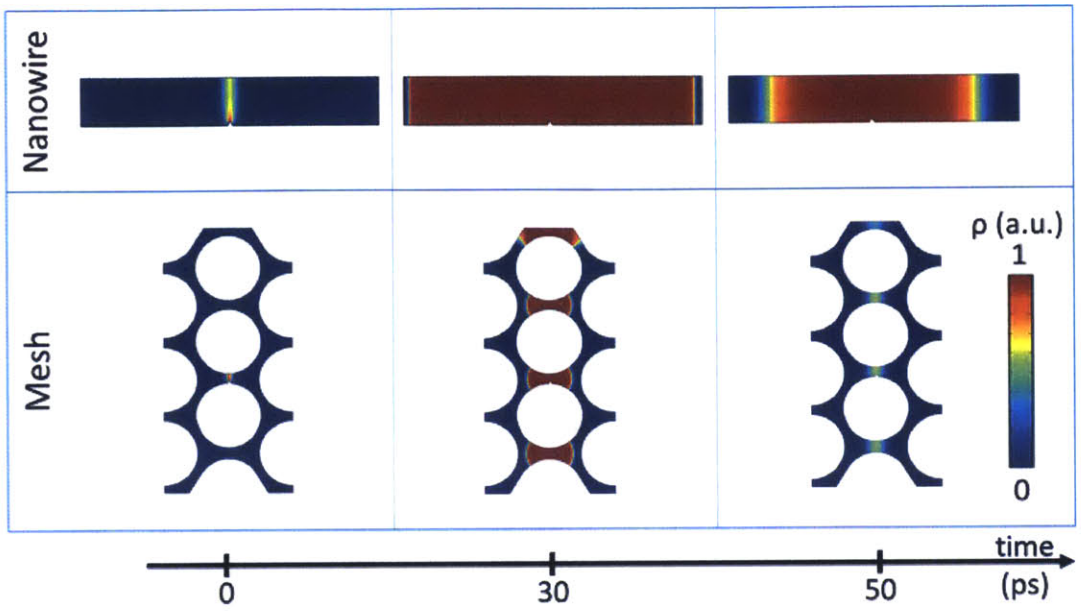

Figure 4-5: Simulated spatial distribution of the resistivity for a 4-section NMSPD based on ~30-nm-wide sections compared to a nanowire of equivalent width $w = w_1 + w_2 + w_3 + w_4$ carrying the same initial steady-state current.

maximum size of the resistive region.

4.3 Fabrication process

HSQ was used to fabricate the NMSPD etch masks. The exposure had to be optimized to yield the correct feature size and a continuous pattern. A 125-keV electron beam tool (Elionix) was used to expose the HSQ. Details on the Elionix fabrication and patterning steps and parameters can be found in Ref. [53]. Due to beam drift, the fracturization of the design file and the exposure order of the polygons were key to achieving a continuous NMSPD pattern. Figure 4-6(a) shows a design file with large polygons without a specified writing order. These files typically resulted in discontinuous features as in Fig. 4-6(b). An enhanced fracturization approach is shown in Fig. 4-6(c): The pattern is divided into narrow horizontal strips of 4-10nm and the write order is adjusted so that strips are written from bottom to top. This

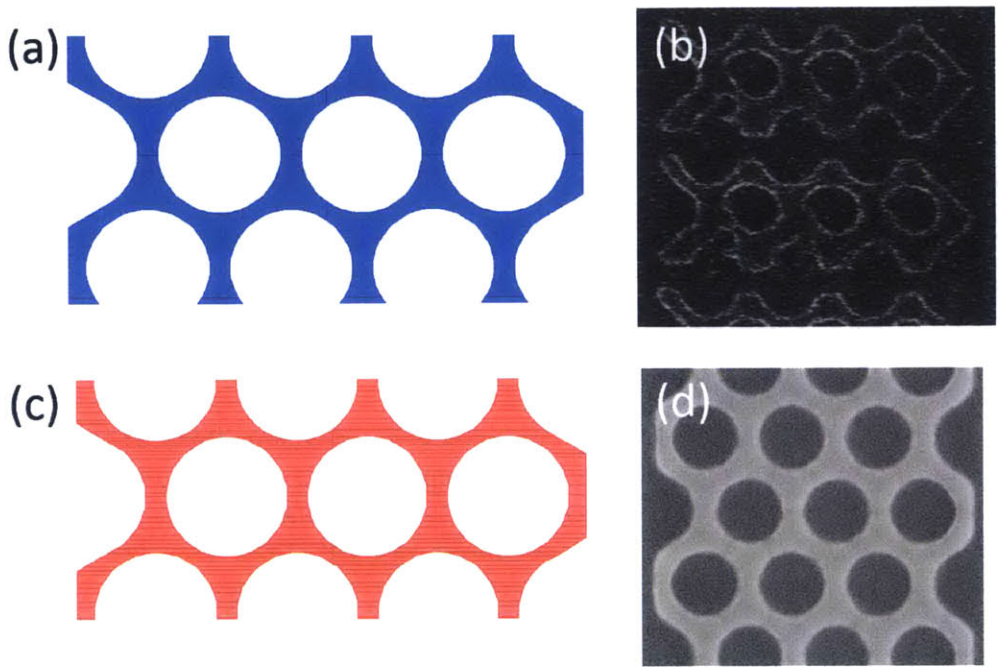

Figure 4-6: (a) NMSPD pattern file for electron-beam exposure divided into large sections using horizontal and vertical cuts. (b) Post-development HSQ mask using an exposed pattern divided as in (a). (c) NMSPD pattern file for electron-beam exposure divided into narrow sections using only horizontal cuts. (d) Post-development HSQ mask using an exposed pattern divided as in (c).

approach yielded continuous patterns as in Fig. 4-6(d).

An SEM of a 12-μm-long NMSPD is shown in Figs.4-7(a,b), comprising 40- to 65-nm-wide unit cells as shown in Figs. 4-7(c,d). The process for pattern transfer into the NbN was analogous to the optimized SNSPD process in chapter 2.

4.4 Experimental results

2- to 14-μm-long NMSPDs based on ~30- to 80-nm-wide unit cells were fabricated on films with R_S=520-600Ω/square. The kinetic inductance can be extracted from the room-temperature device resistance as follows [54]:

$$L_{\mathrm{K}} = \frac{\mu_0 \lambda_{\mathrm{eff}}^2}{\rho_{300\mathrm{K}}} R_{300\mathrm{K}} = \gamma R_{300\mathrm{K}}, \tag{4.1}$$

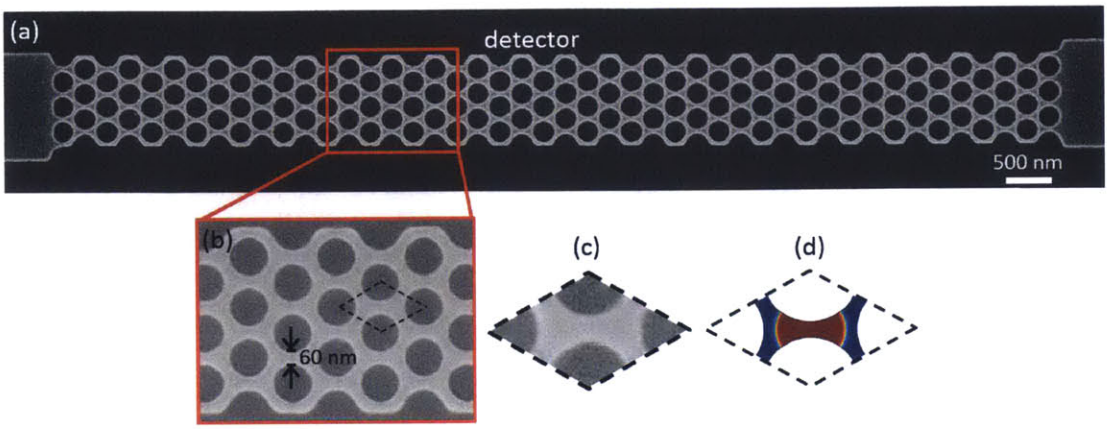

Figure 4-7: (a) Top-down SEM of a 12-μm-long NMSPD. (b) Magnified SEM of NMSPD shown in (a). (c,d) SEM and representative simulated ρ-distribution of an NMSPD unit cell.

where ρ_{300K} is the room-temperature resistivity, $\lambda_{\text{eff}} = \lambda_{\text{NbN}}^2/d$ is the zero-temperature magnetic penetration depth of an NbN film of thickness d. Since the proportionality factor γ is a constant for all devices on the same film, it can be extracted from a simple reference nanowire of known length and width on the same film, by calculating its kinetic inductance and measuring its room-temperature resistance.

The NMSPDs had inductances of 10-60 nH. Furthermore, a 50nH series inductor was added to some detectors to force them into the traditional SNAP avalanche mode with a long resistive region. Several parameters had to be evaluated:

- Dead time: Can the NMSPD operate in the simulated 'quenching mode', where the hotspot growth is limited by diffusive cooling in the structure?
- Detection: Does the NMSPD operate as a non-latching single-photon detector?
- Jitter: Does the timing jitter compare with SNSPDs?

Figure 4-9 shows the count rate of an NMSPD of a design as in Fig. 4-7(a) as a function of incident photon flux. The linearity for the photon flux, which was in the range of 0.1-10 M photons/second, is a strong indication for the operation in single-photon regime.

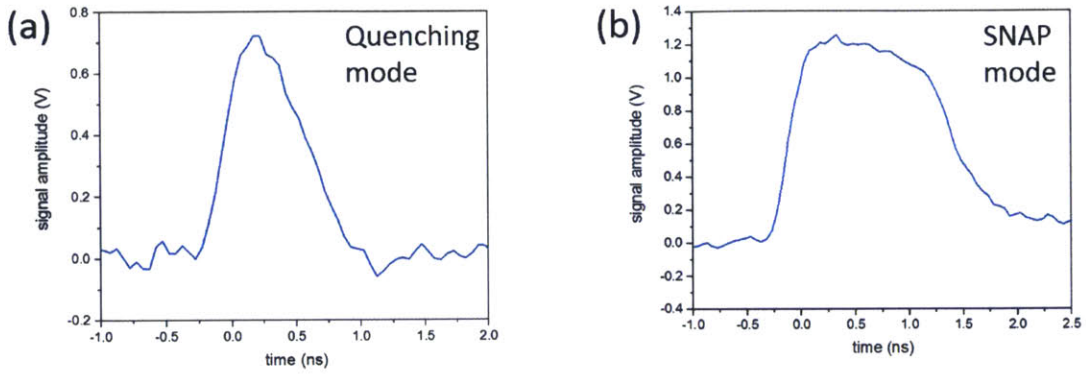

Figure 4-8: (a) Detector pulse of an NMSPD with L_{tot} ~10 nH, operating in quenching mode with ~500 ps reset time. (b) An NMSPD with the same dimensions as (a) connected to a series inductor with $L_{\text{S}} = 50$ nH operating in SNAP mode with a reset time of ~1.5 ns.

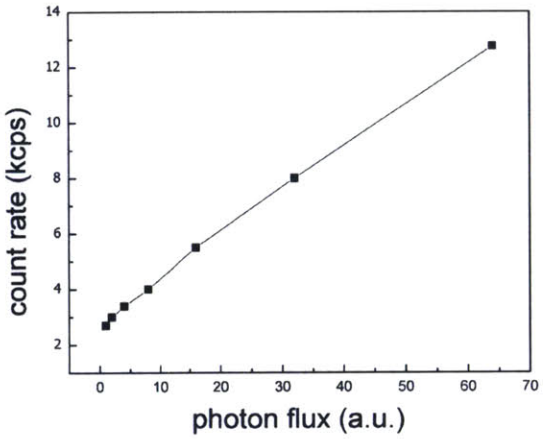

Figure 4-9: Count rate vs. photon flux for an NMSPD with the same design as in Fig. 4-7(a) operated in quenching mode.

Dead time and jitter results are shown in Fig. 4-10. The NMSPDs reached 400-600 ps non-latching dead times and 50-60 ps FWHM timing jitter.

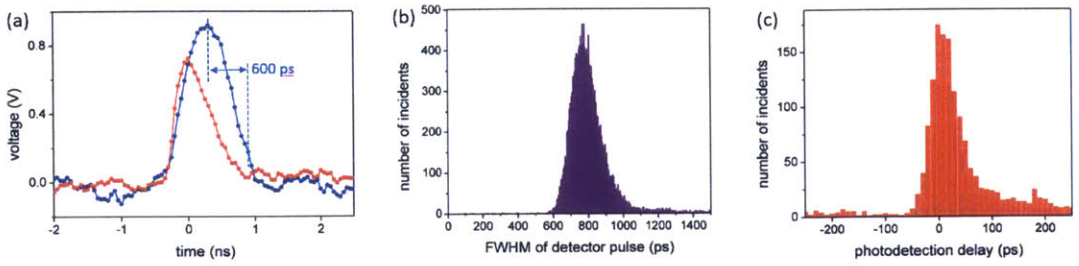

Figure 4-10: NMSPD detection pulses (a), histogram of detection pulse FWHM (b) and histophotodetection delay (c) measured with 3-GHz-bandwidth room-temperature amplifiers (MiniCircuits ZX60-3018G).

4.5 Summary and outlook

We examined the limits to the reset time of superconducting nanowire-based single-photon detectors in NbN. Joule heating, responsible for the growth of the resistive region inside the nanowire after an HSN event, was identified as a major bottleneck to the detector speed. In order to limit the growth of the resistive region, we proposed a new detector design, the nano-mesh single-photon detector (NMSPD). Preliminary experimental results showed that NMSPDs can operate with sub-1-ns dead times and ~60-ps timing jitter. While the operation of these detectors in single-photon (avalanche) regime (e.g., see Fig. 4-9) indicates that the prototype devices do not exhibit severe constrictions, further characterization and design optimization are likely needed in the future to characterize device efficiency and yield due to less severe constrictions (e.g., due to fabrication or current crowding).

Chapter 5

Membrane-integrated detectors

We developed a micron-scale flip-chip process that allowed the scalable integration of SNSPDs with photonic chips. Figure 5-1 shows a sketch of the integration concept: A detector is aligned to a waveguide and placed face-down onto the photonic chip using a micro-manipulated probe and a membrane as detector carrier. Electrical pads on the photonic chip make surface contact with matching detector pads on the membrane. The contact pads on the photonic chip are connected to larger pads appropriate for wire bonding.

5.1 Membrane design

Figure 5-2(a) shows a basic (initial) design of a suspended membrane-detector. The small size of the membrane allowed for the selective transfer of detectors. In the initial design, the membrane was connected to the bulk substrate via four $\sim$15-μm-long microbridges. These bridges had an unpredictable breaking pattern (Fig. 5-2(b)), resulting in fractured SiN_x pieces that could fall in between the membrane and the photonic chip surface and prevent tight contact between the detector and the waveguide. In order to avoid residual SiN_x pieces we modified the bridge design as shown in Fig. 5-2(c): the bridges were shorter ($\sim$3 μm long) with a $\sim$0.8 - to 1.5-μm-wide constriction in the middle section of the bridge, resulting in a preferred breaking region marked by the dashed red lines. With this improved design most membranes

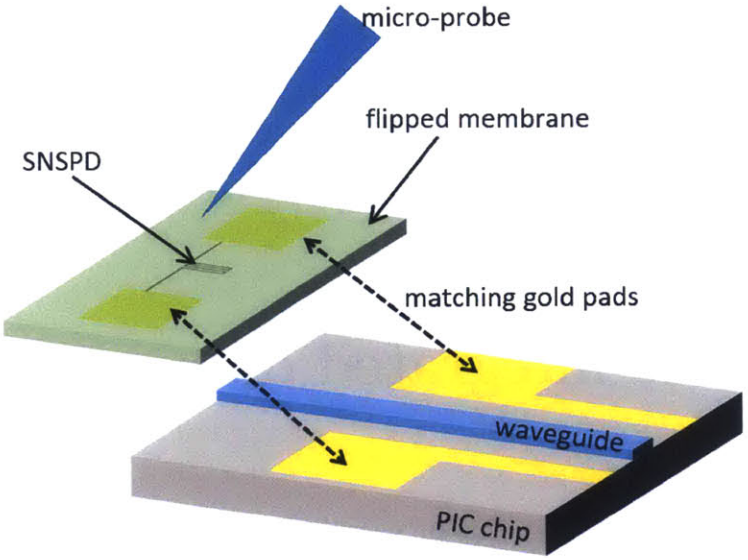

Figure 5-1: Sketch of waveguide-integrated SNSPD assembly using a flip-chip approach.

could be removed from the bulk substrate (Fig. 5-2(d)) without substantial residual SiN_x pieces. Other etch-chemistry-specific aspects of the membrane design shown in Fig. 5-2(c) will be discussed in the next section.

5.2 Membrane fabrication process

There are two basic approaches to fabricating membrane-detectors:

(1) The membrane is fabricated first, and the detectors are patterned directly on top of the membrane; or

(2) The detector is fabricated on a solid substrate first, and a thin layer under the detector is suspended in a second step.

Direct nano-fabrication on top of a membrane would be very demanding due to challenging resist application and uniformity on membranes and a challenging liftoff process since no contact masks, sonication and solvent squirts can be used. A more

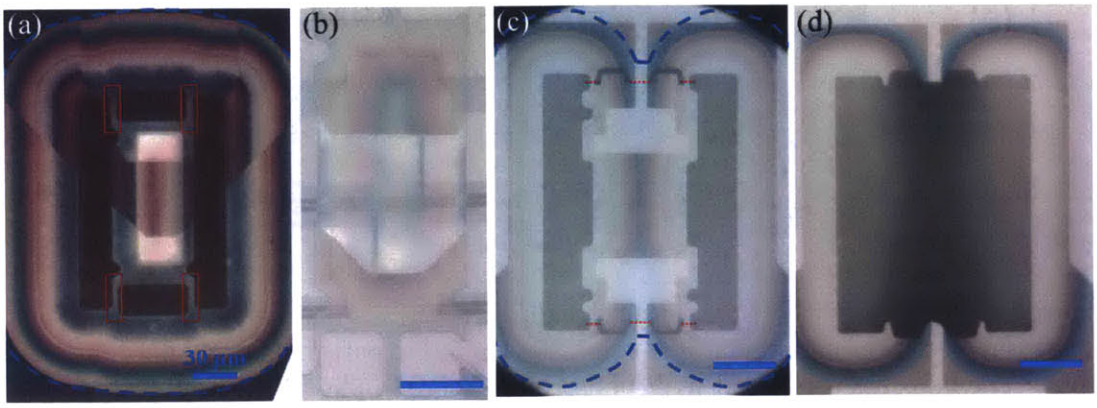

Figure 5-2: Top-down optical micrographs of membrane-detectors. (a) Suspended membrane held with long microbridges (enclosed with red lines) and surrounded by four large trenches. The dashed blue lines separate the undercut SiN_x region from the bulk substrate. (b) Transferred membrane with a design similar to the membrane shown in (a). (c) Suspended membrane with only two large trenches and short microbridges with constrictions. (d) Remaining structures on the primary SiN_x chip after the membrane identical to the membrane shown in (c) has been removed. The equivalent length of the blue scale bar is 30 μm.

straight-forward membrane fabrication approach is the fabrication of an SNSPD on top of a thin layer, e.g. SiN_x or SiO_x, covering a bulk substrate (e.g. silicon) that can be selectively removed after detector fabrication while leaving the thin layer on top intact. We followed this approach using a SiN_x-on-Si substrate due to the wide range of selective etchants for this material combination. A SiN_x layer (typically 200- to 300-nm-thick) was grown via plasma-enhanced chemical vapor deposition (PECVD) on double-polished silicon substrates. The NbN film was deposited on top of the SiN_x layer via reactive magnetron sputtering (AJA system) at a substrate holder temperature of 800 °C. The sheet resistance of the 4-nm-thick NbN films (thickness estimated from the deposition time) was 515 Ω/square and the critical temperature was 10.9 K. Electrical contact pads were defined via bi-layer liftoff as discussed in chapter 2. The pads comprised an e-beam-evaporated 15-nm-thick gold layer on top of a 10-nm-thick titanium layer. The SNSPD was fabricated using the process outlined in chapter 2.

One of the most common processes for SiN_x membrane fabrication involves a wet etch process using Tetramethylammonium hydroxide (TMAH). In order to expose

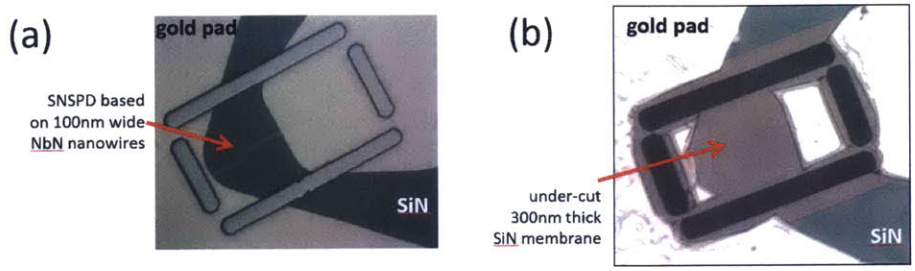

Figure 5-3: (a) An SNSPD on top of a bulk SiN_x-on-Si substrate. The underlying silicon is exposed through rectangular trenches that were patterned using photolithography. The detector is covered with photoresist from the photolithography step. (b) A detector from the same chip shown in (a) after $\sim$ 3 hours in 5% TMAH at 90°C

the silicon we patterned trenches around the detector area using photoresist (Fig. 5-3(a)) as mask. The detector chip was then dipped in a TMAH solution (5%) that was heated to 90°C. After $\sim$ 3 hours the entire detector region was undercut. The resulting membrane, shown in Fig. 5-3(b), exhibited two issues:

(1) The Au pads were lifting off due to etching of the underlying NbN from the sides.

(2) The detector layer was mostly etched away, and the photoresist on top did not offer sufficient protection.

