Neuromorphic Computing Based on Superconductive Quantum Phase-Slip Junctions

by

Ran Cheng

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
December 11, 2021

Keywords: Quantum phase-slip junctions, neuromorphic computing, superconductive neuron, superconductive synapse, non-von Neumann

Approved by

Michael C. Hamilton, Chair, Professor of Electrical and Computer Engineering
Mark L. Adams, Associate Professor of Electrical and Computer Engineering
Masoud Mahjouri-Samani, Assistant Professor of Electrical and Computer Engineering
Peng Li, Assistant Professor of Electrical and Computer Engineering
Luke Oeding, Associate Professor of Mathematics

Abstract

Superconductive electronics that exhibit ultra-low power consumption and high speed are good candidates for neuromorphic computing, which aims to emulate the human brain, as CMOS-based systems are approaching the bottleneck of Moore's law. Quantum phase-slip junctions (QPSJs) are 1-D superconducting nanowires that have been identified as exact duals to Josephson junctions (JJs), based on charge-flux duality in Maxwell's equations. These superconductive circuits that operate by the propagation of small voltage or current pulses, corresponding to propagation of single flux or charge quantum, are naturally suited for implementing spiking neuron circuits.

Superconductive circuits based on QPSJs can conduct quantized charge pulses, which naturally resemble action potentials generated by biological neurons. Synaptic circuits, which work as dynamic weighted connections between two neurons, can also be realized by circuits comprised of QPSJs and magnetic Josephson junctions (MJJs) or only using QPSJs as a means of charge modulation for quantized charge propagation. A fan-out circuit uses charge-flux converters to emulate dendrites that allow a neuron to connect with many other neurons. Unlike a JJ splitter circuit that provides very limited fan-out, charge-flux converters, along with the corresponding circuitry, can provide significantly more fan-out. We present basic neuromorphic computing circuitry components, including neuron, synaptic, fan-out, and axon circuits. In addition to that, a learning circuit is introduced to explore learning functions in these systems. We use a simplified spike timing dependent plasticity (STDP) learning rule to automatically update the synaptic weight between a presynaptic and postsynaptic neuron according to their relative spike timings. Using a SPICE model developed for QPSJs, circuits were simulated in WRspice to demonstrate corresponding functionalities.

An important step for the experimental realization of QPSJ-based circuits is to fabricate superconducting nanowires that exhibit coherent quantum phase-slip. Since niobium nitride (NbN) was identified as an appropriate material for QPSJ, we optimized an NbN deposition

process in our available equipment and fabricated ultra-narrow NbN nanowires in search of quantum phase-slips. We investigated low-temperature transport behavior in our NbN superconducting nanowires. NbN nanowires with different dimensions on different substrates were fabricated and tested at temperatures down to 1.5 K. Resistive tails were observed for ultra-thin and narrow nanowires below the superconducting critical temperature T_{C}. These results suggest that phase-slips may exist in these test structures.

This project and work presented in this dissertation provide not only multiple neuromorphic circuits using QPSJs and other superconducting technologies, such as JJs, for high-speed, low power dissipation neuromorphic computing, but also provide experimental results of NbN nanowire fabrication and characterization that show potential evidence of quantum phase-slips. These results are critical for the future development of the fabrication process of reproducible and controllable QPSJs and physical implementations of QPSJ-based neuromorphic circuits.

Acknowledgments

I would like to express my gratitude to my academic advisor, Dr. Michael C. Hamilton for providing significant guidance, assistance and financial support throughout my PhD study at Auburn University. I thank him not only for introducing me to the exciting field of superconductors, but also for welcoming me as part of his group and providing me so many useful resources to build up my knowledge and experience towards a PhD degree.

I would also like to thank Dr. Mark Adams, Dr. Masoud Mahjouri-Samani and Dr. Peng Li for their review and commentary, and Dr. Luke Oeding for serving as the University Reader for this work. Their comments and advice are very valuable and helpful.

I would also like to thank Drew Sellers, and my group members Dr. Vaibhav Gupta, Bhargav Yelamanchili, Archit Shah, George Hughes, Stephen Bankson, Harry Goyal and Sherman Peek for their cooperative support and encouragement throughout my PhD study. I sincerely recognize their contribution and advice to my research and this dissertation. I thank Dr. Uday S. Goteti for his encouragement and many enlightening discussion.

Finally, my deepest gratitude goes to my wife Qiong Hu, to my parents Xiaowen Cheng and Qin Yang, for their unending love, encouragement and support throughout my life.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AC	Alternating current
AFM	Atomic force microscope
ANN	Artificial neural network
AP	Action potential
CMOS	Complementary metal-oxide-semiconductor
DC	Direct current
EBL	Electron beam lithography
EHT	Electron high tension
EPSC	Excitatory postsynaptic potential
GPU	Graphics processing unit
HSQ	Hydrogen silsesquioxane
IFN	Integrate-and-fire neuron
IPSP	Inhibitory postsynaptic potential
JJ	Josephson junction
LTD	Long term depression
LTP	Long term potentiation

MCM Multi-chip module

MJJ Magnetic Josephson junction

MQT Macroscopic quantum tunneling

NDRO Non-destructive readout

PCB Printed circuit board

PMMA Polymethyl methacrylate

PPMS Physical Property Measurement System

PVD Physical vapor deposition

QPS Quantum phase-slip

QPSJ Quantum phase-slip junction

RNN Recurrent neural network

RSFQ Rapid single flux quantum

SEM Scanning electron microscope

SFQ Single flux quantum

SNN Spiking neural network

SNSPD Superconducting single-photon detector

SPICE Simulation program with integrated circuit emphasis

SQUID Superconducting quantum interference device

SRDP Spike rate dependent plasticity

STDP Spike time dependent plasticity

TAPS Thermally activated phase-slip

VLSI Very large-scale integration

XRD X-ray diffraction

Chapter 1

Introduction

1.1 Background and motivation

The imbalance between data processing speed and memory access speed restricts the further development of von Neumann architecture [18]. While nearing the impending end of Moore's Law scaling [19], the saturation in the power and operation frequency of CMOS processors has seen a bottleneck, which has driven the development of new approaches, such as GPU computing [20], quantum computing [21] and neuromorphic computing [22], all of which are considered to improve computational performance.

Neuromorphic computing is a brain-inspired, non-von Neumann architecture that has been realized in CMOS technologies [22, 23, 24, 18, 25]. The average human brain has a level of approximately 10^{11} neurons, and each neuron is connected to up to 15,000 adjacent neurons via synapses, resulting in a total of 10^{15} synapses [26, 27]. For approximately 10^{15} operations per second, the brain dissipates only approximately 10^{-16} J per operation [3]. Mapping conventional neural networks running on a von Neumann machine to a neuromorphic platform designed specifically for neural networks is expected to significantly reduce power consumption and improve processing efficiency [28]. Neuromorphic systems, proposed as adaptive analog technology, can process massive data represented in the form of relative analog values in parallel and are expected to be important platforms for information processing moving forward.

Solid-state neuromorphic hardware has recently experienced remarkable improvement due to advanced, highly-scaled CMOS technology and emerging memory devices [23, 24, 29, 18,

30, 31, 25, 32]. These non-von Neumann architectures show great promise for solving complex problems, such as object recognition and decision-making, faster and more efficiently than conventional von Neumann architectures [33, 34]. However, a bottleneck arises when scaling CMOS technology to smaller nodes, as its energy efficiency is several orders of magnitude lower than the human brain. This is evident in the case of Spiking Neural Network Architecture (SpiNNaker), which is one of the largest projects within the Human Brain Project [35]. SpiNNaker can simulate 1% of the human brain's 85 billion neurons but dissipates over 2500 times more power than a human brain [36].

Superconducting circuits that exhibit ultra-low power dissipation and high-speed operation are interesting candidates for brain-inspired computation. The massive interconnection required for neuromorphic circuits presents a significant challenge for systems with lossy interconnects, such as CMOS integrated circuits, while superconducting interconnects are nearly lossless at the chip, multi-chip module (MCM) and printed circuit board (PCB) levels [37]. Research on Josephson junction (JJ)-based superconducting neuromorphic circuits has made great progress over the past decade [38, 5, 39, 40, 41, 42, 43, 44]. With the realization of magnetic Josephson junctions (MJJs)[45, 46], the research of Russek, Schneider, et al. [42, 43, 44] has recently demonstrated learning functions in a superconducting neuromorphic circuit, consisting of JJs and MJJs, which improved the speed to the order of ps/spike and energy dissipation to the order of aJ/spike. Comparable achievements have been realized by a hybrid platform with semiconducting few-photon light-emitting diodes and superconducting nanowire single-photon detectors (SNSPDs)[47, 7]. A superconducting nanowire in parallel with a shunt resistor is viewed as a relaxation oscillator that also has spiking behavior, which emulates a neuron action potential [48, 8]. Furthermore, these systems perform well, even when taking into consideration the power required for cooling the superconducting hardware to required temperatures of $\lesssim 4$ K [49]. Additional state-of-the-art superconducting neuromorphic technologies will be introduced in Section 1.3.

Quantum phase-slip junctions, which rely on quantum phase-slip (QPS) in superconducting nanowires, are an emerging superconducting device family that shows promise for high-speed and low-power superconductive electronics [50, 12, 10, 51, 52, 53, 54, 13, 14, 15].

Quantum phase-slips can be observed as a resistive tail at a temperature sufficiently below the critical temperature T_c [55]. The phase of superconducting order parameter is shifted by 2π in a one-dimensional superconducting nanowire due to quantum tunneling [10]. Multiple experiments were performed to demonstrate the existence of quantum phase-slip during the past few years [11, 51, 56]. Applications were suggested in digital computing, quantum computing, and current standards [50, 10, 51], but related neuromorphic computing applications have not been particularly explored. Coherent quantum phase-slips, as exact duals to Josephson junctions [12], can generate quantized current pulses that behave like neuron spiking events [57, 53]. QPSJ-based neuromorphic hardware systems are expected to provide lower power consumption, similar operation speed and smaller chip area compared to JJ-based neuromorphic systems [13], and could outperform or complement other existing superconducting neuromorphic hardware systems. With a learning circuit proposed in this work, this QPSJ-based neuromorphic hardware system is expected to perform complex computational functions, e.g., pattern recognition. Furthermore, taking the advantage of loss-less superconducting interconnections along with the scalability of QPSJ-based neuromorphic hardware could help simulate neuron spiking dynamics in large spiking neural networks. In this work, we explored the use of QPSJ-based circuits for high-speed and low-power neuromorphic computing. To begin, we give an introduction to brain-inspired neuromorphic computing and artificial neural networks (including spiking neural networks), and review state-of-the-art superconducting neuromorphic computing hardware in the following subsections.

1.2 Brain-inspired neuromorphic computing

In this section, we will explain some of the known biological processes in human brains and corresponding mathematical models that are used to emulate them. Synaptic plasticity, as a means of learning functions, will also be introduced. A biological neural network has multiple functional components to receive, process, and transmit signals. Artificial neural network (ANN) is a computational model inspired by biological neural network for applications in machine learning and pattern recognition [3, 58]. As the third generation of ANNs, spiking neural networks (SNNs) are the core of brain-inspired neuromorphic computing systems. A biological

neural network, also known as the nervous system, is modeled by mathematical equations and realized by circuits to emulate brain functions. Neuromorphic chips are physical implementations of spiking neural networks at a low energy cost to perform computational functions.

1.2.1 Biological nervous system

A biological nervous system has a large number of interconnected cells that coordinate voluntary and involuntary actions by transmitting electrical signals to and from different parts of the body [1]. Neurons are the fundamental and primary functional units of a biological nervous system, as shown in Figure 1.1(a). The cell body of a neuron is called soma (typically known as "neuron"), which can receive, process, and transmit information in the forms of cellular signals. The membrane potential of an excitable neuron increases after it receives currents from other neurons or external stimuli. Once the membrane potential is beyond a threshold, it will rapidly rise and fall to a resting potential followed by a refractory period, during which the neuron does not respond to any stimulus. The resulting nerve impulse signal, which occurs in neuron cells and propagates along an axon towards synaptic boutons, is called an action potential (AP) or "spike" [59], as shown in Figure 1.1(b).

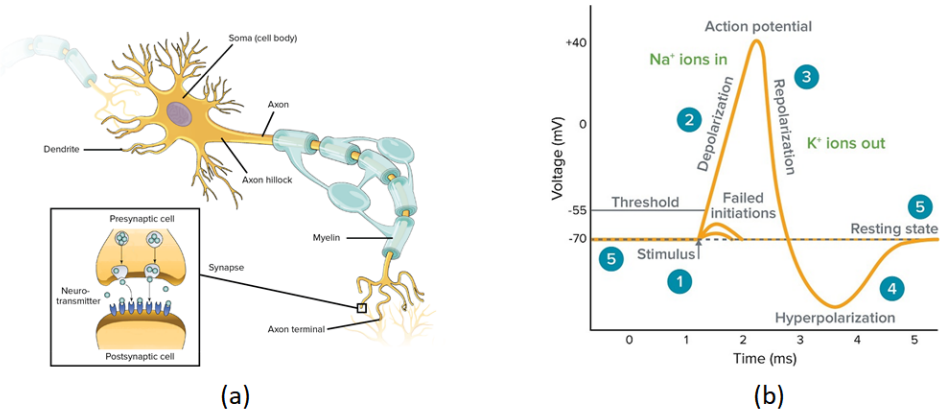


Figure 1.1: (a) Biological neuron anatomy, from [1]. A neuron is comprised of soma, synapse, dendrite, axon, etc. (b) Action potential, from [2].

A synapse connects two neurons, which are referred to as presynaptic and postsynaptic neurons, by diffusing neurotransmitter molecules produced by presynaptic neuron to postsynaptic neuron [60]. The type of receptor and neurotransmitter determines whether the postsynaptic neuron should be excited or inhibited. The synapse does not directly store information to be processed, but its strength affects the changing speed of the postsynaptic neuron's membrane potential. This kind of memory is well-known as synaptic plasticity, which can be modified through learning [61].

The main neuronal functions, receiving and processing incoming information, generally take place in the dendrites and cell body [1]. A neuron can receive up to thousands of input signals, either excitatory or inhibitory signals, throughout its dendrite trees. A single neuron could have more than one set of dendrites. These signals are summed up at the dendrites, which then determine whether the neuron should fire an action potential or not.

Action potentials are generally transmitted through axons to postsynaptic neurons. Many axons are covered with a particular insulating substance called myelin, which helps them propagate signals fast [1]. At the end of an axon, it is split into many branches (known as axon terminals), which allow multiple connections with other neurons.

A biological nervous system is a dynamic system that extracts signals from the internal and external environments using sensory receptors. The central part of the biological nervous system is the brain, which processes signals and sends signals to other parts of the body. While a human brain is so powerful to solve complex problems through learning, mathematical models and corresponding hardware systems have been developed to emulate a human brain for complex computation.

1.2.2 Artificial neural networks

The human brain is a highly complex, nonlinear, and parallel computing system. Artificial neural networks, commonly referred to as "neural networks", are mathematical models to emulate brain functions. A neural network is comprised of massive interconnections of simple computing cells referred to as "neurons" or "processing units" to model the way in which the brain

performs a particular task or function of interest [3]. The procedure used to perform the learning process is called a learning algorithm, the function to modify the synaptic weights of the network to achieve desired goals. Neural networks have many advantages over traditional computing systems. A neural network has a built-in capability to adapt synaptic weights to changes in the surrounding environment. Another feature is that neural networks have excellent error tolerance and are capable of robust computation. This feature is beneficial from the distributed nature of the information stored in the network, where the errors usually do not cause severe degradation in performance.

In a neural network, neurons are fundamental elements used to process signals. A nonlinear neuron model is illustrated in Figure 1.2. Generally, an artificial neuron can be modeled by the following equations:

$$v_k = \sum_{j=1}^{m} x_j w_{kj} + b_k \tag{1.1}$$

$$y_k = \phi(v_k) \tag{1.2}$$

where x_1, x_2, ..., x_m are input signals, w_1, w_2, ..., w_m are the respective synaptic weights of neuron k, b_k is the bias, v_k is the activation potential, $\phi(\cdot)$ is the activation function, and y_k is the output signal of the neuron.

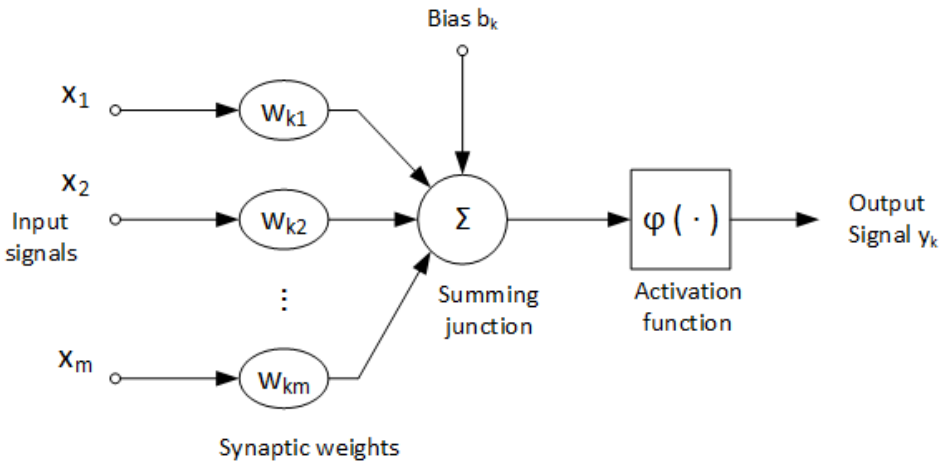


Figure 1.2: An artificial neuron model. Weighted signals with a bias b_k are summed up and injected to an activation function, which generates an output signal y_k (adapted from [3]).

The activation function $\phi(\cdot)$ that defines the output y_k is a function of v_k. Two basic activation functions are the threshold function and the sigmoid function. In engineering, a

threshold function is often referred to as a Heaviside function, as shown in Figure 1.3. The output of a neuron is given by:

$$y_k = \begin{cases} 1, & \text{if } v_k \geq 0 \\ 0, & \text{if } v_k < 0 \end{cases} \tag{1.3}$$

This equation defines a simple binary neuron, which is either active or inactive. This is the first computational model of artificial neurons, proposed by McCulloch and Pitts [62].

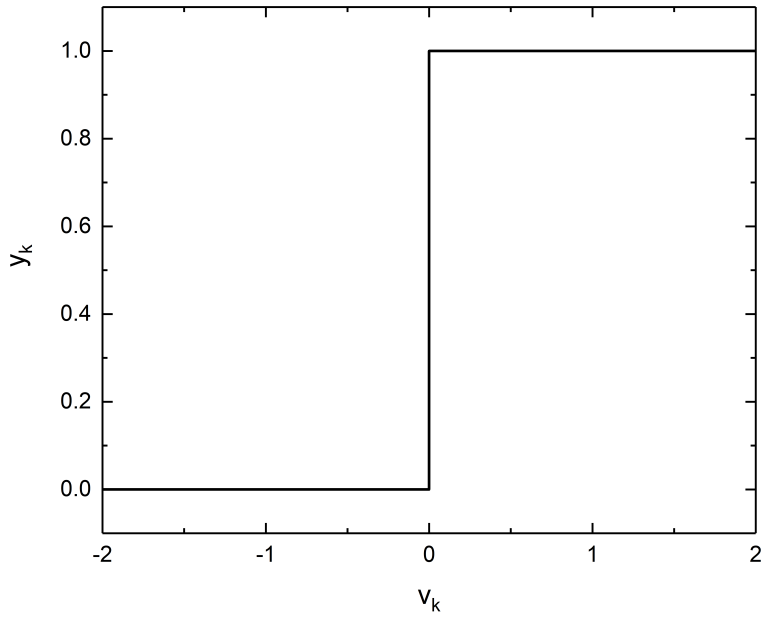


Figure 1.3: The threshold function used to describe a binary neuron behavior (adapted from [3]).

The most commonly used activation function is an "s" shaped sigmoid function. An example of the sigmoid function is a logistic function, given by:

$$\phi(v) = \frac{1}{1 + e^{-av}} \tag{1.4}$$

where a is the slope parameter. Taking different values of a, we can obtain different slopes for the sigmoid function, which exhibits a graceful balance between linear and nonlinear behavior,

as shown in Figure 1.4. The sigmoid function becomes a threshold function if parameter a approaches infinity. However, the sigmoid function is differentiable while the threshold function is not.

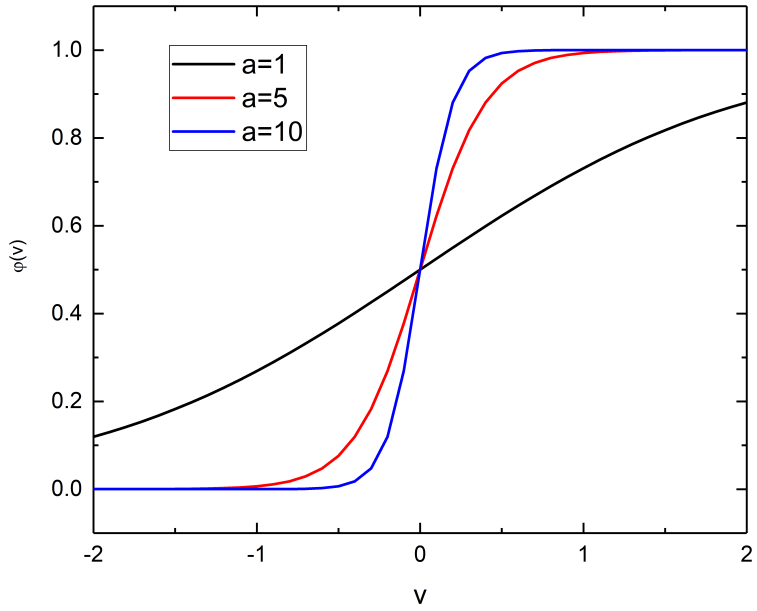


Figure 1.4: The sigmoid function with a varying slope parameter a (adapted from [3]).

Sometimes, when -1 to 1 are the desired output range, activation functions are changed to odd functions. Specifically, a modified threshold function can be defined as:

$$\phi(v) = \begin{cases} 1, & \text{if } v > 0 \\ 0, & \text{if } v = 0 \\ -1, & \text{if } v < 0 \end{cases} \tag{1.5}$$

An example for an odd sigmoid function is the hyperbolic tangent function, defined by:

$$\phi(v) = \tanh(v) \tag{1.6}$$

This function may exhibit more practical benefits (e.g., to provide inhibitory signals) compared to a logistic function.

Neurons are combined to form a network in different architectures, which is also intimately linked with the learning algorithms used to train the network. There are generally two types of network architectures: feedforward networks and recurrent networks.

A neural network is usually organized as different layers of neurons with connections between different layers. A feedforward network generally has an input layer, one or more hidden layers, and an output layer, as shown in Figure 1.5. The input layer takes input signals from source nodes and transmits them to the hidden layer. Since there are no actual neurons and no computations in the input layer, this layer is usually not counted. The hidden layer cannot be seen directly from either the input or output of the network. In a single-layer feedforward network, there are no hidden layers. The function of hidden neurons is to intervene between the external input and the network output in some useful manner. Adding more hidden layers helps the network extract high-order statistics from its inputs. The network shown in Figure 1.5 is a fully connected feedforward network. Every neuron in one layer is connected to every other neuron in adjacent layers. However, the numbers of neurons in each layer do not have to be identical.

A recurrent neural network (RNN) has at least one feedback loop, as shown in Figure 1.6. A simple example of RNNs is a single layer of neurons, where each neuron feeds its output signal back to the inputs of all the other neurons. An RNN may or may not have hidden layers. The presence of feedback loops has a profound impact on the learning capability of the network and on its performance. An RNN considers the current input and also what it has learned from the inputs it received previously to make a decision. This feature allows RNNs to show temporal dynamic behavior.

1.2.3 Spiking neural networks

Spiking neural networks (SNNs) are the third generation ANNs that more closely mimic natural neural networks [63]. Traditional ANNs do not contain temporal information of spikes and thus are not able to emulate the real biological neural networks. An SNN has been developed to process precise timing information of spikes. Neuromorphic computing is generally the hardware realization of SNNs at a chip level, which is specifically built for SNN algorithms.

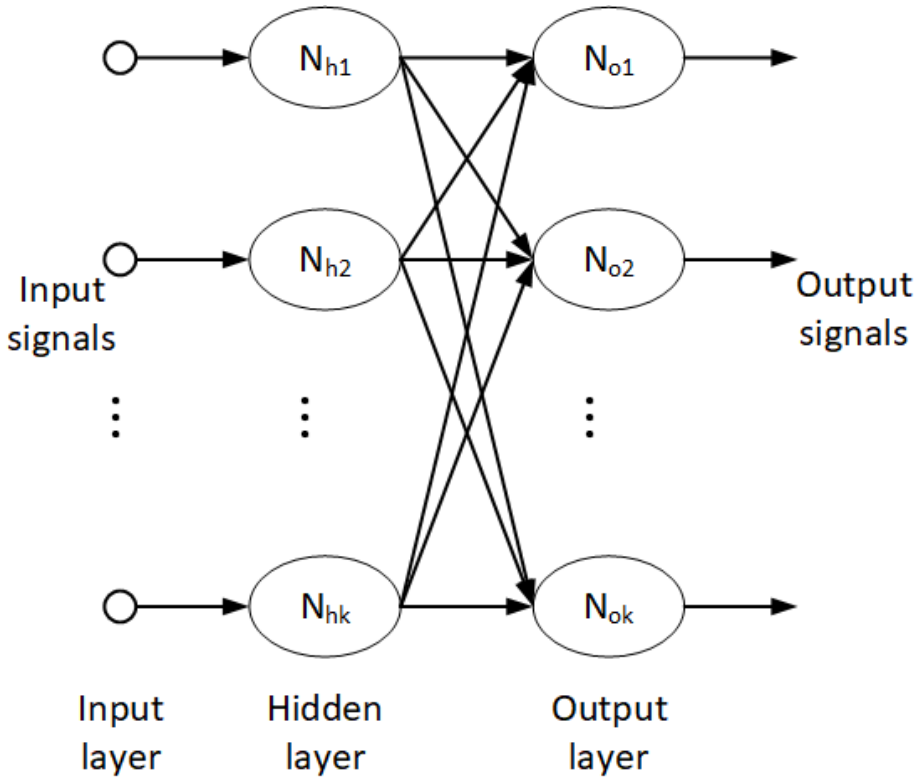


Figure 1.5: A fully connected feedforward neural network with one hidden layer and one output layer (adapted from [3]).

Unlike the conventional ANNs that process neural information with real-valued numbers, SNNs exploit both the presence and timing of individual spikes as the means of communication among neurons in different layers. An SNN utilizes spike trains as input and output, as shown in Figure 1.7. In an SNN, neurons generate spiking signals (or spikes) at the moment of threshold crossing. These spikes are transmitted to other neurons, which in turn changes their membrane potentials. Spiking information is encoded in the form of spike timings (or frequencies, etc.) rather than the amplitudes of spikes. The output signals are also decoded to interpret the result of the network. There are various input and output coding schemes for SNNs, such as rate coding [64], temporal coding [65], population coding [66] and sparse coding [67], to interpret spike trains as real-valued numbers. By using either the frequencies of the spikes or the timings between the spikes, spiking information can be processed more efficiently [68].

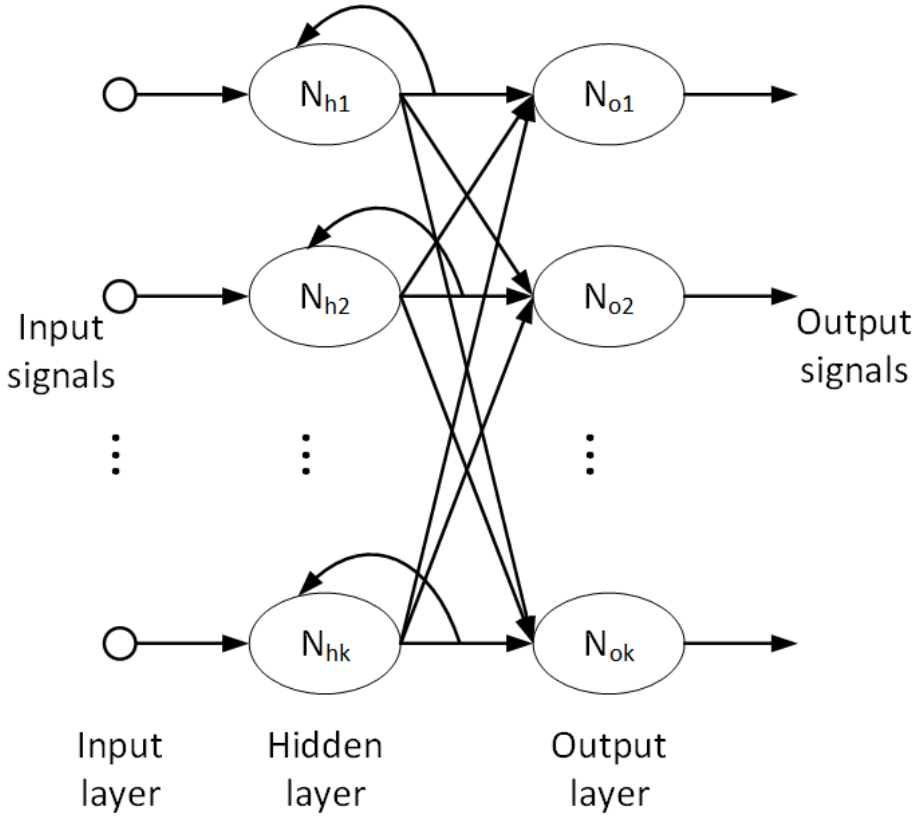


Figure 1.6: A recurrent neural network with one hidden layer and one output layer. The neurons in the hidden layer have feedback loops (adapted from [3]).

1.2.4 Synaptic plasticity in neuromorphic computing

The major difference between a computer system and a human brain is that the human brain can learn through training while the computer system cannot. While weight information is stored in synapses, the learning function of a neural network is determined by synaptic plasticity. Biologists discovered that learning behaviors such as long term potentiation (LTP), long term depression (LTD) and spike timing dependent plasticity (STDP) existed in cortical neurons [69]. Besides the broad implementation of STDP, spike rate dependent plasticity (SRDP), as an extension of Hebbian learning rule [70], was also implemented in neuromorphic hardware systems [71, 72, 73]. There are other learning rules based on STDP tailored for neuromorphic hardware systems, such as spike-driven synaptic dynamics, which takes both spike timing and recent spiking activity into account and has comparable performance to state-of-the-art learning rules [74].

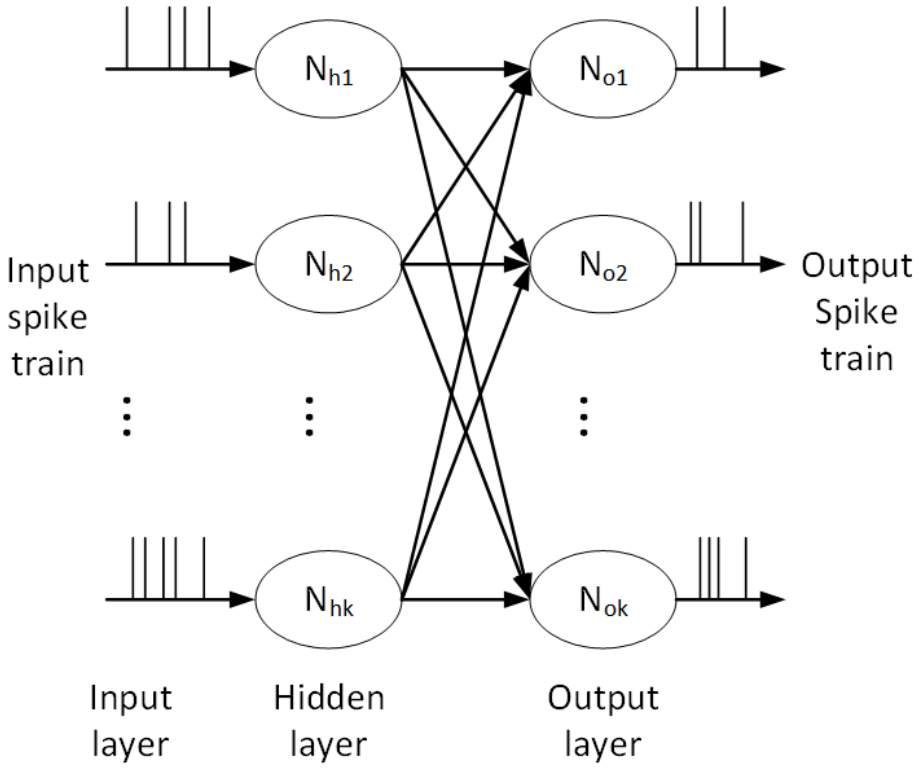


Figure 1.7: A spiking neural network that utilizes spike trains as input and output (adapted from [3]).

Spike timing dependent plasticity is most commonly used in neuromorphic computing. In 1983, biologists observed that relative timing of presynaptic and pos-synaptic action potentials at milliseconds had an effect on the synaptic plasticity [75]. After that, a series of experiments carried out by different research groups showed evidence of long-term depression induced by pairing presynaptic and postsynaptic activities [68, 76]. The subsequent research on this topic demonstrated the existence of STDP [77] and clearly recorded the entire synaptic change within 5-20 ms relative timing [4, 78]. It has also been demonstrated that STDP is more efficient than Hebbian learning rule [79]. The experimental result of STDP is shown in Figure 1.8.

Spiking rates of presynaptic and postsynaptic neurons were observed as another key factor to determine synaptic plasticity [71]. During the last few years, a memristor-based circuit that exhibits SRDP has been designed and fabricated [80, 73]. Furthermore, experiments have been conducted to demonstrate the successful application of memristor-based synapses with SRDP in neuromorphic hardware systems [81, 82].

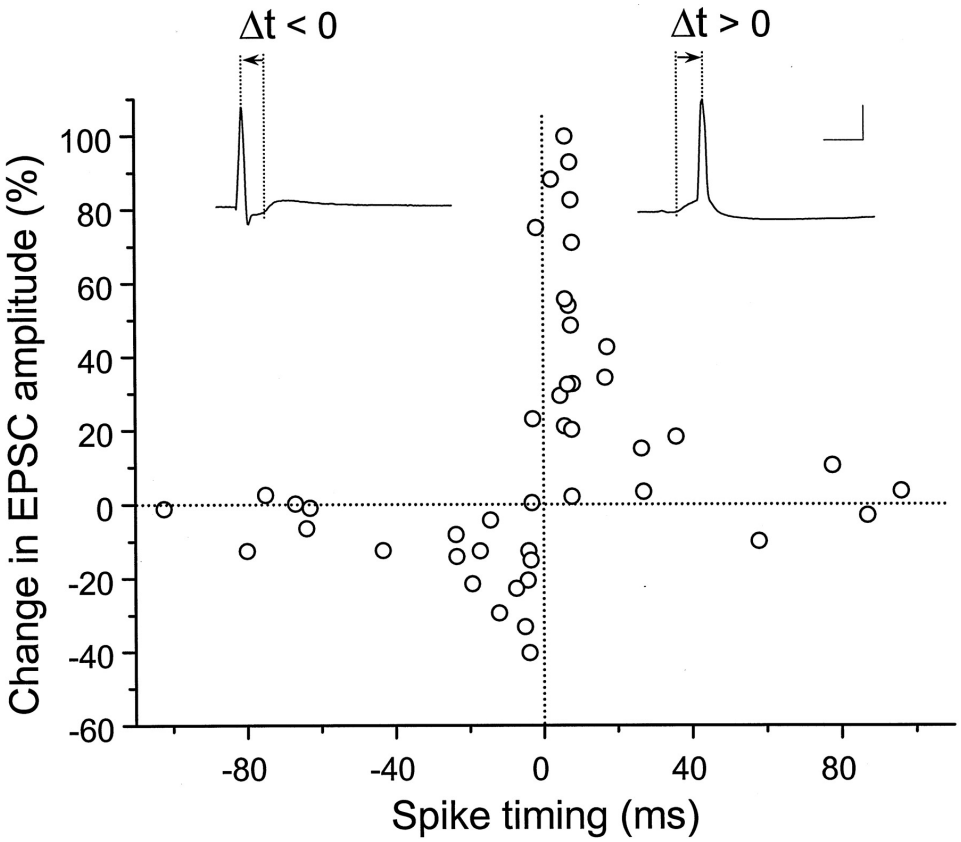


Figure 1.8: Experimental result of spike timing dependent plasticity, from [4]. EPSC is potentiated if $\Delta t > 0$ while EPSC is depressed if $\Delta t < 0$.

Although STDP and SRDP are biologically plausible, they cannot account for all the observed learning behaviors existing in a nervous system [83]. To improve the learning performance, a novel synaptic plasticity model called spike-driven synaptic dynamics was proposed by Brader in 2007 [74]. This learning rule was designed for electronic hardware and proved to be compatible with VLSI neuromorphic systems [84, 85]. The weight update rule is defined as:

$$\Delta w = a, \text{if } V(t_{\text{pre}}) > V_{\text{th}} \text{ and } \theta^l_{\text{up}} < C(t_{\text{pre}}) < \theta^h_{\text{up}}, \tag{1.7}$$

$$\Delta w = -b, \text{if } V(t_{\text{pre}}) \leq V_{\text{th}} \text{ and } \theta^l_{\text{down}} < C(t_{\text{pre}}) < \theta^h_{\text{down}} \tag{1.8}$$

where the calcium variable $C(t)$ is a function of postsynaptic activity; θ^l_{up}, θ^h_{up}, θ^l_{down} and θ^h_{down} are thresholds of the calcium variable. During the absence of presynaptic spikes, weight w

drifts towards one of two stable states:

$$\Delta w = \alpha, \text{if } w > \theta, \tag{1.9}$$

$$\Delta w = \beta, \text{if } w \leq \theta \tag{1.10}$$

Homeostasis is the capacity of neurons to regulate firing activities and maintain a stable state [86]. Neuromorphic systems with homeostatic plasticity can adjust each neuron's firing rate and prevent minor neurons from dominating the entire system [87]. It has been demonstrated that neurons with homeostasis are more robust to threshold variation since they can mitigate the negative impact the threshold variation has on the neuron. [88]. Biologists discovered that synaptic scaling mechanism existing in cortical synapses was used for homeostasis [86]. The synapses connected to frequently fired neurons or infrequently fired neurons are scaled up or down accordingly. However, when implemented in a neuromorphic system, this mechanism is different from that in a biological system. For instance, some neuromorphic systems used adjustable firing thresholds to control neuron firing rates, as seen in [88, 87]. The proposed adaptive neuron has negative feedback to control leaky current. Upon the arrival of each spike, the leaky current increases slightly, which makes it more difficult to generate more spikes subsequently[89].

1.3 A review for state-of-the-art superconducting neuromorphic hardware

Superconducting technologies are featured by very low power dissipation and very fast processing speed, which could be ideal candidates for neuromorphic computing. There are a few potential superconducting technologies for neuromorphic computing, proposed during the last few decades [90, 91, 6, 92, 47, 7, 8, 13, 14, 15]. In this section, we will introduce several state-of-the-art superconducting neuromorphic hardware systems, such as Josephson junction based neuromorphic systems, neuromorphic hardware based on superconducting optoelectronics and neuromorphic hardware based on superconducting nanowires.

1.3.1 Neuromorphic hardware based on Josephson junctions

Josephson junctions are suitable candidates for neuromorphic computing due to their intrinsic spiking behavior. An example of JJ neurons is shown in Figure 1.9. This neuron circuit provides capabilities to explore neural interactions such as synchronization, long term dynamics and bifurcations, feature identification, and information processing [5]. This circuit includes two Josephson junctions connected in a loop, which is basically a "DC-to-SFQ" converter from rapid single flux quantum (RSFQ) circuitry family [93]. Since the switching time of a JJ is within tens of picoseconds, it only takes 50 ps to generate an action potential in a JJ neuron, which is several orders of magnitude faster than typical CMOS-based neurons. In addition, the JJ neuron can reproduce some neuron behaviors such as firing threshold and refractory period, as shown in Figure 1.10. The synapse in this circuit mimics a chemical synapse that can be either excitatory or inhibitory. The voltage across the junction ϕ_p is coupled to a synaptic loop comprised of L_{syn}, R_{syn} and C_{syn} in Figure 1.9. An output current I_{12} can be generated and propagated to the next neuron. One of the biggest challenges for this JJ neuron is the limitation of integration level and device tolerance, as seen from many other JJ-based circuits [94]. Additional circuitry is necessary to provide weighted connections for synaptic plasticity.

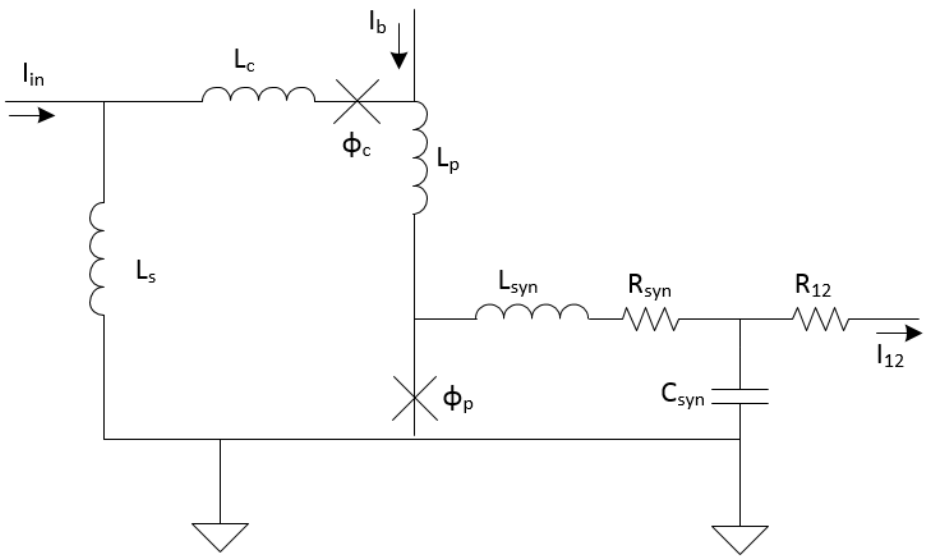


Figure 1.9: Schematic of a JJ neuron (left loop) connected to a synapse (right loop), from [5].

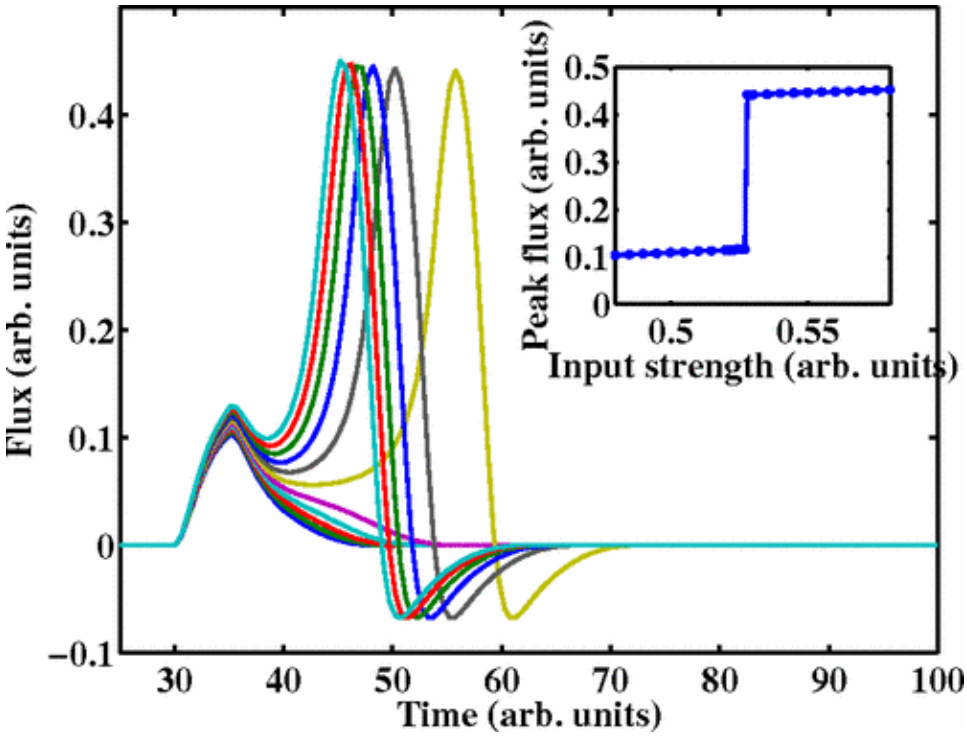


Figure 1.10: Simulation results of the JJ neuron shown in Figure 1.9, from [5].

A more complex neuromorphic system based on Josephson junctions and magnetic Josephson junctions (MJJ) was proposed by Schneider [43, 6]. The nonlinear neuron shown in Figure 1.11 has a similar structure to a superconducting quantum interference device (SQUID). The SQUID neuron is characterized by the total inductance L_{SQ}, the coupled magnetic flux Φ_{cpl} through the inductively coupled synaptic inputs, the DC bias current I_{bias}, and the critical current I_{C} of JJs in this loop. Similar to the JJ neuron circuit in Figure 1.9, this SQUID neuron can generate SFQ pulses to emulate the neuron's spiking behavior. During the zero-voltage state (i.e., $d\phi/dt = 0$ for both junctions), the current through each JJ changes smoothly according to the change of coupled flux Φ_{cpl}, until the current through one of the JJs reaches its critical current. This JJ generates an SFQ pulse, which results in increments or decrements of circulating current I_{cir}. If Φ_{cpl} decreases to zero or its original value, the other JJ will generate an SFQ pulse, returning I_{cir} to its initial state. Figure 1.12 shows the simulation results of a SQUID neuron with coupled flux ramp. One of the JJs outputs an SFQ pulse once Φ_{cpl} passes a threshold, and the other JJ generates an SFQ as Φ_{cpl} decreases.

A significant improvement of the neuromorphic hardware proposed by Schneider's group was that non-volatile devices called magnetic Josephson junctions (MJJs) were first used to

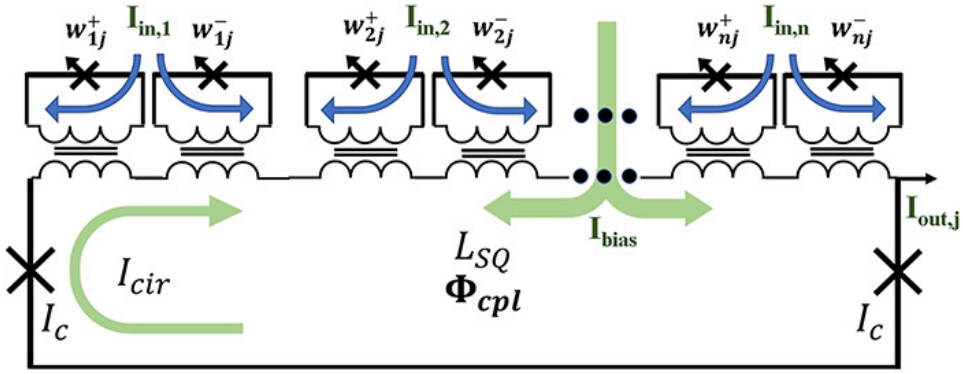


Figure 1.11: Schematic of multiple synapses inductively coupled to an SQUID neuron, from [6].

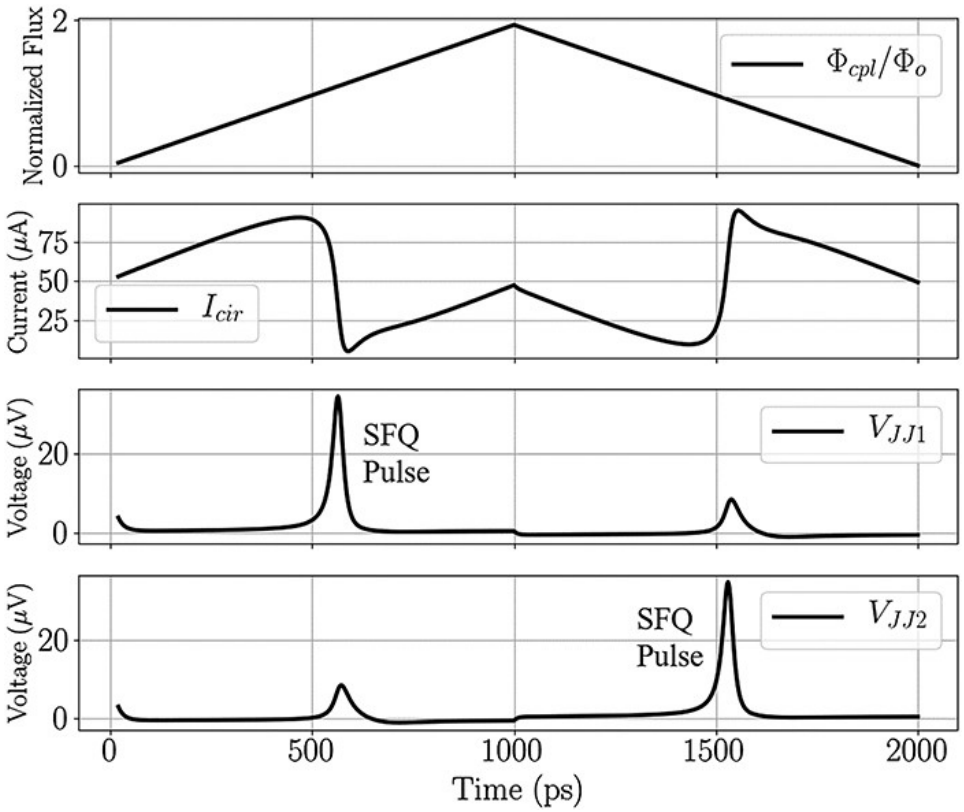


Figure 1.12: Simulation results of a SQUID neuron with coupled flux ramp, from [6].

emulate biological synapses [42]. This type of JJs doped with magnetic material has tunable critical currents, which can be used as non-volatile memory. The circuit shown in Figure 1.13 is a synaptic circuit consisting of an MJJ in parallel with a fixed inductor. An MJJ can be viewed as a nonlinear inductor that has an inductance of $\Phi_0/2\pi I_{\mathrm{C}}$, in parallel with a resistance R_{n}. The input current I_{in} through this synapse is split into two inductive branches. The current through the inductor branch is coupled to a SQUID neuron, as seen in Figure 1.11. A change of the

MJJ's inductance causes a change of current distribution in these two branches. Therefore, the current (or flux) coupled to the SQUID neuron can be adjusted by tuning the critical current of the MJJ. The adjustable synaptic weight is thus represented by the coupled current (or flux) in the JJ loop. There are also some limitations for MJJ-based superconducting synapses, e.g., the requirement of high energy and a local magnetic field for programming. [42, 44, 6, 95, 96].

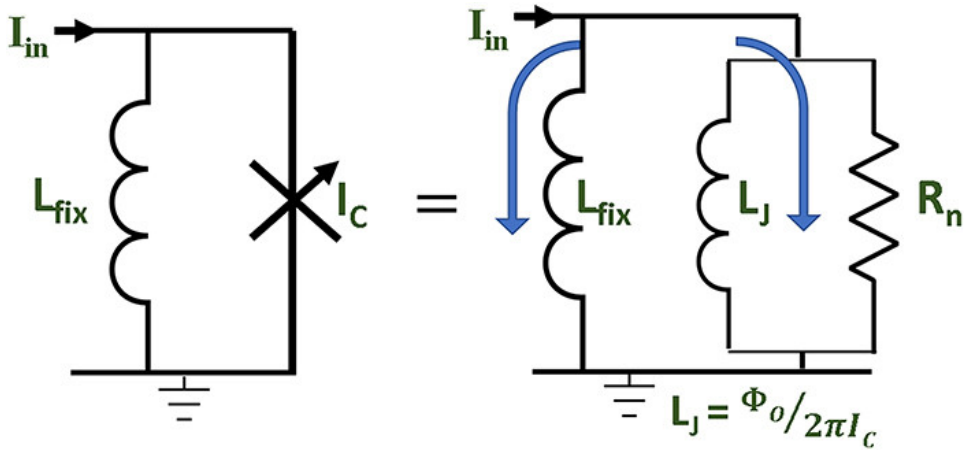


Figure 1.13: A synaptic circuit consisting of an MJJ in parallel with a fixed inductor, from [6].

1.3.2 Neuromorphic hardware based on superconducting optoelectronics

Superconducting optoelectronic hardware is also a promising candidate for large-scale neuromorphic computing. Circuits using superconducting single-photon detectors (SNSPDs) and Josephson junctions have been developed to perform neuromorphic functions such as signal reception, synaptic weighting and integration [47]. This system provides a large fan-in so that a neuron can receive input from thousands of synaptic connections. Light signals, generated from semiconductor diodes, are used to communicate between neurons in the network. They also demonstrated spike timing dependent plasticity using two photons to strengthen and two photons to weaken the synaptic weight via Hebbian learning rules. The firing speed of this type of neuron is typically up to 20 MHz. The scalability is also impressive, and a system with 8,100 neurons and 330,430 synapses can fit onto a 1 cm $\times$ 1 cm die.

Figure 1.14 shows the basic operations of the optoelectronic system. An SPD is viewed as a variable resistor r_{spd} in series with an inductor L_{spd}, which exhibits zero resistance in the steady state and switches to a high-resistance state upon the absorption of a photon. When a

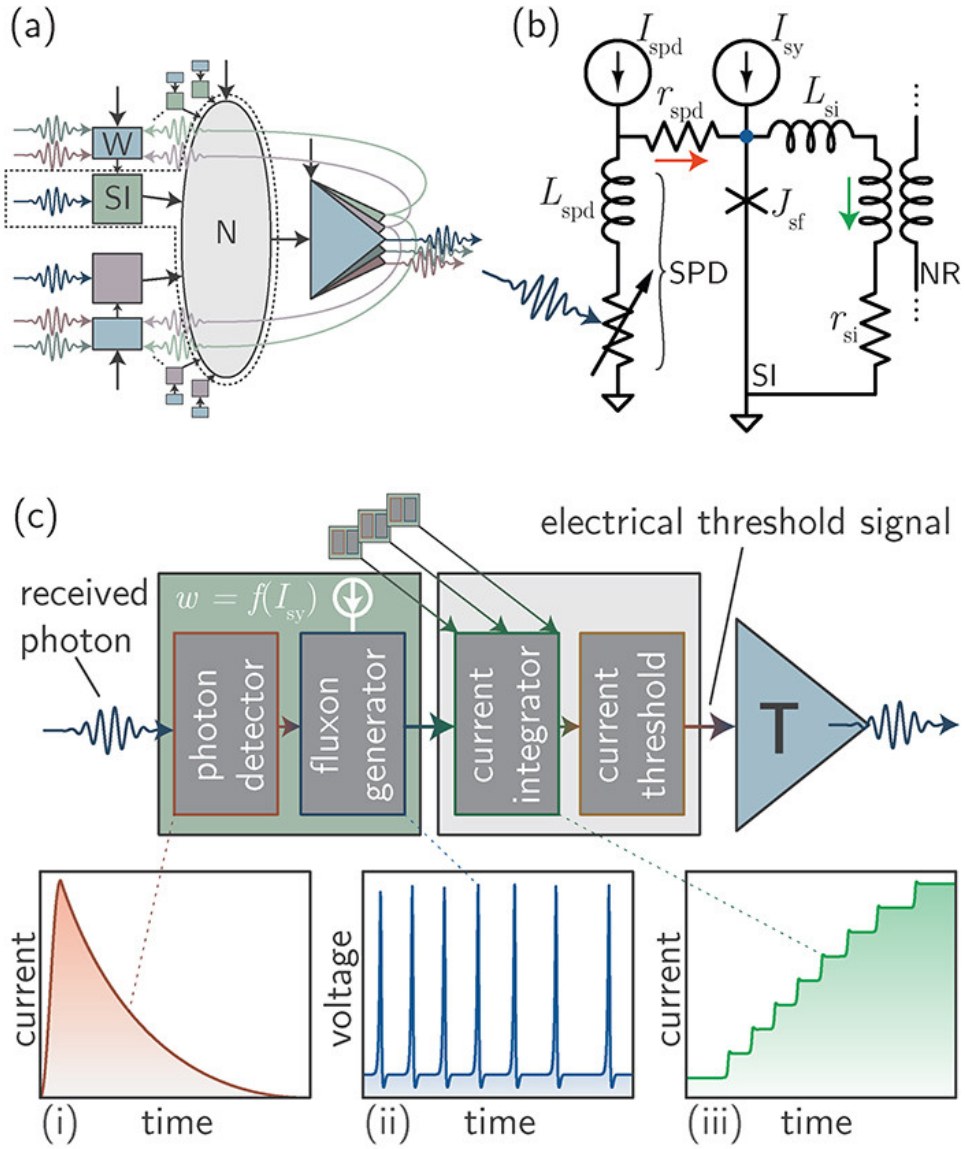


Figure 1.14: A neuromorphic hardware system based on superconducting optoelectronics, from [7]. (a) Schematic of the neuron. (b) Schematic of a simple photon-to-fluxon transducer combining a single-photon detector, Josephson junction, and flux-storage loop. (c) Signal processing sequence during the synaptic firing event. (i) Single-photo detector transduces photon to electrical current. (ii) SFQ pulses are generated when SPD diverts current to JJ. (iii) Current is integrated in the flux-storage loop.

photon is detected, current flows into J_{sf}, causing the total current in the JJ loop to exceed I_C. Depending on the value of I_{sy}, different numbers of fluxons are generated by J_{sf}. Therefore, the synaptic weight is implemented via the current bias across J_{sf}, controlled by I_{sy}. Whether the number of photons present is one or greater than one, the response of SPD is virtually identical. Furthermore, the SPD response is not strongly dependent on the frequency of light across a bandwidth broad enough for multiplexing. Implementing synaptic weight in this manner takes

advantage of both the speed and energy efficiency of JJs, while also leveraging the strengths of light for communication.