We addressed (1) by fabricating smaller gold pads that did not reach the edges of the membrane, as shown in Figure 5-4. Small gold leads connected the membrane-detectors electrically to larger pads on the substrate (see Fig. 5-4(b)), allowing for pre-transfer electro-optical characterization of the suspended detectors.

In order to address (2) we searched for alternative protective coatings that could withstand the TMAH etch step. Figure 5-5 outlines a process based on Protek, a commercial protective layer by Brewer Science Inc. used in the MEMS community. A $\sim$4-μm-thick Protek layer was spun on top of the SNSPD sample, followed by a $\sim$2-μm-thick S1813 photoresist layer, as shown in Figure 5-5(a). The trenches were transferred into the Protek and the SiN_x layers via photolithography and RIE (CF_4 at 150 W for 18 minutes). The resulting stack, shown in Figure 5-5(b), was exposed to TMAH as outlined above. The Protek was resistant to the TMAH (Fig. 5-5(c)) and sufficient to protect the detector. However, subsequent removal of Protek in

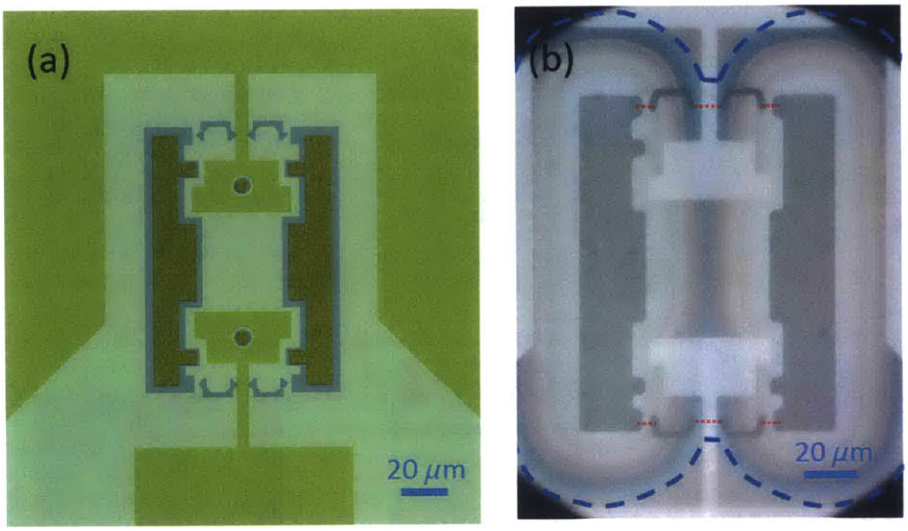

Figure 5-4: (a) Sketch of SiN_x membrane with Au pads on top. The yellow layer represents the gold pads. The grey regions represent the trenches surrounding the membrane. The brown regions are exposed during the selective silicon etch. Some designs included holes within the membrane to speed up the membrane undercut process. (b) Top-down optical micrograph of a suspended membrane.

solvents (see appendix) proved to be highly unreliable due to hardening of the resist during the RIE step and membrane collapse under the weight of the Protek during the removal step. Faced with these challenges we decided to transition to a different etch chemistry.

The final suspension process is outlined in Figure 5-6. The detector was covered with S1813 and a trench pattern was exposed in the photoresist (Fig. 5-6(a)). This pattern was then used as an etch mask to define trenches around the detector through the SiN_x layer via RIE with CF_4 (Fig. 5-6(b)). This trench pattern left the underlying silicon substrate exposed. The silicon under the SiN_x layer was removed using XeF_2, a selective isotropic etch gas (Fig. 5-6(c)). In the final step, the photoresist was removed in an NMP solution (see appendix), resulting in a detector on a suspended SiN_x membrane.

During the membrane-detector fabrication process, illustrated in Figure 5-6, photoresist layers covering the detector are used to define the outline of the membrane with trenches (Fig. 5-6(b)) and to protect the detector during the Si etch step (Fig.

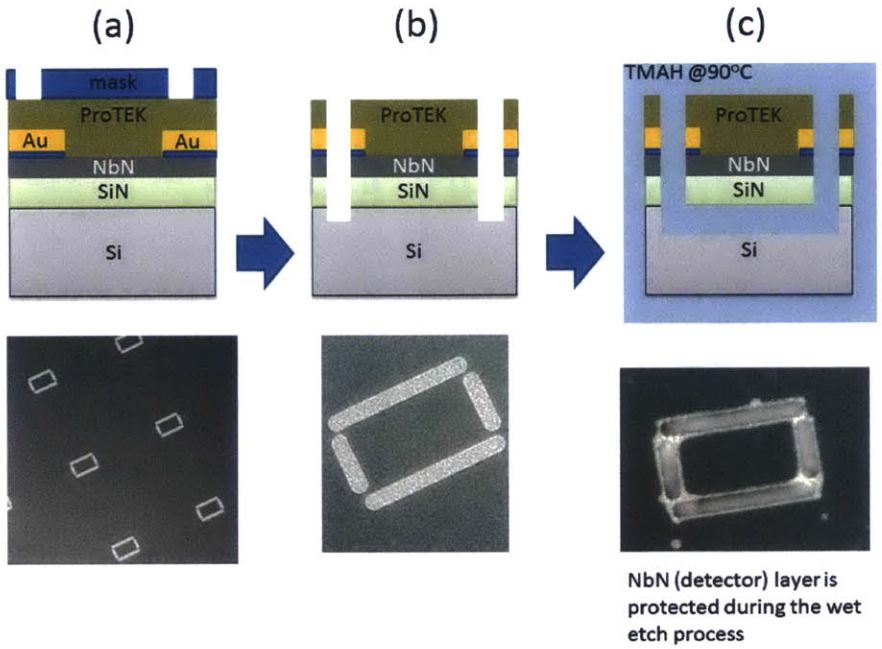

Figure 5-5: (a) Detector sample covered with Protek. (b) Trenches patterned through Protek and the SiN_x layers via photolithography and RIE. (c) Suspended membrane after silicon removal in TMAH.

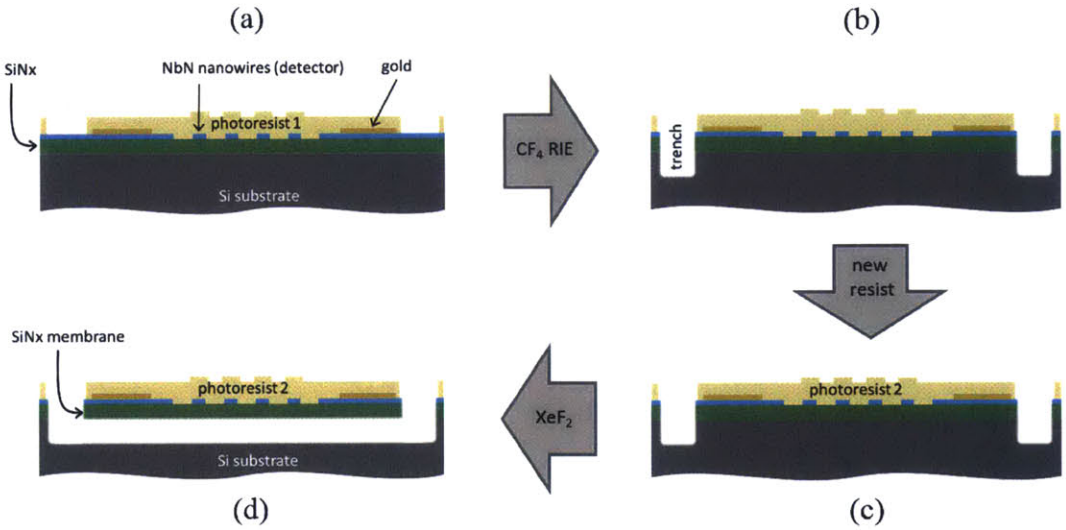

Figure 5-6: Schematic cross-section illustrating the membrane-detector fabrication process.

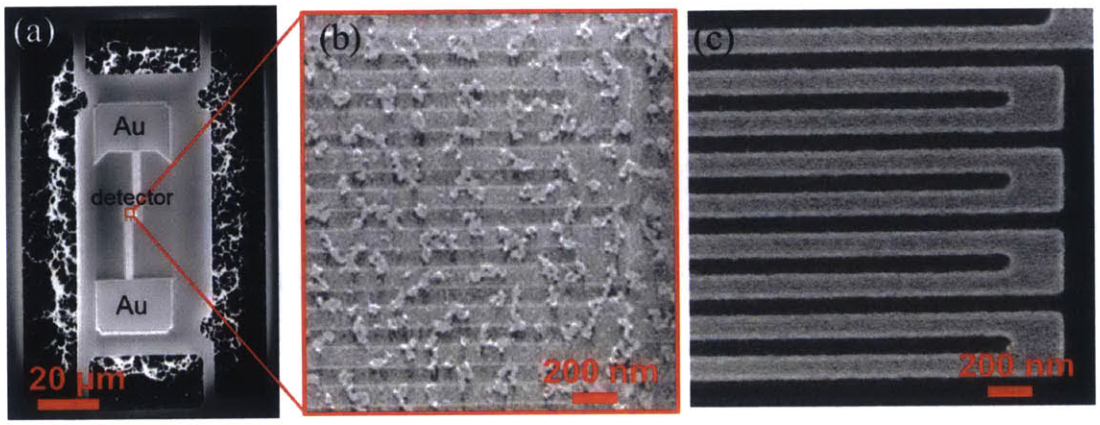

Figure 5-7: (a, b) SEMs of membrane-detector after the protective photoresist was stripped in an oxygen plasma. (c) SEM of membrane-detector after the photoresist was stripped in an NMP solution.

5-6(c)). Initially the protective etch mask that was used to fabricate the trenches via reactive ion etch (RIE) with CF_4 was also used as a protective layer in the subsequent etch step with XeF_2. The fluorine gas (plasma) treatment during the RIE fluorinated the surface and hard-baked the resist, making it irremovable in solvents unless ultrasonic agitation was used. However, sonication could not be used after membrane undercut since it was found to cause membrane collapse. Oxygen-helium plasma (ashing) was the remaining option, but we could not remove the hard-baked residue after ashing, shown in Figure 5-7(b). We solved this issue by removing the resist mask after the trenches were fabricated via sonication (photoresist 1 in Figure 5-6(b)), and coating the detectors with a new resist mask for the silicon removal step (photoresist 2 in Figure 5-6(c)). Since the second mask was not exposed to a long CF_4 etch, we were able to remove it in an NMP-based resist stripper followed by an Acetone and IPA rinse. While requiring an additional photolithography step, this stripping process did not leave a visible residue on the nanowires, as shown in Figure 5-7(c).

5.3 Pre-transfer electrical testing

Before transferring membranes onto the photonic chip, we characterized the room-temperature resistance R_{after} of detectors suspended on membranes and compared to detector resistance values R_{before} before the substrate was removed. Figure 5-8(a) shows that the relative detector resistance change $(R_{\text{after}} - R_{\text{before}})/R_{\text{before}}$ was 1-2%, indicating no significant material damage to the detectors due to the membrane fabrication process. Figure 5-8(b) shows the critical current of membrane-detectors that were successfully transferred onto a secondary substrate. The membranes here consisted of ~300-nm-thick SiN_x. The critical currents of detectors on 300- to 400-nm-thick membranes were suppressed by ~10% compared to values measured on the solid substrate before undercut, while critical currents of detectors on sub-200-nm-thick membranes were suppressed by ~10-20% (see chapter 6). Due to the small change in room temperature resistance values (Fig. 5-8(a)) we attribute the critical current suppression to the lower thermal capacity of the membranes compared to a solid substrate. Thermal cycling did not result in a measurable degradation (within the measurement accuracy of ~0.5 μA) of the critical current of the transferred detectors.

5.4 Summary

We developed a fabrication process that allowed the suspension of SNSPDs on top of sub-micron-thick SiN_x membranes. The central challenges were performing the undercut of SiN_x layer without significantly damaging the detector, and developing a process that allowed for reproducible removal of the layer that protected the SNSPDs during the etch. The final process included a selective dry-etch step based on XeF_2. Furthermore, we modified the membrane design in order to:

(1) Protect the gold pads during the etch step;

(2) Allow for reproducible bridge breaking behavior during transfer and for reduced residual SiN_x pieces; and

(3) Enable pre-transfer electro-optical characterization of the suspended detectors.

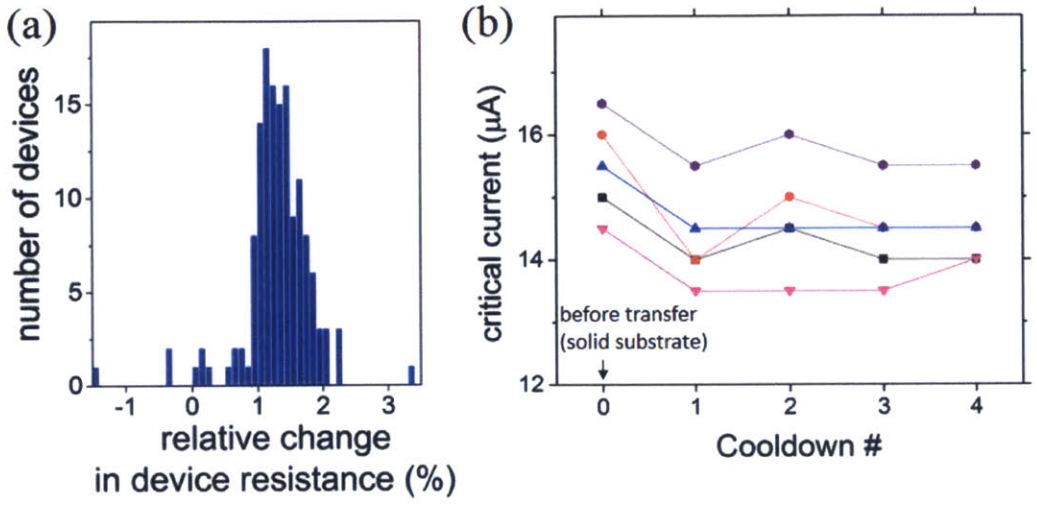

Figure 5-8: (a) Histogram of relative change in room-temperature detector resistance after membrane undercut compared to the resistance values before membrane undercut (suspension). (b) Critical current of detectors that were successfully transferred onto a secondary substrate on ~300-nm-thick SiN_x membranes. Up to four thermal cycles were performed between ~2.8 K and room temperature.

In the next chapter we will discuss the design of the photonic chip and the membrane transfer process.

Chapter 6

Integration of SNSPDs on photonic chips

The fabrication of SNSPDs on top of SiN_x membranes was discussed in detail in the previous chapter. The central goal of the membrane-SNSPD was to enable the scalable integration of NbN detectors onto photonic waveguides. The focus of the work presented here was the integration of SNSPDs with silicon photonic integrated circuits. In this chapter we will discuss the optical design of the detector and the waveguide.

6.1 Waveguide-detector design

Figure 6-1 shows a waveguide-detector assembly scheme: Hairpin-shaped SNSPDs [55] on top of membranes (transparent green) are placed on top of waveguides (red) and absorb photons travelling in the waveguide. The silicon waveguides used here had a cross section of 220×500 nm^2 and were single-mode for 1550-nm-wavelength light.

We used a 2D finite-element model in COMSOL to calculate the detector-to-waveguide coupling length that would provide sufficient optical absorption in the detector. Figure 6-2(a) shows the cross-section of the simulated geometry superposed by the electromagnetic TE mode profile. In order to achieve significant optical absorp-

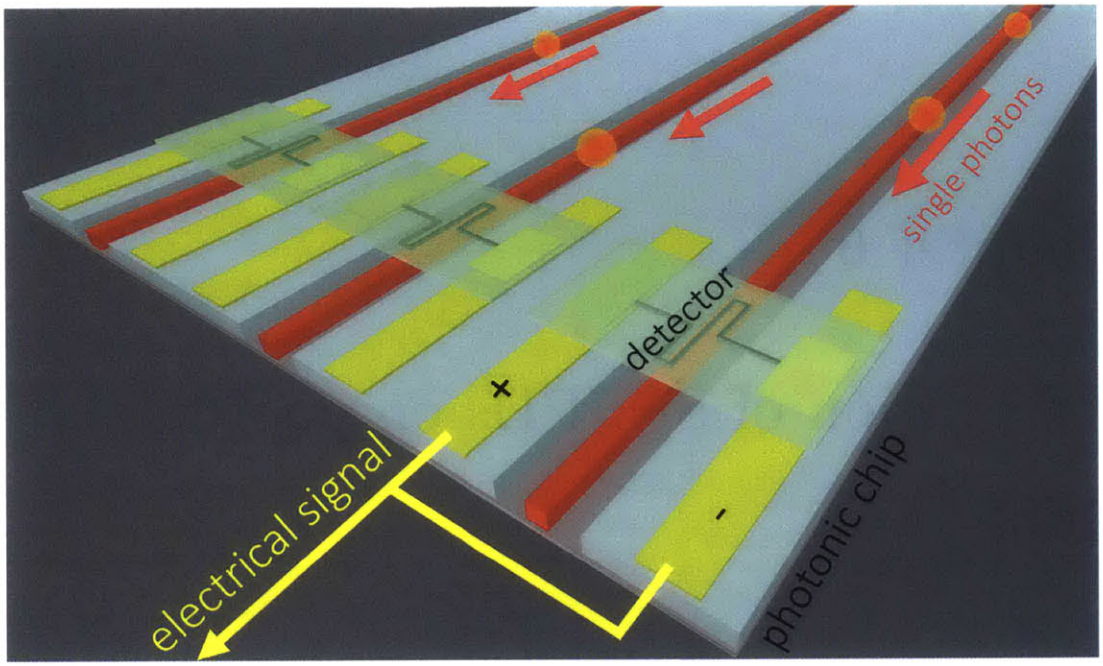

Figure 6-1: Waveguide-detector assembly scheme: Hairpin-shaped SNSPDs on top of membranes are aligned and placed onto an array of waveguides.

tion it is crucial that the evanescent field of the optical mode has significant overlap with the SNSPD. Following the procedure in Hu et al. [55], we used the COMSOL model to calculate the complex effective refractive index n_{eff} of the structure shown in Fig. 6-2(a). A larger imaginary component of the index means a larger optical absorption in the nanowire. The optical absorption A as a function of coupling (detector) length L_C is then given by

$$A = 1 - exp\,(-\,(2\pi/\lambda) \cdot \mathrm{Im}(n_{texteff}) \cdot 2 \cdot L_{\mathrm{C}}) \tag{6.1}$$

where λ is the light wavelength. Figure 6-2(b) shows the optical absorption in an SNSPD based on 80-nm-wide nanowires as a function of L_{C} for a separation of 20, 60 and 80 nm between the detector and the waveguide. The separation is due to residual resist remaining after the SNSPD and waveguide fabrication processes.

For large enough L_{C}, eventually most of the light will be absorbed by the detector. In fact, waveguide-integrated SNSPDs with up to 90% on-chip detection efficiency

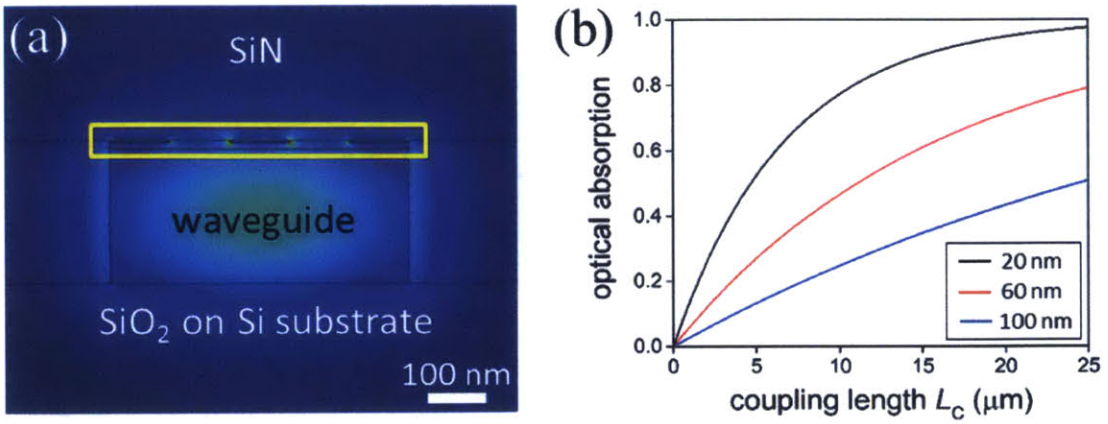

Figure 6-2: (a) Cross-sectional geometry of waveguide-integrated detector (marked in yellow) superposed by the simulated spatial distribution of the intensity of the waveguide eigenmode. The detector consists of 80-nm-wide, 4-nm-thick NbN nanowires arranged in a 200 nm pitch. The 500-nm-wide silicon waveguide was designed for 1550 nm center wavelength. (b) Calculated optical absorption in the detector vs. coupling length for a residual resist thickness of 20, 60 and 100 nm.

have been demonstrated by Pernice et al. [16]. We considered two additional factors besides detector length that could limit the practically achievable absorption: detector-to-waveguide misalignment and scattering at the interface of the SiN_x membrane due to index mismatch. Figure 6-3(a) shows the effect of misalignment on the optical absorption: Horizontal misalignment between the detector and the waveguide results in a sub-1% change in A. For this simulation, we assumed 80 nm resist and a 200 nm SiN_x membrane. We calculate the transmission κ at the SiN_x membrane edge using coupled mode theory. κ is therefore calculated as in Ref. [40], which gives the fraction of power transmitted from the silicon waveguide with air cladding to the silicon waveguide with SiN_x cladding. For our experimental conditions, $\kappa<1\%$. The perturbation is this small because the silicon nitride is thin, the index of silicon nitride is significantly lower than that of silicon, and the resist spacer layer on the silicon waveguide decreases the mode of overlap with the silicon nitride. We plot the mode pattern with and without the SiN_x cladding in Fig. 6-3(b).

The waveguide-detectors, shown in Fig. 6-4, consisted of four units connected in series, with each unit comprising two ~80-nm-wide nanowires (200 nm pitch) in

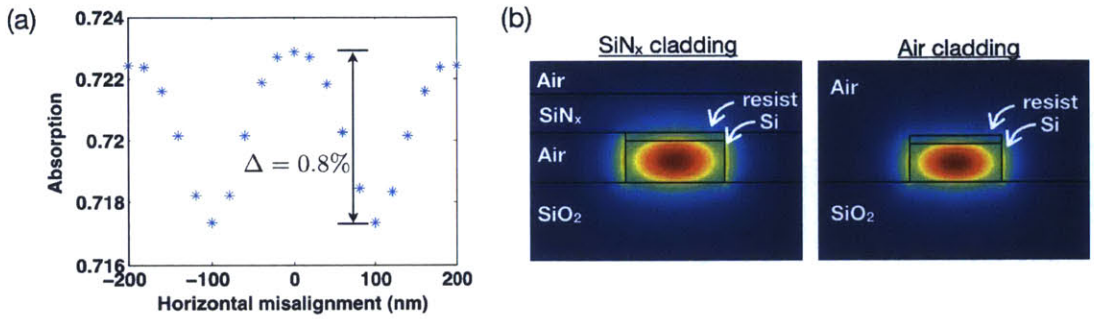

Figure 6-3: Imperfections in the detector-membrane integration. (a) Simulation of the absorption of the detector given in-plane misalignment orthogonal to the direction of light propagation. This simulates the tolerance of the detector performance to misalignment smaller than the size of the detector. The absorption is decreased by at most 0.8 percent. (b) Magnitude of the electric field of the fundamental TE mode of the waveguide with (left) and without (right) cladding of the waveguide by the SiN_x membrane. Black lines were drawn on the figure to more clearly display the simulated device geometry.

parallel. This series-SNAP design is similar to the designs discussed in chapter 3. The value of the series inductor L_S was generally chosen as $\sim$ 50 nH so that the total inductance in series with a single parallel-nanowire unit was about 3- to 7-times the series inductance of a single nanowire (see chapter 3). For each 2-SNAP, we need $> 3 \cdot L_{kin}$ of a single section in series to ensure that the detectors have a broad avalanche regime of at least 20% of the switching current of an unconstricted SNAP. Since every 2-SNAP has already three 2-SNAPs in series ($3 \cdot 0.5 \cdot L_{kin}$ of single section), we only need to add $L_S \leq 1.5 \cdot L_{kin}$ of a single section as series inductor. Further details on the inductance values are shown in table 6.1.

6.2 Fabrication of the photonic chip

The waveguide chip was fabricated in a CMOS processing facility at the IBM TJ Watson Research Center. The substrate was a 10 Ω-cm, p-doped, 200-mm silicon-on-insulator (SOI) wafer from SOITEC. The wafer had a 220-nm-thick silicon device layer on top of a 2 μm buried oxide layer. In a first step, the ~500-nm-wide silicon waveguides were fabricated via electron-beam lithography. In a subsequent optical

Nanowire width (nm)	Detector length (μm)	Nanowire length per SNAP section (μm)	Number of squares, L_{kin} per section (nH) assuming 80pH/square	$L_S \geq 1.5 \times L_{kin}$ per section	Total inductance L of detector (nH) = $L_S + 2 \cdot$ (L_{kin} per section)	Estimated reset time $=3 \cdot L/50\Omega$ (ns)
80	17	17	213, 17	50[2.9x]	84	5

Table 6.1: Calculated inductance values for series-2-SNAPs based on 80-nm-wide nanowires.

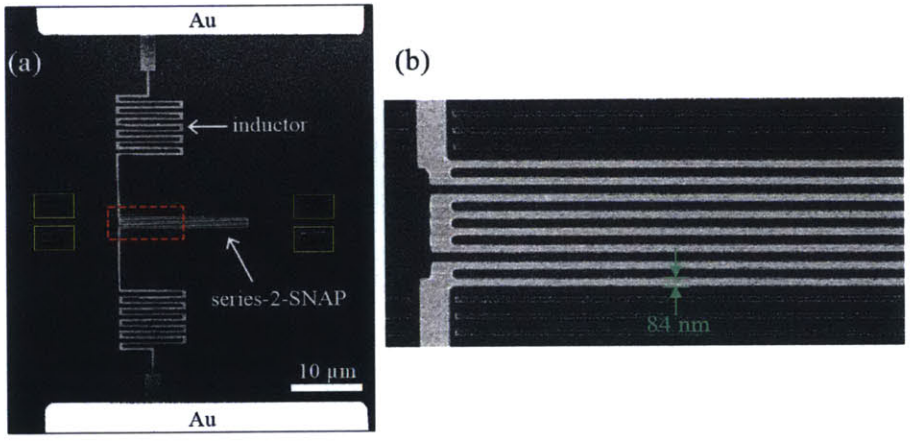

Figure 6-4: (a,b) Top-down SEMs of a hairpin-shaped series-2-SNAP with design parameters as in Table 6.1. The rectangular alignment marks are highlighted with yellow lines.

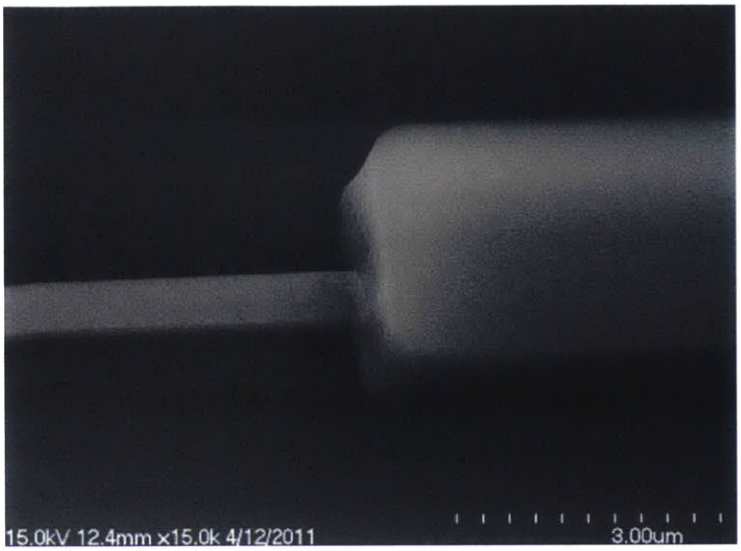

Figure 6-5: Angled SEM of SU8 coupler (left) on top of a silicon waveguide (right).

lithography step, SU8 polymer couplers (see Fig. 6-5) were fabricated to allow sub-4-dB coupling loss from a lensed fiber to the silicon waveguide (see Ref. [56] for further details).

The gold pads were added to the waveguide chip at MIT. There were several challenges when fabricating the gold contact pads:
(1) The width of the waveguide chip was ~ 7 mm, making the application of a uniform resist layer challenging;
(2) No sonication was allowed due to the potential damage to the SU8 couplers; and
(3) Nanometer-scale smoothness of the Au pad edge height profiles was required in order to ensure close contact between the detector and the waveguide.
The Au pads were defined using a bi-layer fabrication process as in chapter 2 (see appendix for further details), but without the sonication step, which could lead to delamination of the polymer couplers (see Fig. 6-6).

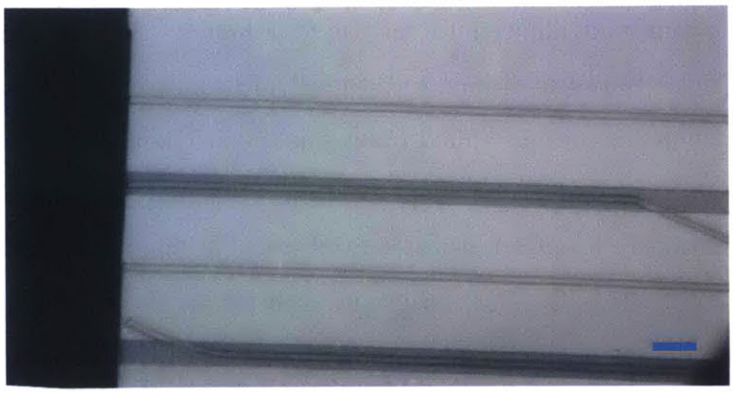

Figure 6-6: Top-down optical micrograph of the edge of a photonic chip with partially-delaminated SU8 couplers. The equivalent length of the blue scale bar is ~20 μm.

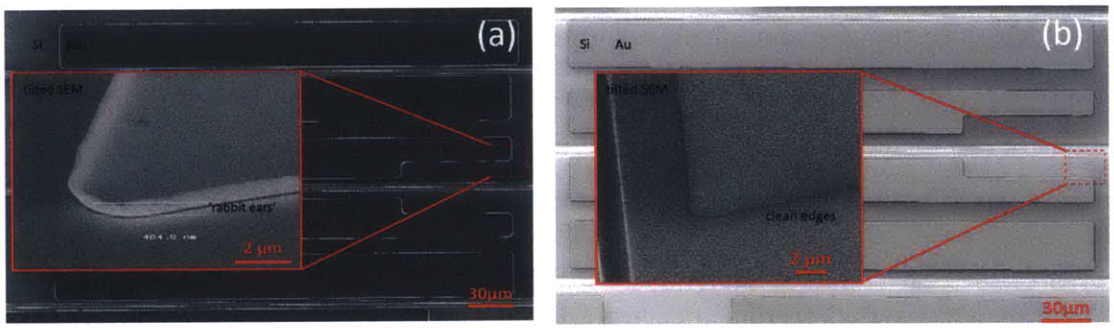

Figure 6-7: SEMs of gold contact pads fabricated using a single-layer (a) and a bi-layer (b) liftoff process.

6.3 Membrane transfer process

Polydimethylsiloxane (PDMS), an adhesive polymer, was mixed in a ten-to-one ratio with the curing agent and allowed to set for four hours. A tungsten microprobe (Ted Pella Autoprobe 100) was dipped into the PDMS solution, resulting in a PDMS droplet near the tip of the probe. This droplet was used to hold the membrane during transfer. The PDMS-covered probe was baked on a hot plate at 100°C for 8 hours, followed by sonication in an ethanol-water mixture. The baking and cleaning steps are essential to avoid residual PDMS after the transfer, as shown in Fig. 6-9, which can prevent tight contact of the membrane-SNSPD to the target chip surface.

To remove the detector membranes from the substrate, three of the six microbridges (shown in Fig. 6-8(a)) were broken using a plain tungsten probe. This probe was then placed under the membrane and used to bend the membrane upwards, as shown in Figure 6-8(b). A second tungsten probe, covered with PDMS droplet and mounted on a 6-axis micromanipulator, was then used to lift the membrane from the substrate, touching only the passive (back) side of the membrane (Figure 6-8(c)). The PDMS served as an adhesive surface during the transfer (Figure 6-8(d)) from the fabrication (carrier) chip to the secondary (PIC) chip. Using this probe, the membrane was flipped, aligned and placed (with the detector facing down) on top of a waveguide. The transfer and alignment were performed under an optical microscope (Figure 6-8(e)). After placement, the PDMS probe was used to press down on any regions of the membrane that exhibited interference fringes, indicating a separation between the PIC and membrane. The detector surface was not in contact with any PDMS or other surfaces during membrane pickup, minimizing contamination risk.

Figure 6-10 shows detectors aligned to silicon waveguide on a photonic chip using the alignment marks highlighted in red. The arrows in Figure 6-10(a) mark the boundaries between which the waveguide must be located for efficient detection. Of the four membrane-detectors placed, all were aligned to the waveguide.

Efficient detection requires close contact between the detector and the waveguides. Interference fringes served as a first visual indicator of the closeness of contact. The

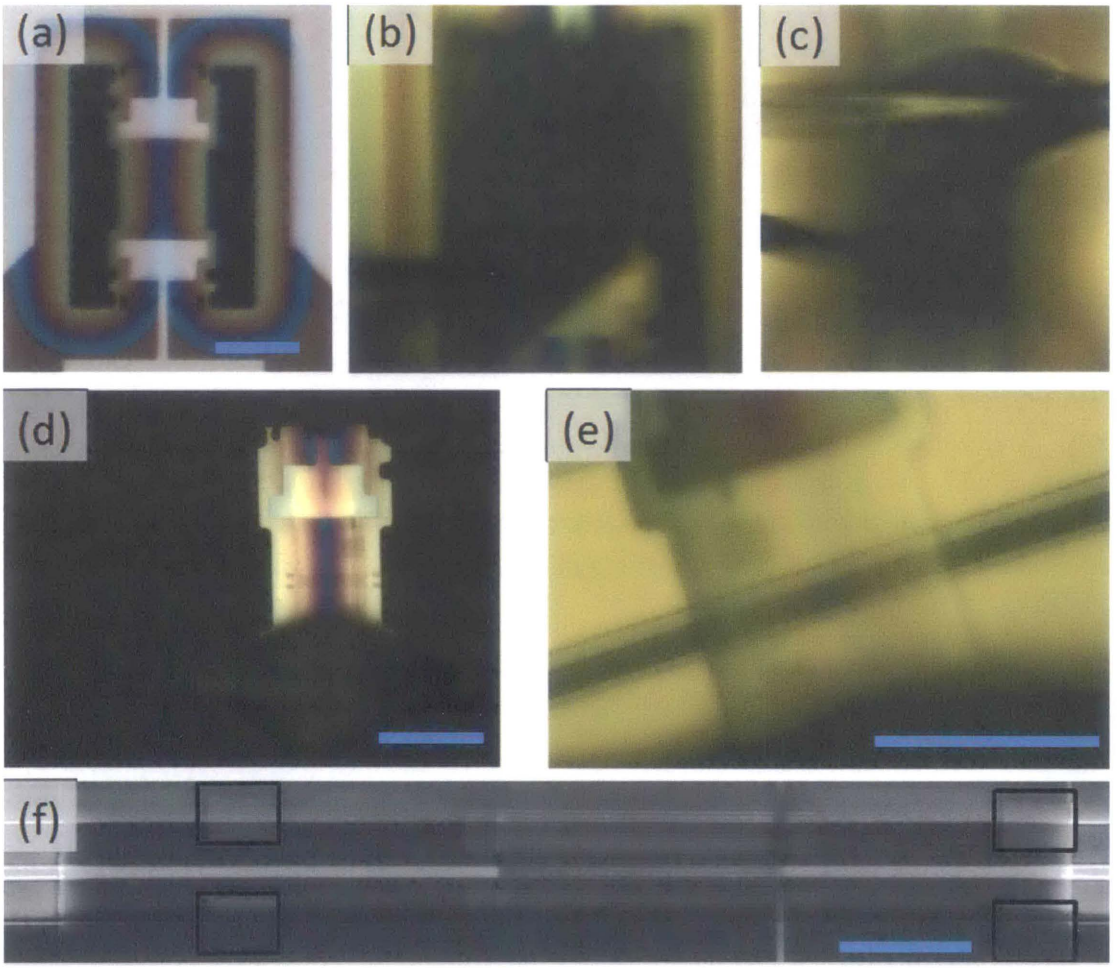

Figure 6-8: (a) Optical micrograph of a suspended membrane carrying a detector. (b) The top three microbridges are broken using a clean tungsten probe. The probe is then used to lift up the membrane. (c) A second probe with PDMS makes contact with the back of the membrane. While the second probe is holding the membrane, the first probe is used to break the remaining microbridges. (d, e) The second PDMS probe is used to pick up, flip and align the membrane to the optical waveguide. (f) Top-down SEM of resulting waveguide-integrated SNSPD. Alignment marks, highlighted with black lines, are used to align the SNSPD to the waveguide. The equivalent length of the scale bars in (a,d,e) is 40 μm. The equivalent length of the scale bar in (f) is 6 μm.

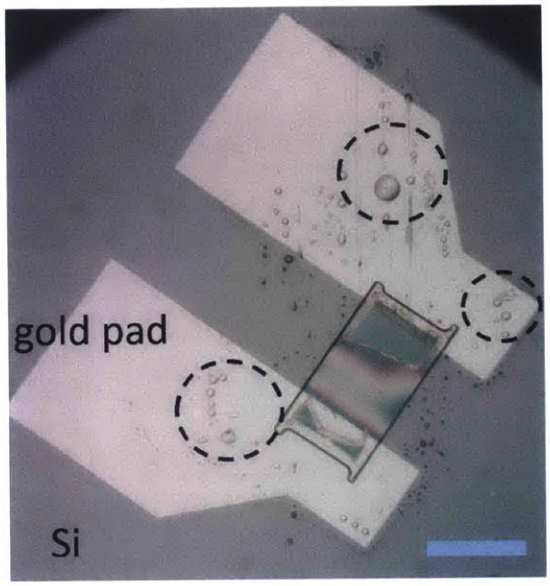

Figure 6-9: Residual PDMS (enclosed with black dashed lines) on the surface of the secondary chip after membrane transfer. The equivalent length of the blue scale bar is 50 μm.

detector shown in Figure 6-11(a) shows little to no visible interference fringes, implying close contact between the membrane and waveguide chip surface. In contrast, the membrane in Fig. 6-11(b) shows significant fringing in the central region above the waveguide as well as near the gold pads. The detector shown in Figure 6-11(b) would, in the best case, have poor detection efficiency and electrical properties.

Vacuum low-heat annealing is a possible path to improving membrane-to-PIC bonding. Figure 6-12 shows the detection efficiency and dark count rate before (a) and after (b) heating a waveguide-integrated detector at $\sim$ 100°C in vacuum for 4 hours, showing an increase in both the critical current and the maximum efficiency of the detector. This method was not part of the standard waveguide-detector assembly procedure, but one could consider adding a controlled heating step to future waveguide chips.

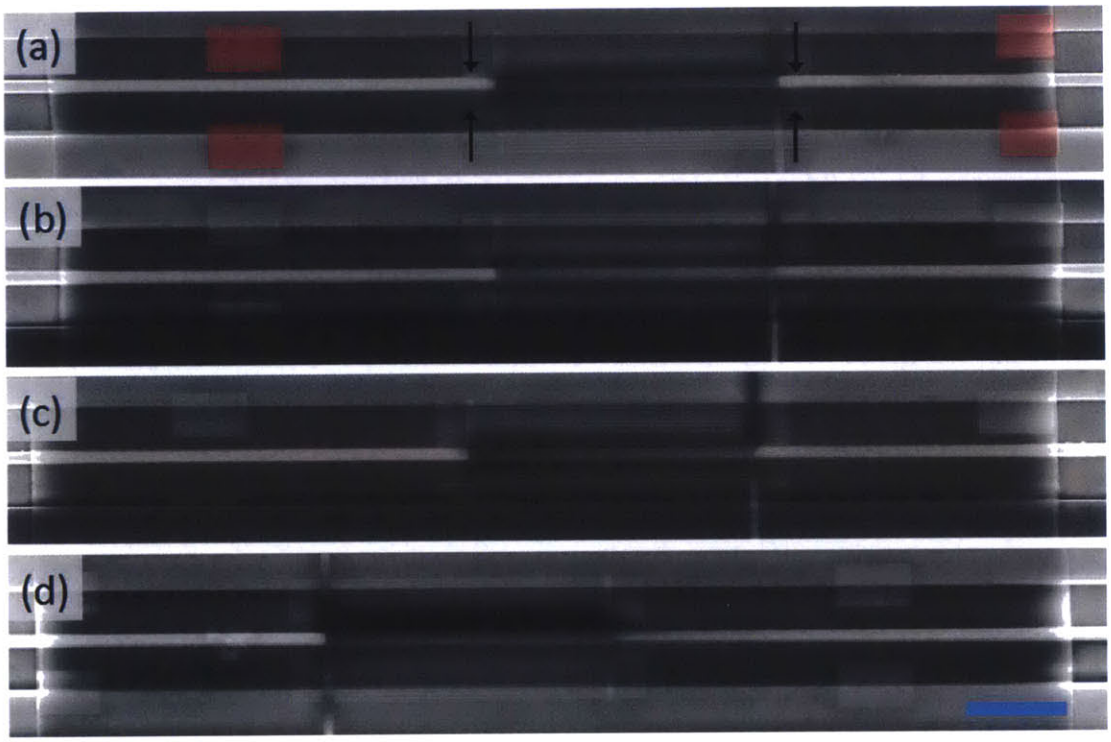

Figure 6-10: SEM of four detectors (out of a total of four transfers) aligned to 500-nm-wide silicon waveguides. The equivalent length of the blue scale bar is ~5 μm.

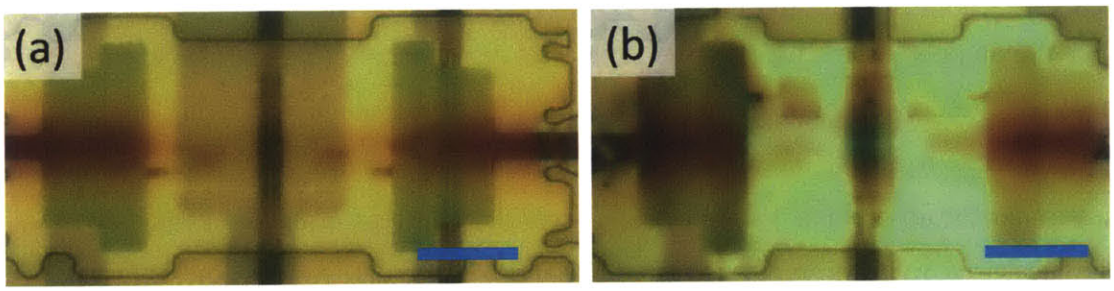

Figure 6-11: Optical micrographs of two different membrane-detectors after transfer. (a) Membrane-detector with negligible interference fringes. (b) Membrane-detector with visible fringing indicating poor contact. The equivalent length of the blue scale bars is 18 μm

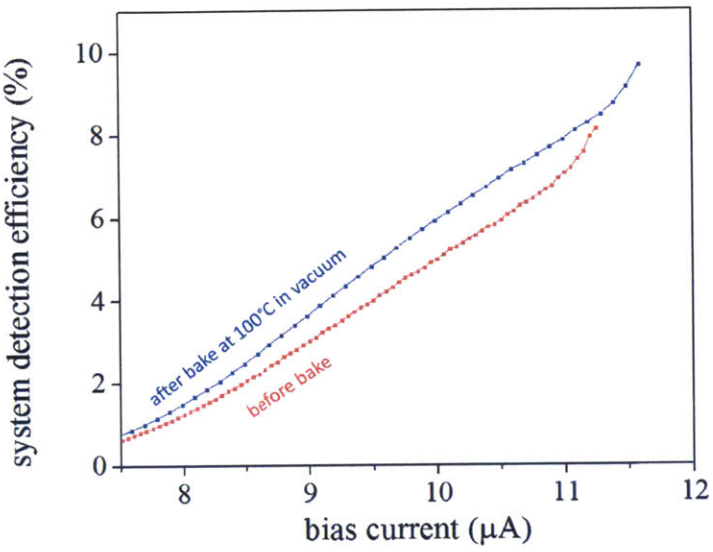

Figure 6-12: System detection efficiency vs. bias current for a membrane-detector on top of a waveguide before (red curve) and after a bake at 100°C in vacuum (blue curve).