1.3.3 Neuromorphic hardware based on superconducting nanowires

Superconducting nanowires have intrinsic non-linearity and spiking behavior [48] similar to Josephson junctions. Based on the architecture proposed for a JJ neuron [5], two coupled nanowires were used to emulate spiking neuron behavior [8]. The nanowire neuron circuit shown in Figure 1.15 has a similar design to the JJ neuron circuit shown in Figure 1.9. When the current flowing into a superconducting nanowire exceeds its critical current, superconductivity is suppressed, and the nanowire becomes resistive, producing a voltage across it. The nanowire switches back to the superconducting state once the bias current reduces below the retrapping current (I_r). With a parallel shunt resistor, the switching process of the nanowire can be viewed as relaxation oscillations. The two nanowires viewed as two oscillators (main oscillator and control oscillator) can emulate the biological neuron's Na^+ ion channel and K^+ ion channel. Like the original JJ neuron circuit [5], the nanowire neuron circuit also exhibits other neuron behaviors such as adjustable firing threshold and refractory period.

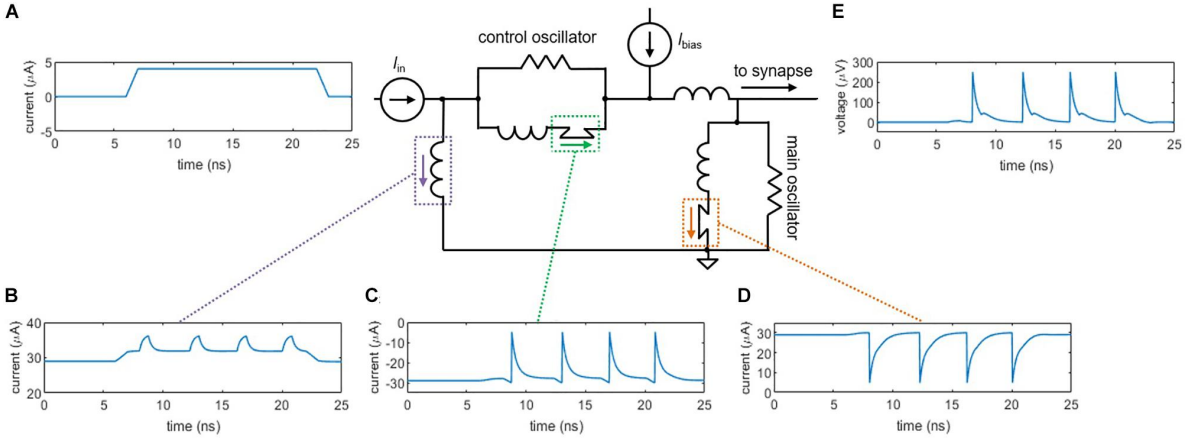


Figure 1.15: Simulation results of the nanowire neuron, from [8]. (A) Input current $I_{in} = 4\ \mu$A. (B) Current through the loop inductor. (C) Current through the control nanowire. (D) Current through the main nanowire. (E) Output voltage.

An inductive synapse is coupled with the nanowire neuron to provide excitatory or inhibitory outputs, as shown in Figure 1.16. The output from the nanowire neuron charges a large inductor L_{syn}, resulting in a slow discharging current through downstream neurons. The

excitatory or inhibitory effects are generally based on the sign of the bias current applied to the upstream neuron (the neuron shown in Figure 1.16). The synapse strength can be modulated by adding a high inductance nanowire with an ideal current source in parallel with the synapse inductor, as shown in Figure 1.17. By injecting current through the nanowire, the kinetic inductance changes, reflecting a change of overall parallel inductance. Therefore, the output current from the synapse is dependent on the modulating current. The energy dissipation per spiking event for this design is as low as 0.05 fJ, while the synapse consumes energy of $\sim$ 0.005 fJ. This suggests that the nanowire-based neuromorphic hardware is a good candidate for low-power neuromorphic computing.

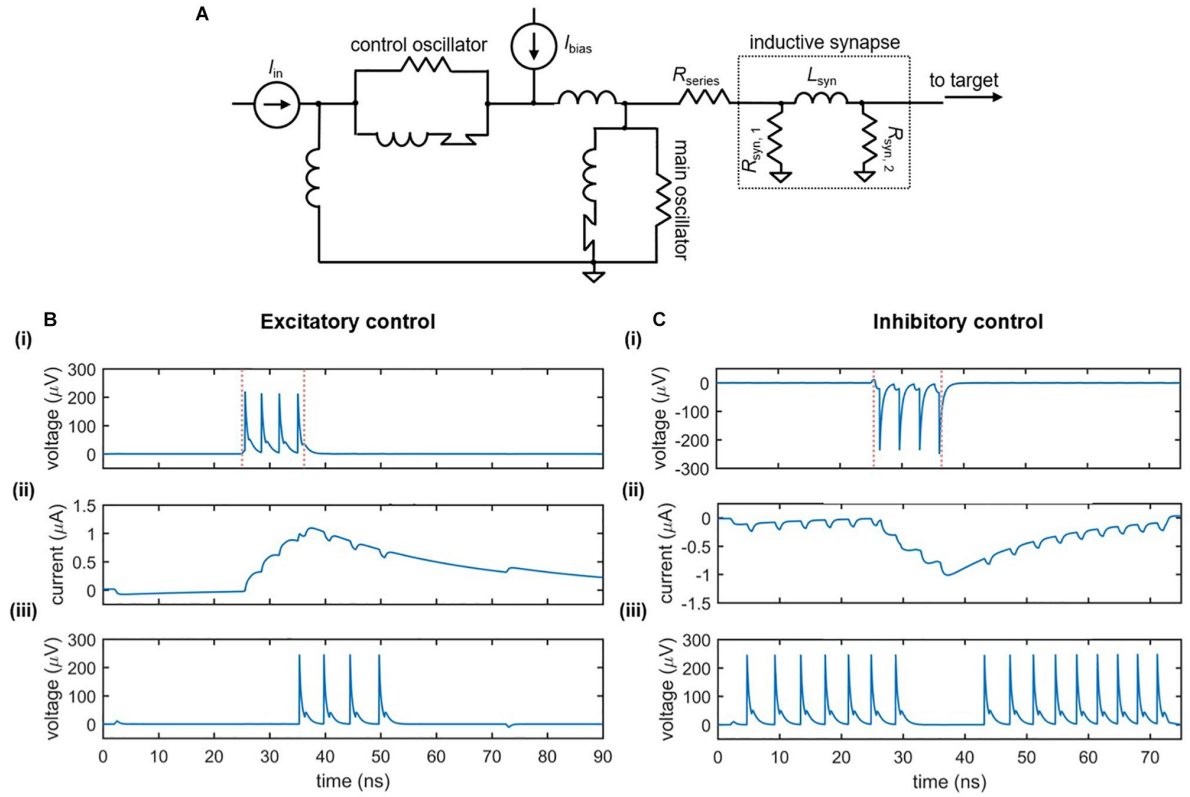


Figure 1.16: Nanowire neuron coupled with an inductive synapse, from [8]. (A) Circuit schematic. (B) Excitatory coupling. (C) Inhibitory coupling.

1.4 Outline of this dissertation

The dissertation is organized as follows. In Chapter 2, the theoretical background of phase-slip phenomena, including thermally activated phase-slip (TAPS) and quantum phase-slip (QPS),

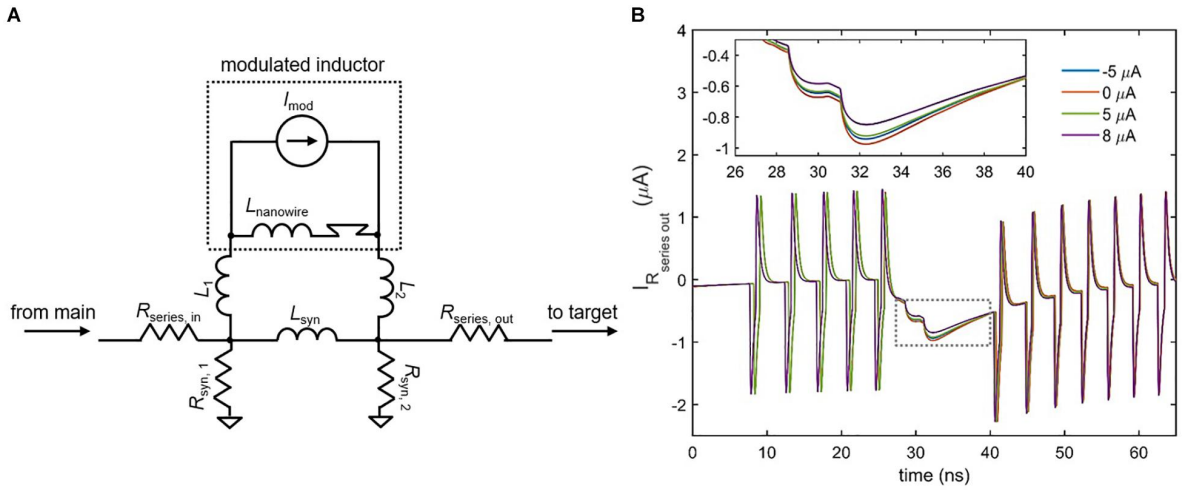


Figure 1.17: Synapse modulating, from [8]. (A) Circuit schematic of a high inductance nanowire with an ideal current source in parallel with the synapse inductor. (B) Simulation results of synaptic current as a function of different modulation currents. (Inset) Zoom-in view of the boxed area.

will be described in detail. A single QPSJ SPICE model will be introduced as the fundamental element for QPSJ-based neuromorphic circuits. In Chapter 3, QPSJ-based neuromorphic circuits that are used to emulate biological neuron cells will be introduced, and corresponding WRspice simulation results will be presented. The proposed neuromorphic circuits include spiking neuron circuits, an integrate-and-fire-neuron (IFN) circuit, synaptic circuits, a fan-out circuit and an axon circuit. In Chapter 4, we will focus on QPSJ-based neuromorphic circuits towards learning. A simplified STDP learning circuit consisting of an LTD circuit and an LTP circuit will be introduced to update synaptic weight based on relative timings between presynaptic and postsynaptic spikes. The proposed simplified STDP learning circuit is the starting point to provide on-chip learning functions, and corresponding simulation results will be presented. Chapter 5 describes the power dissipation and operation speed of QPSJ-based circuits. We provide estimates of the switching energy of a single QPSJ at its operating temperature and analyze power/energy dissipation in QPSJ-based circuits. Then we discuss the operating temperature, experimental challenges, scalability and tolerance of QPSJ-based circuits. In Chapter 6, the method of fabricating NbN nanowires will be introduced, and the experimental results of NbN thin films and NbN nanowires will be presented. We will show the characterization results of NbN thin films and the characterization results of NbN nanowires of different geometries on

different substrates. Chapter 7 is the final chapter that summarizes the work presented in this dissertation and proposes future works.

Chapter 2

Quantum Phase-slip Phenomena

A superconducting wire whose transverse dimensions are $\lesssim \xi$, where ξ is the superconducting coherence length, is generally viewed as a quasi-1D system. In quasi-1D systems, as the cross-section of a superconductor reduces, finite resistance can be observed at temperatures below $T_{\rm C}$, where $T_{\rm C}$ is the critical temperature of the superconductor. The finite resistance is attributed to the fluctuation of superconducting order parameter, also known as phase-slips [97, 98]. The phase-slip phenomenon in quasi-1D superconducting nanowires is in great agreement with the Mermin-Wagner theorem, which prohibits phase transitions in 1D systems [99]. During a phase-slip, the order parameter fluctuates to zero in a short segment of the channel, allowing the phase difference between the ends to be 2π. Phase-slips that were observed at temperatures just below critical temperatures in superconducting nanowires [100, 101] are referred to as thermally activated phase-slips (TAPS). Quantum phase-slips are experimentally observed at temperatures much lower than $T_{\rm C}$ [102, 11, 103, 104, 51].

2.1 Thermally activated phase-slips

In TAPS, thermal fluctuations provide free energy that can suppress the order parameter locally. A theory known as LAMH theory was proposed by Langer and Ambegaokar [105], followed by McCumber and Halperin to explain TAPS in the 1960s. At the position where TAPS occurs, the order parameter $\psi = |\psi|e^{i\phi}$ is driven to zero and phase change is $d\phi = \pm 2\pi$. The free energy barrier of a TAPS event is defined by:

$$\Delta F(T) = (8\sqrt{2}/3)V_{\rm PS}H_{\rm C}^2(T)/8\pi \tag{2.1}$$

where V_{PS} is the volume across the phase-slip and $H_{\mathrm{C}}(T)$ is the critical field. $V_{\mathrm{PS}} = \xi(T)A$, where $\xi(T)$ is the Ginzburg-Laudau (GL) coherence length [106] and A is the cross-sectional area of the wire. The phase-slip attempt frequency is given by:

$$\Omega = \frac{8k_{\mathrm{B}}(T_{\mathrm{C}} - T)}{\pi\hbar}\frac{L}{\xi(T)}\sqrt{\frac{\Delta F}{k_{\mathrm{B}}T}} \tag{2.2}$$

where k_{B} is the Boltzmann constant, $\hbar$ is the Planck constant and L is the length of the wire. The attempt frequency is inversely proportional to the relaxation time $\tau_{\mathrm{GL}} = \pi\hbar/(8k_{\mathrm{B}}(T_{\mathrm{C}} - T))$ of GL theory. The TAPS rate is thus given by

$$\Gamma_{\mathrm{TAPS}} = \Omega e^{-\Delta F(T)/k_{\mathrm{B}}T} \tag{2.3}$$

With a non-zero bias current I, the voltage is defined by

$$V = \frac{\hbar\Omega}{e}e^{(-\Delta F/k_{\mathrm{B}}T)}\sinh(I/I_0) \tag{2.4}$$

where $I_0 = 4ek_{\mathrm{B}}T/h$. Differentiation of this equation with respect to bias current I results in differential resistance:

$$dV/dI = \frac{\hbar\Omega}{e}e^{(-\Delta F/k_{\mathrm{B}}T)}\cosh(I/I_0) \tag{2.5}$$

In the limit of low current $I \ll I_0$, the temperature-dependent resistance is given by:

$$R_{\mathrm{LAMH}}(T) = R_{\mathrm{q}}\frac{\hbar\Omega}{k_{\mathrm{B}}T}e^{(-\Delta F/k_{\mathrm{B}}T)} \tag{2.6}$$

where $R_{\mathrm{q}} = h/(2e)^2 = 6.5$ kΩ is the quantum resistance.

2.2 Quantum phase-slips

2.2.1 Macroscopic quantum tunneling model for quantum phase-slips

At temperatures much below T_{C}, thermal fluctuations are significantly suppressed, and phase-slips are resulted from quantum tunneling, as shown in Figure 2.1. This phenomenon was first

suggested by Mooij et al. [107] that an analogous quantum phase-slip process may exist, similar to macroscopic quantum tunneling (MQT) in Josephson junctions (JJs) [108, 109, 110]. Like TAPS, QPS can also suppress the order parameter over a length of ξ. QPS event is expected to exist in extremely narrow superconducting nanowires. QPS can be dissipative or non-dissipative, which corresponds to incoherent QPS or coherent QPS, respectively. Early experimental results for QPS from Giordano [55] suggested that finite resistance persisted at temperatures much below T_C for indium wires that were $\lesssim 50$ nm wide. In his model, the QPS rate is given by:

$$\Gamma_{\mathrm{QPS}} = B\Omega_{\mathrm{QPS}} e^{(-a\tau_{\mathrm{GL}}\Delta F_{\mathrm{QPS}})/h} \tag{2.7}$$

where $\Omega_{\mathrm{QPS}} = (L/\xi_{\mathrm{GL}})\sqrt{\Delta F_{\mathrm{QPS}}/k_{\mathrm{B}}T}$ is the quantum tunneling attempt rate, and a and B are fitting constant parameters. The temperature-dependent resistance as a result of QPS is given by:

$$R_{\mathrm{QPS}} = BR_{\mathrm{q}}\tau_{\mathrm{GL}}\Omega_{\mathrm{QPS}} e^{(-a\tau_{\mathrm{GL}}\Delta F_{\mathrm{QPS}})/h} \tag{2.8}$$

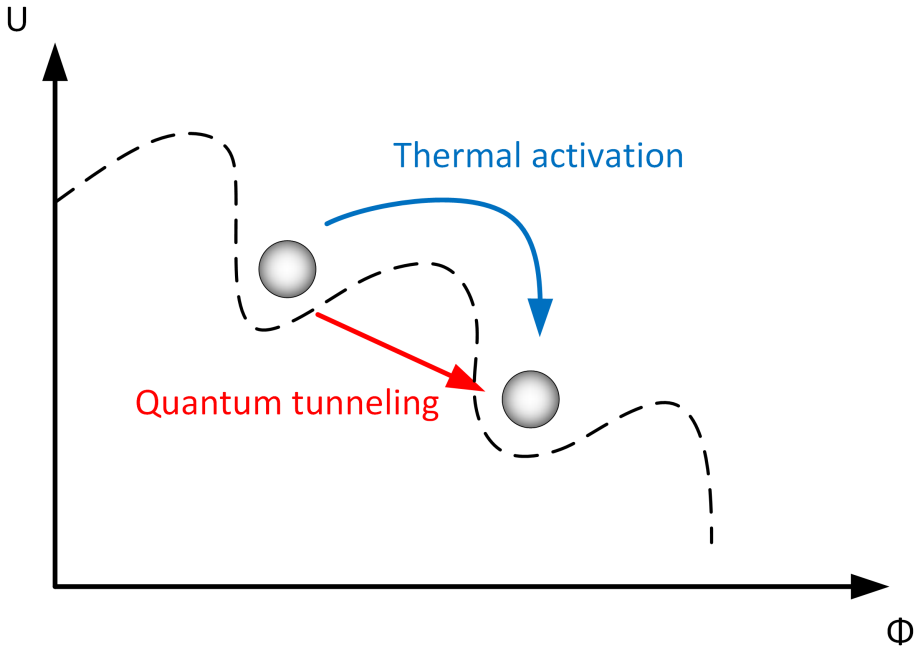


Figure 2.1: The particle can overcome the energy barrier through thermal activation or quantum tunneling (adapted from [9]).

2.2.2 Flux-charge duality

Charge and phase are dual quantum variables for superconductors. A coherent QPS is a dual process of Cooper-pair tunneling in a JJ, as demonstrated by Mooij [12]. Flux-charge duality based on Maxwell's equations can be explained by the context of planar lumped-element circuits [111, 112, 113, 114, 115, 116]. Figure 2.2 illustrates the duality transformation for planar, lumped-element circuits. The classical quantities can be interpreted as arising from a sum of free and bound current densities:

$$J_{\mathrm{Q}} = \rho_{\mathrm{Q}} v_{\mathrm{Q}} + \frac{dD}{dt} \tag{2.9}$$

$$J_{\Phi} = v_{\Phi} \times B_{\mathrm{f}} - \frac{dA}{dt} \tag{2.10}$$

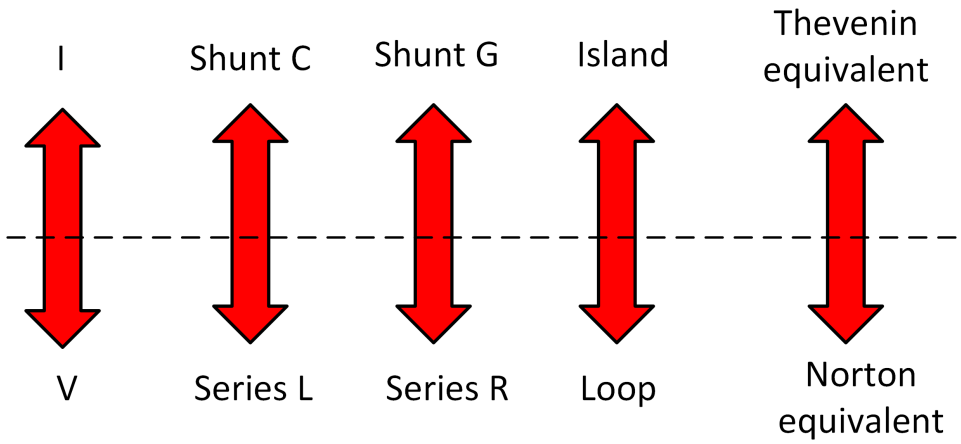


Figure 2.2: Duality transformation for planar, lumped-element circuits (adapted from [10]).

where J_{Q} is the current density corresponding to charge, ρ_{Q} is the ordinary density of free charge moving at velocity v_{Q}, D is the electric displacement, J_{Φ} is the current density corresponding to flux, B_{f} is the magnetic flux density moving at velocity v_{Φ} and A is the vector potential. Using equation $A = -\Lambda \rho_{\mathrm{Q}} v_{\mathrm{Q}}$ for a superconductor and equation $D = \epsilon E$ for an insulator, we obtain:

$$\Lambda \frac{dJ}{dt} = E \rightarrow L_{\mathrm{k}} \frac{d^2 Q(\Sigma)}{dt^2} = V_{\Gamma} \tag{2.11}$$

$$\epsilon \frac{dE}{dt} = J \rightarrow C \frac{d^2 \Phi(\Gamma)}{dt^2} = I_{\Sigma} \tag{2.12}$$

where V_Γ is the voltage difference between the two ends of Γ, I_Σ is the current flow through Σ, L_k is the kinetic inductance and C is the kinetic capacitance. Now we consider the duality of a JJ and a QPSJ. A QPSJ is typically an ultra-narrow nanowire that shows coherent QPS. A JJ has two superconducting layers separated by an insulating barrier, where Cooper pairs can tunnel through. A QPSJ can be viewed as two insulating layers separated by a superconducting barrier, where flux quanta (or fluxons) can tunnel through, as shown in Figure 2.3. The charge and flux quantities in a JJ and QPSJ can be defined by:

$$Q = n2e + C_\mathrm{J}V, \Phi = \frac{\Phi_0}{2\pi}\theta + \oint_\Gamma A \cdot d\Gamma = m\Phi_0 + L_\mathrm{J}I \tag{2.13}$$

$$\Phi = m\Phi_0 + L_\mathrm{k}I, Q = Q_\mathrm{f} + \int_\Sigma D \cdot d\Sigma = n2e + C_\mathrm{k}V \tag{2.14}$$

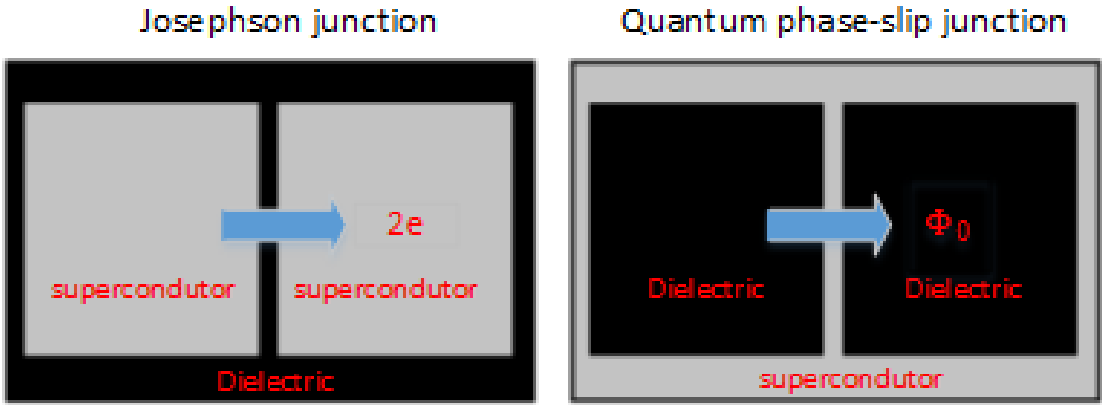


Figure 2.3: Duality between the charge tunneling in a JJ and the flux tunneling in a QPSJ (adapted from [11]).

For JJs, Q is the junction quasicharge [115, 116, 117, 118], n is the number of Cooper pairs that have passed through it and $C_\mathrm{J}V$ is the charge on the capacitor C_J induced by voltage V. The quantity Φ is due to the phase difference between two superconducting layers and magnetic fields inside the junction. For QPSJs, quantity Φ consists of the total bound flux of a nanowire having kinetic inductance L_k and fluxons that passed through the wire. The quasicharge Q is due to free charge Q_f that has passed through the wire and the electric fields on the wire.

2.2.3 QPSJ SPICE model

The operation of a single QPSJ can be illustrated through a simulation of the current-voltage (I-V) characteristics of a QPSJ using a compact equivalent SPICE model [12, 52]. The equivalent circuit is comprised of an intrinsic QPSJ, a series inductor and a series resistor, as shown in Figure 2.4. The I-V characteristics can be described by Equations 2.15 and 2.16, where L is the geometrical inductance, R is the normal resistance, and V_{C} is the critical voltage. The specific values of the model parameters are determined by the superconducting material and junction geometry.

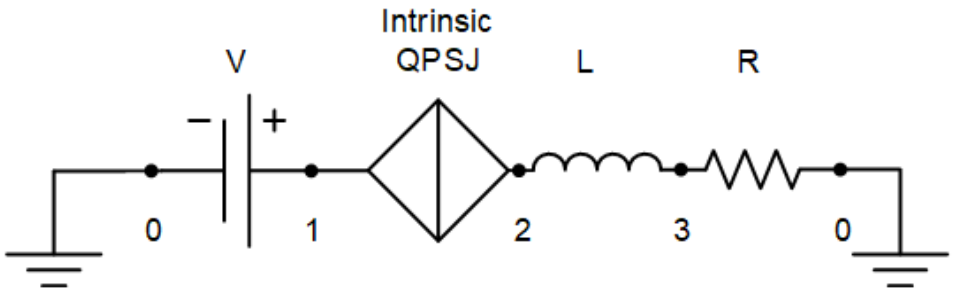


Figure 2.4: Voltage biased QPSJ in SPICE model. A QPSJ SPICE model comprises of an intrinsic QPSJ in series with an inductor and a resistor (adapted from [12]).