6.4 Post-transfer electro-optical testing

Before transferring detectors, we measured the transmission of the silicon waveguides at room temperature (see setup in Fig. 7-11(a)) using lensed fibers identical to the fibers in the low-temperature setup (cryostat). The fibers in the room-temperature setup were mounted on high-precision piezo scanning stages, which are more precise than the stepper stages used in the cryostat. After transferring the detectors onto these waveguides, we measured the room-temperature transmission again to obtain the amount of light either scattered or absorbed by the detector. Our simulations indicate that the absorption is significantly larger than the scattering (see next chapter). The measured values were 74%, 74%, 65% and 62% for A1, A2, B1 and B2 respectively, with errors typically less than 5%. We note that the error and expected value of these transmission measurements were not used, and therefore do not contribute to subsequent detection efficiency calculations outlined in the next chapters. We performed the room-temperature optical absorption measurements simply to confirm intimate detector-to-waveguide contact before proceeding to later rounds of testing.

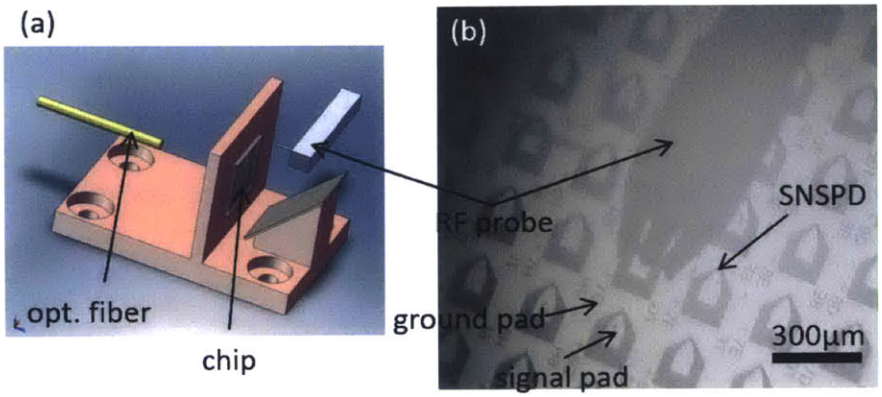

Figure 6-13: (a) Sketch of chip mount, optical fiber and RF probe in the cryogenic probe station setup used for quick electro-optical characterization of transferred SNSPDs. The detectors are illuminated from the top using a high-NA fiber. The chip is then placed inside a closed-cycle cryostat that allows coupling through the individual waveguides. (b) RF probe used to make electrical contact with SNSPDs for rapid pre- and post-transfer electro-optical characterization.

The photonic chip was then mounted into a closed-cycle cryostat and the detectors operated at ~3 K base temperature.

The critical currents after detector undercut and transfer were measured in a cryogenic probe station using the measurement scheme illustrated in Fig. 6-13. For the prototype chip discussed in the next chapter, the switching current values (15.2 µA, 16.8 µA, 16.4 µA, and 14.8 µA) were about 20% lower compared to pre-undercut values measured on the solid silicon substrate, possibly arising from the small thermal capacitance of the membranes, as noted in chapter 5.

Chapter 7

On-chip single-photon experiments

In this chapter we will review the experimental setup, methods and prototype chip used to perform rapid on-chip photon correlation measurements. The scalable integration of SNSPDs onto photonic circuits enabled us to perform the first detection of non-classical light directly on a photonic chip.

7.1 Experimental setup

We performed the measurements using four simultaneously-operated waveguide-detectors operated at ~3 K base temperature inside a cryostat with optical access. In this section we will describe the cryogenic, optical and electronic parts of the experimental setup.

7.1.1 The cryostat

Figure 7-1(a) shows the sample chamber (black cube) and RF cables (blue) used to read out the SNSPDs. An optical window allows for optical access to the photonic chip inside the cryostat. An optical microscope allows for a top-down view onto the photonic chip and is used to align the input optical fibers, shown in Fig. 7-1(b), to the waveguides on the chip.

Figure 7-2 shows the thermal cross section of the cryostat. The cryostat used a

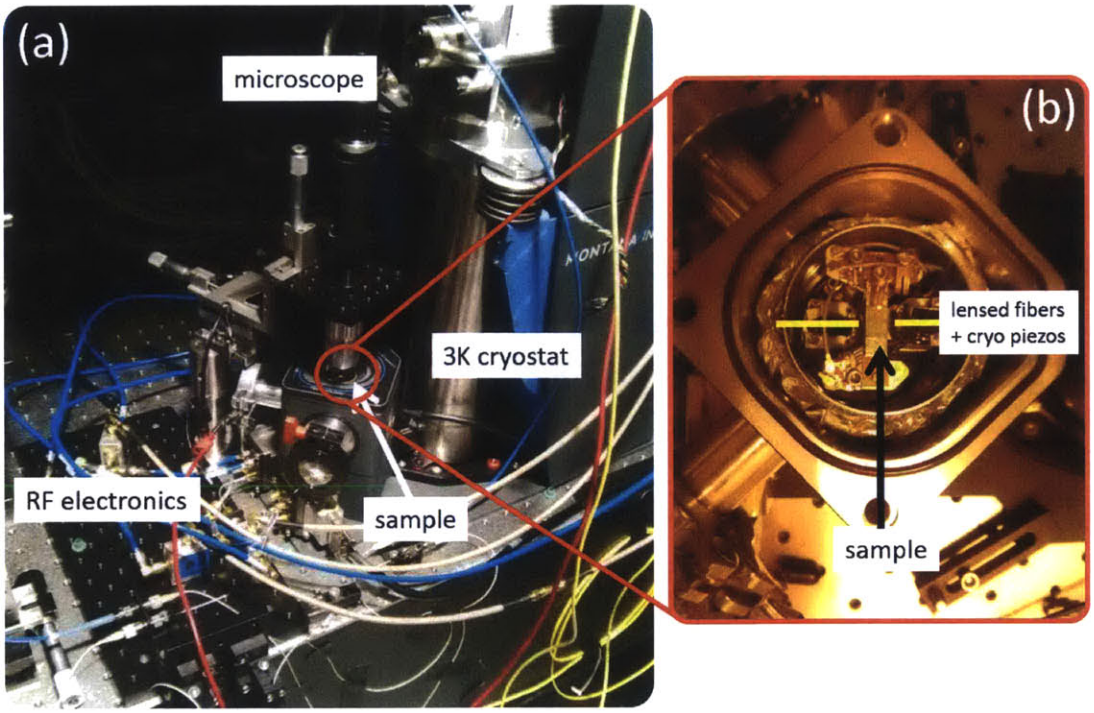

Figure 7-1: (a) Photograph of the cryostat used to operate the waveguide-integrated detectors. Active RF electronics and lasers are located outside the cryostat. (b) Top-down view of the sample chamber. The lensed fibers are marked with yellow lines.

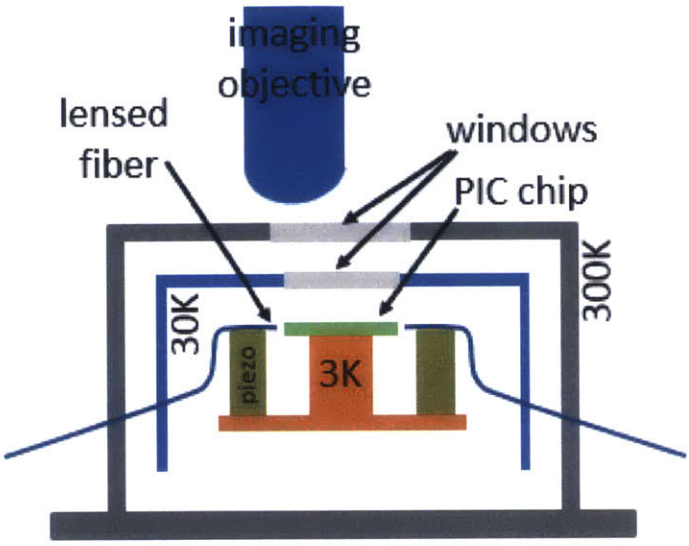

Figure 7-2: Cross-sectional sketch of the cryostat used to operate the waveguide-SNSPD chip.

commercial design by Montana Instruments to minimize mechanical vibrations: The cold head was fixed on the optical table and only connected to the Gifford-McMahon stage through thin copper braids. This connection provided a strong thermal link but a weak mechanical link, effectively shielding the vibrations in the pulse-tube stage from the cold head. The low vibration (<100 nm peak-to-peak, estimated at the top of the piezos) was key to keeping the optical coupling to the photonic chip stable, as shown in Fig. 7-2(d). The sample is mounted on top of a 3 K platform and surrounded by a radiation shield operated at ~30 K. In addition to the sample platform the cold head holds the two piezo nano-positioner stacks.

7.1.2 Optical coupling to the chip

Optical coupling to the chip was performed using SMF lensed fibers provided by OZ Optics. The lensed fibers are fixed onto the piezo stack using a detachable magnetic mount (see Fig. 7-3(a)). The coupling to the detector is as shown in Fig. 7-4(b): Infrared light (red arrow) was coupled from a lensed fiber with a spot diameter of

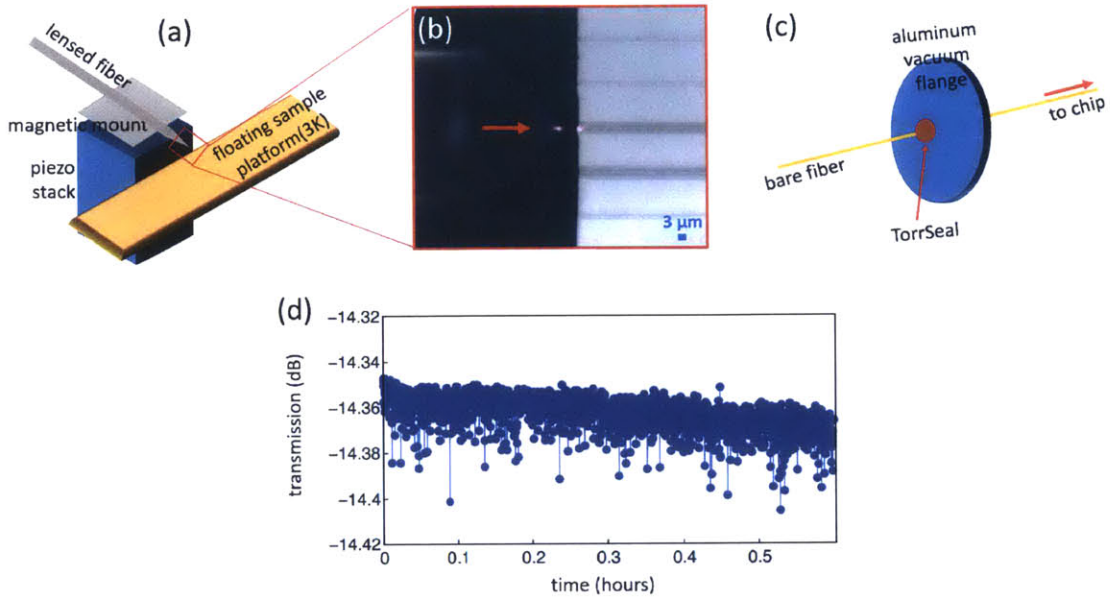

Figure 7-3: (a) Sketch of fiber-to-sample coupling scheme. A lensed fiber is mounted onto a cryogenic 3-axis piezo stack using a removable magnetic mount. The photonic chip is placed on top of a floating copper sample platform. (b) Top-down optical micrograph of a lensed fiber aligned to an SU8 coupler at the edge of the photonic chip. (c) Sketch of fiber feedthrough: A bare fiber is fed through a ~1 mm-diameter hole through an aluminum vacuum flange and sealed with TorrSeal epoxy. (d) Transmission stability throgh a waveguide at base temperature measured with two lensed fibers at each end of the waveguide.

2.5 μm into a 2 μm × 3 μm polymer coupler (Fig. 7-4(a)). The coupler overlapped with a 50- to 500-nm-wide inverse-tapered section of a silicon waveguide (Fig. 7-4(b)). The input light traveled along the 500-nm-wide waveguide (Fig. 7-4(c)) over a distance of 2 mm. For some of the experiments the waveguide fed into a 50:50 beamsplitter (directional coupler in Fig. 7-4(d)) followed by the waveguide-integrated detectors (Fig. 7-4(e)). The equivalent length of the scale bar (blue) is 3 μm. This coupling scheme provided a fiber-to-waveguide coupling loss of ~3.7dB (see section 7.2.1).

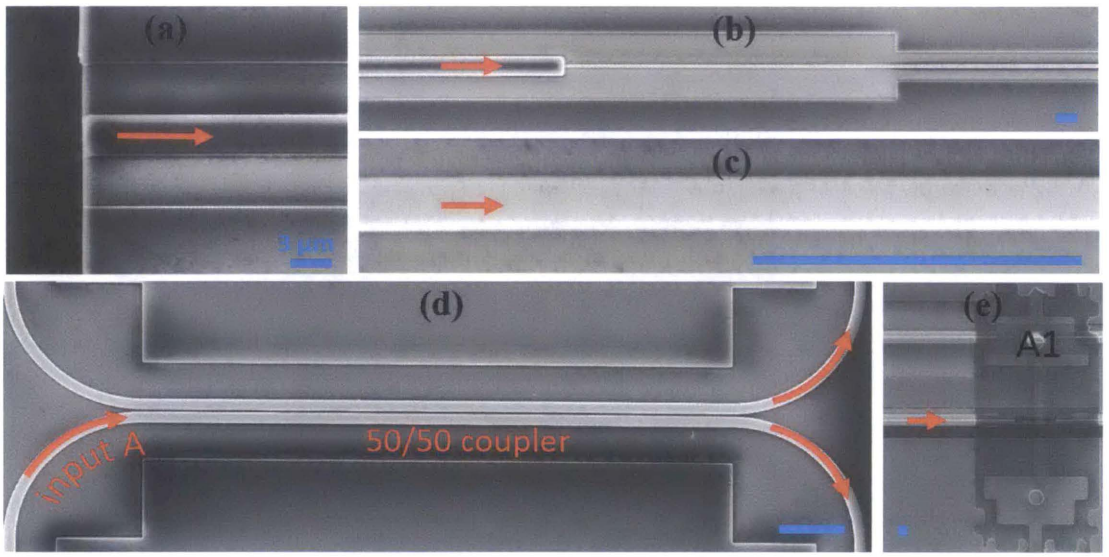

Figure 7-4: SEMs of different sections of the photonic chip. Infrared light (red arrows) was coupled from a lensed fiber (a) with a spot diameter of 2.5 μm into a 2 μm × 3 μm polymer coupler. The coupler overlapped with a 50- to 500-nm-wide inverse-tapered section of a silicon waveguide (b). The input light traveled along the 500-nm-wide waveguide (c) over a distance of 2 mm before reaching a 50:50 beamsplitter (directional coupler in (d)) followed by the waveguide-integrated detectors (e). The equivalent length of the blue scale bar is 3 μm.

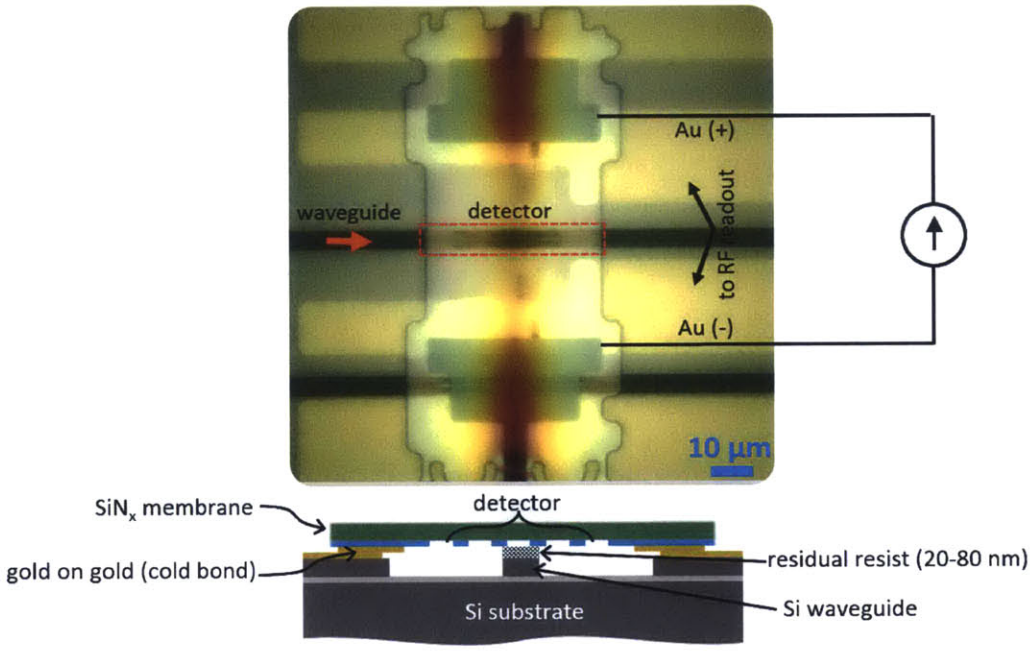

Figure 7-5: Top-down optical micrograph (top) and cross-sectional sketch (bottom) of a membrane-detector on top of a silicon waveguide. The detector was current-biased and read out through Au electrical contact pads on the PIC that matched contact pads on the flipped membrane.

7.1.3 Electrical readout

The SNSPD made electrical contact with matching gold-titanium pads on the photonic chip (see Fig. 7-5). The contact pads were wire-bonded to electrical traces on a printed circuit board (PCB) that was floating on top of the photonic chip (see Fig. 7-6). The PCB traces were connected to low-profile UMC RF connectors via 6-GHz-bandwidth miniature RF cables. The sample mount could be removed from the 3-K-platform of the cryostat to perform the sample mounting and wire bonding. Inside the cryostat, the UMC cables were connected to room-temperature SMA feedthroughs via low-thermal-conductivity (steel-core) RF cables. The surface area of the contact pads on the photonic chip was on the order of hundreds of μm^2 to allow for (1) significant distance between the bonding stub and the membrane to avoid damage to the membrane, (2) distance between the bonding stub and the waveguide to avoid scattered metallic debree on top of the waveguide that could result in additional scattering loss, and (3) a significant conductive area in order to avoid electrostatic discharge damage to the membrane-detectors, as shown in Fig. 7-7.

7.2 Waveguide-detector testing

Before performing on-chip detection of non-classical light we characterized the efficiency and timing performance of the waveguide-integrated detectors.

7.2.1 Detection efficiency

The detection of photon pairs on a chip requires the controllable integration of multiple high-efficiency single-photon detectors within the same circuit. Using the process outlined in the previous chapters, we integrated four detectors (labeled A1, A2, B1 and B2) on a PIC and characterized the performance of the assembled system shown in Fig. 7-8 using four parameters: system detection efficiency (SDE), on-chip detection efficiency (ODE), FWHM timing jitter (TJ), and noise-equivalent incident power (NEIP). The SDE includes all losses (i.e., coupling and transmission) between

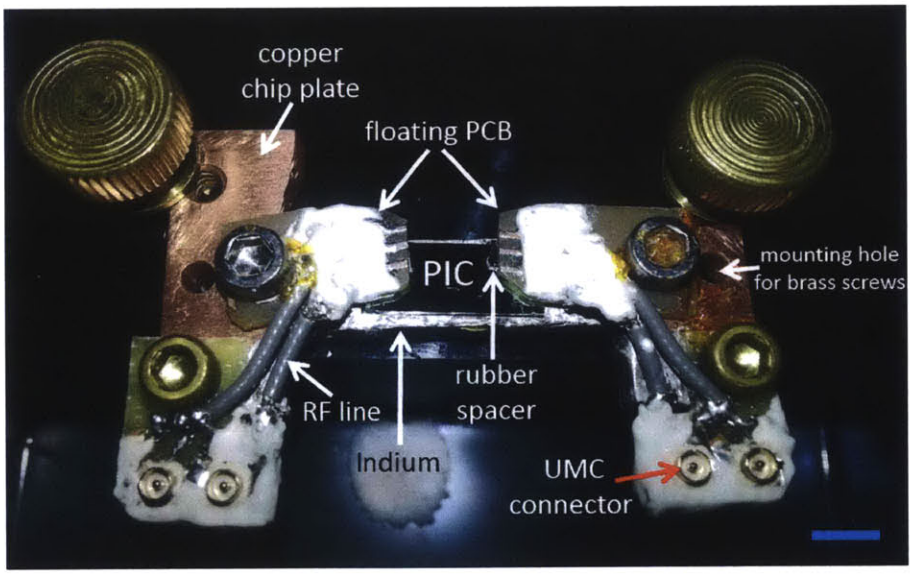

Figure 7-6: Top-down view of the cryogenic sample plate. The plate is removed from the cryostat to place a PIC chip on top of it. The PIC chip is wire-bonded onto traces on a floating PCB which also acts as a mechanical clamp that presses down the PIC onto a thermal-grease-covered indium layer. The length of the blue scale bar is $\sim$7 mm.