The equations that govern the QPSJ SPICE model are derived from:

$$V = V_{\mathrm{C}}\sin(q) + L\frac{dI}{dt} + RI \tag{2.15}$$

$$I = \frac{2e}{2\pi}\frac{dq}{dt} \tag{2.16}$$

where q is the charge equivalent in the QPSJ normalized to the charge of a Cooper pair $(2e)$. The critical voltage V_{C} is defined by:

$$V_{\mathrm{C}} = \frac{2\pi E_{\mathrm{S}}}{2e} \tag{2.17}$$

where E_{S} is the phase-slip energy, which can be calculated using the model by Mooij et al. [119] and e is the elementary charge of an electron. The normal resistance R is calculated from normal-state resistivity of the given material at operating temperature and the physical

dimensions of the nanowire that forms the QPSJ. The inductance L of the junction is related to inductive energy E_{L}, which is a function of normal resistance R and critical temperature of the material [119]. Simulated I-V characteristics of a single QPSJ are shown in Figure 2.5. The junction exhibits a Coulomb blockade [120, 51] when the voltage is below its critical voltage, and is resistive when the voltage is above the critical voltage. The parameters selected are from estimates to demonstrate a QPSJ and all its switching characteristics; the model may be tuned to represent a practically-realized QPSJ. These parameters depend on the specific material properties and device geometry for generation of quantum phase-slips [119].

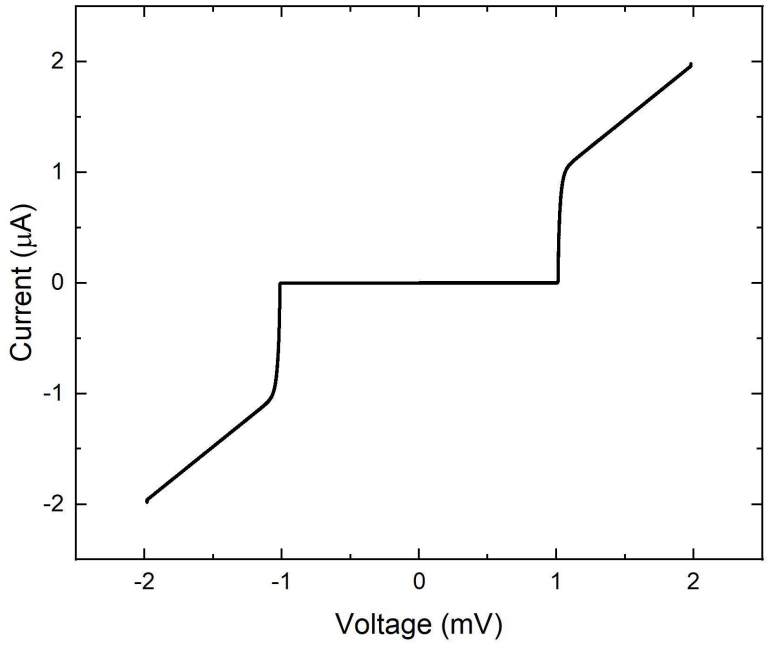


Figure 2.5: I-V characteristics of a single QPSJ with $V_{\mathrm{C}} = 1$ mV, $L = 20$ nH and $R = 1$ kΩ, from [13].

A QPSJ circuit can be designed and operated in an appropriate configuration to produce quantized-area current pulses corresponding to tunneling of a Cooper pair per switching event. The QPSJ can be treated as a series RLC oscillator with a damping parameter defined as:

$$\beta_{\mathrm{L}} = \frac{2\pi V_{\mathrm{C}} L}{2eR^2} \tag{2.18}$$

where β_{L} is the damping parameter, V_{C} is the critical voltage of the QPSJ, L is the geometrical inductance and R is the normal resistance. The QPSJ is overdamped if $\beta_{\mathrm{L}} \ll 1$ and underdamped if $\beta_{\mathrm{L}} \gg 1$. Therefore, by appropriately choosing material and junction (nanowire) geometry, the QPSJ can be made to be overdamped or underdamped [13]. This aspect can be verified by simulation of a QPSJ connected with voltage pulse input and appropriate DC bias voltage, as shown in Figure 2.6. A quantized current pulse with area equal to $2e$ can only be produced when the junction is overdamped, as shown in Figure 2.7. This operation is analogous to an overdamped JJ circuit producing a single voltage pulse with area corresponding to a single magnetic flux quantum. The overdamped QPSJ circuit can be achieved by either choosing a superconducting material with sufficient normal resistance or adding a series resistor to each junction to increase the value of R in Equation 2.18. In the underdamped case, the integrated area under the output pulse curve is more than $2e$ and thus not properly quantized, as shown in Figure 2.8.

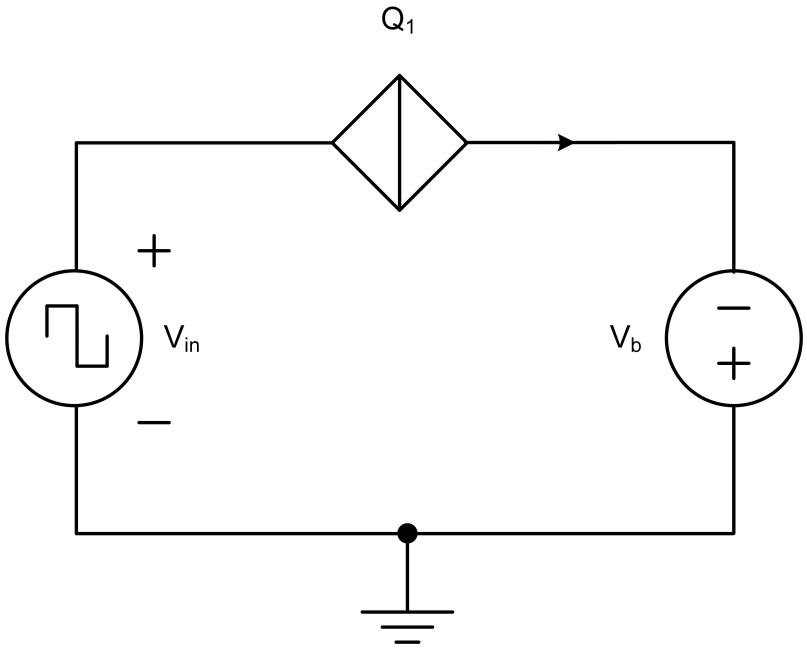


Figure 2.6: Single QPSJ circuit with voltage pulse input and DC voltage bias, from [13].

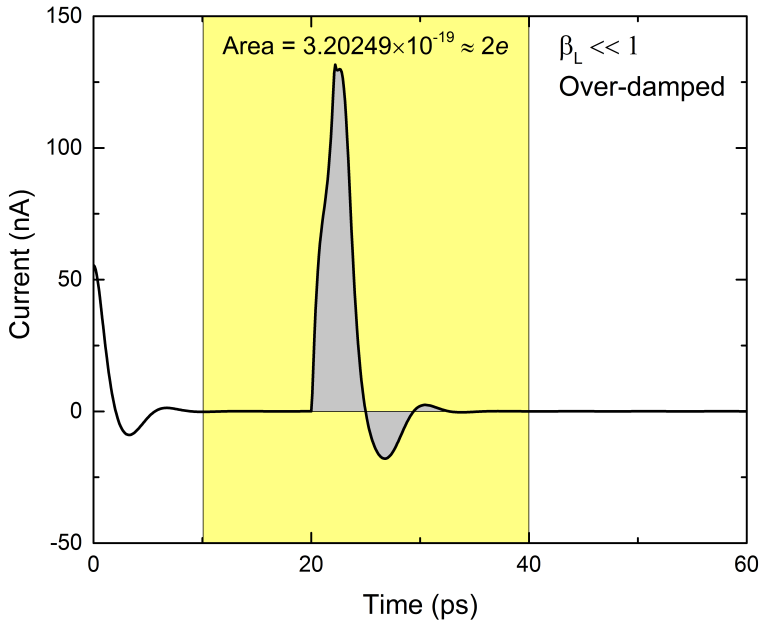


Figure 2.7: Simulation result of switching of a single QPSJ circuit with an overdamped junction, from [13]. The current shown is the current that flows through the QPSJ. The critical voltage of the junction Q_1 is 0.7 mV. Voltage bias V_b is 0.5 mV and magnitude of input pulse V_in is 1.1 mV. The normal resistance of Q_1 is 1 kΩ, and series resistance of Q_1 is 8 kΩ.

2.3 Summary

We presented the physics of thermally activated phase-slips and quantum phase-slips and the single QPSJ operations in this chapter. The superconducting order parameter can overcome the energy barrier through thermal activation (TAPS) or through quantum tunneling (QPS). A coherent QPSJ is a dual to a JJ, explained by charge-flux duality. The research on QPSJs are at an early stage, where materials, dimensions, and other critical conditions for fabricating reproducible QPSJs are still under investigation. Using the QPSJ SPICE model, we showed the nonlinear characteristics of QPSJs. The spiking behavior of overdamped QPSJs makes them a potential candidate for digital and neuromorphic computing. We made many efforts to design and simulate neuromorphic circuits using QPSJ technology. In the following chapters, we will show our neuromorphic circuit designs and simulations using QPSJs, JJs, and lumped elements, to emulate neuron spiking and learning behaviors and to potentially form a large scalable spiking neural network for high-speed and low-power neuromorphic computing.

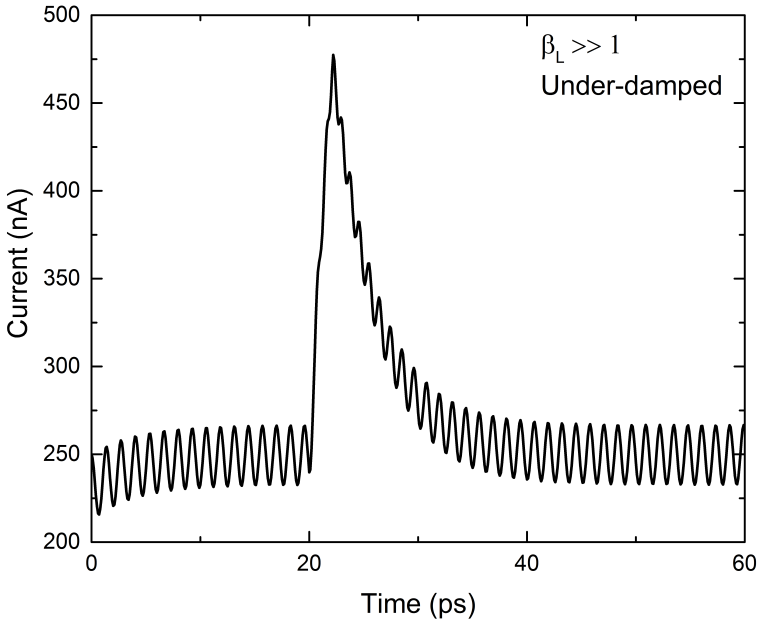


Figure 2.8: Simulation result of switching of a single QPSJ circuit with an under-damped junction, from [13]. The current shown is the current that flows through the QPSJ. The critical voltage of the junction Q_1 is 0.7 mV. Voltage bias V_b is 0.5 mV and magnitude of input pulse V_in is 1.1 mV. Normal resistance of Q_1 is 1 kΩ, and series resistance of Q_1 is 1 kΩ.

Chapter 3

Neuromorphic Circuits Design and Simulation

3.1 Introduction

A biological neural network is a cognitive system with self-learning functions and is mainly comprised of neurons, synapses, axons, dendrites, etc. [121]. A neuromorphic system that emulates a biological neural network also has multiple functional components. We use QPSJs, JJs, and other lumped elements to build circuits that function like biological neurons (somas), synapses, axons, dendrites, etc. In this chapter, we introduce QPSJ-based neuromorphic circuits that mimic biological neural networks to perform neuromorphic functions. These circuits are well-designed to be able to form a large scalable network through the operation of quantized charge and fluxons. The spiking signals, represented by quantized current pulses, are generated, processed, and transmitted among these circuits. The simulations were carried out in an SPICE-based platform (i.e., WRspice) using a QPSJ SPICE model with several tunable parameters.

3.2 QPSJ-based spiking neuron circuit

Neurons, as the primary components in a nervous system, can receive, process, and transmit information in the form of cellular signals. The membrane potential of an excitable neuron increases after it receives currents from other neurons or external stimuli. Once the membrane potential is beyond a threshold, it will rapidly rise and fall to a resting potential followed by a refractory period, during which the neuron does not respond to any stimulus. The resulting nerve impulse, which occurs in neuron cells and propagates along an axon towards synaptic boutons, is called action potential (AP) or "spike" [59]. Spiking neuron circuits that emulate

biological neuron functions play an important role in neuromorphic hardware systems [122]. The neuron circuits can process electrical signals and propagate them to other neuron circuits through synaptic devices. Depending on the system requirement, neuron circuits may have adjustable firing rate, threshold voltage, and refractory period.

One implementation of a QPSJ-based neuron circuit is comprised of two QPSJs and a small capacitor [53], forming a charge island, as shown in Figure 3.1. The QPSJ neuron conducts individual Cooper pairs that tunnel across the phase-slip center in the superconducting nanowire [13]. When phase-slip occurs in both the junctions, the node 1 between both the QPSJs is isolated from the rest of the circuit acting as an island that can hold charge up to $Q = CV_{\mathrm{C}}$. Both the junctions Q_1 and Q_2 are biased by a DC voltage V_{b} such that the voltage across each junction does not exceed the critical voltage V_{C} of either junction. Bias voltage V_{b} typically provides approximately 70% of the critical voltage for each junction and therefore is 1.4 V_{C}. The input voltage V_{in} is a pulse signal that can drive the junction Q_1 above its critical voltage V_{C}. The capacitor functions as a membrane capacitor in a neuron circuit. If the capacitance $C < 2e/V_{\mathrm{C}}$, the capacitor cannot hold the charge generated by exciting Q_1 above its critical voltage, and therefore immediately switches the junction Q_2. But if the capacitance $C > 2e/V_{\mathrm{C}}$, then the island does not trigger Q_2 when pulses arrives until $C < Q/V_{\mathrm{C}}$, after which, it fires multiple pulses as quickly as the QPSJs can switch. Here, Q is the total charge accumulated on the capacitor. Depending on the value of the capacitor, the QPSJ neuron circuit emulates a biological neuron response by exhibiting tonic spiking and tonic bursting [123]. If Q_1 and Q_2 are identical, the critical voltage of both the junctions is V_{C}, the magnitude of V_{in} is V_a and bias voltage is V_{b}, the circuit works only if $V_a + V_{\mathrm{b}} > 2V_{\mathrm{C}}$. The circuit operates in a tonic spiking mode if $C < 2e/V_{\mathrm{C}}$ as in Figure 3.2, and it operates in a tonic bursting mode if $C > 2e/V_{\mathrm{C}}$ as in Figure 3.4. In each of these figures, the panels show input voltage pulses, output current pulses and voltage across the capacitor, for top, middle and bottom, respectively. Figure 3.3 shows a zoom-in view of the input voltage and output current shown in Figure 3.2. A zoom-in view of these two types of pulses is shown in Figure 3.5. These responses, which resemble certain types of neuron firing responses, are described in more detail in the following paragraphs.

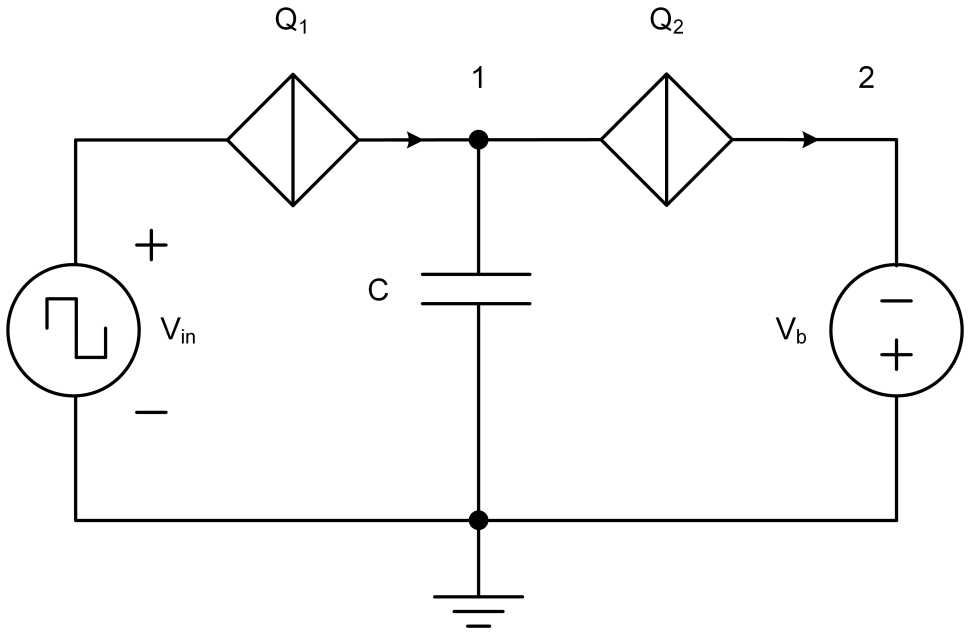


Figure 3.1: The QPSJ circuit comprised of two QPSJs and one capacitor, along with corresponding voltage pulse input and DC voltage bias, from [13].

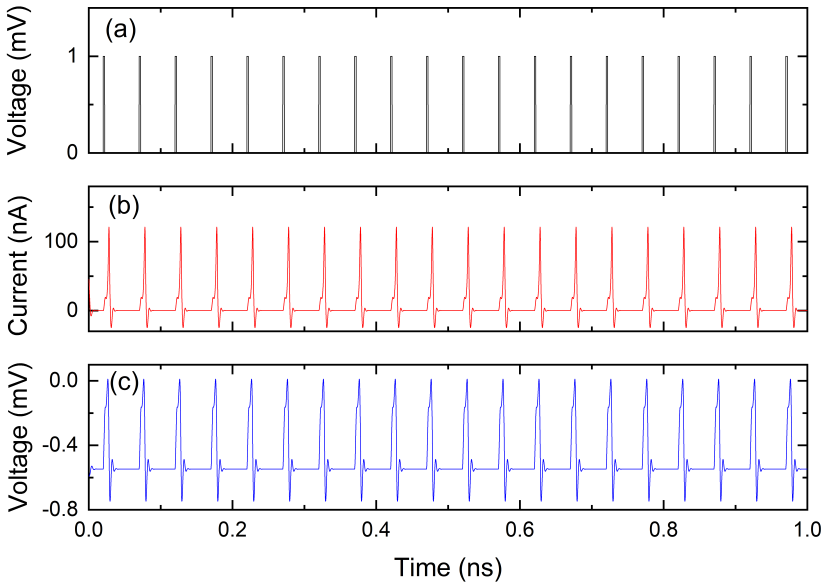


Figure 3.2: Simulation results of a QPSJ spiking neuron circuit operated in tonic spiking mode, from [13]. The critical voltage of junctions Q_1 and Q_2 is 0.7 mV. Capacitance $C = 0.23$ fF, voltage bias $V_\mathrm{b} = 1$ mV and magnitude of input pulses $V_\mathrm{in} = 1$ mV. (a) Input voltage pulses. (b) Output current. (c) Capacitor voltage at node 1.

In Figure 3.2, junction Q_2 is switched immediately and fires a single pulse after a Cooper pair is propagated to the capacitor C, since the capacitor cannot hold the charge. In Figure 3.4, the voltage at node 1 increases by an amount $2e/C$ (64 μV) upon the arrival of each

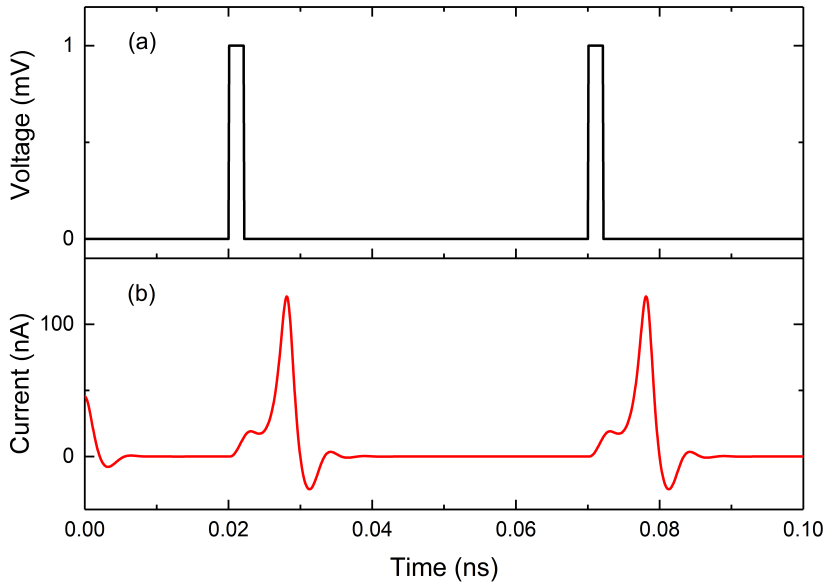


Figure 3.3: A zoom-in view of the input voltage and output current shown in Figure 3.2. (a) Input voltage. (b) Output current.

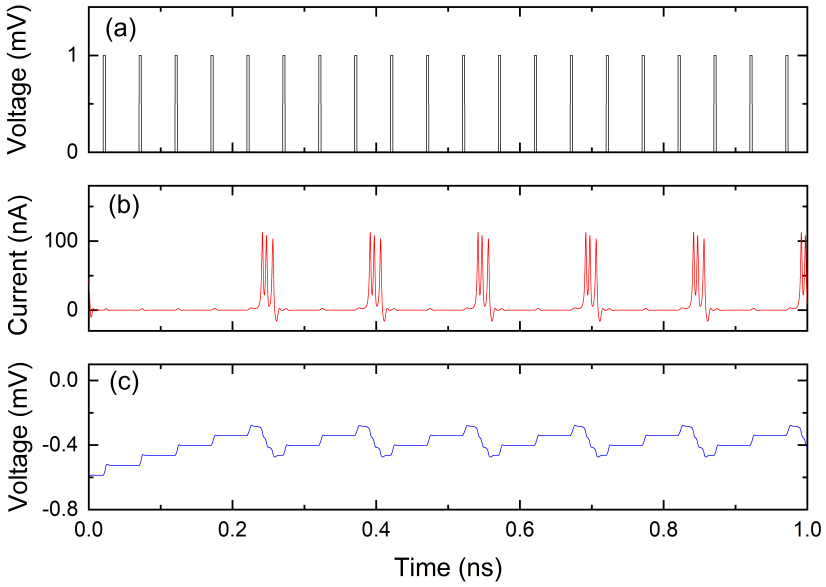


Figure 3.4: Simulation results of a QPSJ spiking neuron circuit operated in tonic bursting mode, from [13]. The critical voltage of junctions Q_1 and Q_2 is 0.7 mV. Capacitance $C = 5$ fF, voltage bias $V_{\rm b} = 1$ mV and magnitude of input pulses $V_{\rm in} = 1$ mV. (a) Input voltage pulses. (b) Output current. (c) Capacitor voltage at node 1.

input voltage pulse. The 5 fF capacitor can hold a charge up to $10e$ before the voltage is large enough to switch junction Q_2. Thus the capacitance defines the firing threshold for this neuron

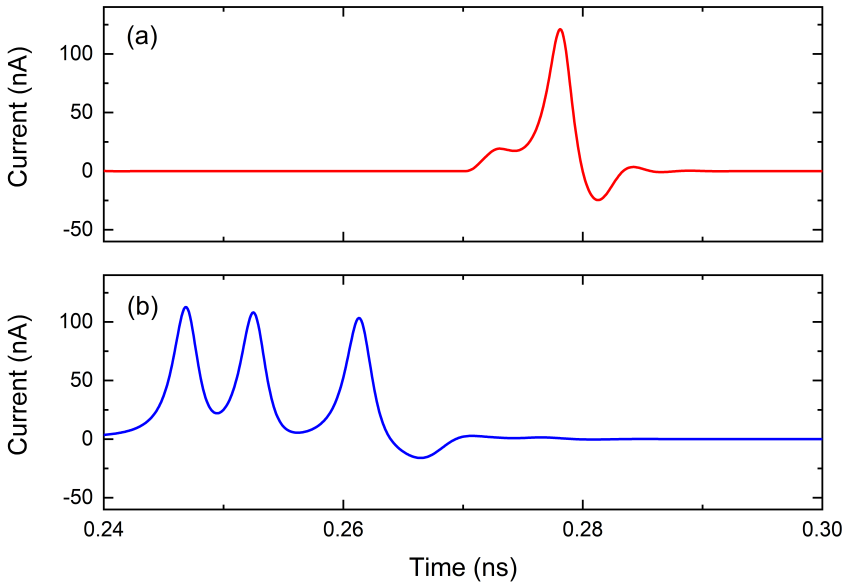


Figure 3.5: A zoom-in view of the output current pulses shown in Figure 3.2 and Figure 3.4, from [13]. (a) Tonic spiking pulse from Figure 3.2. (b) Tonic bursting pulses from Figure 3.4.

circuit. After junction Q_2 is switched, three subsequent current pulses that contain $6e$ are generated, which resembles a tonic bursting behavior in a cortical neuron [123]. The membrane voltage gradually drops three steps but is not fully reset since the membrane capacitor does not discharge all the charges it stores. It is also different from a biological neuron, which has a refractory period after a firing event during which the neuron does not respond to any stimulus. By adjusting the resistance in Equation 2.18, the damping condition can be changed and therefore the number of electrons discharged each time changes. This technique can be used to adjust the firing period of a tonic bursting neuron as shown in Figure 3.6.

In Figure 3.6, the firing periods and pulse shapes change according to different values of the resistances in series with each QPSJ (or, alternatively, with different values of QPSJ normal resistance). When the series resistance is small, junction Q_2 is heavily under-damped and the capacitor is fully discharged; when the resistance is large, junction Q_2 is heavily overdamped and the capacitor can only discharge one pair of electrons per switching event. With appropriate choice of capacitor and resistor, the circuit operates similar to a digital integrate and fire neuron (IFN) that has a pre-defined firing threshold, as described in the next section. These

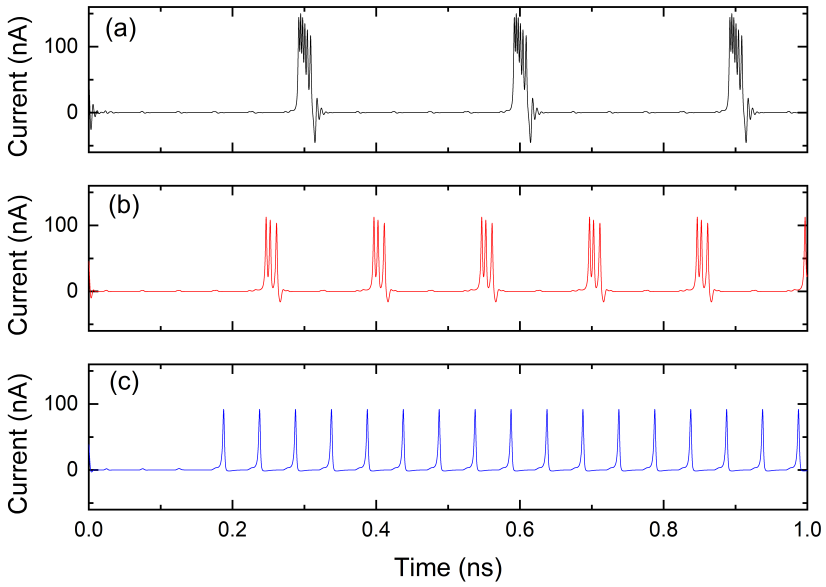


Figure 3.6: Simulation results of a QPSJ-based tonic bursting neuron with different firing frequencies, from [13]. The critical voltage of junctions Q_1 and Q_2 is 0.7 mV. Capacitance $C = 5$ fF, voltage bias $V_{\text{b}} = 1$ mV, normal resistance of Q_1 and Q_2 is 1 kΩ, series resistance of Q_1 is 12 kΩ, and magnitude of input pulses $V_{\text{in}} = 1$ mV. (a) Series resistance of Q_2 is 4 kΩ. (b) Series resistance of Q_2 is 8 kΩ. (c) Series resistance of Q_2 is 12 kΩ.

responses demonstrate that QPSJ-based circuits hold promise for realizing certain functions of neuromorphic circuits.