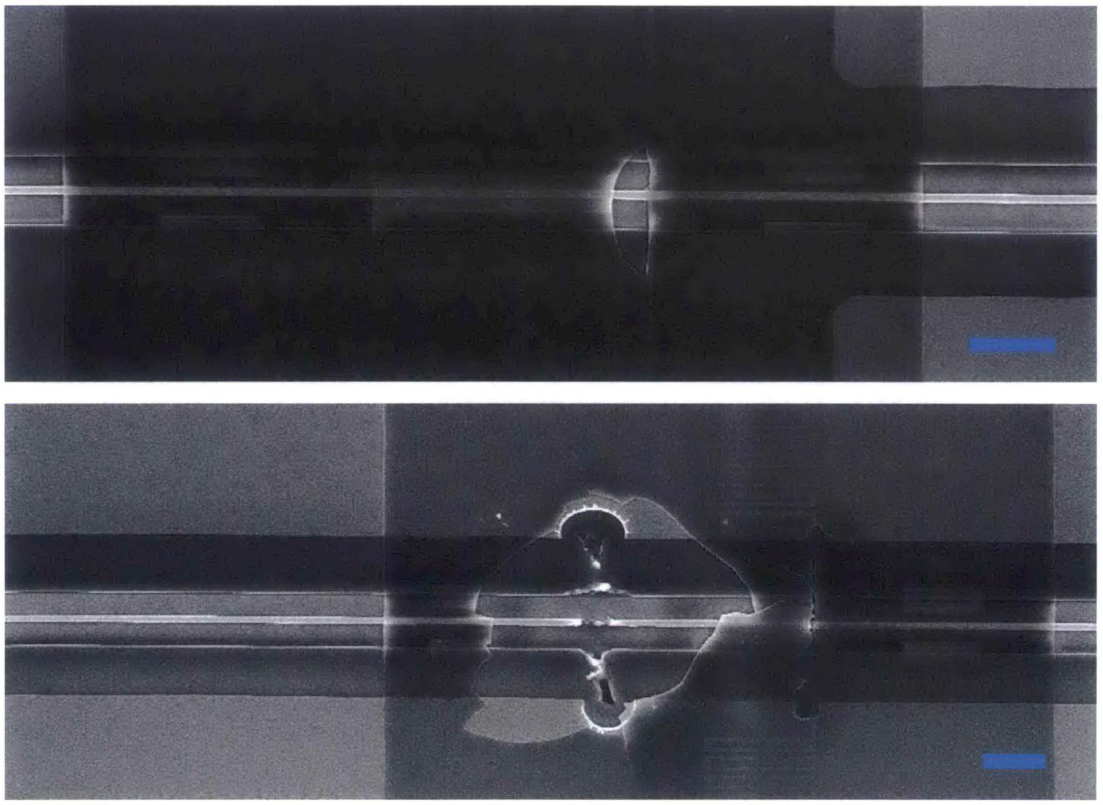

Figure 7-7: Top-down SEMs showing damaged membrane-detectors and waveguides due to electrostatic discharge. The equivalent length of the scale bars in blue is 5 μm.

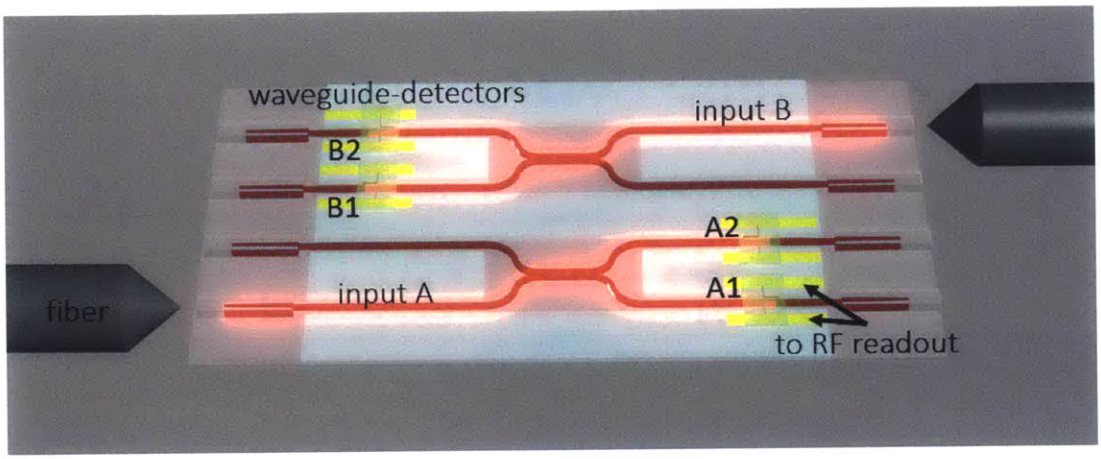

Figure 7-8: Sketch of proof-of-concept photonic chip with four waveguide-integrated detectors coupled to two optical inputs.

the fiber port outside the cryostat and the detector. We determined the SDE from the ratio of the SNSPD photocount rate to the photon flux coupled into the fiber port.

Two optical inputs (A and B) are each coupled to two waveguide-integrated detectors (A1,2 and B1,2) through a directional coupler acting as a 50/50 beamsplitter. A sketch of the experimental setup used to measure the SDE of the waveguide-integrated SNSPDs is shown in Figure 7-9(a). Light from a fiber-coupled CW laser (Thorlabs S3FC1550, emitting at $\lambda = 1550$ nm, output power 1 mW) was split into two outputs using a calibrated, fiber-coupled 50/50 splitter (Thorlabs 10202A-50-FC). One output, used to monitor the power directly, was coupled to an InGaAs Photodiode (Thorlabs S154C), calibrated with a NIST-traceable curve down to 100 pW input power. Light in the second output passed through a variable attenuator (JDS Uniphase HA9, manually calibrated), a polarization controller and an SMF28 fiber feedthrough to couple to the PIC in the cryostat.

We confirmed the calibration of the variable HA9 attenuator beyond the sensitivity of the photodiode (powermeter) as follows: we recorded the SNSPD count rate under a given HA9 attenuation value, typically 50 - 80 dB, then replaced the HA9 with fixed fiber optic attenuators of the same attenuation value. The fixed fiber optic attenuators

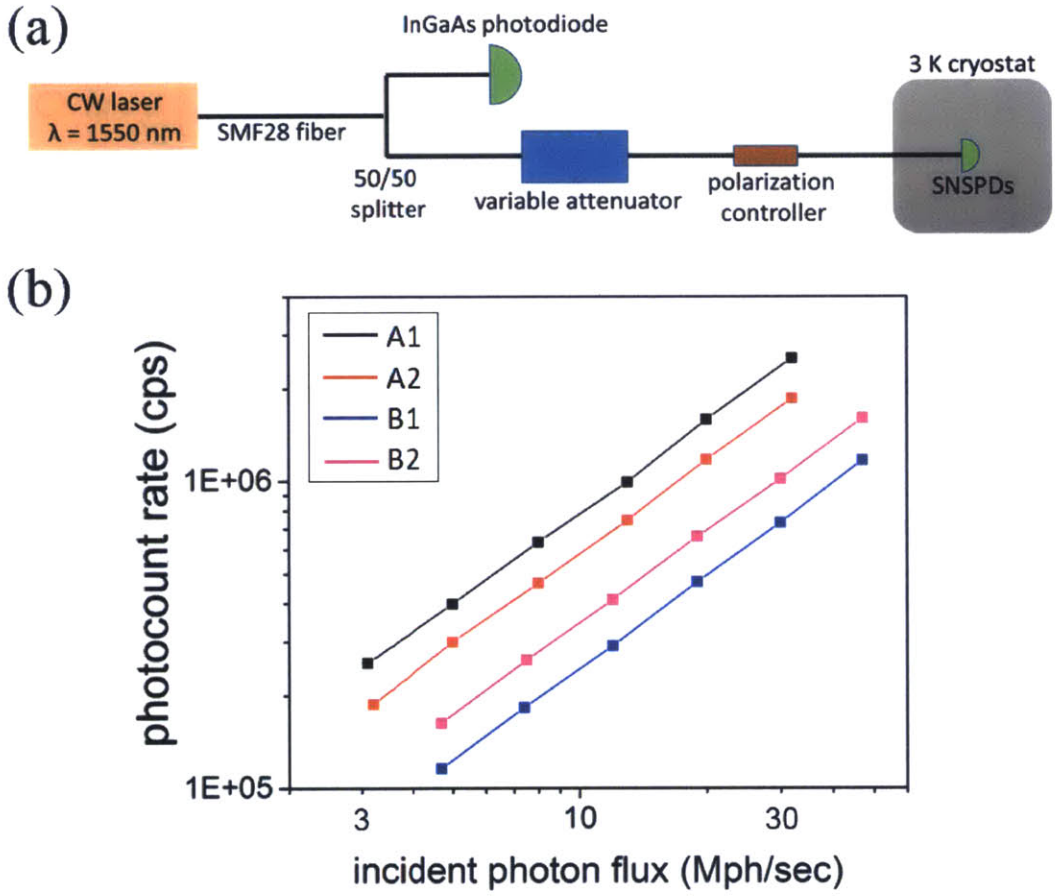

Figure 7-9: (a) Schematic depiction of experimental setup used to measure the system detection efficiency of waveguide-integrated detectors. (b) Photon count rate vs. incident photon flux for the waveguide-integrated detectors A1,A2, B1 and B2.

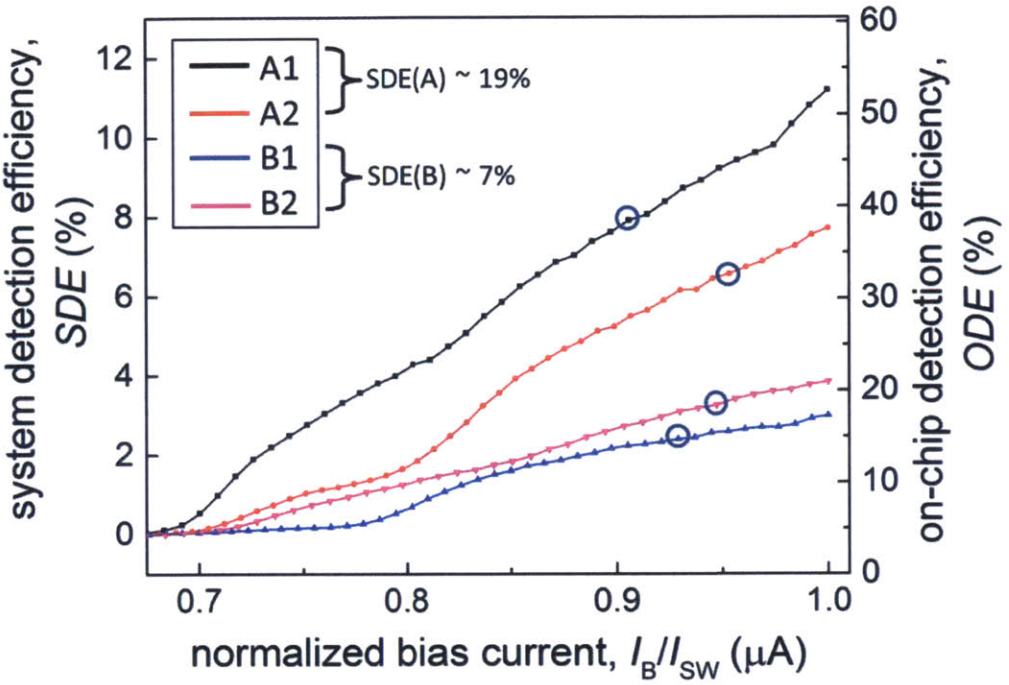

Figure 7-10: System detection efficiency (SDE) vs. normalized bias current of the waveguide-integrated detectors shown in Figure 7-8. The bias current (I_B) on the horizontal axis was normalized by the maximum bias current (switching current I_{SW}) of the detector. The relative error of the SDE value is ±10% and the relative error of the ODE values is ±11.4%.

used here – Thorlabs FA attenuators connected in series, with an attenuation value of 10 to 25 dB per unit – were calibrated at high laser power using the InGaAs photodiode. The detector count rate measured with the HA9, set to a given nominal attenuation value, was within $\delta = 10\%$ (relative error) of the count rate measured at the same attenuation value set with the FA attenuators. Since the SDE includes all losses in the system, except for the variable attenuator, the overall relative error of the SDE value can also be estimated as $\delta_{SDE} \approx \delta$. The measured SDE for detectors A1, A2, B1 and B2 is shown in Fig. 7-10. This yields an SDE of 19±2% for input A (11% for A1 and 8% for A2) and 7±1% for input B (3% for B1 and 4% B2). These SDE values represent a significant improvement by two orders of magnitude compared with previous reports of waveguide-integrated multi-SNSPD systems [57].

The ODE is defined as the probability that a photon already coupled into the waveguide is detected [16, 57]. We extracted the on-chip detection efficiency (ODE) as

$$\text{ODE} = \frac{1}{\eta_{\text{fiber-WG}} \cdot \eta_{\text{DC}}} (\text{SDE per detector}), \tag{7.1}$$

where $\eta_{\text{fiber-WG}}$ is the fiber-to-waveguide coupling efficiency and η_{DC} is the on-chip transmission of 47±4.6% due to the nominal 3 dB splitting ratio of the on-chip directional coupler (beam splitter). $\eta_{\text{fiber-WG}}$ and η_{DC} were estimated from room-temperature measurements using high-precision scanning piezos (Fig. 7-11(a)), and the results are shown in Figs. 7-11(b,c). From these measurements, the on-chip coupling loss and propagation loss were calculated from a linear regression, resulting in an estimated loss due to coupling on and off the chip of 7.37±0.39 dB. Assuming defects in the structures for on- and off-chip coupling are uncorrelated, we estimate the distribution of the coupling efficiency for each coupler is 43±3%. All other on-chip losses are included in the ODE estimate (e.g. we do not normalize the ODE to account for ~2.15 dB/cm propagation loss in the waveguide nor do we normalize to account for the loss in the polarization controller shown in Figure 7-9(a) and therefore they do not contribute to the ODE error. The overall ODE error is estimated as $\delta_{\text{ODE}} = \sqrt{10^2 + 3^2 + 4.6^2} = 11.4\%$. We confirmed that the detector operated in single-photon regime during the system efficiency measurements, as demonstrated by the linearity of the photodetection count rate vs. incident photon flux shown in Fig. 7-9(b). The assembled waveguide-integrated detectors showed maximum ODE values between 14±2% and 52±6%, as shown in Fig. 7-12. We note that, since the fiber coupling at base temperature in the cryostat was performed with slip-stick stepper stages with worse resolution than room-temperature piezo scanners used to estimate the overall coupling efficiency $\eta_c = \eta_{\text{fiber-WG}} \cdot \eta_{\text{DC}} = 0.22$ and its error, we expect the low-temperature η_c to be significantly smaller than room-temperature value, and the ODE values provided here should be considered as pessimistic values. However, the SDE, which is largely not affected by loss estimations, is of higher significance for the non-classical measurements that are presented in the following sections.

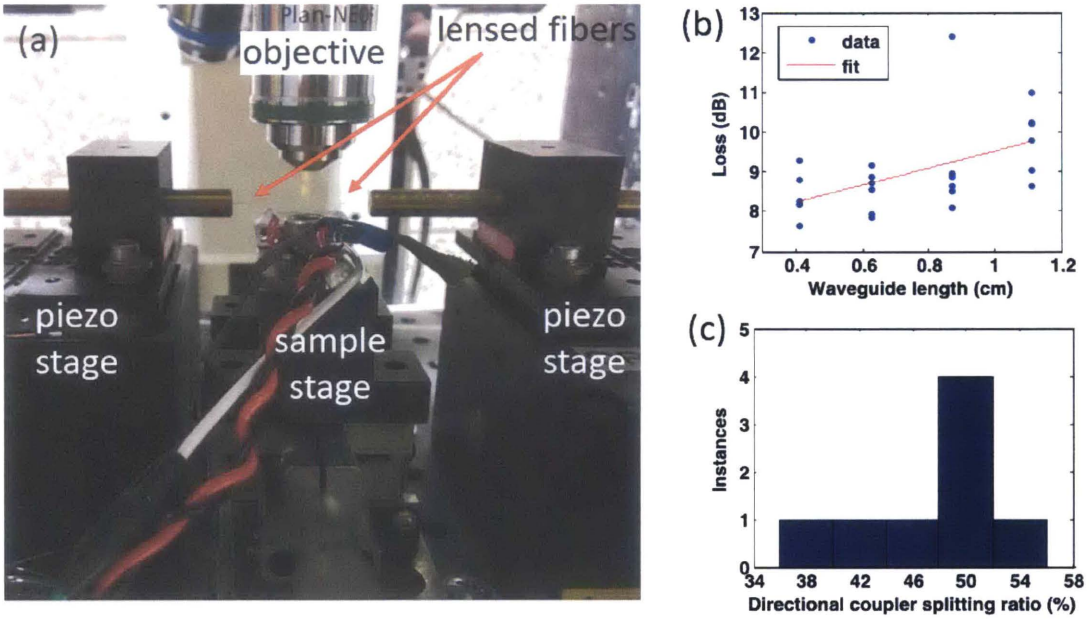

Figure 7-11: (a) Photo of setup used to measure the room-temperature transmission of optical waveguides. (b) Propagation loss of single-mode silicon waveguides of different lengths. Extrapolation to the zero length value gives two times the on-chip coupling loss. The slope of the line gives the propagation loss in the waveguide. (c) Histogram of measured directional coupler splitting ratios.

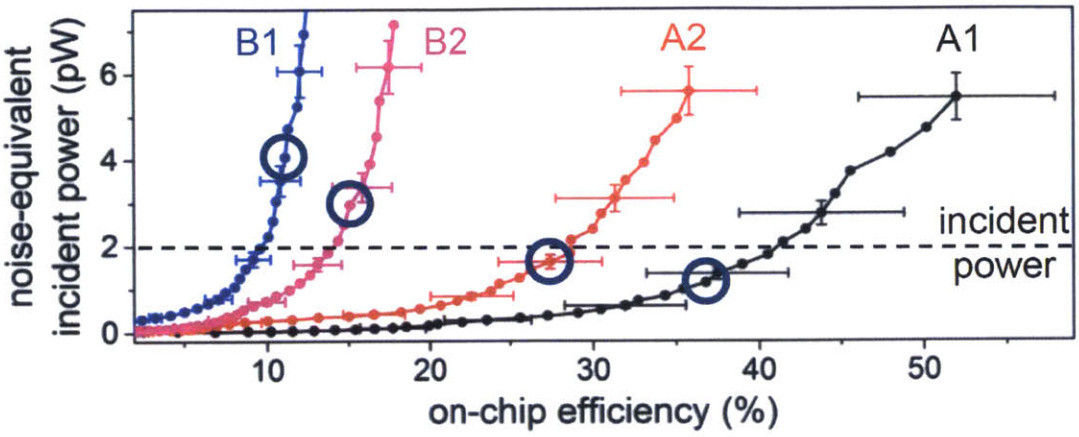

Figure 7-12: Measured noise-equivalent incident power as a function of on-chip detection efficiency for detectors A1, A2, B1 and B2.

We define the noise-equivalent incident power as

$$\text{NEIP} = \frac{\text{SDCR} \cdot \hbar\omega}{\text{SDE}}, \tag{7.2}$$

where SDCR is the system dark count rate and $\hbar\omega = 0.81$ eV. Fig. 7-12 shows the NEIP vs. ODE for the waveguide detectors on couplers A and B. The ratio of the power incident onto the detectors (IP) and the NEIP characterizes the signal-to-noise ratio for single-shot measurements. In this work, the NEIP was limited by radiation leakage (Ref. [1]) through a cryostat window used to image and align the lensed fibers to the polymer couplers. Hence, for subsequent measurements, we operated the detectors at lower ODEs of 12±1% to 37±4% (circled points in Fig. 7-12), which reduced the dark count rate ($\sim$ 800 k counts per second, on the same order as the PCR) and resulted in a ratio of IP/NEIP $\sim$ 0.5 - 1.7. The low NEIP of these detectors is crucial for characterizing picowatt-level optical signals, which can be the case for non-classical light sources.

7.2.2 Dark count rate

The system dark count rate crucially depends on the shielding conditions. We used a closed-cycle cryostat with optical access to operate the chip shown in Figure 7-8. The schematic cross-section of the cryostat is shown in Figure 7-2. The PIC chip and platform holding the micro-manipulated lensed fibers were kept at 3 K base temperature. In order to couple light from the lensed fibers into the waveguides, the edges of the chip, containing the polymer couplers (Figs. 7-4(a,b)), were imaged through the windows using a 50× long-working-distance objective. The direct imaging greatly simplified pre-alignment, while finer fiber-to-coupler alignment was performed using feedback from the on-chip detectors. However, the optical access ports in this prototyping setup resulted in radiation leakage and therefore increased the system dark count rate of the detectors significantly, as shown in Figure 7-13. When we replaced the 30 K window in the cryostat with a solid copper plate, we observed a significantly lower dark count rate of $\sim$ 5 kcps instead of $\sim$ 800 kcps at the operation point.

7.2.3 Timing jitter

To characterize the timing jitter, we used a mode-locked, sub-ps-pulse-width laser emitting at 1550 nm wavelength and 38 MHz repetition rate. The laser output was split into two SMF28 fibers, which we coupled to the detector under test and to a low-timing-jitter InGaAs photodiode (Thorlabs S1R5). A sketch of the experimental setup is shown in Fig. 7-14.