3.3 Integrate and fire neuron circuit

In addition to have a neuron circuit that can generate spikes to mimic biological neuron dynamics, computational neuron circuit is also an essential component for neuromorphic computing hardware. One of the simplest and most widely used neuron models for computation is the integrate and fire neuron model [124]. An example of IFN models is the leaky integrate and fire (LIF) neuron, which can be defined by:

$$\frac{dV}{dt} = I + a - bV \tag{3.1}$$

$$V = c, V \geq V_{\text{th}} \tag{3.2}$$

where V is the membrane potential, I is the input current, V_{th} is threshold voltage and a, b and c are relevant parameters. This model includes a leaky path that allows voltage to slowly recover to the resting potential. The integrate and fire neuron circuit we propose here does not have a leaky path, although it can be easily realized by adding a dissipative path. The integrate and fire neuron integrates input signals and fires an action potential once the membrane voltage reaches a threshold value. The circuit shown in Figure 3.7, which is based on the circuit in Figure 3.1, restricts one pulse for each firing event to imitate the behavior of an IFN circuit. This circuit has a flexible configuration with tunable threshold and firing frequency. The capacitor C, in Figure 3.7, is chosen such that it can store a charge $Q = 2Ne$ and all the junctions are nominally identical and overdamped. The value of C is given by:

$$C = \frac{2Ne}{V_{\text{C}} - V_{\text{b}} - V_1} \tag{3.3}$$

where V_1 is the initial voltage at node 1 and V_{C} is the critical voltage of all junctions. The number of parallel QPSJs N determines the threshold for the firing. The equivalent resistance of multiple parallel QPSJs is much smaller than the normal resistance of Q_0, especially when N is large. Therefore, bias voltage V_b cannot provide equal bias voltage for all QPSJs without using a bias resistor. The bias resistor R_{b} is thus used to balance the bias voltages of Q_0 and multiple parallel QPSJs, i.e. Q_1 to Q_N. Similar to the previous design in Figure 3.1, the first firing event occurs when the voltage at the capacitor C is large enough to switch junctions Q_1 to Q_N. A pair of electrons are then propagated through each parallel junction where a current pulse is generated. Meanwhile, the voltage at the capacitor gradually drops N steps with each step equal to $2e/C$.

Figure 3.8 shows the simulation results of a QPSJ neuron circuit with threshold N equal to 10. Input rectangular voltage pulses have a period of 120 ps and a width of 3 ps. The voltage at capacitor C keeps increasing as a result of quantized charge accumulation. Once the voltage applied on the parallel QPSJs reaches the critical voltage, the capacitor immediately discharges $20e$ charge. Each $2e$ charge is transported through a parallel QPSJ, resulting in a spike-like current pulse. We can also observe that the magnitude of current pulse through Q_0 decreases

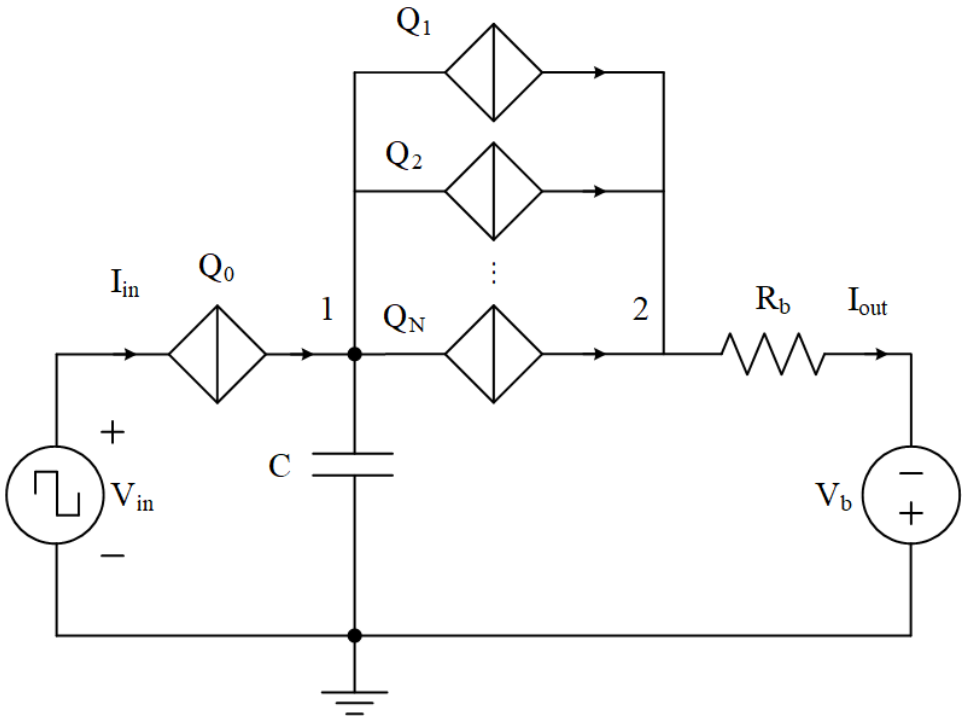


Figure 3.7: A QPSJ-based integrate and fire neuron circuit with integration capacitor C and multiple parallel output QPSJs, from [13].

periodically, because the voltage across Q_0 decreases when the voltage at the capacitor keeps increasing. The input voltage pulse width and magnitude must be appropriately selected so that a single quantized current pulse can be generated during the voltage integration process. Otherwise, multiple pulses might be generated after one input voltage pulse.

In Figure 3.9, the output current is the current that flows into node 2 and has a $2Ne$ area under each pulse. Since junctions Q_1 to Q_N are identical and parallel, they are switched at the moment when voltage between node 1 and 2 is above the critical voltage so that a $2e$ current pulse is generated at each branch and sums up at node 2. All the simulations were performed with a 50 ps periodic rectangular pulse input and therefore the neuron firing period is N times the source period. If the input is a random pulse sequence, the neuron circuit starts to generate output pulses as soon as it receives N pulses. The circuit only works when the capacitor is large enough, otherwise quantized charge-based current cannot be generated in each parallel QPSJ branch.

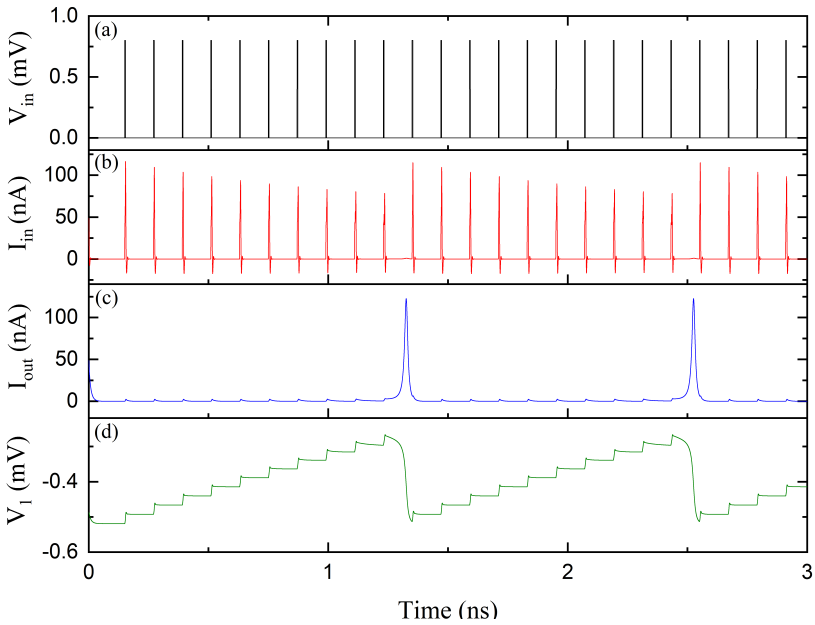


Figure 3.8: Simulation results of an IFN circuit with $N = 10$, from [13]. (a) Input voltage. (b) Input current. (c) Output current. (d) Capacitor voltage at node 1.

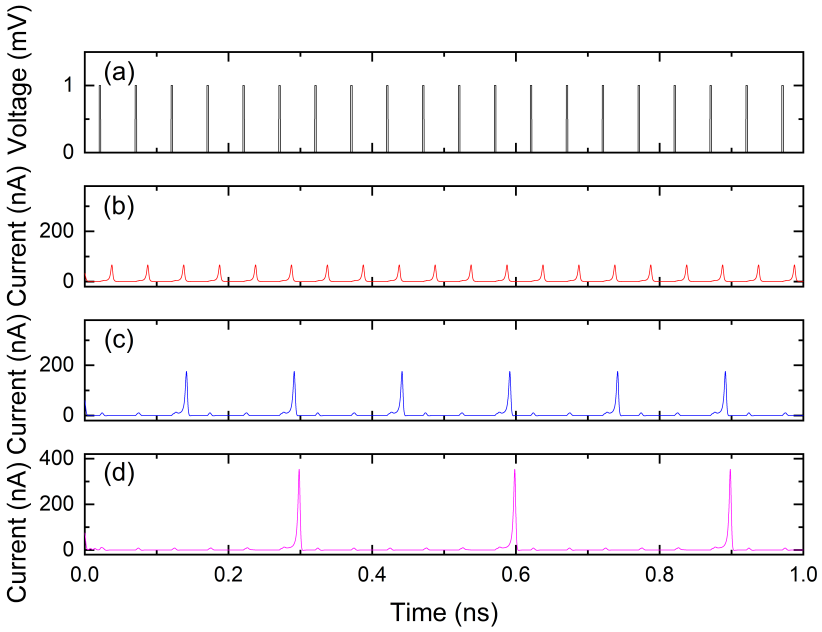


Figure 3.9: Simulation results of IFN circuit showing output currents for different N values, from [13]. (a) Input voltage pulses. Output current when (b) $N = 1$, (c) $N = 3$, and (d) $N = 6$.

3.4 Synaptic circuits

A synapse connects two neurons in a brain and determines the signal strength transmitted from a presynaptic neuron to a postsynaptic neuron. Similarly, a synaptic circuit should be able to

adjust the connection strength between two neuron circuits. In CMOS neuromorphic systems, non-volatile memory cells are usually used to implement synaptic circuits [125]. A lack of non-volatile superconducting devices/circuits made superconducting neuromorphic implementation more challenging until the recent realization of magnetic Josephson junctions (MJJs) for this purpose [45, 46, 126, 127]. An MJJ has a tunable critical current that can control the switching threshold [43] to function as a binary synapse or control the circulating current in a superconducting loop to function as an analog synapse [44]. Although a corresponding tunable critical voltage of a QPSJ has not yet been theoretically or experimentally demonstrated, we were able to combine MJJs and QPSJs to realize synaptic functions. To overcome the MJJ programming challenges, we also introduce new synaptic circuits that are essentially non-destructive readout (NDRO) memories based on QPSJs. These synaptic circuits can be configured as binary synapses and multi-weight synapses, which work seamlessly with the neuron circuit to transmit a weighted signal between two neurons.

3.4.1 Synaptic circuits based on QPSJs and MJJs

A simple binary synaptic circuit using an MJJ and a QPSJ is shown in Figure 3.10 [14]. This circuit includes an MJJ to store the synaptic weight information. Initially, MJJ J_1 is biased by a DC current I_b while QPSJ Q_1 is biased by V_b. A short voltage pulse from input V_in, through R_1 and L_1, provides a current pulse for J_1. Inductor L_1, which can be omitted in this circuit operation, is used to trap an SFQ pulse from a previous stage if there are multiple JJ stages. If J_1 is designed to have two distinct critical currents controlled by the magnetic order parameters [42], the junction can be either switched or not upon the arrival of an input current pulse. For instance, in the case of low critical current of J_1, the input current pulse is large enough to switch J_1. An SFQ pulse is generated at node 1, which can, in turn, switch Q_1, resulting in a current pulse at the output. However, if the critical current of J_1 is high, the total current injected to J_1 is insufficient to switch it. Voltage at node 1 stays at zero and there will be no quantized current pulses at the output.

The full operation of this binary synaptic circuit is illustrated in the simulation results shown in Figure 3.11, where the critical current of J_1 is either 200 μA or 300 μA, to represent

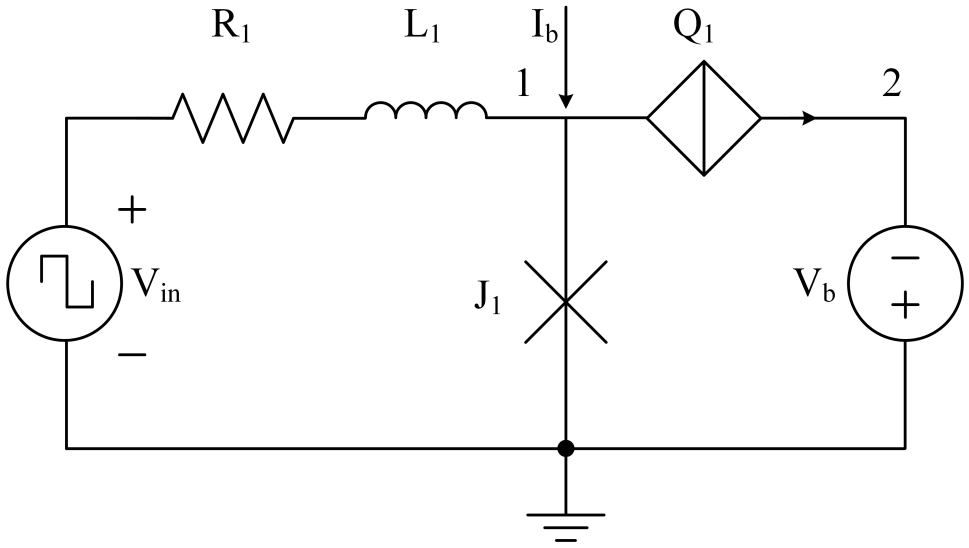


Figure 3.10: A binary synaptic circuit based on a superconducting MJJ (J_1) and QPSJ (Q_1), from [14].

a weight of 1 or 0. In Figure 3.11(a), an input current pulse is generated by an input voltage pulse from V_{in}. The current pulse is then injected into J_1, which results in an SFQ pulse across the junction. The SFQ pulse can switch Q_1 to produce a quantized current pulse at output. However, if we increase the critical current value of J_1 to 300 μA, the current from input is insufficient to switch J_1 and therefore no SFQ pulses and quantized current pulses are generated.

In the neuromorphic system described here, signals are generated and propagated in the form of quantized charge current pulses. The multi-state synaptic circuit shown in Figure 3.12 can generate multiple current pulses that contain a charge of $2Ne$, depending on the state of MJJ J_2. Josephson junction J_1 is biased by I_{b} and switched after an input current pulse from V_{in}. Inductor L_2 is chosen to store a single flux quantum, which in turn switches MJJ J_2 after a short delay. The resulting voltage pulse at node 2 switches Q_1 and generates current pulses at the output. Since the magnitude of SFQ pulse at node 2 is proportional to the critical current I_{C} of J_2, the width of this SFQ pulse is inversely proportional to I_{C}. As a result, more current pulses are generated when I_{C} is low because the switching speed of Q_1 is fast (i.e. the normal resistance of Q_1 is relatively low), which allows Q_1 to be switched multiple times within a short time.

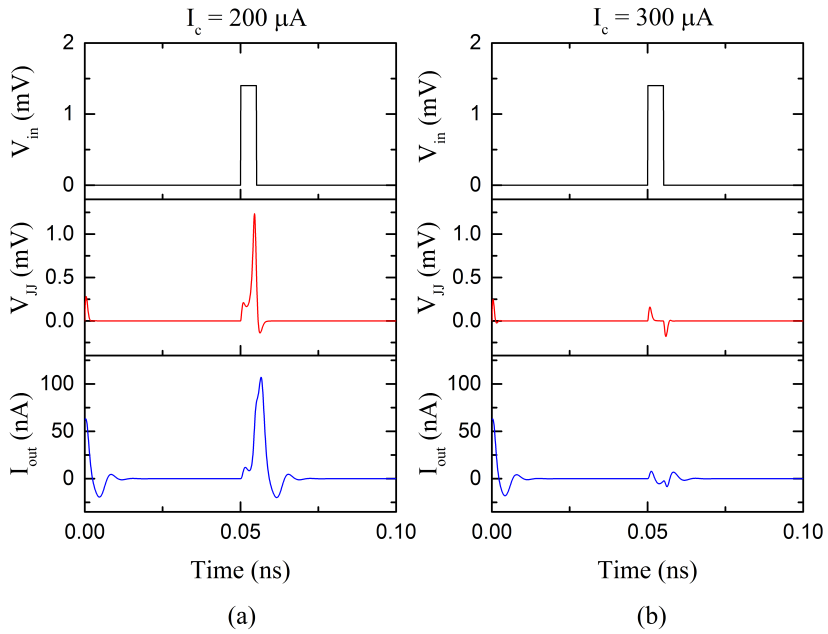


Figure 3.11: Simulation results of input voltage, voltage at J_1, and output current in a binary synaptic circuit with different J_1 critical current I_C, from [14]. Critical voltage of Q_1 is 0.7 mV, bias current I_b is 140 μA, and magnitude of input pulse V_{in} is 1.4 mV. (a) I_C is 200 μA. (b) I_C is 300 μA.

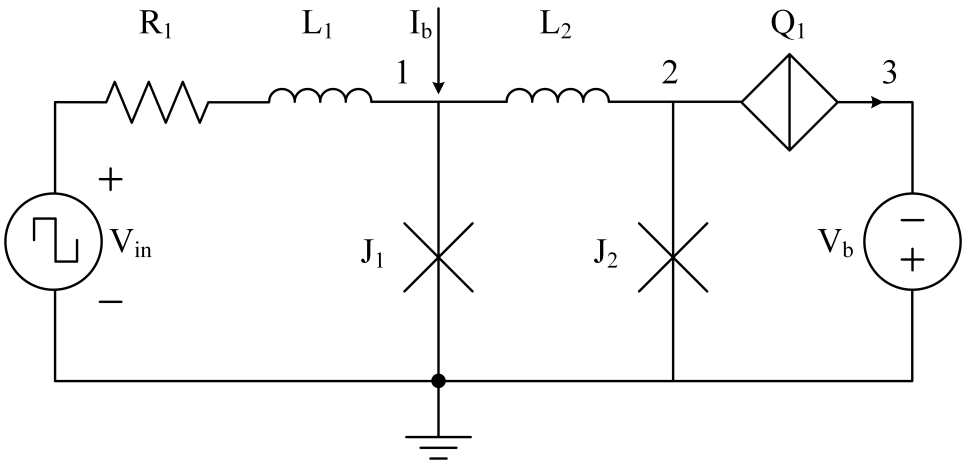


Figure 3.12: A multi-state synaptic circuit based on a superconducting JJ (J_1), MJJ (J_2), and QPSJ (Q_1), from [14].

The simulation results in Figure 3.13 show that different numbers of current pulses are generated when I_C varies from 10 μA to 400 μA. When $I_C = 10$ μA shown in Figure 3.13(a), each input voltage pulse from V_{in} can cause two sequential pulses at output. As I_C increases, the synaptic weight decreases. No output current pulses can be generated once I_C is more than

400 μA. Although the change is not linear, this can be used to represent a synaptic weight of 2, 1, 0.5, or 0, according to the number of current pulses generated within a set time interval.

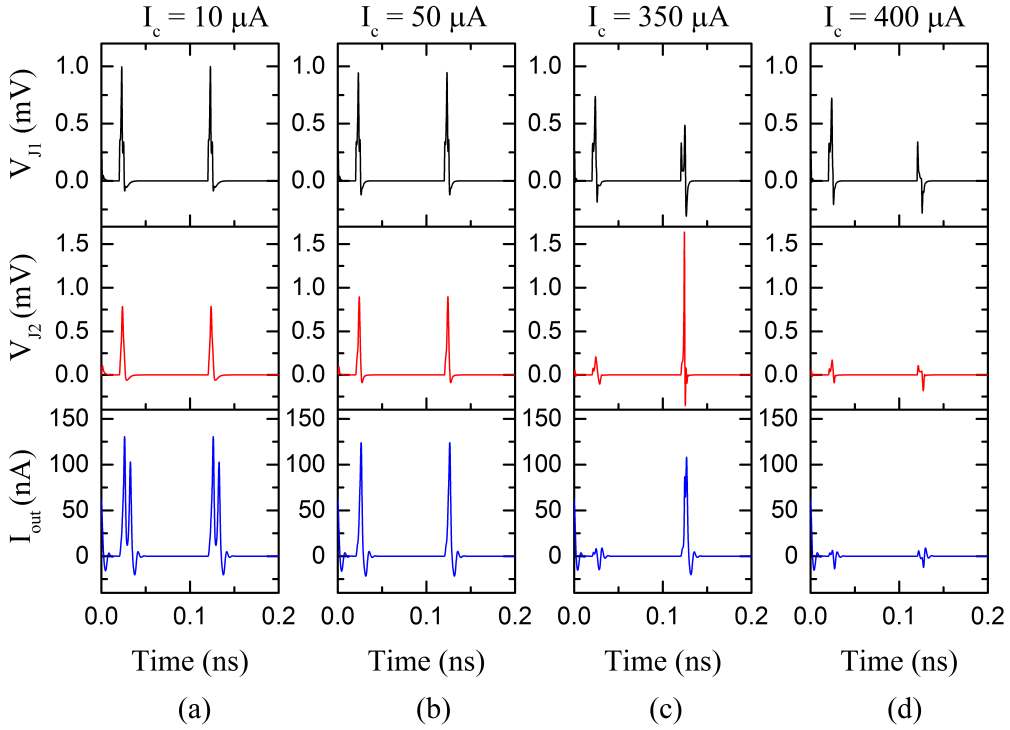


Figure 3.13: Simulation results of voltage at J_1, voltage at J_2, and output current in a multi-state synaptic circuit with different J_1 critical current I_C, from [14]. Critical voltage of Q_1 is 0.7 mV, critical current of J_2 is 200 μA, and bias current I_b is 160 μA. (a) I_C is 10 μA. (b) I_C is 50 μA. (c) I_C is 350 μA. (d) I_C is 400 μA.

3.4.2 Synaptic circuits based on QPSJ memories

The tunable critical current of an MJJ, incorporated into a circuit with QPSJs, made it possible to realize either binary or multi-state memory. However, programming, or changing the state, of an MJJ requires not only a larger number of energy in comparison to the traditional JJ operation but also a magnetic field (global or potentially local) to polarize the magnetic materials within the junction [42, 44, 6, 95, 96]. An alternative QPSJ-based approach could potentially be more energy efficient than MJJs and could potentially allow on-chip learning.

We introduce a non-destructive readout (NDRO) memory circuit based on QPSJs, as shown in Figure 3.14, to overcome the programming problem described above [15]. Q_2 and Q_3 are identical while the other QPSJs do not have to be identical. For simplicity, we used

identical parameters for Q_0, Q_1 and Q_4 in our simulations. Read signal $V_{\rm r}$ and write signal $V_{\rm w}$ were rectangular pulses with a duration of 5 ps and a magnitude of 0.6 mV.

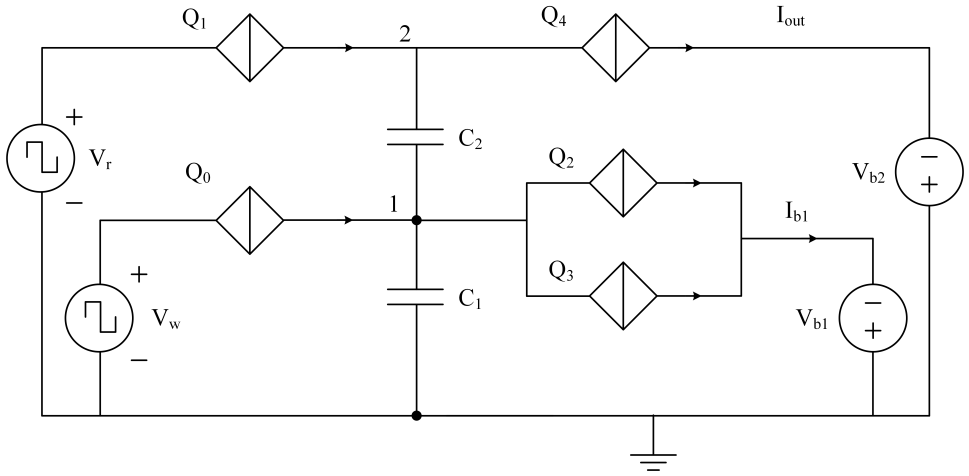


Figure 3.14: A non-destructive readout (NDRO) memory circuit based on QPSJs as a synaptic circuit, from [15]. Specific device parameters and bias values are included with simulation result plots.

The writing process for this circuit can be described as follows. For the case where there is initially no charge on capacitor C_1 or C_2. A short voltage pulse from $V_{\rm w}$ can switch Q_0, resulting in a quantized current pulse that contains a charge of $2e$. The $2e$ charge is then moved onto C_1. This is a "set" process and the weight is set to 1. One more pulse from $V_{\rm w}$ moves another $2e$ charge onto C_1. At that point, the voltage across Q_2 and Q_3 exceeds their critical voltages, which switches Q_2 and Q_3. The charge on C_1 is thus released to the bias voltage supply $V_{\rm b1}$. This is a "reset" process and the weight is reset to 0. The operation that comprises a "set" and a "reset" process is a complete writing cycle of the QPSJ-based memory cell, as shown in Figure 3.15.

Every time Q_1 is switched by a pulse from $V_{\rm r}$, a $2e$ charge flows onto C_2, which induces another $2e$ charge on C_1. So the voltage change on node 2 is the summation of voltage change on both of the two capacitors. If there is no charge on C_1, Q_1 can only be switched once because the potential on node 2 becomes too high for another incoming voltage pulse from $V_{\rm r}$ to switch Q_1 again. As a result, there is no output at $I_{\rm out}$. However, if there is a charge on C_1 while no charge on C_2, the initial voltages on node 1 and node 2 are slightly higher. Once Q_1 is switched, the voltage on node 2 increases by $2e/C_1 + 2e/C_2$ and is high enough to switch Q_4

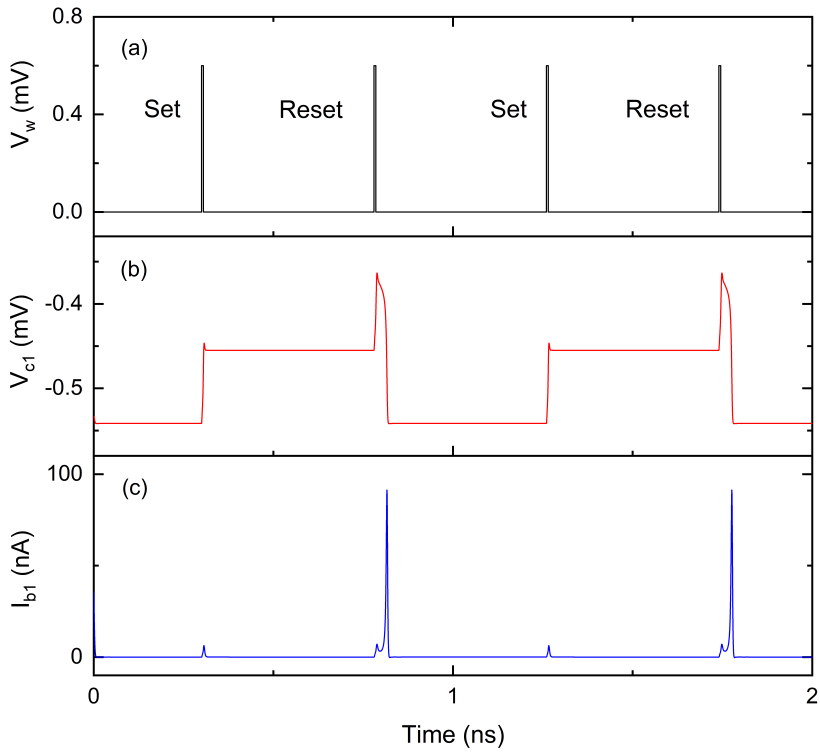


Figure 3.15: Simulation results of set and reset operation of the synaptic circuit shown in Figure 3.14, from [15]. The critical voltage values used for Q_0, Q_1 and Q_4 were 0.7 mV and for Q_2 and Q_3 were 0.4 mV. The capacitance values used for C_1 and C_2 were 3 fF. Bias voltage V_{b1} was 0.8 mV and V_{b2} was 1 mV. The magnitude of V_{r} and V_{w} was 0.6 mV. (a) Write pulse signal. (b) Voltage on capacitor C_1. (c) Current through V_{b1}.

before Q_2 and Q_3 are switched. Therefore, every read signal from V_{r} can trigger a current pulse at I_{out}. In Figure 3.16, read signals cannot trigger any current pulses at the output initially. After a set signal, a current pulse is immediately generated at the output, indicating that the synapse is set to 1. The upcoming read pulse is able to read the state 1 in terms of a current pulse at the output until the synapse is reset to 0 by a reset pulse from V_{w}.

We present a multi-weight synaptic circuit that is a more flexible synaptic circuit compared to the binary version. For the case of spiking neuromorphic systems being described here, this multi-weight synaptic circuit can generate multiple pulses for each triggering event. Generally speaking, a multi-weight synaptic circuit can store more weight information rather than just 0 or 1. Therefore, the weight between two neurons can be more accurate, potentially providing more complex computational capacity. The circuit shown in Figure 3.17 can store up to three

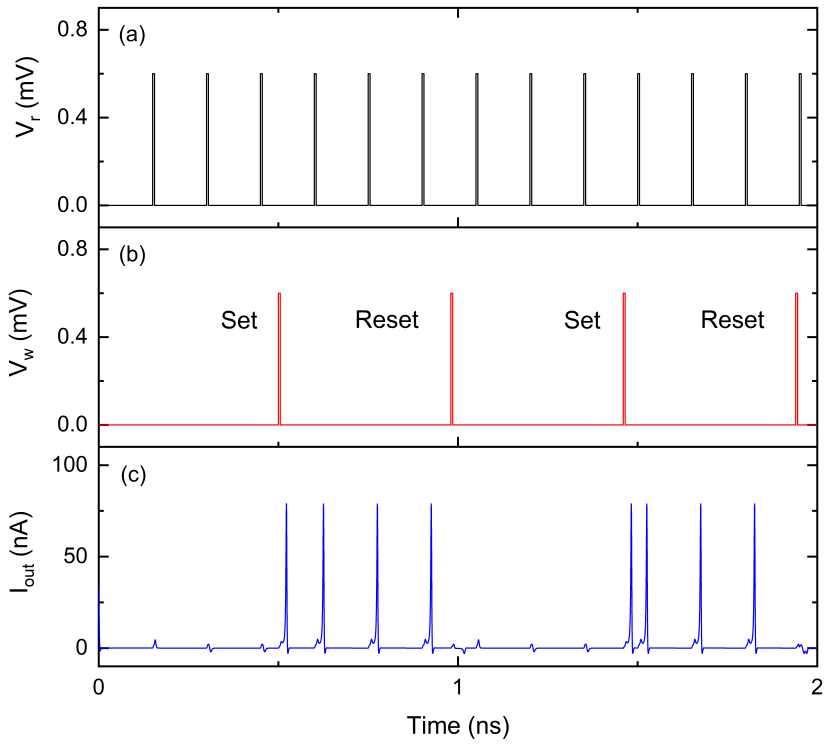


Figure 3.16: Simulation results of the synaptic circuit shown in Figure 3.14, from [15]. The critical voltage values used for Q_0, Q_1 and Q_4 were 0.7 mV and for Q_2 and Q_3 were 0.4 mV. The capacitance values used for C_1 and C_2 were 3 fF. Bias voltage V_{b1} was 0.8 mV and V_{b2} was 1 mV. The magnitude of V_r and V_w was 0.6 mV. (a) Read pulse signal. (b) Write pulse signal. (c) Output current.

pairs of electrons, which correspond to a weight of 3. QPSJ Q_1, Q_2 and Q_3 have different critical voltages. The critical voltages of Q_1, Q_2 and Q_3 are V_{C1}, V_{C2} and V_{C3} respectively while $V_{C1} > V_{C2} > V_{C3}$. Initially, a read pulse from V_r is sufficient to switch Q_1, Q_2 and Q_3, providing a current pulse that carries six electrons at I_{out}. Parallel QPSJs Q_5 to Q_8 are identical. The capacitor C_1 is appropriately chosen to store up to three pairs of electrons before the voltage at node 1 is large enough to switch Q_5 to Q_8. A write pulse from V_w can switch Q_0 and store a pair of electrons at C_1, which also increases the voltage at node 2. The voltage change at node 2 makes one of the parallel QPSJs (Q_1 in this case) not to be switched upon the arrival of another upcoming read pulse. Therefore, the weight of the synapse is changed to 2. Each write pulse can reduce the weight by 1 until it reaches 0. Then the following write pulse

would push another pair of electrons onto C_1, which in turn switches Q_5 to Q_8 and releases all the electrons at C_1.

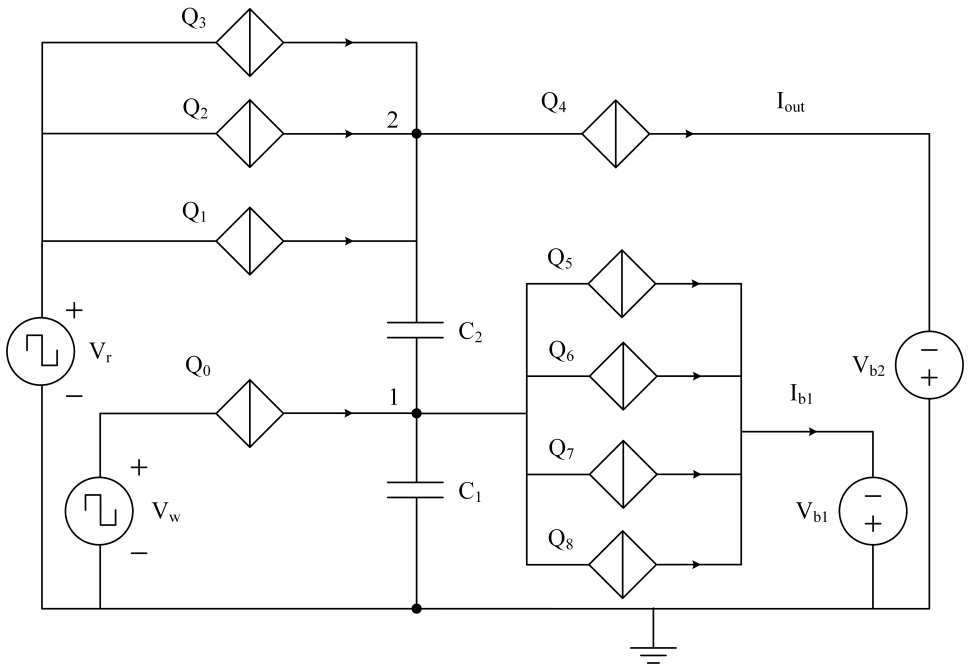


Figure 3.17: A multi-weight synaptic circuit based on QPSJs, from [15]. This synaptic circuit can have a weight of 0, 1, 2 or 3, which can be modified by applying pulses at V_{w}.

In Figure 3.18, periodic pulses are applied at V_{r} and V_{w} to show how the number of output pulses changes according to the write pulses. Initially, a read pulse from V_{r} is able to switch Q_1, Q_2 and Q_3, resulting in three sequential pulses that contain a charge of $6e$ at I_{out}. Once the first pulse from V_{w} is introduced, the read pulse could only switch Q_2 and Q_3 while the number of sequential pulses at I_{out} reduces to two. Further applying pulses at V_{w} makes the number of sequential pulses at I_{out} reduce to one, zero and then be reset to three. The multi-weight synaptic circuit we have shown here can exhibit multiple quantized weights other than 0 and 1, and thus provides a more flexible configuration for large-scale and complex neural networks.

3.4.3 Inhibitory synaptic circuit

The inhibitory synapse is a type of synapse that makes a postsynaptic neuron less likely to generate an action potential [128]. Inhibitory presynaptic neurons can release neurotransmitters that bind to the postsynaptic receptors, which induces a change in the permeability of the postsynaptic neuron membrane to particular ions. As a result, an electric current that changes the

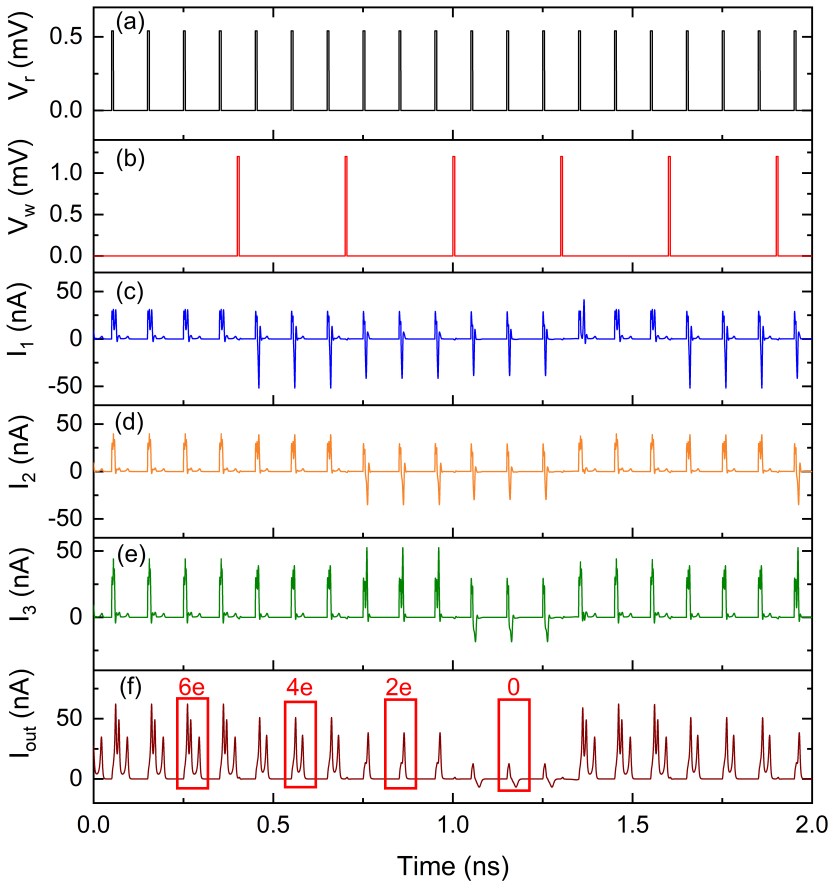


Figure 3.18: Simulation results of the multi-weight synaptic circuit shown in Figure 3.17, from [15]. The critical voltages values used for Q_0, Q_1, Q_2, Q_3 and Q_4 were 0.6 mV, 0.54 mV, 0.52 mV, 0.51 mV and 0.3 mV, respectively. The critical voltage values used for Q_5 to Q_8 were 0.09 mV. The capacitance value used for C_1 was 7.8 fF and the capacitance value used for C_2 was 1.2 fF. Bias voltage V_{b1} was 0.2 mV and V_{b2} was 0.5 mV. The magnitude of V_r was 0.54 mV and the magnitude of V_w was 1.2 mV. (a) Read signal. (b) Write signal. (c) Current through Q_1. (d) Current through Q_2. (e) Current through Q_3. (f) Output current I_{out}.

postsynaptic membrane potential to create a more negative postsynaptic potential is generated, i.e., the postsynaptic membrane potential becomes more negative than the resting potential. To generate an action potential, the postsynaptic neuron membrane must reach a voltage threshold more positive than the resting potential, which makes the postsynaptic neuron less likely to generate an action potential. In neuromorphic computing, inhibitory synapses are also necessary to process certain types of input signals. We emulate inhibitory postsynaptic potential (IPSP) by sending negative pulses to the neuron's membrane capacitor and discharging the membrane capacitor accordingly.

Here we introduce a binary inhibitory synaptic circuit in Figure 3.19. This circuit comprises a binary synaptic circuit and a current inverter circuit. The binary synaptic circuit uses V_{w} to switch the weight between 0 and 1, similar to the circuit shown in Figure 3.14. Voltage pulses applied at V_{r} can read the synaptic weight by switching Q_2. If the synaptic weight is 1, every pulse from V_{r} can trigger an output current pulse at I_1. The slight voltage change at node 3 allows Q_2 and Q_3 to be switched simultaneously, resulting in a current pulse that contains a charge of $4e$. Since there are only two electrons coming from V_{r}, the voltage drop at node 3 allows Q_6 to be switched and to provide another pair of electrons. Therefore, the output current I_{out} is a negative current pulse containing a charge of $2e$. The capacitor C_4 shown in the schematic diagram is considered as the membrane capacitor from an IFN circuit. Every negative current pulse from I_{out} could take $2e$ from C_4. The simulation results are shown in Figure 3.20 and 3.21.

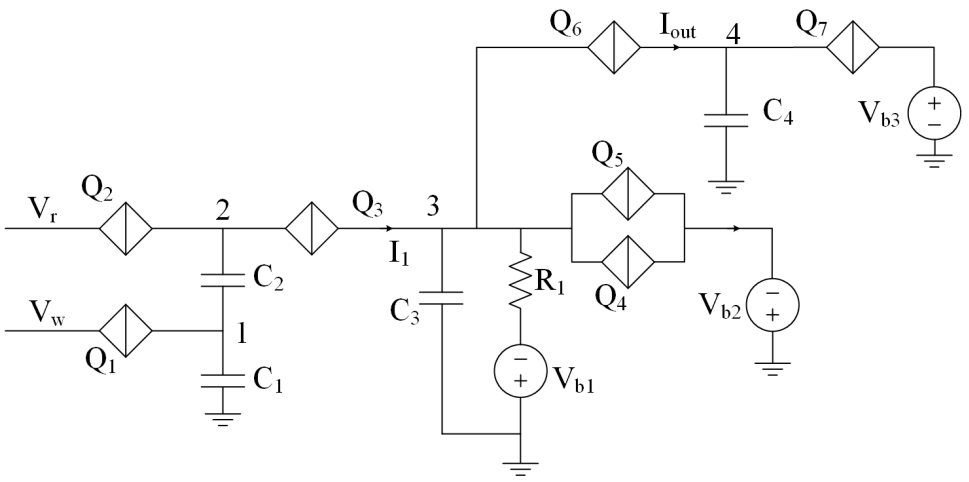


Figure 3.19: An inhibitory synaptic circuit based on QPSJs. This circuit has a binary synaptic circuit combined with a current inverter circuit to store and transmit weighted signals of 0 or -1.

In Figure 3.20, the initial weight of the synaptic circuit is -1. Each pulse from V_{r} can trigger a positive current pulse at I_1, although there is a delay of $\sim$ 40 ps. The QPSJ Q_6 will be switched shortly, providing a negative current pulse through capacitor C_4. The voltage drop at node 4 is $\sim$ 35 μV, which is $2e/V_{\mathrm{C4}}$. In this design, we choose appropriate parameters so that the voltage at node 4 keeps decreasing as negative current pulses I_{out} are generating but will not drop below the minimum value, which is the resting potential of a neuron. Figure 3.21

shows the simulation results when the weight is changed from -1 to 0. This simulation has the same initial condition as the previous one shown in Figure 3.20. Before there is a voltage pulse applied at V_{w}, the read pulse can cause a voltage drop at V_4. However, once a positive voltage pulse is applied at V_{w}, the synaptic weight is changed to 0, and the upcoming voltage pulses from V_{r} are no longer able to switch Q_3. Therefore, no current pulses are observed at I_{out} and no voltage drops are observed at node 4 even there is a voltage pulse at V_{r}.

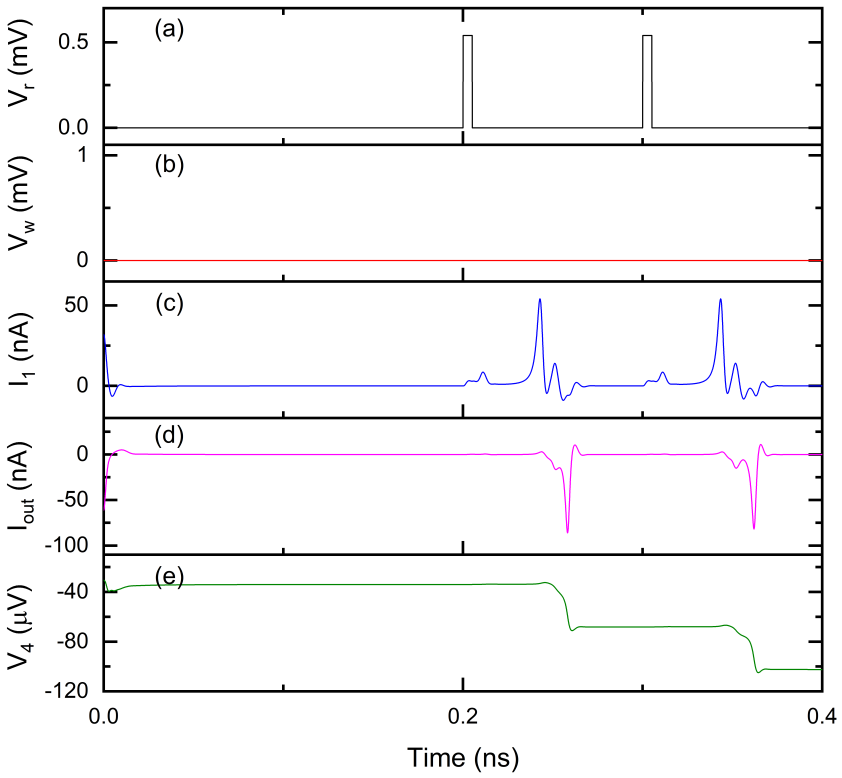


Figure 3.20: Simulation results of the inhibitory synaptic circuit shown in Figure 3.19 when the weight is -1. (a) Read signal. (b) Write signal. (c) Current at I_1. (d) Current at I_{out}. (e) Voltage at node 4.

This circuit is a simple example of inhibitory synapse applications in QPSJ-based neuromorphic computing systems. For complex computational tasks, multi-weight inhibitory synapses are desired. However, the same design idea may not be applied to multi-weight synaptic circuits. This is because that the output of the proposed multi-weight synaptic circuits is a fast pulse sequence while the charging and discharging speed of capacitor C_3 in Figure 3.19

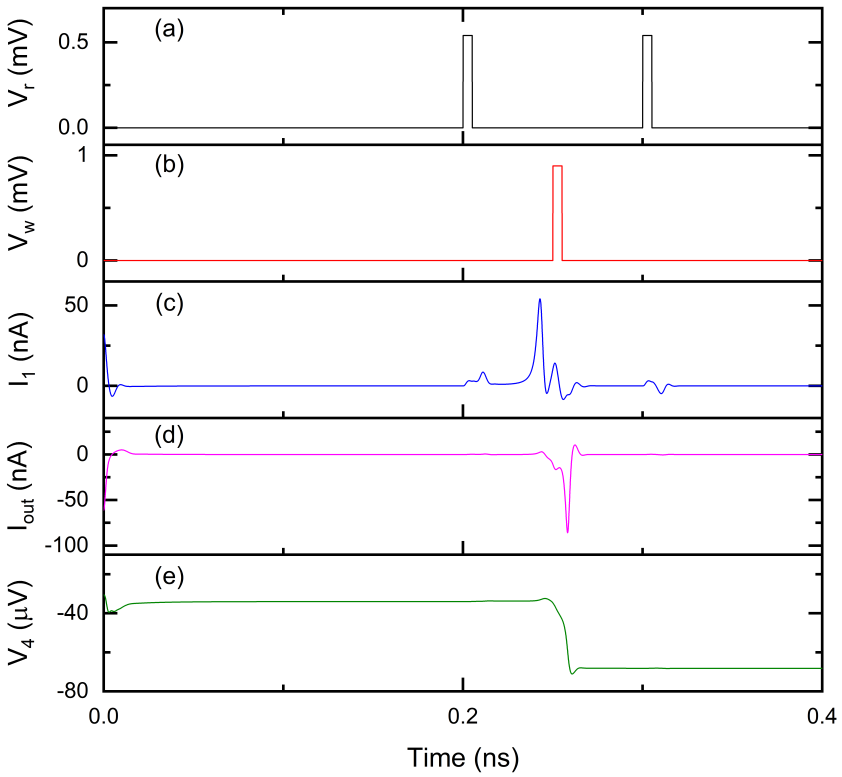


Figure 3.21: Simulation results of the inhibitory synaptic circuit shown in Figure 3.19 when the weight is changed from -1 to 0. (a) Read signal. (b) Write signal. (c) Current at I_1. (d) Current at I_{out}. (e) Voltage at node 4.

are relatively slow. Therefore, capacitor C_3 and QPSJs Q_4 and Q_5 cannot respond promptly to the pulse sequences from a multi-weight synaptic circuit. We envision a future improvement on this circuit that allows inhibitory multi-weight outputs for more complex computations.

3.5 Fan-out circuit consisting of charge-flux converters

A typical biological neuron has thousands of synaptic connections to other neurons [129], which allow the formation of complex networks. Therefore, it is important to have a large fan-out for neuromorphic circuits to emulate biological counterparts. As described previously, the operation of QPSJ-based circuits relies on the propagation of quantized current pulses. JJ-based (single flux quantum, SFQ) circuits face a similar difficulty due to quantized flux. The difficulty of signal fan-out in these circuits arises from the need to equally split a current pulse

to deliver to as many neuron inputs as needed. One way to handle neuron fan-out is by using charge-to-flux and flux-to-charge converters [54]. We note that there is no maximum limit for the fan-out in simulation, though a real circuit may have fabrication-related issues.

In Figure 3.22, the current pulse I_{in} from an IFN circuit flows into an inductor L_1. The current pulse is then coupled to a mutual inductor L_2 and injected to junction J_1 and J_2. Since J_2 is biased close to its critical current by bias current I_{b}, the additional current pulse from L_2 can switch J_2 and generate a single flux quantum (SFQ) pulse that is able to switch multiple parallel QPSJs. Once J_2 is switched and in the resistive state, J_1 can be switched by I_{b} and the system is recovered to its initial state [93]. To avoid directly connecting a QPSJ and a JJ in series with a bias voltage, we use small mutual inductors to transmit current from a QPSJ-based IFN circuit to a JJ loop.

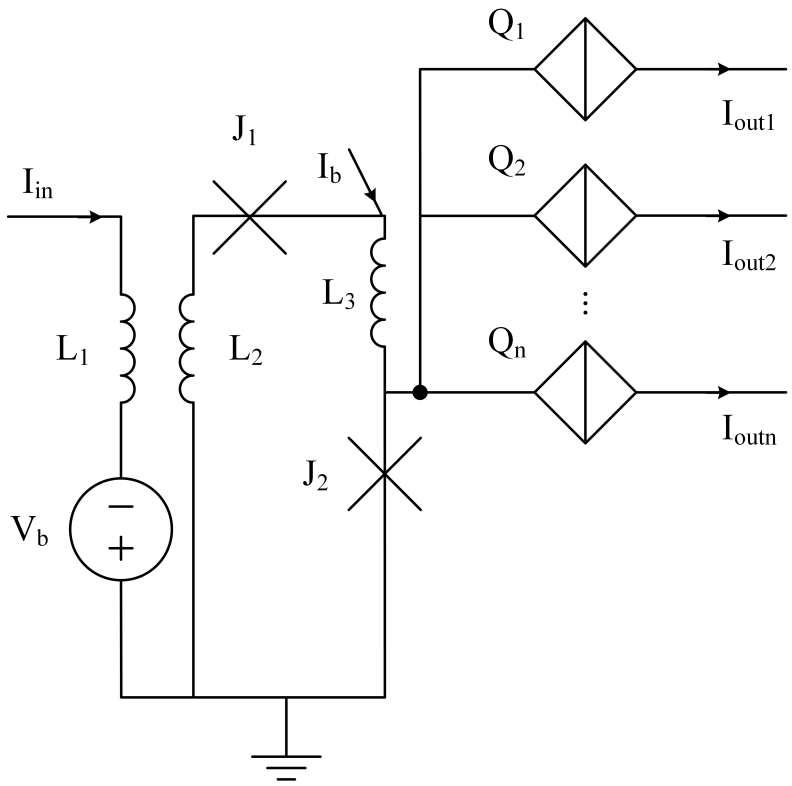


Figure 3.22: A fan-out circuit is comprised of flux-charge and charge-flux circuits, from [15]. The number of fan-out shown in this schematic is n.

The simulation results of the fan-out circuit that has a fan-out of 10 (i.e., ten parallel QPSJs) are shown in Figure 3.23. We use an IFN circuit that has a threshold of 500 to generate current pulses flowing into I_{in}. The induced current pulses from mutual inductors are injected

to J_2, resulted in SFQ pulses across J_2. Since all the parallel QPSJs (Q_1 to Q_{10}) are switched at the same time, we can see identical output current pulses in Figure3.23(c)(d).

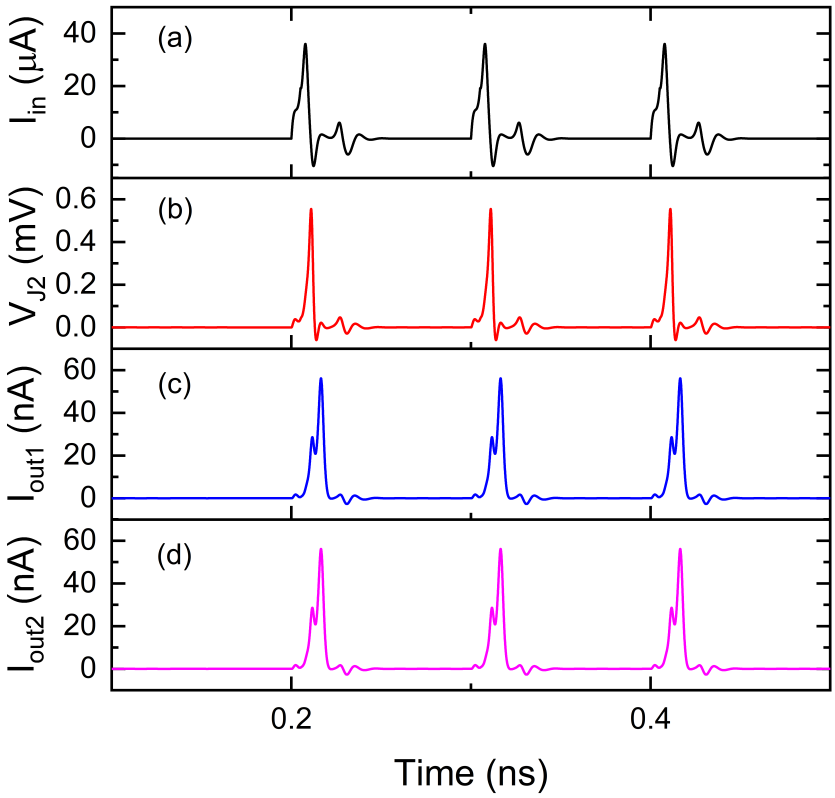


Figure 3.23: Simulation results of the fan-out circuit shown in Figure 3.23, from [15]. The fan-out was set to be 10 in this simulation. Only output 1 and output 2 are shown here while the other outputs are identical. The critical voltage values used for Q_1 to Q_{10} were 0.5 mV. The critical current value used for J_1 was 40 μA and the critical current value used for J_2 was 50 μA. The inductance values used for L_1 and L_2 were 0.1 nH with a coupling coefficient of 0.9. The inductance value used for L_3 was 0.01 nH. Bias current I_b was 70 μA and bias voltage V_b was 0.5 mV. Input current I_{in} is from the output of a QPSJ-based IFN circuit that has a threshold of 500. (a) Input current. (b) Voltage at J_2. (c) Output current 1. (d) Output current 2.