The light coupled to the detector was attenuated to < 5 pW and operation of the detector in single-photon regime was checked by confirming the linearity of the photocount rate as a function of incident photon flux (see Figure 7-9(b)). The electrical output from the detector and from the photodiode were sent to a 6-GHz-bandwidth, 40-GSamples/s oscilloscope. We measured time delay t_D between the detector pulse (start signal) and the pulse from the fast photodiode (stop signal). We acquired the instrument response function (IRF), a histogram of > 2000 samples of t_D, and

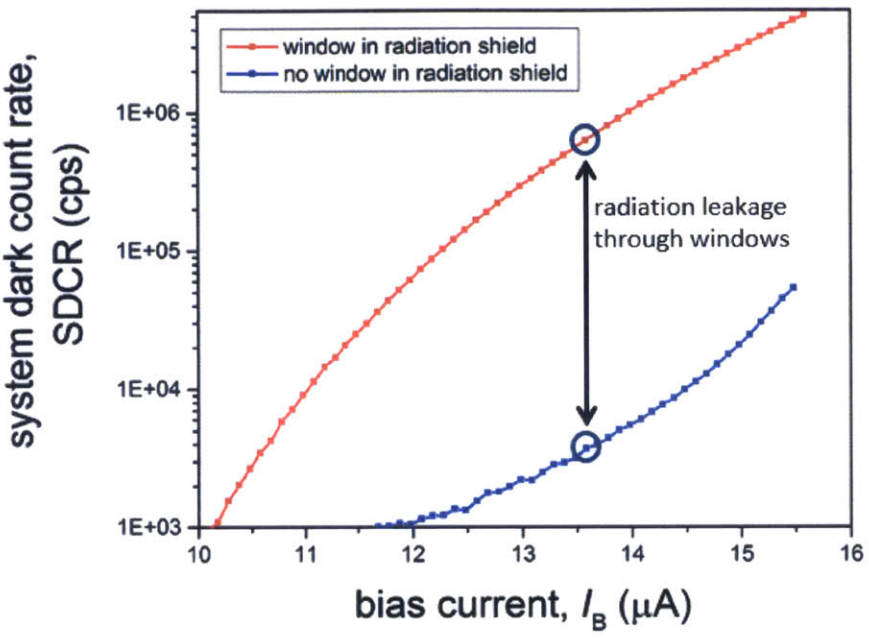

Figure 7-13: System dark count rate (SDCR) curves representative of waveguide-integrated detectors operated in the cryostat shown in Fig. 7-1. The red curve shows the SDCR during the regular operation of the cryostat with windows, and the blue curve shows the SDCR with the windows replaced with copper plates.

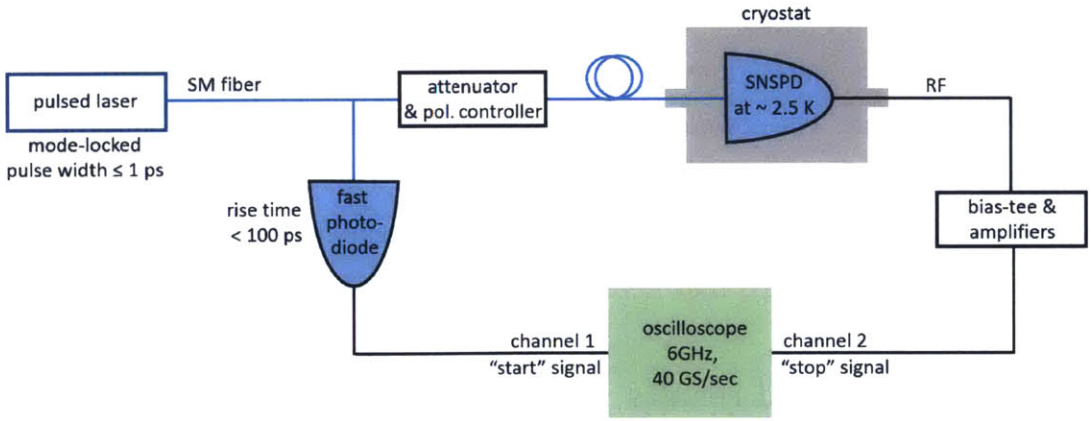

Figure 7-14: Sketch of experimental setup used to measure the timing jitter.

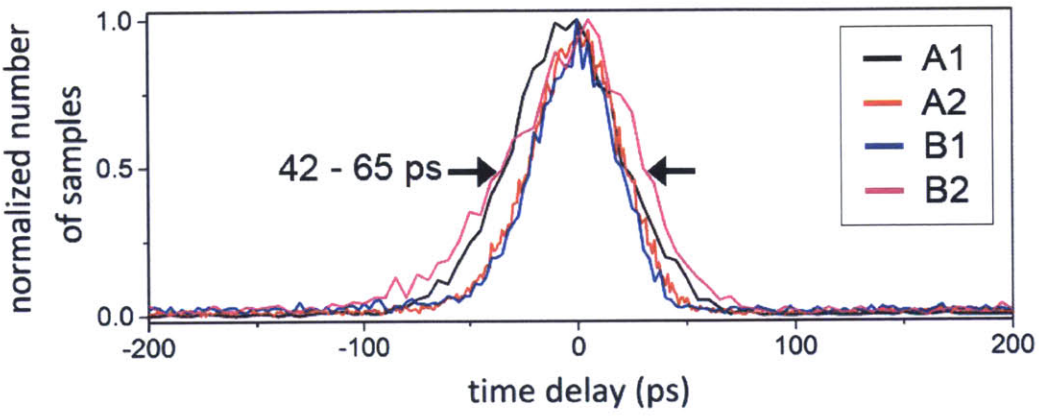

Figure 7-15: Instrument response function of waveguide-detectors A1, A2, B1 and B2, showing a FWHM jitter of 42-65 ps.

measured the timing jitter of the detector, which was defined as the FWHM of the IRF. The results are shown in Figure 7-15. All detectors, when biased close to the switching current, showed TJ values of 42-65 ps.

7.3 Correlation measurements of entangled photons

We used a PPKTP waveguide source to generate entangled photons pairs at 1561 nm wavelength via type-II spontaneous parametric down conversion. A pump beam (781 nm wavelength, 50 mW power) was focused on the PPKTP waveguide with cross section 2 μm × 4 μm. The down-converted signal and idler photons were coupled into a single fiber and split with a fiber polarizing beam splitter. The output fibers were coupled to polarization controllers, which were each connected to separate optical fiber feedthroughs (A and B) that were leading into the cryostat.

7.3.1 Experimental results

We used these high-SDE SNSPDs to characterize time-energy entangled photon pairs entirely on the PIC. Photon pairs were generated by type-II spontaneous parametric down conversion (SPDC) from a 1-cm periodically poled potassium titanyl phosphate

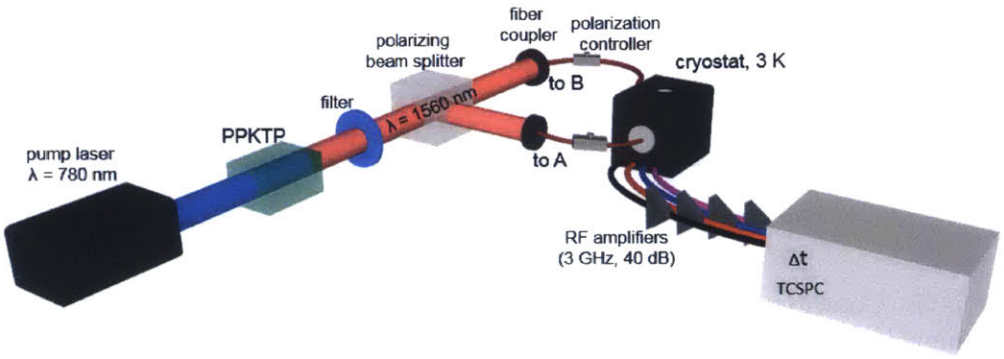

Figure 7-16: Experimental setup for on-chip $g^{(2)}_{AB}(\tau)$-measurements of an entangled-photon source coupled into the PIC (cooled to 3 K).

(PPKTP) waveguide, as shown in Fig. 7-16. Signal and idler photons of ~ 1 ps duration and orthogonal polarization were separated using a polarizing beam splitter and sent into inputs A and B of the PIC. The SPDC pump power was adjusted to generate pairs at $\sim 1.5 \cdot 10^8$ Hz, corresponding to a multi-pair probability of $\sim 4 \cdot 10^{-4}$ over the correlation timing uncertainty of 200 ps. We obtained the second-order correlation function from $g^{(2)}_{AB}(\tau_i) = N_{AB}(\tau_i)/(r_A r_B \Delta\tau T)$, where $N_{AB}(\tau_i)$ is the measured number of coincidences between inputs A and B at time difference τ_i, r_A (r_B) is the count rate from input A (B), $\Delta\tau$ is the coincidence bin duration, and T is the integration time. The left panel in Fig. 7-17 shows the resulting $g^{(2)}_{AB}(\tau_i)$ function. Photon bunching is evident between inputs A and B, but not within individual channels (i.e., between A1 and A2 or B1 and B2), as expected for a photon pair source. The observed peak heights of $g^{(2)}_{AB}(0) \sim 4$ and $g^{(2)}_{AB}(0) \sim 6$ are lower than the theoretical value of infinity for ideal detectors due to the finite IP/NEIP ratio of our detectors and the non-zero multi-pair probability. By contrast, when pulses from a mode-locked laser were injected into inputs A and B with an average photon number per pulse greater than one, bunching was observed between all detector pairs (Fig. 7-17, right panel), as expected for a pulsed classical light source.

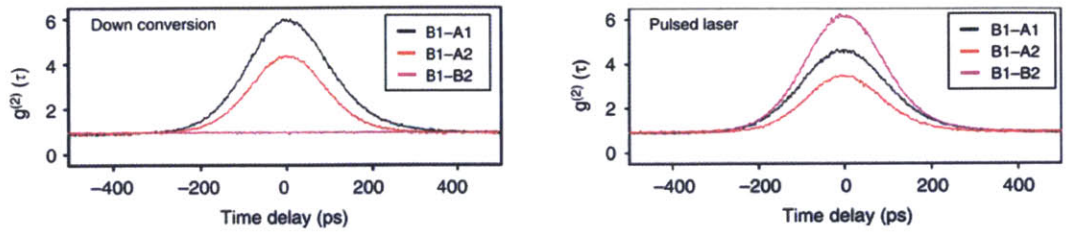

Figure 7-17: Coincidence counts versus time delay between B1 and A1, A2, B2 for the entangled-photon-pair source (left) and for a mode-locked sub-picosecond-pulsed laser (right). The average laser power was adjusted to match that of the photon-pair source. The data was acquired with a time-correlating counter (TCSPC, HydraHarp 400)

7.3.2 Discussion

$g^{(2)}_{AB}(\tau)$ can be calculated from experimental data using the formula given in the main text. To incorporate detector dark counts, we define rates r^Y_X, where $X \in \{A, B\}$ (for channels A and B, respectively) and $Y \in \{P, D\}$ (corresponding to a 'photon' and 'dark count', respectively). r^D_A, for example, is the rate at which channel A registers dark counts, and $r_A \equiv r^D_A + r^P_A$ is the count rate on channel A. Now $g^{(2)}_{AB}(0)$ is

$$g^{(2)}_{AB}(0) = \frac{r^P_A\left(\eta_H + r^D_B\Delta\tau\right) + r^D_A\Delta\tau \cdot r_B}{r_A r_B \Delta\tau}, \tag{7.3}$$

where η_H is the probability that channel B registers a photon given that channel A also registers a photon (i.e. the heralding efficiency) and $\Delta\tau$ is the bin duration. For $r^Y_A = r^Y_B \equiv r^Y$ and the ratio $K \equiv r^P/r^D$,

$$g^{(2)}_{AB}(0) = \left(\frac{K}{K+1}\right)^2 \frac{\eta_H}{r^P\Delta\tau} + \frac{2K+1}{(K+1)^2}. \tag{7.4}$$

In our experiment, $g^{(2)}_{AB}(0) \approx 5$, which gives an estimate of the heralding efficiency, $\eta_H = 4 \cdot 10^{-3}$.

7.4 Improving system performance

The system efficiency of the devices presented here could be improved with several changes to the PIC. Propagation loss in the waveguide could be reduced from 2-3 dB/cm to 0.3 dB/cm using ridge waveguides [58]. On-chip coupling loss can also be reduced from 3.7 dB using either high-performance grating couplers, which can achieve 0.6 dB loss [59], or edge couplers, which can achieve 1 dB [60]. In the cryostat, fiber-to-chip coupling losses could be improved using piezo scanners or by permanently bonding the chip to the fiber. Scattering at the SiN_x membrane edge is small (<1%), but can be improved by making this transition in a wider region of the Si waveguide where the evanescent field above the waveguide would be reduced. Lastly, the absorption into the SNSPD increases with device length; a tapered waveguide with stronger evanescent overlap can also lead to greater absorption. An optical cavity could be used to increase the detector-waveguide interaction length, but at the expense of loss of bandwidth. As shown in Figure 7-18, an increase in detector coupling length from 17 to 28 μm increases system efficiency to 24±2%, an improvement by ~26±3% compared to the previous detector design with shorter coupling length.

The system dark count rate could be reduced significantly by eliminating the cryostat windows, and without optical access this could be accomplished through fiber bonding to the photonic chip. The timing jitter of the on-chip detectors can be improved to 33 ps by reducing the length of the RF line by 1 cm and the length of our wire-bonds onto the Au pads by 3 mm, which reduces the overall RF loss. The time delay histogram is shown in Figure 7-18(c) and represents a significant improvement compared to ≥42 ps jitter previously measured in the same cryostat. On-chip amplification electronics, e.g. Ref. [61], could be used the further reduce jitter to 24 ps. To speed up the manual assembly process currently employed, a high-throughput assembly scheme could be adopted [62].

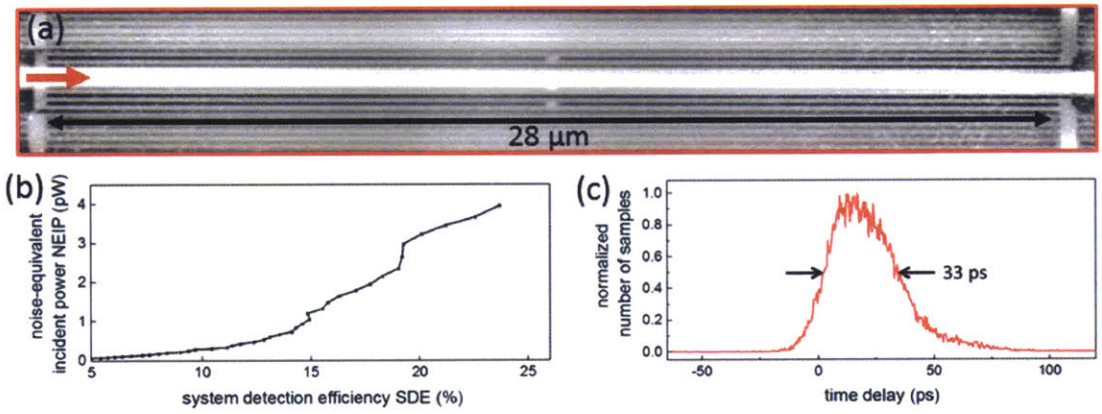

Figure 7-18: (a) Top-down SEM of a membrane-detector integrated with a Si waveguide. The length of the detector is ~28 µm. (b) System detection efficiency (SDE) vs. noise-equivalent incident power for a directional coupler integrated with two large-coupling-length detectors as shown in (a). The relative error of the SDE values is ±10 percent. (c) Instrument response function of a waveguide-integrated detector measured with reduced length of the electrical path.

Chapter 8

Scaling to on-chip detector arrays and new material systems

Many aspects of the work presented in this thesis are directed at the development of high-performance single-photon detectors for compact optical quantum processors. Harnessing the improved speed and scalability presented in previous chapters, we will show a path to larger-scale integration of SNSPDs with different material systems and on-chip single-photon sources.

8.1 Large-scale integration of on-chip detectors

The pick and place approach presented in the previous chapters enables the assembly of large arrays of on-chip high-speed single-photon detectors with unity yield. We define yield as the ratio of detectors that operate in the high-efficiency single-photon regime (also referred to as avalanche regime, see Ref. [1]). In this regime the detectors show sub-100-ps timing jitter [1]. Figure 8-1(a) shows ten SNSPDs (D1-10) on adjacent waveguides with timing jitter values of 39 ps - 57 ps for 1550-nm-wavelength light. For rapid characterization, these devices were measured by top illumination in a cryogenic probe station. The photodetection delay histograms for all detectors are shown in Fig. 8-1(b).

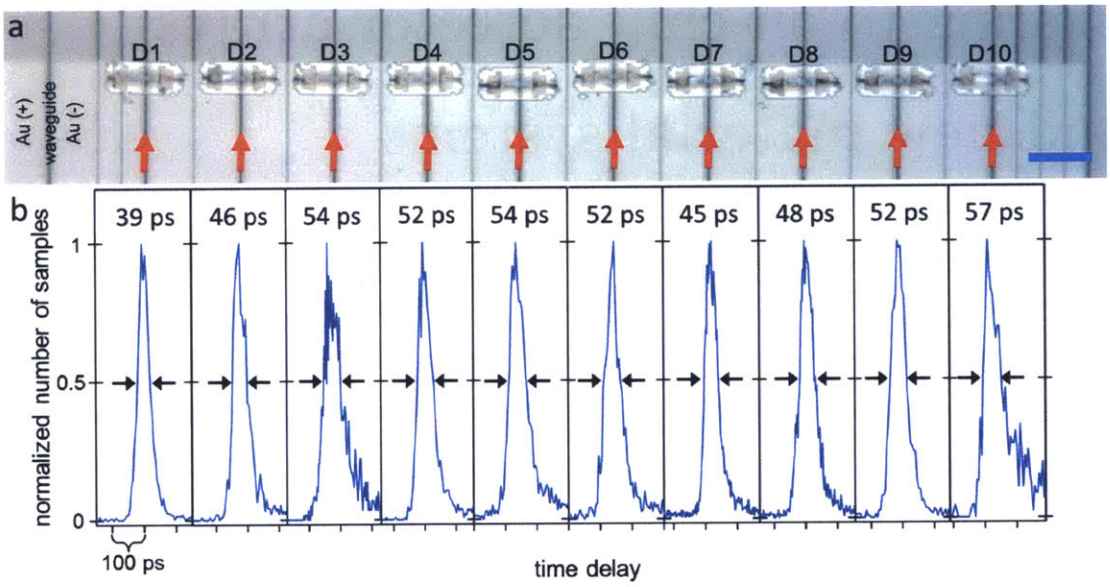

Figure 8-1: (a) Optical micrograph of ten waveguide-integrated detectors D1-D10. The detectors were assembled on the same photonic chip and integrated with silicon waveguides, marked by red arrows. The equivalent length of the blue scale bar is 100 μm. (b) Top-illuminated photodetection delay histogram of the detectors shown in (a) measured in a cryogenic probe station at $\sim$ 2.8 K base temperature. The FWHM timing jitter, extracted from the histograms, is listed above each histogram.

8.2 Integration with different material systems

The pick-and-place method effectively separates the fabrication processes of the detector and the photonic chip. As a result, this approach can be used to integrate SNSPDs with photonic chips based on diverse material systems. A promising material for photonic circuits is aluminum nitride (AlN). Compared to silicon, AlN offers several properties of interest [63, 64, 65]:

(1) The large bandgap (6.2 eV), resulting in a large transparency window ranging from the UV to the infrared.

(2) The significant second-order susceptibility has the potential to enable fast electro-optic modulation and wavelength conversion.

(3) The high piezoelectric transduction efficiency could be interesting for on-chip opto-mechanical applications.

Traditional attempts at integrating SNSPDs with new photonic materials relied on developing a processes for the growth of the superconducting film and the fabrication of SNSPDs. These efforts have often proven lengthy and challenging [66, 57], whereas the pick-and-place method can be rapid and scalable. Fig. 8-2(a) shows an SNSPD integrated with an AlN-on-sapphire waveguide. The detector also showed good jitter performance (see Figure 8-2(c)). The membrane transfer process could be used to integrate other electro-optic devices, such as III-V lasers or single-photon sources, onto PICs, therefore enabling the ground-up assembly of a quantum (photonic) circuit using pre-selected high-performance components. As shown in Fig. 8-2, the flexible membrane can conform to small non-uniformities on the surface of the target chip. Furthermore, due to the relatively small size of the membrane, the process is more tolerant to defects on the target chip processes involving large-area flip-chip bonding (e.g., see Ref. [67]), which require both surfaces to be free of defects.

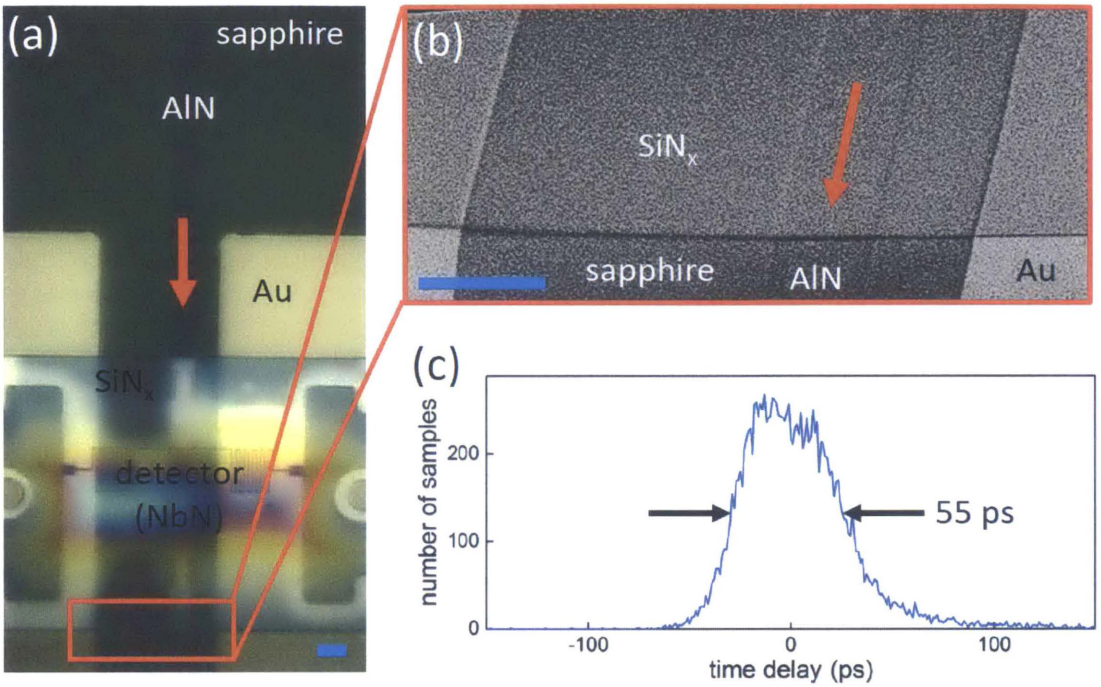

Figure 8-2: (a) Single-photon detector integrated with a multi-mode AlN-on-sapphire waveguide. The equivalent length of the blue scale bar is 5 μm. (b) Angled SEM showing the membrane conforming to waveguide and Au pad surfaces. The equivalent length of the scale bar (blue) is 5 μm. (c) Top-illuminated photodetection delay histogram of the detector shown in (a,b).