In order to provide sufficient current for J_1, current I_{in} from the QPSJ neuron circuit should be at the same level of the critical current of J_1. For practical application, we may need JJs with small critical currents and QPSJs with small critical voltages. Otherwise, we instead use a neuron that has a large threshold and can fire a large current pulse that contains up to 1,000 Cooper pairs [54].

3.6 QPSJ transmission line circuit as an axon circuit

Like the biological axon that transmits action potentials, the QPSJ-based axon circuit performs charge/current transmission between neurons and synapses over a long distance. We use a simple QPSJ transmission line circuit to handle the signal transmission and control the signal delays. This circuit is based on the charge island circuit proposed in [53]. The number of charge island stages in the axon circuit, shown in Figure 3.24, is flexible and depends on how much signal delay is required between neurons in different layers.

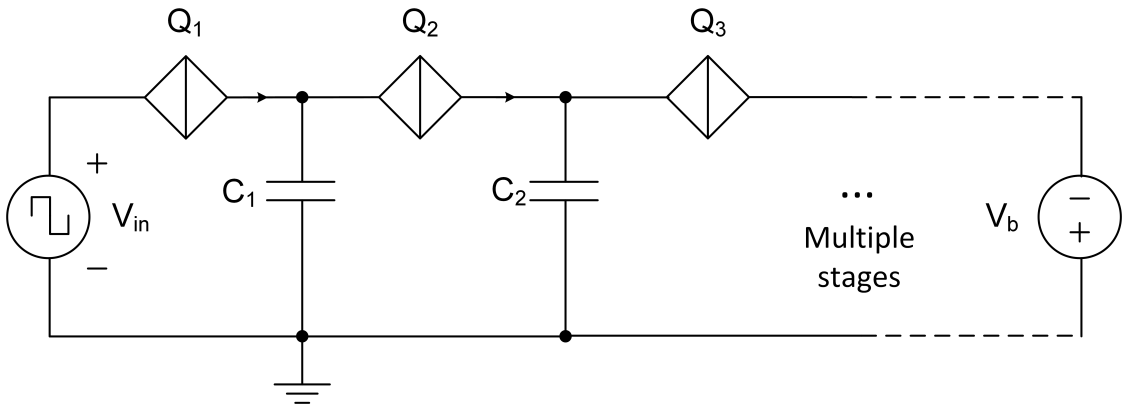


Figure 3.24: Schematic of the axon circuit that has multiple stages of charge island. Signals are transmitted in one direction, depending on the bias voltage and source voltage.

In the axon circuit, signals can only travel in one direction, controlled by the bias voltage and source voltage. Generally speaking, a positive pulse propagated in the axon circuit requires a negative bias voltage. However, a negative pulse (e.g., from an inhibitory synapse) requires a positive bias voltage. This behaves like a biological neuron that only allows action potential to travel in one direction. Figure 3.25 shows the simulation results of the axon circuit with different charge island stages. The capacitance value of the island capacitor is 0.3 fF in this simulation, resulting in less than 200 ps delay after 50 charge island stages. These results indicate that the signal propagation delay in QPSJ-based circuits is expected to be very small. Furthermore, manipulating signal delay could improve circuit performance, especially when the timing is a big concern (e.g., in the time-dependent learning circuits).

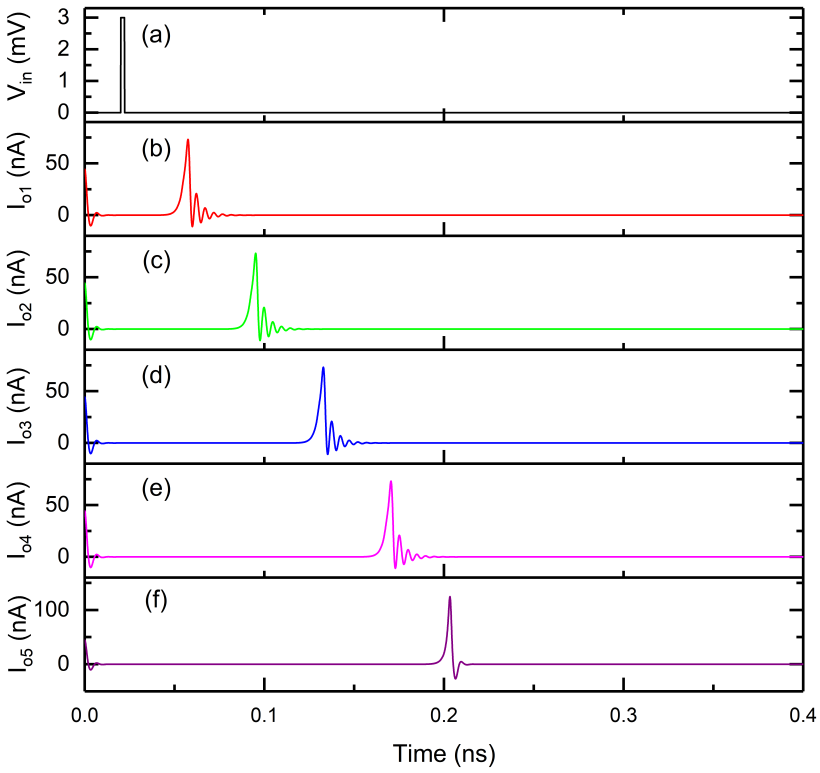


Figure 3.25: Simulation results of the axon circuit shown in Figure 3.24 with different charge island stages. (a) Input voltage. (b) Output current after 10 stages. (c) Output current after 20 stages. (d) Output current after 30 stages. (e) Output current after 40 stages. (f) Output current after 50 stages.

3.7 Simulation of small neural networks

To verify the functions of our neuron and synaptic circuits and demonstrate extension to more complex circuits, we combined neuron and synaptic circuits and simulated neural networks. We show different network configurations based on choices of different synaptic circuits proposed in previous sections. In the first configuration, we used synaptic circuits based on QPSJs and MJJs. The synaptic circuits were assigned with different weight matrices initially. Those synaptic weights did not change during the simulation. In the second configuration, synaptic circuits based on QPSJ memories were used to show a more complex application. Although the synaptic circuits were also assigned with different weights initially, the weights changed during

the simulation through different write signals. The functionality of these neural networks were demonstrated through WRspice simulation results.

3.7.1 A network comprised of IFN circuits and synaptic circuits based on QPSJs and MJJs

A basic 3×2 network architecture is shown in Figure 3.26, for simplicity. We simulated the 3×2 network consisting of three voltage sources and two neuron circuits connected by six binary synaptic circuits. Voltage sources provide input pulses with different frequencies, which represent three different external signals or signals from upstream neurons. The output neuron fires a pulse/spike as soon as it receives a set number (in this case 10) of quantized current pulses from the three synaptic circuits, as the threshold is set to 10.

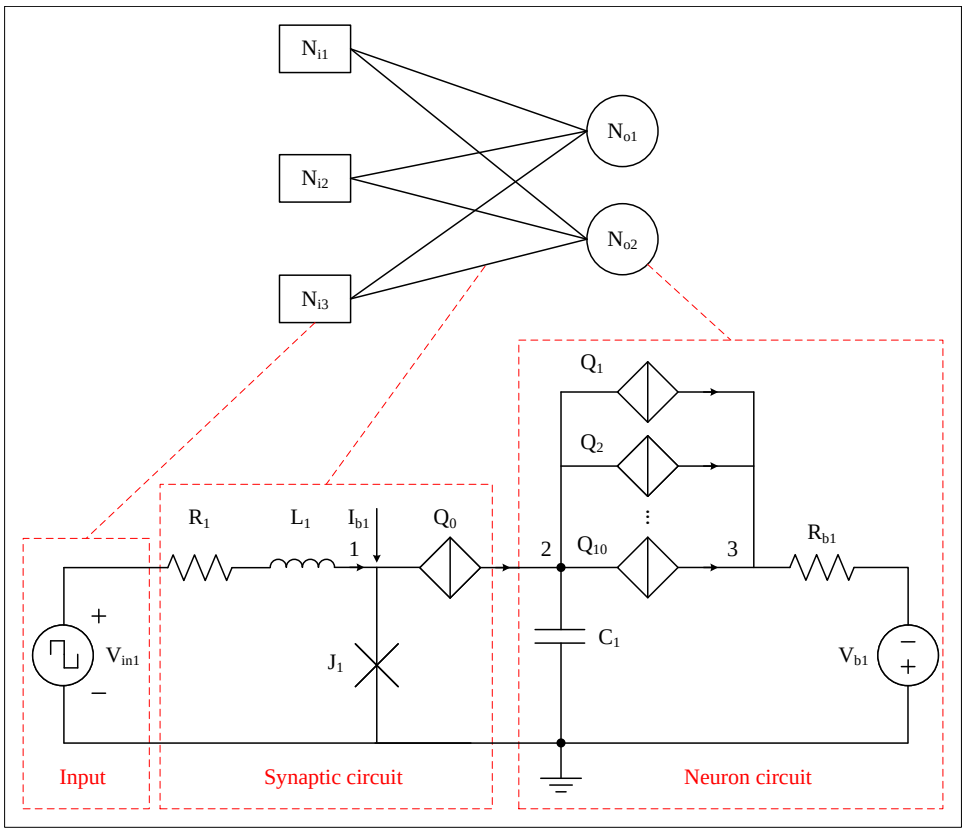


Figure 3.26: Network architecture of a 3×2 neural network based on superconductive synaptic and neuron circuits, from [14]. N_{ix} are input neurons and N_{oy} are output neurons.

The simulation was conducted with different initial values of each synaptic weight (i.e. critical current of J_1), that reflected different firing frequencies observed for each output neuron. Weight matrices for neuron 1 and neuron 2 were [1 1 1] and [0 1 1], respectively, in Figure

3.27, and were [1 0 1] and [0 0 1], respectively, in Figure 3.28. The results show that pulses from a synapse of weight 1 is counted while those from a synapse of weight 0 are not counted, and the neuron fires immediately after the number of total pulses received reaches the threshold (= 10 here). As a result, this circuit works similar to a digital neuromorphic system that can process binary signals. If we replace the synaptic circuit in Figure 3.26 with a multi-state synaptic circuit, the network will be more complex, but also, more flexible.

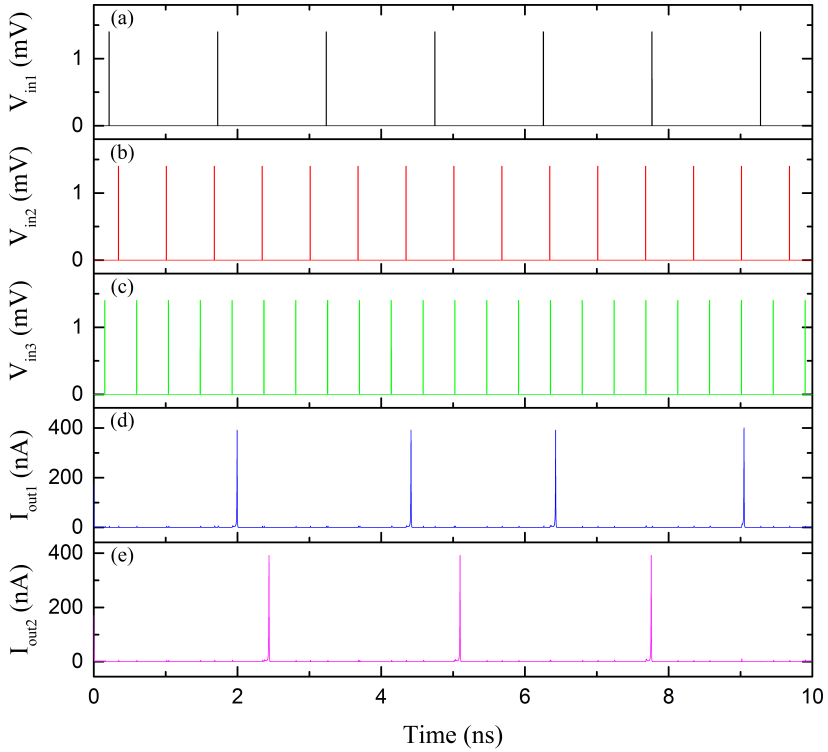


Figure 3.27: Simulation results of the 3×2 network illustrated in Figure 3.26 with weight matrices [1 1 1] and [0 1 1], from [14]. (a) Input voltage 1. (b) Input voltage 2. (c) Input voltage 3. (d) Output current 1. (e) Output current 2.

3.7.2 A network comprised of IFN circuits and synaptic circuits based on QPSJ memories

We conducted a simulation of an example network comprised of two inputs and one output neuron, connected by two synapses proposed in Figure 3.14. The neuron circuit is a QPSJ-based IFN circuit that was proposed earlier [13]. For simplicity, the threshold of the neuron is set to three, which means the neuron will fire a current pulse as soon as it receives three pulses from two synapses. The schematic of the network is shown in Figure 3.29. We apply periodic

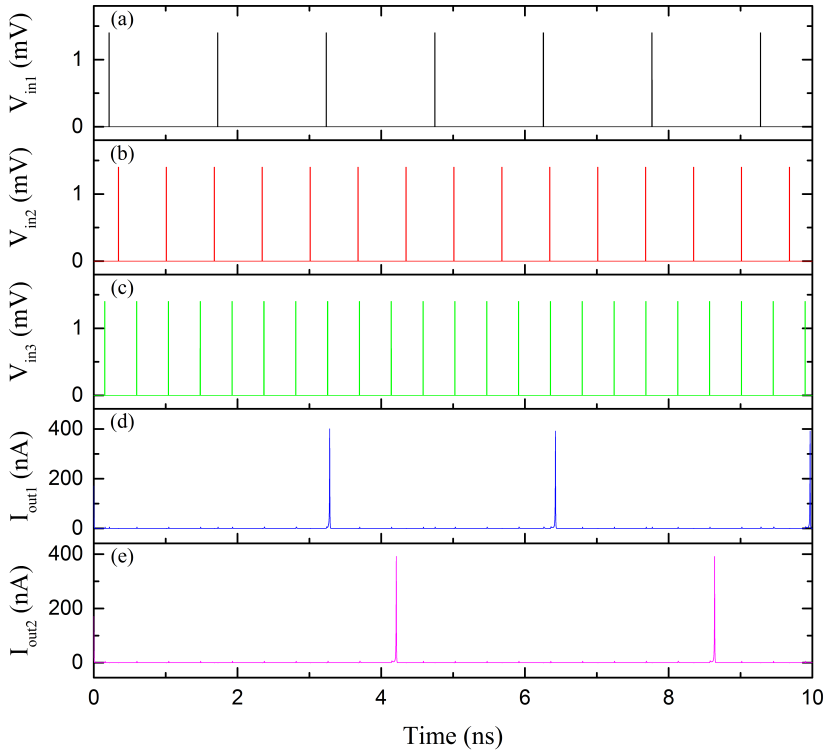


Figure 3.28: Simulation results of the 3×2 network illustrated in Figure 3.26 with weight matrices [1 0 1] and [0 0 1], from [14]. (a) Input voltage 1. (b) Input voltage 2. (c) Input voltage 3. (d) Output current 1. (e) Output current 2.

pulse signals at each read and write terminals of the two synaptic circuits to show how the input pulses are propagated to the output neuron through synapses and make the neuron fire pulses. The write signals from V_{w1} and V_{w2} are chosen randomly to demonstrate circuit functionality but not for any learning purposes.

As we can see from Figure 3.30, once a synapse is set to 1, a current pulse is generated and injected into the output neuron every time a read signal arrives at the synapse. The voltage on capacitor C_5 gradually increases until it can simultaneously switch three parallel QPSJs Q_{10}, Q_{11} and Q_{12}, resulting in a quantized current pulse that contains a charge of $6e$. Figure 3.30 shows that a pulse is generated at I_{out} after the output neuron receives a total amount of three pulses from I_{syn1} and I_{syn2}. The operation of this network is similar to a digital neuromorphic system that comprises digital IFN circuits and binary memories.

We also replace the binary synapses in Figure 3.29 with multi-weight synapses to demonstrate the functionality of multi-weight synapses in a neural network, as shown in Figure 3.32.

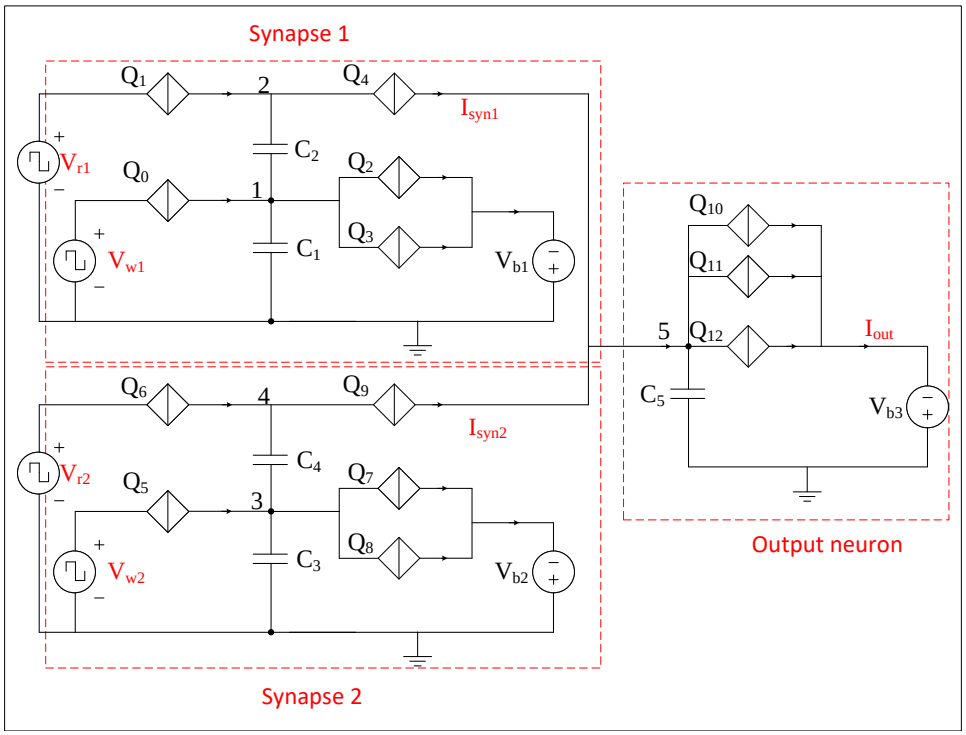


Figure 3.29: A 2×1 network consisting of QPSJ-based neurons and binary synapses, from [15].

We increase the neuron firing threshold to 10 to reduce output neuron firing frequency that is potentially affected by the multi-weight synapses in the network. The simulation results are shown in Figure 3.33 and Figure 3.34. The weights of synapses 1 and 2 are modified by V_{w1} and V_{w2}, respectively.

In Figure 3.33, the initial weights of both synapses are set to 3. As soon as a positive pulse arrives at V_{w1} or V_{w2}, the weight of synapse 1 or synapse 2 is reduced by 1 until it reaches 0. Once the weight of either synapse is 0, one more positive pulse from either write terminal could reset the corresponding synaptic weight to 3. As the weights of both synapses change during the simulation, the firing frequency of the output neuron also changes.

Sometimes, when the output neuron receives more pulses (or electrons) than it requires to fire a pulse, it will only release 20 electrons and the remaining electrons will still be stored on capacitor C_5 after firing. As a result, the neuron will require fewer pulses from two synapses to fire another pulse. This operation could be an issue in real circuit operations because the IFN circuit we designed previously only allows a fixed number of electrons to be released at a time. We believe that future improvement or modification of IFN circuit (for example by

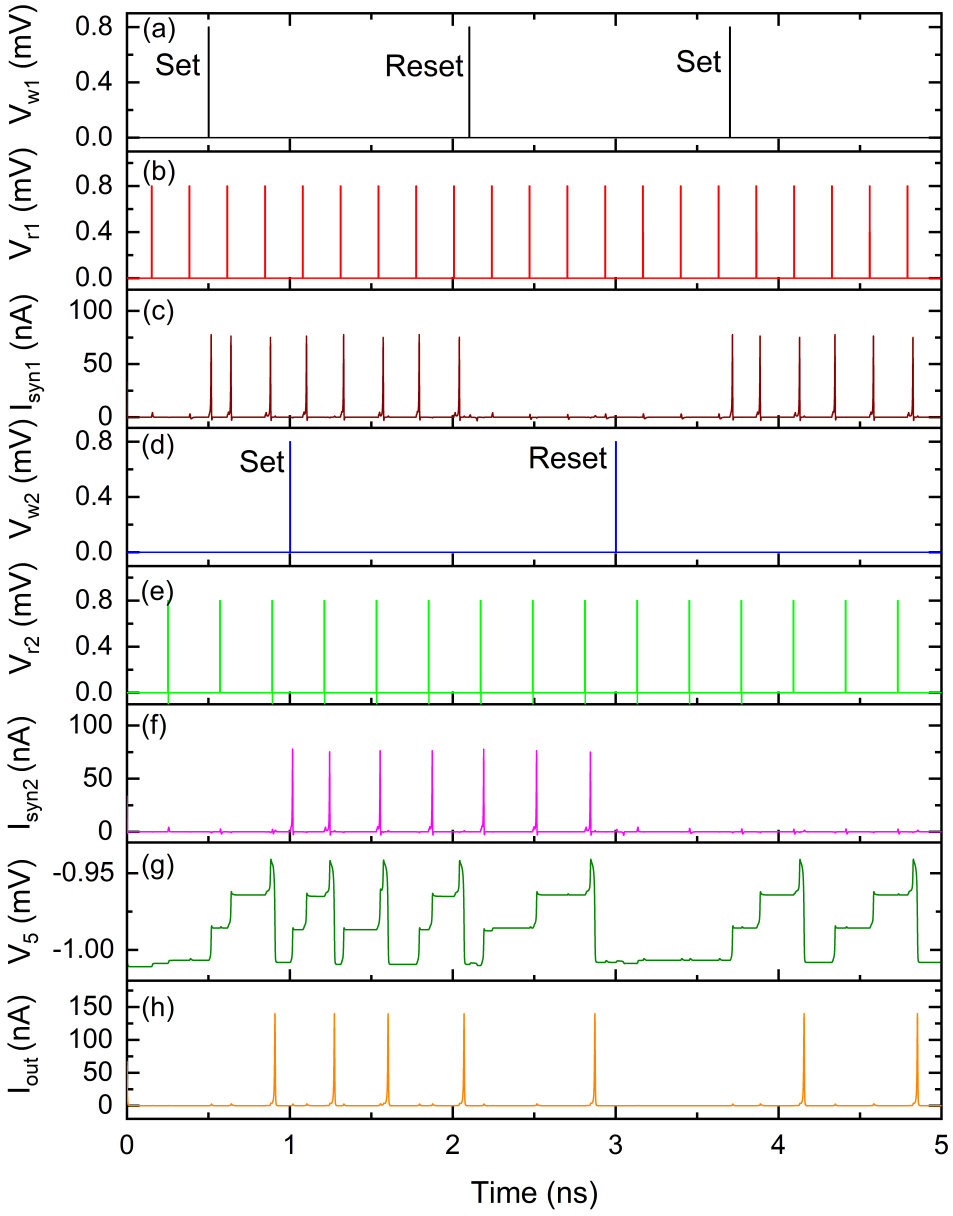


Figure 3.30: Simulation results of the 2×1 network shown in Figure 3.29, from [15]. (a) Write signal of input 1. (b) Read signal of input 1. (c) Synapse 1 output. (d) Write signal of input 2. (e) Read signal of input 2. (f) Synapse 2 output. (g) Voltage at node 5. (h) Neuron output.

incorporating a reset or inhibitory mechanism) could solve this problem. The simulation results have demonstrated the functionality of combined neuron and synaptic circuits based on QPSJs.

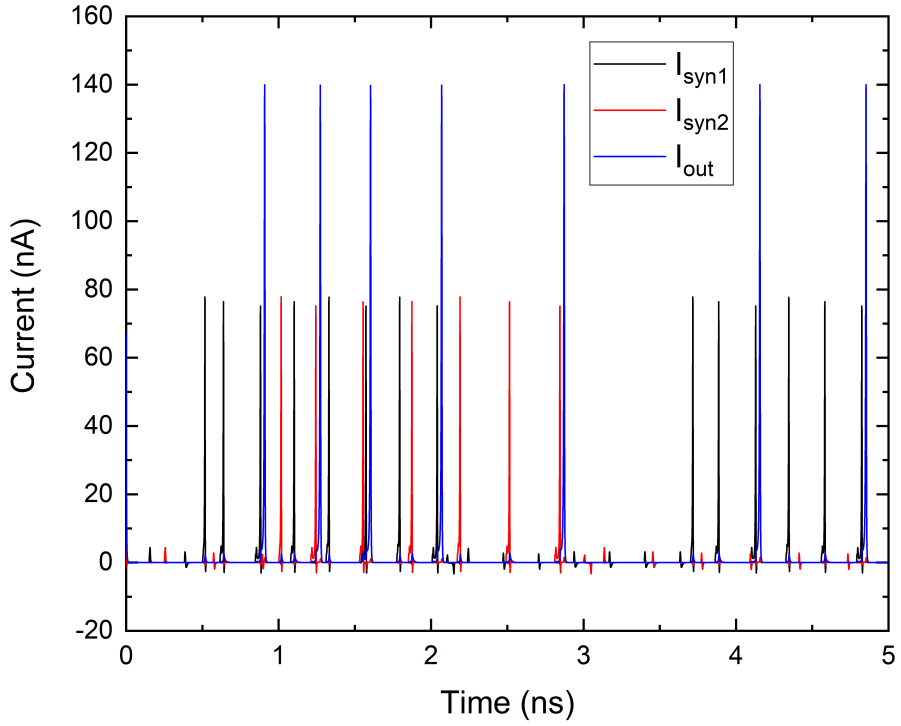


Figure 3.31: Synaptic current I_{syn1}, I_{syn2} and output current I_{out} re-plotted from Figure 3.30, from [15].

3.8 Summary

Neuromorphic circuits based on superconductive QPSJs are introduced in this chapter. These circuits are designed to emulate biological neuron components to perform neuromorphic computing through charge and fluxon operations. The intrinsic spiking behaviors of QPSJs make them an ideal candidate for high-speed and low-power neuromorphic computing. We have demonstrated the functionalities of each circuit and some combined neuron and synaptic circuits through simulations in WRspice.