8.3 Towards a fully-integrated solution

Nitrogen vacancy centers (NV-centers) are defects in diamond comprising one carbon atom in the grid has been replaced by a nitrogen atom, and a second adjacent carbon location in the lattice that is vacant. These centers have garnered interest as solid-state optical quantum repeaters, single-photon sources and spin-based quantum memories. Recent progress [68] based on pre-selected pick-and-place waveguide-NVs has enabled the scalable integration of NVs with photonic circuits. Figure 8-3 shows a cross-sectional sketch that illustrates one approach to building a fully-integrated photonic circuit. A pre-selected single-mode diamond waveguide carrying an NV center is placed on top of an AlN waveguide. An air gap below the diamond waveguide and tapered waveguide ends ensure the adiabatic transition of the mode from the diamond waveguide into the AlN waveguide [68]. The emitted photons from the NV

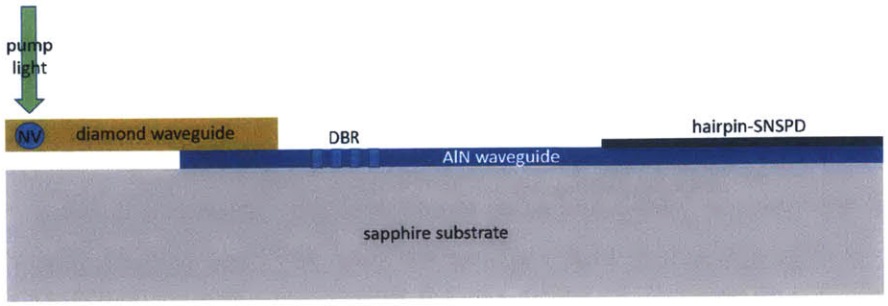

Figure 8-3: Cross-sectional sketch of an NV center inside a diamond waveguide coupled to an AlN waveguide and an on-chip single-photon detector.

center are then detected by a hairpin-shaped SNSPD on top of the AlN waveguide. Residual pump light is filtered using a DBR grating built into the AlN waveguide. Since the fluorescence lifetime of NV centers is several nanoseconds, the pump light can also be filtered using a pulsed pump laser and by gating the SNSPD.

Since the detectors are fabricated directly on top of the waveguide, limited nano-fabrication yield is a concern. While we have solved the yield problem for detectors using the pick-and-place approach outlined in this thesis, the recent nano-fabrication improvements that resulted in a direct-fabrication-yield of 70-90% (see chapters 2 and 3) are encouraging. This approach is further promising for this specific application since the photoemission wavelength range is λ = 630-770 nm and the effective yield of SNSPDs is higher for decreasing wavelengths (see discussion of cutoff current in chapter 1). Figure 8-4 shows a top-down multi-layer pattern for the fully-integrated system. Two hairpin-shaped NbN detectors are located on top of an AlN waveguide. The NV waveguide is placed on top of an airgap (diamond-shaped black gap). A close-by RF transmission line generates a magnetic field that is used for NV spin control.

NbN films (R_S = 405 Ω/sq., T_C = 11.8 K) were grown on a ~300-nm-thick AlN-on-sapphire substrate. NbN series-2-SNAPs were fabricated using the Elionix process outlined in the appendix. Figure 8-5 shows an SEM of waveguide-2-SNAPs based on ~70-nm-wide nanowires on top of AlN.

The critical current distribution of the first fabricated sample, shown in Fig. 8-6,

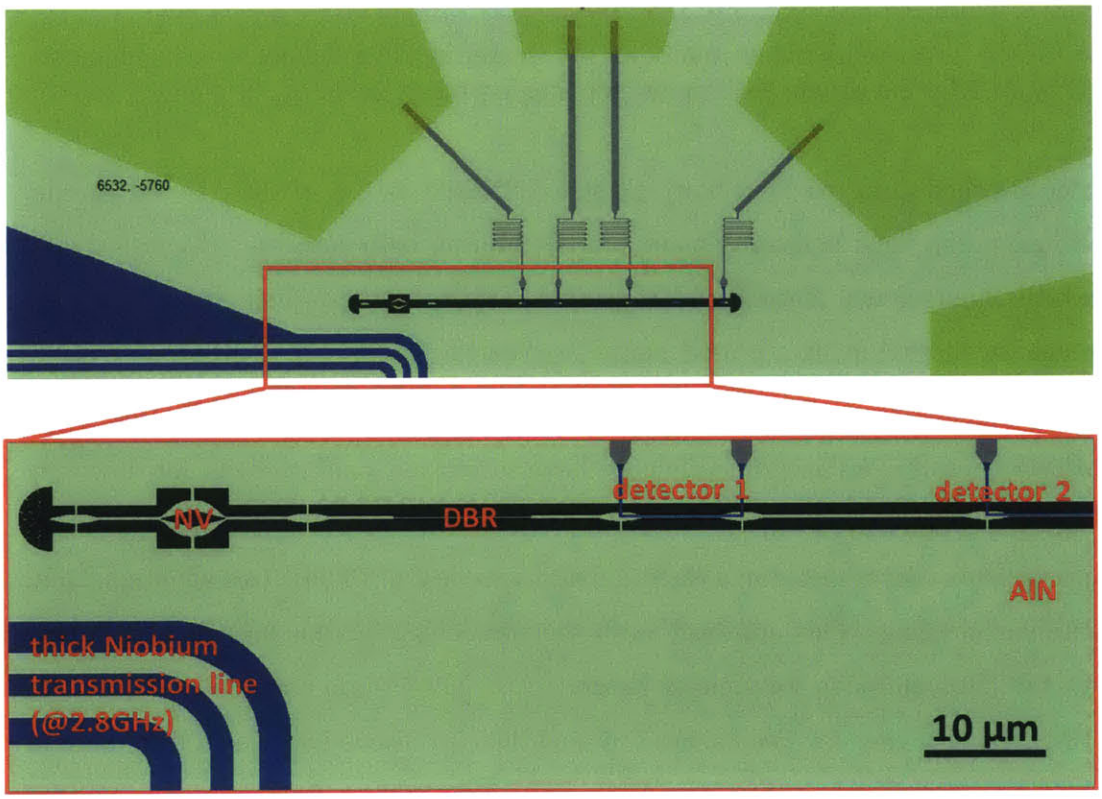

Figure 8-4: Top-down view of the PIC pattern used to integrate NV centers with on-chip SNSPDs. Detector 1 is shorter than detector 2 and results in a partial absorption of light travelling in the wavegide. This configuration is effectively a beamsplitter coupled to two SNSPDs and can be used for on-chip correlation measurements of photoemission from the NV center.

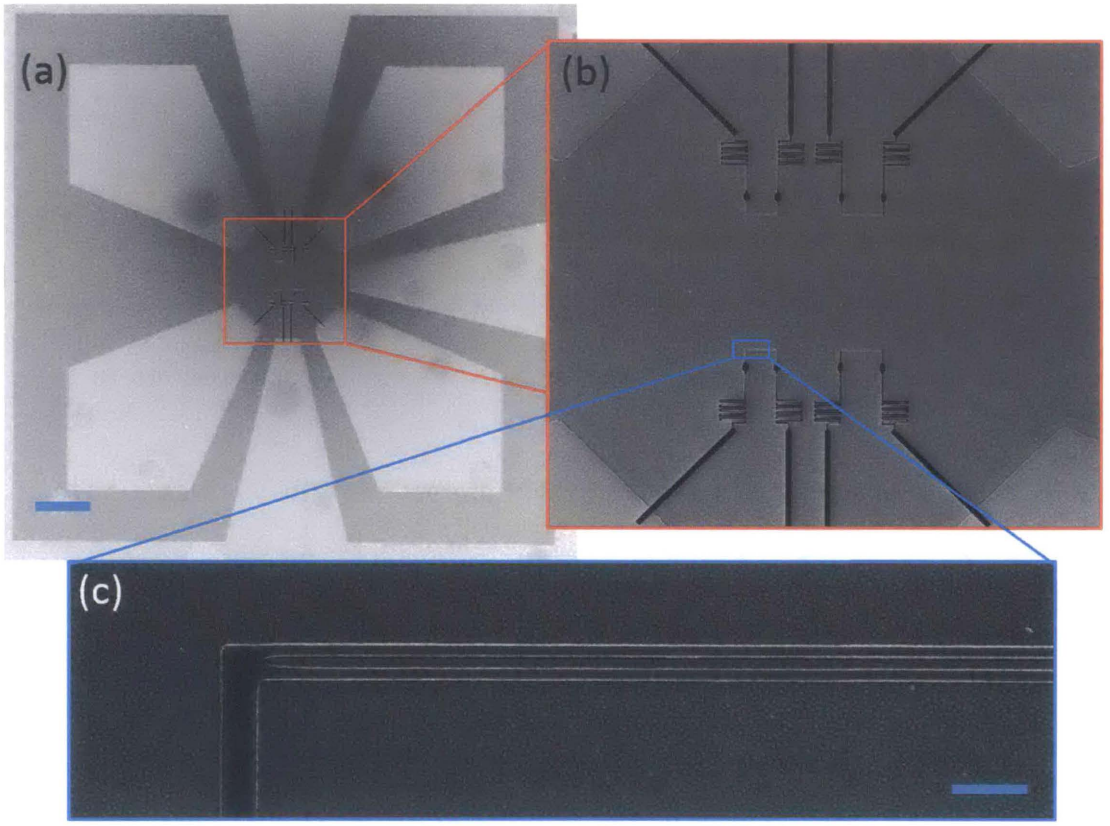

Figure 8-5: SEMs of SNSPD portion of the pattern shown in 8-4. The equivalent length of the scale bars in (a), (b) and (c) is 80 μm, 21 μm and 400 nm respectively.

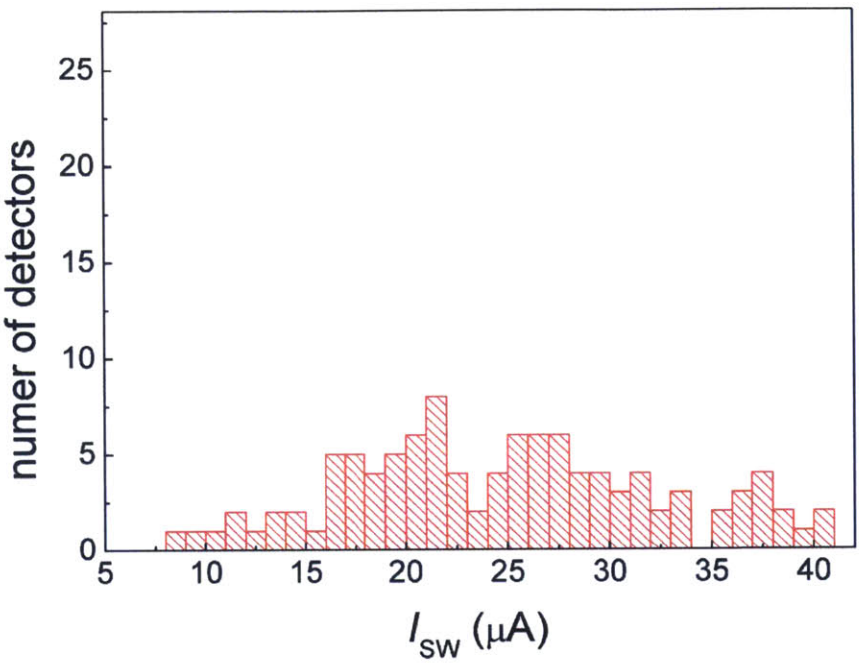

Figure 8-6: Switching current histogram for 2-SNAPs on the sample shown in Fig. 8-5

is fairly broad and implies that the fabrication process has to be optimized further (dose, development time, etch times) for for the AlN sample.

The second step is the fabrication of AlN waveguides. A ~400-nm-thick layer of ZEP, a positive electron-beam resist, was used to pattern the mask for the waveguides. The process is outlined in the appendix. Figure 8-7 shows a developed ZEP test mask on top of silicon. The mask withstands ~6 minutes of CF_4 RIE at 150 W without visible degradation, which should be sufficient [69] for pattern transfer into the AlN.

A remaining challenge is the alignment of the waveguide pattern to the SNSPD pattern. The patterns have an alignment tolerance of ±50nm.

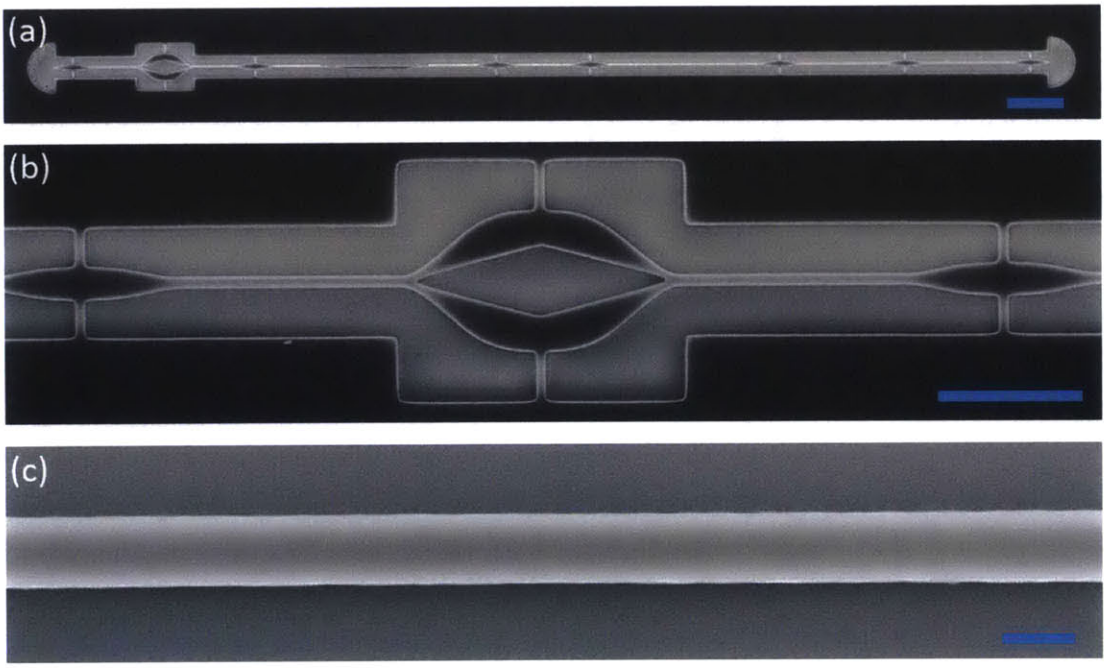

Figure 8-7: SEMs of ZEP mask on top of silicon for the photonic chip portion of the pattern shown in Fig. 8-4. The equivalent length of the scale bars in (a), (b) and (c) is 6 μm, 3 μm and ~370 nm respectively.

Chapter 9

Summary and outlook

We reviewed detector architectures based on superconducting nanowires and the different performance metrics. For chip-scale applications, which would benefit from a large number of high-speed on-chip single-photon detectors, we first focused on improving the fabrication process.

Improvements to the detector fabrication process included an optimized growth process, the removal of steps that could degrade the NbN film, a bi-layer liftoff process and an optimized RIE step. These modifications increased the detector yield to ~70% and simplified the overall fabrication process. The yield could be likely further improved by reducing defects in the detectors.

Besides introducing an improved fabrication process, a modified detector design, the series-SNAP, was presented. This design allowed sub-10-ns reset times and sub-30-ps timing jitter for single-photon detectors with active areas on the order of several tens of μm^2. The device speed remained ultimately limited by the latching limit due to Joule heating.

In order to overcome the latching limit, we introduced a new detector design, the nano-mesh single-photon detector (NMSPD), that was designed to limit the growth of the resistive region through diffusive cooling within the NbN film. Prototype NMSPDs showed sub-1-ns dead times and ~60-ps timing jitter. We believe that the NMSPDs are a promising path to high-speed single-photon detectors. However, further device efficiency measurements and design optimization are likely required to

reduce the timing jitter, improve the signal-to-noise ratio and ensure operation in a saturated efficiency regime.

In addition to improvements to the detector design and fabrication processes, a scalable method had to be developed for the integration of SNSPDs with photonic integrated circuits (PICs). Two main challenges had to be overcome: the incompatibility of the SNSPD fabrication process with a variety of PIC processes and the insufficient direct fabrication yield of high-performance SNSPDs. We addressed these challenges by developing a micrometer-scale flip-chip process that separated the fabrication processes of the SNSPDs and the photonic chip. Detectors were fabricated on sub-400-nm-thick membranes, characterized and only high-performance SNSPDs were selected for the subsequent integration with the PIC via a pick-and-place approach.

A fabrication process that allowed the suspension of SNSPDs on top of sub-micrometer-thick SiN_x membranes was developed. Central fabrication challenges were the suspension of the SiN_x layer without introducing defects into the detectors, and the development of a final cleaning process that enabled the reproducible removal of the layer that protected the SNSPDs during the membrane suspension etch.

Based on the developed pick-and-place approach several prototype chips were assembled. Using a photonic chip with four simultaneously-operated SNSPDs with average system detection efficiencies beyond 10% per optical input channel, we demonstrated the first on-chip correlation measurements of photon pairs. We managed to assemble waveguide-integrated SNSPDs on top of a variety of material systems with unity effective yield. While this method is scalable in research settings, increasing the number of on-chip detectors to hundreds or thousands of SNSPDs would likely require a new approach and the transition to other materials such as WSi. In the meantime, our pick-and-place technology will help bridge the gap, and hopefully enable a wide range of applications including on-chip optical quantum computing and quantum simulation.

Appendix A

Membrane-detector fabrication checklist

A.1 Bi-layer Liftoff

A.1.1 Exposure

- In-spin clean with Acetone+IPA
- Spin PMGI SF9 at 2krpm, 4 k acceleration (~700-nm-thick)
- Bake at 90°C for 1 min
- Spin S1813 at 5krpm and 1k acceleration
- Bake at 90°C for 1 min
- Expose for 12sec at 2500μW/cm^2) in scale B in TAMRACK, clean mask with Acetone+Methanol+IPA between exposures
- Develop in CD-26 for 24 sec
- 1 min DI dip, nitrogen blowdry

A.1.2 Liftoff

- Evaporation: 10nm Ti, 15nm Au
- 2min in sonicator at P=3
- Dip in CD-26 for 45 sec (while in sonicator at P=3 for first 30sec)
- DI dip for 1 min, nitrogen blowdry

A.2 SNSPD fabrication

A.2.1 E-beam lithography

- Warm up HSQ (4%) bottle for 60 minutes (10 min handwarm)
- Spin at 4.0 krpm, 10k acceleration
- Raith: dose 1.75-2.1 $\times$ 500μC/cm^2
- exposed GDS
- Develop in TMAH (25%) at 27-28°C for 3 min

A.2.2 RIE

- newclean (no glass on hole)
- CF_4 at 50 W for 2min 45sec
- reference voltage: 60-70V, reference power: 47-50W

A.3 Membrane fabrication

A.3.1 Trench fabrication

- Spin S1813 at 5.5 krpm, 1krpm acceleration

- Bake 1min at 90°C
- Remove edge-beads with razor
- Exposure in new mask aligner (EML, MA4): 7 sec exposure (make sure multiple exposure light is turned off; mask Faraz-16)
- develop 17 sec in CD-26
- dip 1 min in DI
- blowdry
- RIE: CF4@150W for 3×6 min
- S1813 removal
  - Place face-down in Acetone
  - Sonicate for 4 min at P=3
  - Squirt Acetone+Methanol+IPA
  - blowdry

A.3.2 Etch-protective resist

- Spin S1813: at 5.5 krpm, 1krpm acceleration
- Bake 1min at 90°C
- Exposure in new mask aligner (EML, MA4): 5 sec exposure (make sure multiple exposure light is turned off; mask Faraz-17)
- develop 15 sec in CD-26
- dip 45 sec in DI
- blowdry

A.3.3 Membrane under-cut: dummy sample

- XeF_2 etch
- Dummy etch: run purge with empty chamber to measure expansion/XeF_2 pressure
- Etch 2×40sec at P=4000mtorr
- Examine sample under microscope, compare to reference undercut picture
- If necessary, etch additional 1×40sec at P=4000mtorr
- Pop membrane with microprobe to confirm full undercut
- S1813 removal:
  - warm up Microposit S1165 at 80°C (must be <90°C!)
  - face-down in S1165 for 16 min
  - face-down in Acetone for 2 min under manual agitation
  - face-down in S1165 for 1 min
  - face-down in Acetone for 1 min under manual agitation
  - IPA dip (face-up)
  - let dry in air

A.3.4 Membrane under-cut: actual sample (only if dummy undercut successful)

- XeF_2 etch
- Dummy etch: run purge with empty chamber to measure expansion/XeF_2 pressure
- Etch 2×40sec at P=4000mtorr

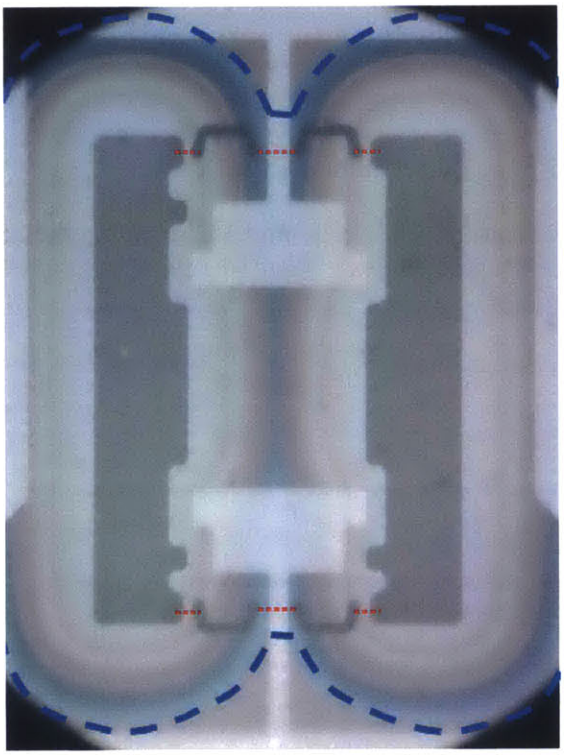

Figure A-1: Undercut reference for 175nm-thick SiN_x. As soon as the sides of the thin 'slots' are free of silicon the etch is done. Do not over-etch.

- Examine sample under microscope, compare to reference undercut picture
- If necessary, etch additional $1\times$40sec at P=4000mtorr
- Pop membrane with microprobe to confirm full undercut
- S1813 removal:
  - warm up Microposit S1165 at 80°C (must be <90°C!)
  - face-down in S1165 for 16 min
  - face-down in Acetone for 2 min under manual agitation
  - face-down in S1165 for 1 min
  - face-down in Acetone for 1 min under manual agitation
  - IPA dip (face-up)
  - let dry in air

Appendix B

AlN-WG-SNSPD fabrication process

Clean

- Rinse in Acetone-IPA
- Blowdry

Spin

- Spin ZEP 520A @ 3krpm, 1 k acceleration
- Bake 3 min @ 120C

Exposure (Elionix)

- 300um field, 240000 dots, scan pitch 1
- beam current = 1nA
- dose time 0.06:0.04:0.3

Development

- Xylene at room temperature for 3 minutes

Etch

- CF_4 RIE at 150W power for ~6 min

Appendix C

Mounting scheme of chip inside cryogenic RF probe station

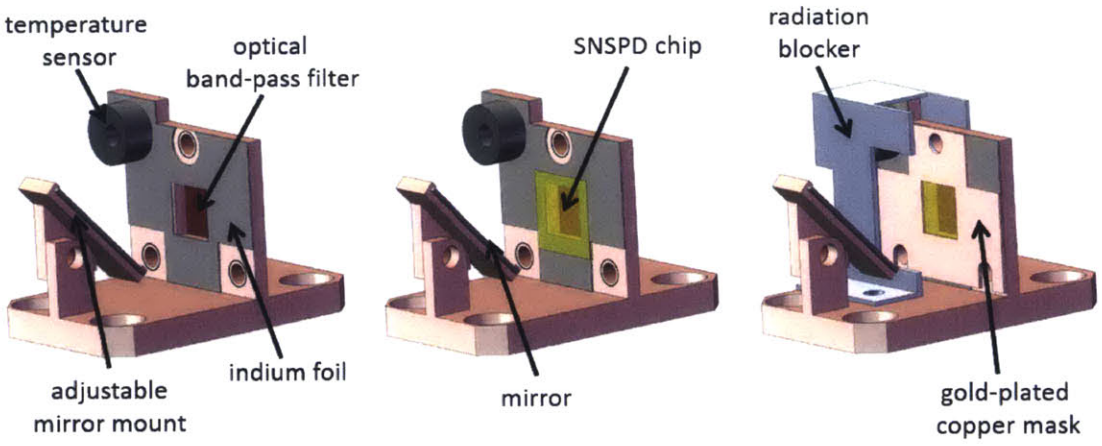

Figure C-1: Sketch of sample mount.

Bibliography

[1] Francesco Marsili, Faraz Najafi, Eric Dauler, Francesco Bellei, Xiaolong Hu, Maria Csete, Richard J. Molnar, and Karl K. Berggren. Single-Photon Detectors Based on Ultranarrow Superconducting Nanowires. *Nano Letters*, 11(5):2048–2053, 2011.

[2] Francesco Marsili, Francesco Bellei, Faraz Najafi, Andrew E. Dane, Eric A. Dauler, Richard J. Molnar, and Karl K. Berggren. Efficient Single Photon Detection from 500 nm to 5 μm Wavelength. *Nano Letters*, 12(9):4799–4804, 2012.

[3] Qingyuan Zhao, Adam N. McCaughan, Andrew E. Dane, Faraz Najafi, Francesco Bellei, Domenico De Fazio, Kristen A. Sunter, Yachin Ivry, and Karl K. Berggren. Eight-fold signal amplification of a superconducting nanowire single-photon detector using a multiple-avalanche architecture. *Optics Express*, 22(20):24574, 2014.

[4] Faraz Najafi. *Timing Performance of Superconducting Nanowire Single-Photon Detectors*. MSc Thesis, MIT, 2015.

[5] F. Marsili, F. Najafi, E. Dauler, R. J. Molnar, and K. K. Berggren. Afterpulsing and instability in superconducting nanowire avalanche photodetectors. *Applied Physics Letters*, 100(11):–, 2012.

[6] F. Najafi, F. Marsili, E. Dauler, R.J. Molnar, and K.K. Berggren. Timing performance of 30-nm-wide superconducting nanowire avalanche photodetectors. *Applied Physics Letters*, 100(15):152602–152602–4, 2012.

[7] Danna Rosenberg, Charles G. Peterson, Jim W. Harrington, Patrick R. Rice, N. Dallman, K. T. Tyagi, K. P. McCabe, Sae Woo Nam, B. Baek, R. H. Hadfield, Richard J. Hughes, and Jane E. Nordholt. Practical long-distance quantum key distribution system using decoy levels . *New J. Phys.*, 11(045009), 2009.

[8] Alan Aspuru-Guzik and Philip Walther. Photonic quantum simulators. *Nat Phys*, 8(4):285–291, 04 2012.

[9] Jeremy L. O'Brien, Akira Furusawa, and Jelena Vuckovic. Photonic quantum technologies. *Nature Photonics*, 3(12):687–695, 2009.

[10] J M Amini, H Uys, J H Wesenberg, S Seidelin, J Britton, J J Bollinger, D Leibfried, C Ospelkaus, A P VanDevender, and D J Wineland. Toward scalable ion traps for quantum information processing. *New Journal of Physics*, 12(3):033031, 2010.

[11] M. E. Grein, M. Willis, A. Kerman, E. Dauler, B. Romkey, D. Rosenberg, J. Yoon, R. Molnar, B. S. Robinson, D. Murphy, and D. M. Boroson. A fiber-coupled photon-counting optical receiver based on nbn superconducting nanowires for the lunar laser communication demonstration. *CLEO:2014, OSA Technical Digest*, page SM4J.5.

[12] M.D. Shaw, J.A. Stern, K. Birnbaum, M. Srinivasan, M. Cheng, K. Quirk, A. Biswas, F. Marsili, V. B. Verma, R. P. Mirin, S. W. Nam, and W. H. Farr. Tungsten silicide superconducting nanowire arrays for the lunar laser octl terminal. *CLEO:2013 Technical Digest*, page QM4L.7.

[13] G. N. Gol'tsman, O. Okunev, G. Chulkova, A. Lipatov, A. Semenov, K. Smirnov, B. Voronov, A. Dzardanov, C. Williams, and Roman Sobolewski. Picosecond superconducting single-photon optical detector. *Applied Physics Letters*, 79(6):705–707, 2001.

[14] F. Marsili, V. B. Verma, J. Stern, S. Harrington, A. Lita, T. Gerrits, I. Vayshenker, B. Baek, M. Shaw, R. Mirin, and S. W. Nam. Detecting sin-

gle infrared photons with 93% system efficiency. *Nat Photon*, 7(3):210–214, 03 2013.

[15] Shigehito Miki, Taro Yamashita, Hirotaka Terai, and Zhen Wang. High performance fiber-coupled nbtin superconducting nanowire single photon detectors with gifford-mcmahon cryocooler. *Optics Express*, 21(8):10208–10214, 2013.

[16] W. Pernice, C. Schuck, O. Minaeva, M. Li, G. N. Gol'tsman, A. V. Sergienko, and H. X. Tang. High Speed and High Efficiency Travelling Wave Single-Photon Detectors Embedded in Nanophotonic Circuits. *Nature Communications*, 3:1325, 2012.

[17] A. Semenov, G. Gol'tsman, and A. Korneev. Quantum detection by current carrying superconducting film. *Physica C*, 351:349–356, 2001.

[18] Francesco Marsili, Faraz Najafi, Eric Dauler, Hasan Korre, Vikas Anant, Kristen Sunter, and Karl K. Berggren. Cavity-Integrated Ultra-Narrow Superconducting Nanowire Single-Photon Detector Based on a Thick Niobium Nitride Film. In *CLEO: QELS-Fundamental Science*, page QTu3E.3. Optical Society of America, 2012.

[19] Andrew J. Kerman, Eric A. Dauler, Joel K. W. Yang, Kristine M. Rosfjord, Vikas Anant, Karl K. Berggren, Gregory N. Gol'tsman, and Boris M. Voronov. Constriction-limited detection efficiency of superconducting nanowire single-photon detectors. *Applied Physics Letters*, 90(10):101110, 2007.

[20] William E. Keicher Joel K. W. Yang Karl K. Berggren Andrew J. Kerman, Eric A. Dauler. Kinetic-inductance-limited reset time of superconducting nanowire photon counters. *Applied Physics Letters*, 88:111116, 2006.

[21] Joel K. W. Yang, Andrew J. Kerman, Eric A. Dauler, Vikas Anant, Kristine M. Rosfjord, and Karl K. Berggren. Modeling the electrical and thermal response of superconducting nanowire single-photon detectors. *IEEE Transactions on Applied Superconductivity*, 17:581–585, 2007.

[22] B. Baek, A. E. Lita, V. Verma, and S. W. Nam. Superconducting a-w(x)si(1-x) nanowire single-photon detector with saturated internal quantum efficiency from visible to 1850 nm. *Applied Physics Letters*, 98:251105, 2011.

[23] V. B. Verma, F. Marsili, S. Harrington, A. E. Lita, R. P. Mirin, and S. W. Nam. A four-pixel single-photon pulse-position array fabricated from wsi superconducting nanowire single-photon detectors. *Applied Physics Letters*, 101:251114, 2012.

[24] Yu P Korneeva, M Yu Mikhailov, Yu P Pershin, N N Manova, A V Divochiy, Yu B Vakhtomin, A A Korneev, K V Smirnov, A G Sivakov, A Yu Devizenko, and G N Goltsman. Superconducting single-photon detector made of mosi film. *Supercond. Sci. Technol.*, page 095012, 2014.

[25] V. B. Verma, A. E. Lita, M. R. Vissers, F. Marsili, D. P. Pappas, R. P. Mirin, and S. W. Nam. Superconducting nanowire single photon detectors fabricated from an amorphous mo0.75ge0.25 thin film. *Optical Engineering*, page 1402.4526, 2014.

[26] F. Najafi, A. Dane, F. Bellei, Q. Zhao, K. Sunter, A. McCaughan, and K. K. Berggren. Fabrication process yielding saturated nanowire single-photon detectors with 24ps jitter. *IEEE Journal on Selected Topics in Quantum Electronics*, 21(2):3800507, 2015.

[27] M. Ejrnaes, R. Cristiano, O. Quaranta, S. Pagano, A. Gaggero, F. Mattioli, R. Leoni, B. Voronov, and G. Gol'tsman. A cascade switching superconducting single photon detector. *Applied Physics Letters*, 91:262509, 2007.

[28] Scott Aaronson and Alex Arkhipov. The computational complexity of linear optics. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, STOC '11, pages 333–342, New York, NY, USA, 2011. ACM.

[29] Jun Chen, Zachary H. Levine, Jingyun Fan, and Alan L. Migdall. Frequency-bin entangled comb of photon pairs from a silicon-on-insulatormicro-resonator. *Opt. Express*, 19(2):1470–1483, Jan 2011.

[30] Hiroshi Fukuda, Koji Yamada, Tetsufumi Shoji, Mitsutoshi Takahashi, Tai Tsuchizawa, Toshifumi Watanabe, Jun ichi Takahashi, and Sei ichi Itabashi. Four-wave mixing in silicon wire waveguides. *Opt. Express*, 13(12):4629–4637, Jun 2005.

[31] Jacob Mower and Dirk Englund. Efficient generation of single and entangled photons on a silicon photonic integrated chip. *Phys. Rev. A*, 84:052326, Nov 2011.

[32] Saikat Guha and Baris I. Erkmen. Gaussian-state quantum-illumination receivers for target detection. *Phys. Rev. A*, 80:052310, Nov 2009.

[33] M. A. Nielsen. Optical Quantum Computation Using Cluster States. *Phys. Rev. Lett.*, 93(4):040503, 2004.

[34] E. Knill, R. Laflamme, and G. J. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409:4652, 2001.

[35] D. Rosenberg, A. J. Kerman, R. J. Molnar, and E. A. Dauler. High-speed and high-efficiency superconducting nanowire single photon detector array. *Optics Express*, 21(2):1440–1447, 2013.

[36] Eric A. Dauler, Andrew J. Kerman, Bryan S. Robinson, Joel K. Yang, Boris Voronov, Gregory Gol'tsman, Scott A. Hamilton, and Karl K. Berggren. Photon-number-resolution with sub-30-ps timing using multi-element superconducting nanowire single photon detectors. *Journal of Modern Optics*, 56(2 & 3):364–373, January 2009.

[37] Z. Wang S. Miki, T. Yamashita and H. Terai. A 64-pixel nbtin superconducting nanowire single-photon detector array for spatially resolved photon detection. *Optics Express*, 22:223–225, 2014.

[38] M.S. Allman, V.B. Verma, R. Horansky, F. Marsili, J.A. Stern, M.D. Shaw, A.D. Beyer, and R.P. Mirin. Progress towards a near ir single-photon superconducting

nanowire camera for free-space imaging of light. *CLEO 2014 Technical Digest*, page AW3P.3, 2014.

[39] V. B. Verma, R. Horansky, F. Marsili, J. A. Stern, M. D. Shaw, A. E. Lita, R. P. Mirin, , and S. W. Nam. A four-pixel single-photon pulse-position array fabricated from wsi superconducting nanowire single-photon detectors. *Applied Physics Letters*, 104:051115, 2014.

[40] Faraz Najafi, Jacob Mower, Nicholas C. Harris, Francesco Bellei, Andrew Dane, Catherine Lee, Xiaolong Hu, Prashanta Kharel, Francesco Marsili, Solomon Assefa, Karl K. Berggren, and Dirk Englund. On-chip detection of non-classical light by scalable integration of single-photon detectors. *Nature Communications*, 6:5873, 2015.

[41] Joel Yang. *Ph.D. Thesis: Advancements in superconducting nanowire single photon detectors and development of fabrication methods for sub-10-nm lithography*. MIT, 2009.

[42] Eric Dauler. *Ph.D. Thesis: Multi-element superconducting nanowire single photon detectors*. MIT, 2009.

[43] Joel KW Yang, Eric Dauler, Antonin Ferri, Aaron Pearlman, Aleksandr Verevkin, Gregory Gol'tsman, Boris Voronov, Roman Sobolewski, William E Keicher, and Karl K Berggren. Fabrication development for nanowire ghz-counting-rate single-photon detectors. *Applied Superconductivity, IEEE Transactions on*, 15(2):626–630, 2005.

[44] F. Marsili, F. Najafi, E. Dauler, R. J. Molnar, and K. K. Berggren. Afterpulsing and instability in superconducting nanowire avalanche photodetectors. *Applied Physics Letters*, 100(11):112601, 2012.

[45] Ryan Murphy, Matthew Grein, Theodore Gudmundsen, Adam McCaughan, Faraz Najafi, Karl K. Berggren, Francesco Marsili, and Eric Dauler. Saturated

photon detection efficiency in nbn superconducting photon detectors. *CLEO 2015 Technical Digest*, page FF2A.3, 2015.

[46] John R. Clem and Karl K. Berggren. Geometry-dependent critical currents in superconducting nanocircuits. *Physical Review B*, 84:174510, 2011.

[47] A. Casaburi, A. Pizzone, and R. H. Hadfield. Large area superconducting nanowire single photon detector arrays. *Proc. Fotonica AEIT Italian Conf. Photon. Technol.*, pages 1–4, 2014.

[48] X. Hu, E. A. Dauler, R. J. Molnar, and K. K. Berggren. Superconducting nanowire single-photon detectors integrated with optical nano-antennae. *Optics Express*, 19:1392, 2011.

[49] Francesco Marsili, Faraz Najafi, Charles Herder, and Karl K. Berggren. Electrothermal simulation of superconducting nanowire avalanche photodetectors. *Applied Physics Letters*, 98:093507, 2011.

[50] W. J. Skocpol, M. R. Beasley, and M. Tinkham. Self-heating hotspots in superconducting thin-film microbridges. *Journal of Applied Physics*, 45:4054–4066, 1974.

[51] Adam McCaughan and Karl Berggren. A superconducting-nanowire three-terminal electrothermal device. *Nano Letters*, (10):5748, 2014.

[52] Andrew J. Kerman, Danna Rosenberg, Richard J. Molnar, and Eric A. Dauler. Readout of superconducting nanowire single-photon detectors at high count rates. *Journal of Applied Physics*, 113:144511, 2013.

[53] Kristen Sunter. *Optical Modeling of Superconducting Nanowire Single Photon Detectors*. PhD Thesis, Harvard University, 2014.

[54] Faraz Najafi. *Characterization of Superconducting Nanowire Single-Photon Detectors*. Diploma Thesis, TU Munich, 2010.

[55] Xiaolong Hu, C.W. Holzwarth, D. Masciarelli, E.A. Dauler, and K.K. Berggren. Efficiently Coupling Light to Superconducting Nanowire Single-Photon Detectors. *Applied Superconductivity, IEEE Transactions on*, 19(3):336–340, june 2009.

[56] Yurii Vlasov and Sharee McNab. Losses in single-mode silicon-on-insulator strip waveguides and bends. *Opt. Express*, 12(8):1622–1631, Apr 2004.

[57] Döndü Sahin, Alessandro Gaggero, Thang Ba Hoang, Giulia Frucci, Francesco Mattioli, Roberto Leoni, Johannes Beetz, Matthias Lermer, Martin Kamp, Sven Höfling, and Andrea Fiore. Integrated autocorrelator based on superconducting nanowires. *Opt. Express*, 21(9):11162–11170, May 2013.

[58] Po Dong, Wei Qian, Shirong Liao, Hong Liang, Cheng-Chih Kung, Ning-Ning Feng, Roshanak Shafiiha, Joan Fong, Dazeng Feng, Ashok V. Krishnamoorthy, and Mehdi Asghari. Low loss shallow-ridge silicon waveguides. *Opt. Express*, 18(14):14474–14479, Jul 2010.

[59] Yunhong Ding, Christophe Peucheret, Haiyan Ou, and Kresten Yvind. Fully etched apodized grating coupler on the soi platform with -0.58db coupling efficiency. *Opt. Lett.*, 39(18):5348–5350, Sep 2014.

[60] Michael Wood, Peng Sun, and Ronald M. Reano. Compact cantilever couplers for low-loss fiber coupling to silicon photonic integrated circuits. *Opt. Express*, 20(1):164–172, Jan 2012.

[61] Adam N. McCaughan and Karl K. Berggren. A superconducting-nanowire three-terminal electrothermal device. *Nano Letters*, (10):5748, 2014.

[62] Matthew A. Meitl, Zheng-Tao Zhu, Vipan Kumar, Keon Jae Lee, Xue Feng, Yonggang Y. Huang, Ilesanmi Adesida, Ralph G. Nuzzo, and John A. Rogers. Transfer printing by kinetic control of adhesion to an elastomeric stamp. *Nat Mater*, 5(1):33–38, 01 2006.

[63] W. H. P. Pernice, C. Xiong, C. Schuck, and H. X. Tang. High-q aluminum nitride photonic crystal nanobeam cavities. *Applied Physics Letters*, 100(9):–, 2012.

[64] C. Xiong, W. H. P. Pernice, and H. X. Tang. Low-loss, silicon integrated, aluminum nitride photonic circuits and their use for electro-optic signal processing. *Nano Letters*, 12:3562–3568, 2012.

[65] Clemens Schriever, Federica Bianco, Massimo Cazzanelli, Mher Ghulinyan, Christian Eisenschmidt, Johannes de Boor, Alexander Schmid, Johannes Heitmann, Lorenzo Pavesi, and Joerg Schilling. Second-order optical nonlinearity in silicon waveguides: Inhomogeneous stress and interfaces. *Advanced Optical Materials*, 3:129–136, 2014.

[66] M G Tanner, L San Emeterio Alvarez, W Jiang, R J Warburton, Z H Barber, and R H Hadfield. A superconducting nanowire single photon detector on lithium niobate. *Nanotechnology*, 23(50):505201, 2012.

[67] Alexander W. Fang, Hyundai Park, Oded Cohen, Richard Jones, Mario J. Paniccia, and John E. Bowers. Electrically pumped hybrid algainas-silicon evanescent laser. *Opt. Express*, 14(20):9203–9210, Oct 2006.

[68] Sara L. Mouradian, Tim Schroeder, Carl B. Poitras, Luozhou Li, Jordan Goldstein, Edward H. Chen, Jaime Cardenas, Matthew L. Markham, Daniel J. Twitchen, Michal Lipson, and Dirk Englund. The scalable integration of long-lived quantum memories into a photonic circuit. *Arxiv*, page 1409.7965, 2014.

[69] Da Chen, Dong Xu, Jingjing Wang, Bo Zhao, and Yafei Zhang. Dry etching of aln films using the plasma generated by fluoride. *Vacuum*, 2:282, 2008.