Furthermore, these circuits can be scaled up to form a large neural network, which is desired to solve complex and practical problems. To enhance scalability, a large network may be broken into multiple smaller blocks where one or a small number of bias voltages can be used for each block. For example, we can bias one neuron and multiple synapses with only one bias voltage, and the downstream neuron and synapses, separated by a fan-out circuit, use

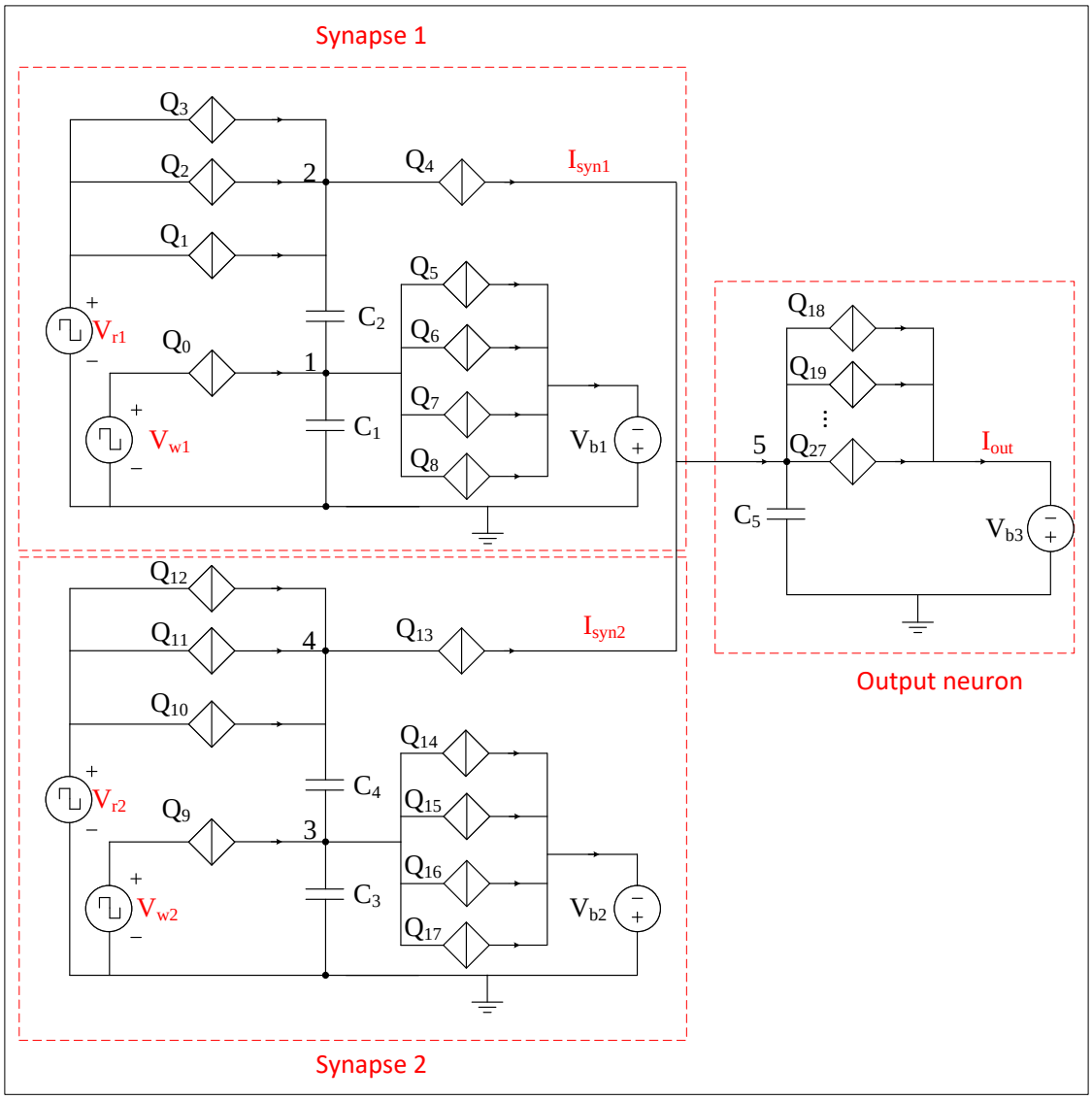


Figure 3.32: A 2×1 network consisting of QPSJ-based neurons and multi-weight synapses.

another bias voltage. The fan-out circuit works as a bridge to connect neurons from adjacent layers. The goal of this work is to build large neural networks using QPSJ technology, which can solve practical problems at a low energy cost and a high speed.

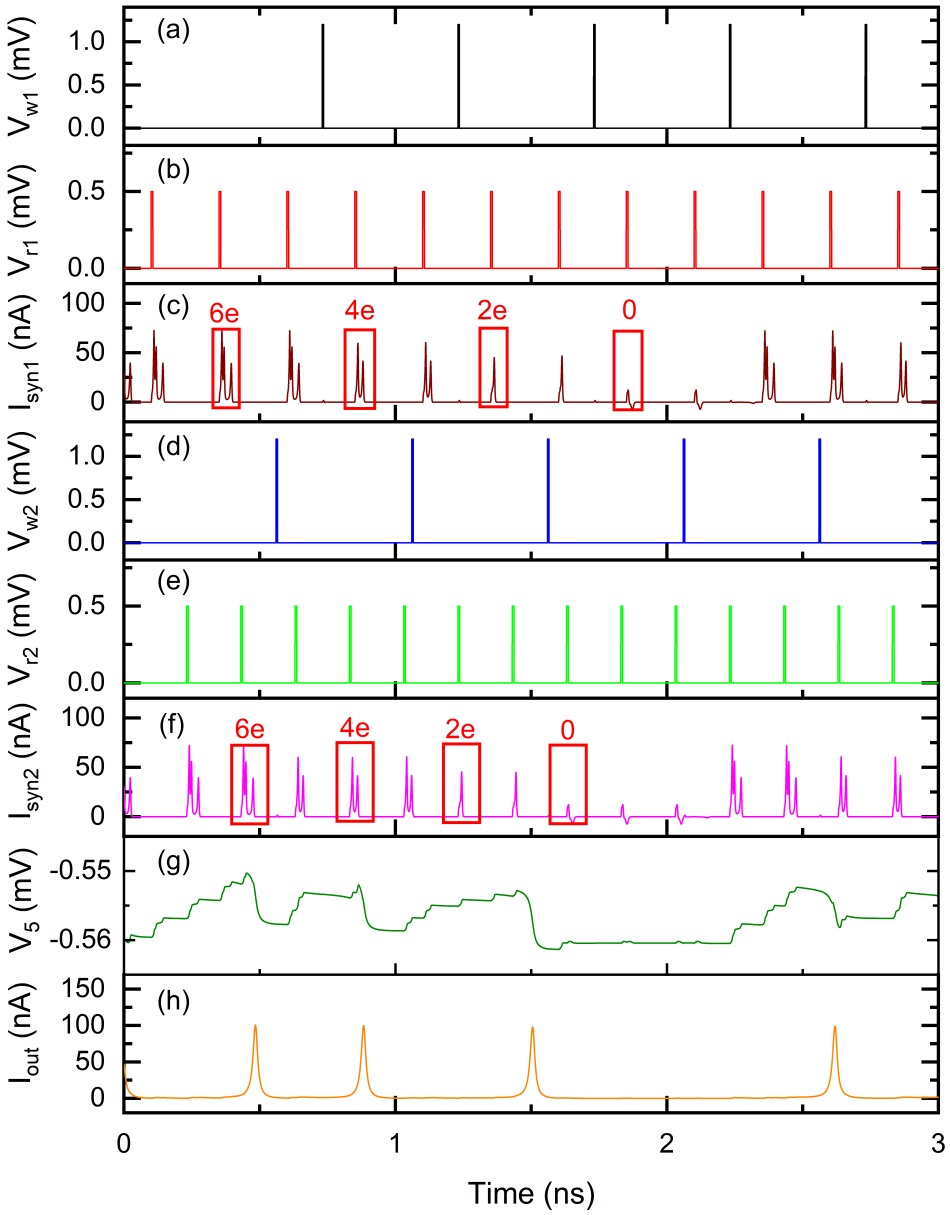


Figure 3.33: Simulation results of the 2×1 network consisting of QPSJ-based neurons and multi-weight synapses, from [15]. (a) Write signal of input 1. (b) Read signal of input 1. (c) Synapse 1 output. (d) Write signal of input 2. (e) Read signal of input 2. (f) Synapse 2 output. (g) Voltage at node 5. (h) Neuron output.

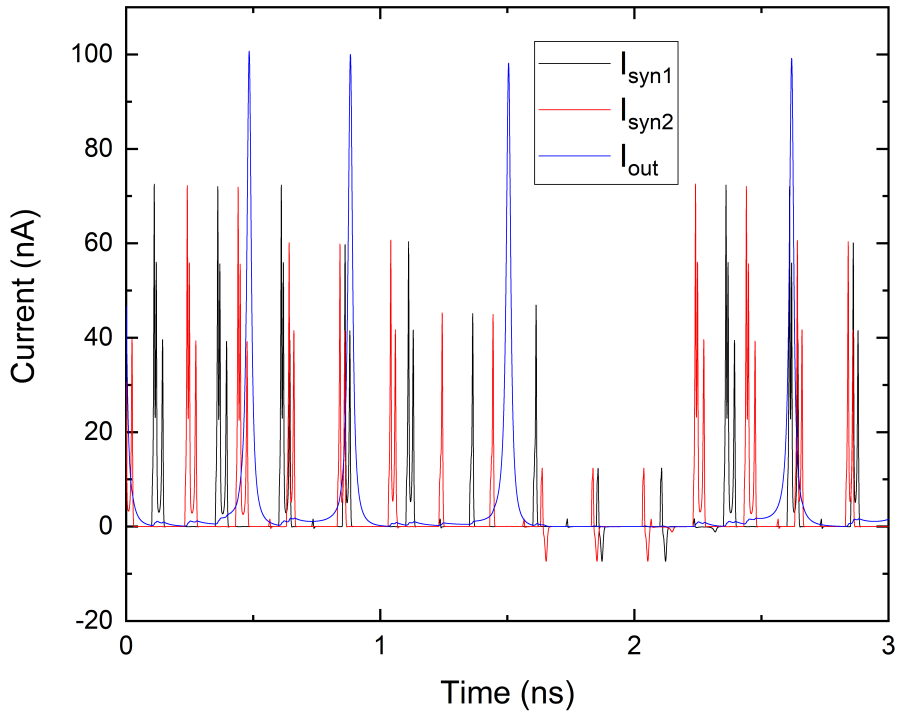


Figure 3.34: Synaptic current I_{syn1}, I_{syn2} and output current I_{out} re-plotted from Figure 3.33, from [15].

Chapter 4

Toward Learning in Neuromorphic Circuits

4.1 Introduction

The human brain serves as a power-efficient learning machine, solving demanding computational tasks while consuming a small number of power. Like a human brain has the ability to adapt to surrounding environment and to solve complex problems by implementing synaptic plasticity, one of the most important functions of neuromorphic computing is to have the ability to adjust synaptic weights through learning. While there are plenty of learning strategies in neuromorphic computing as we discussed in Section 1.2.4, we focused on STDP learning rule in this work to provide potential computational functions for QPSJ-based neuromorphic systems. Early neuroscience experiments on synaptic plasticity suggested that relative timing of presynaptic and postsynaptic action potentials at milliseconds had effects on the plasticity [75]. This is well known as spike timing-dependent plasticity (STDP), which has been observed in cortical neurons [69]. In neuromorphic hardware systems, STDP-type learning rules are widely used as an unsupervised learning method. We introduce a method of realizing a simplified STDP rule by using QPSJ-based circuits. The weight change is either 1 or -1 during each learning event. We modified the synaptic circuit shown in Figure 3.17 to adapt to our learning circuit. The learning circuit is comprised of a long term depression circuit and a long term potentiation circuit, which are combined together to realized a simplified STDP rule [16].

4.2 A multi-weight synaptic circuit for learning

The circuit shown in Figure 4.1 is a multi-weight synaptic circuit that can generate different numbers of sequential pulses, which correspond to a weight of 0, 1, 2 or 3. Here, the weight is defined as the number of pairs of electrons at the output for each input pulse. In general, N sequential current pulses contain N pairs of electrons, although the shapes of these pulses may not look significantly different. This circuit does not have a reset mechanism that was seen in Figure 3.17. Parallel QPSJs Q_1, Q_2 and Q_3 have different critical voltages. The weight can be increased by applying negative pulses at V_{w} or decreased by applying positive pulses at V_{w}. Applying (positive) pulses at V_{r} can read but not destroy the memory state. Different numbers of sequential current pulses will be generated at I_{out} upon the arrival of one short voltage pulse at V_{r}, depending on the number of electrons stored at capacitor C_1. The simulation results of this circuit are shown in Figure 4.2.

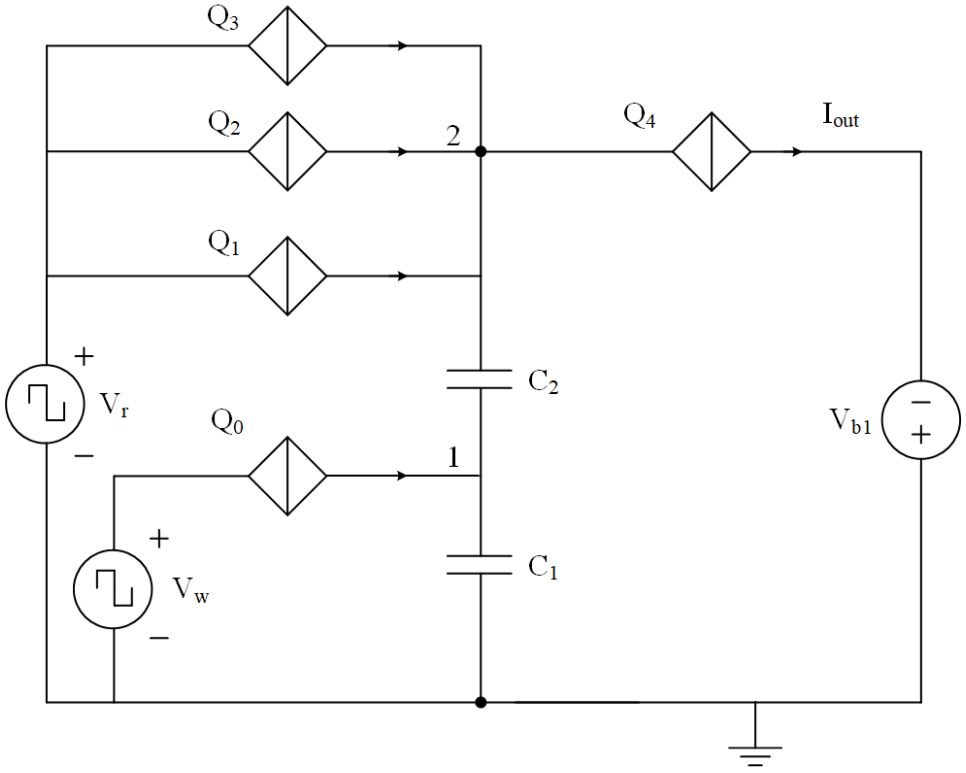


Figure 4.1: A multi-weight synaptic circuit used for learning, from [16]. The synaptic weight can be increased or decreased by applying negative or positive pulses at V_{w}.

In Figure 4.2, the initial weight of the synaptic circuit is set to 3. A voltage pulse from V_r can switch all three parallel QPSJs Q_1, Q_2 and Q_3, resulting in three sequential current pulses at I_{out}. Applying a positive voltage pulse at V_{w} can add two electrons onto capacitor C_1 and the voltages at node 1 and node 2 increase accordingly. In this case, the upcoming voltage pulse from V_r can only switch two out of three parallel QPSJs, which causes two sequential current pulses at I_{out}. Once the synaptic weight reaches 0, it will not decrease any more. Similarly, applying a negative voltage pulse at V_{w} can take two electrons from capacitor C_1 and the voltages at node 1 and node 2 decrease accordingly. Therefore, the synaptic weight is increased by 1. This can be repeated up to reaching the maximum weight. Different weights result in different numbers of sequential current pulses at I_{out} during each read operation. The weight modulation scheme in this circuit allows us to design learning circuits that can generate appropriate positive and negative pulses based on specific learning rules to control the synaptic weight.

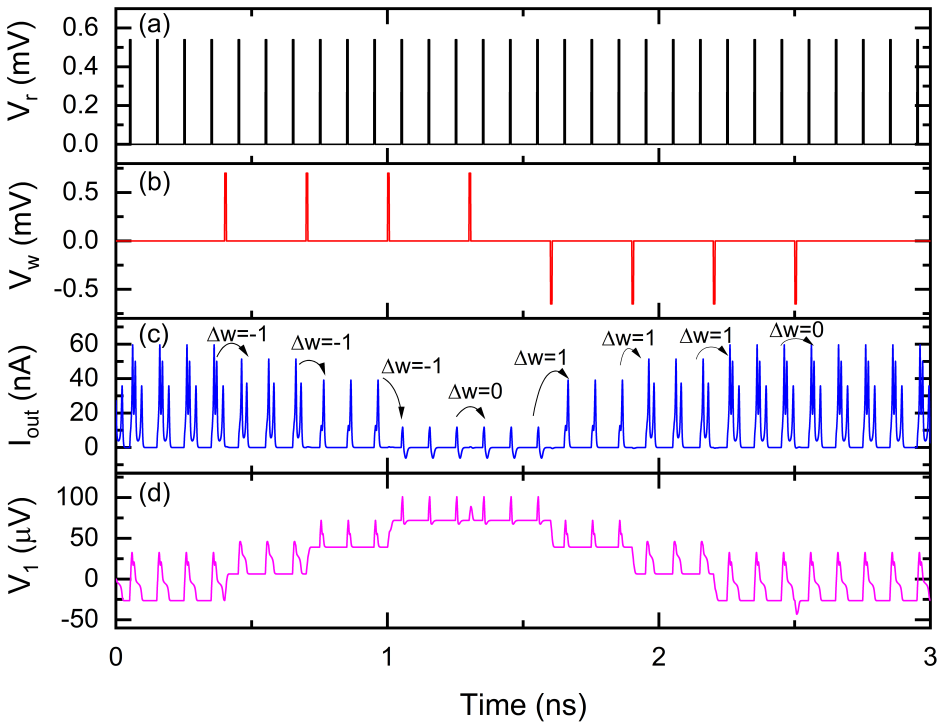


Figure 4.2: Simulation results of the synaptic circuit shown in Figure 4.1, from [16]. (a) Read signal V_{r}. (b) Write signal V_{w}. (c) Output current I_{out}. (d) Voltage at node 1.

4.3 A long term depression circuit

In a biological neural system, long term depression (LTD) occurs when a postsynaptic spike leads a presynaptic spike by $\sim$ 20 to 100 ms [130, 77]. The synaptic weight between these two neurons is thus depressed as they are considered to be uncorrelated. The LTD circuit shown in Figure 4.3 can generate positive pulses used to depress the synaptic weight if the timing difference $\Delta t = t_{\text{post}} - t_{\text{pre}}$ is within a short learning window. This circuit operates at a much faster speed than its biological counterpart, tens of GHz versus kHz, therefore LTD is designed to be effective within a shorter (ps scale) learning window. In Figure 4.3, the initial voltage at node 1 is set by bias voltage V_{b1} when there are no inputs at $\overline{V_{\text{post}}}$. In the circuit design, we choose an appropriate critical voltage value for Q_1 such that Q_1 cannot be switched by V_{pre} should a voltage pulse from V_{pre} arrive first. Therefore, no current pulses are generated at I_{learning}. On the other hand, if a negative voltage pulse from $\overline{V_{\text{post}}}$ arrives first, Q_0 is switched and a pair of electrons are taken from capacitor C_1. The voltages at node 1 and node 2 drop by $2e/C_1$, where C_1 is the capacitance of capacitor C_1. The slight voltage change at node 2 allows the upcoming pulse from V_{pre} to switch Q_1 and in turn switch Q_2, resulting in a positive current pulse at I_{learning}. The voltages at nodes 1 and 2 will recover to their initial states since C_1, R_1 and V_{b1} behave like a series RC circuit with a corresponding voltage decay time. As a result, there will be pulses at I_{learning} only if signals at V_{pre} and $\overline{V_{\text{post}}}$ are close enough in time. The width of the learning window is determined by the resistance value of R_1. The simulation results in Figure 4.4 illustrate how the learning window changes as R_1 changes.

In Figure 4.4, the voltage at node 1 drops upon arrival of a negative pulse into $\overline{V_{\text{post}}}$. A current pulse at I_{learning} is followed by each upcoming pulse from V_{pre} before the voltage at node 1 gradually increases to a stable point. The effective time window over which the circuit responds as intended is viewed as the learning window for this LTD function. In this LTD circuit design, the width of the learning window increases as R_1 increases. This can be explained by the different voltage level recovering speeds due to different RC time constants.

This LTD circuit works seamlessly with a synaptic circuit as shown in Figure 4.5. LTD occurs when the circuit detects $t_1 < \Delta t < 0$, where $t_1 \approx 50$ ps defines the maximum LTD

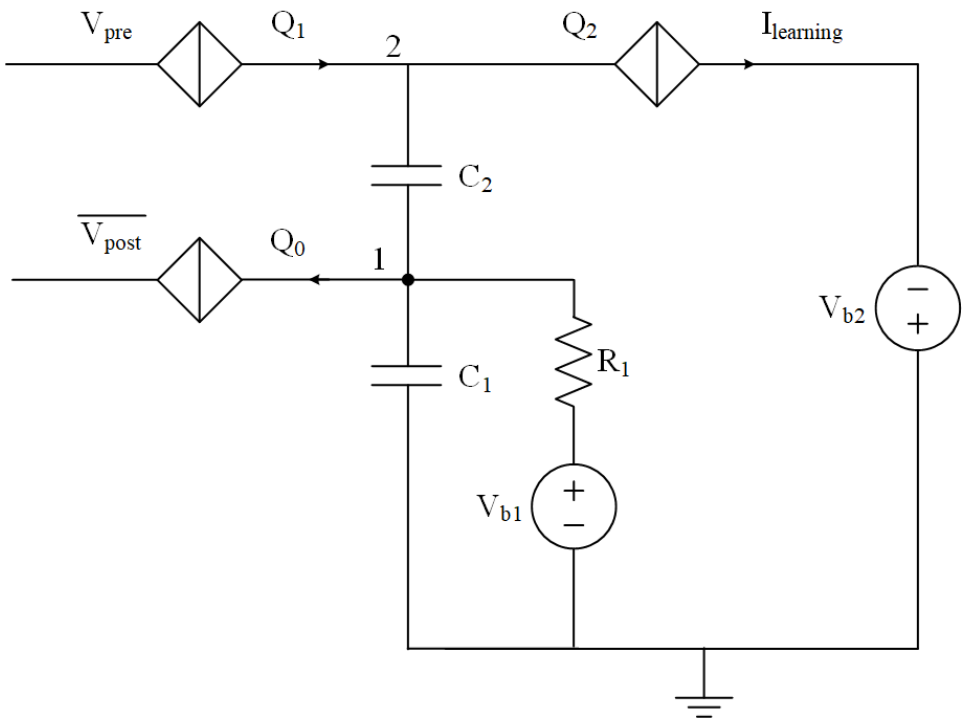


Figure 4.3: An LTD circuit that generates depression pulses to a synapse, from [16]. A pulse will be generated at I_{learning} when the timing difference $\Delta t = t_{\text{post}} - t_{\text{pre}}$ is within a short learning window.

learning window. Charge (electrons) will be injected onto capacitor C_3, which depresses the synaptic weight. A simulation was performed to show how the synaptic weight changes according to the LTD rule. The results are shown in Figure 4.6. The width of LTD learning window was not a concern during this simulation, as the circuit parameters were chosen to demonstrate LTD functions but not for a specific LTD learning window.

In this circuit design, the initial weight was set to 3 based on the device parameters used for this simulation. Each presynaptic pulse could result in three sequential current pulses (containing a charge of six electrons) at I_{syn}. As the first negative voltage pulse from $\overline{V_{\text{post}}}$ is presented, the voltage at node 1 drops due to the switching of Q_0, which takes two electrons from C_1. The voltage at node 2 also drops subsequently, which allows the fourth voltage from V_{pre} to switch Q_1 and in turn switch Q_2 to inject two electrons onto C_3. As a result, the synaptic weight is depressed by 1. The weight change is not immediate but can be observed by the upcoming pulse from V_{pre}, which results in two sequential current pulses (containing a charge of four electrons) at I_{syn}. We can also see that the timings between the third pulse from $\overline{V_{\text{post}}}$ and the tenth pulse

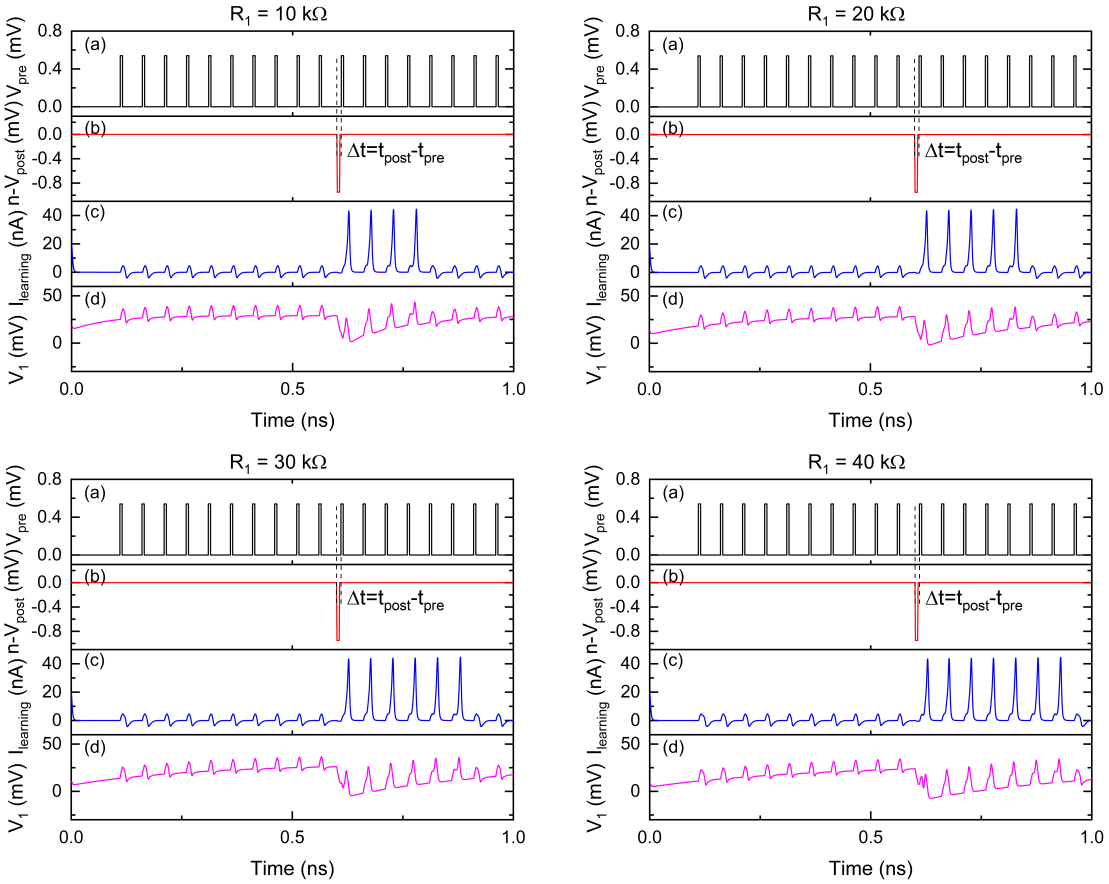


Figure 4.4: Simulation results of the circuit shown in Figure 4.3 with different R_1 values, from [16]. V_{pre} was 0.54 mV and $\overline{V_{\text{post}}}$ was 0.95 mV. The critical voltage values used for Q_0 to Q_2 were 0.75 mV, 0.56 mV and 0.31 mV, respectively. C_1 was 9.2 fF and C_2 was 1.2 fF. V_{b1} was 0.03 mV and V_{b2} was 0.5 mV. R_1 was 10/20/30/40 kΩ. (a) Input signal V_{pre}. (b) Input signal $\overline{V_{\text{post}}}$. (c) Output signal I_{learning}. (d) Voltage at node 1.

from V_{pre} is relatively larger ($\sim$ 100 ps), which does not result in a weight depression. This is because the voltage at node 1 and node 2 recover to their initial states (set by bias voltages) before the tenth pulse from V_{pre} arrives. These simulation results demonstrate that the LTD circuit can realize a weight depression function with respect to the relative timing information between presynaptic and postsynaptic pulses.

4.4 A long term potentiation circuit

In a biological neural system, long term potentiation (LTP) occurs when a presynaptic spike leads a postsynaptic spike by up to 20 ms [131, 77]. The synaptic weight between these two neurons is thus potentiated as they are considered as correlated. In a synaptic circuit shown in

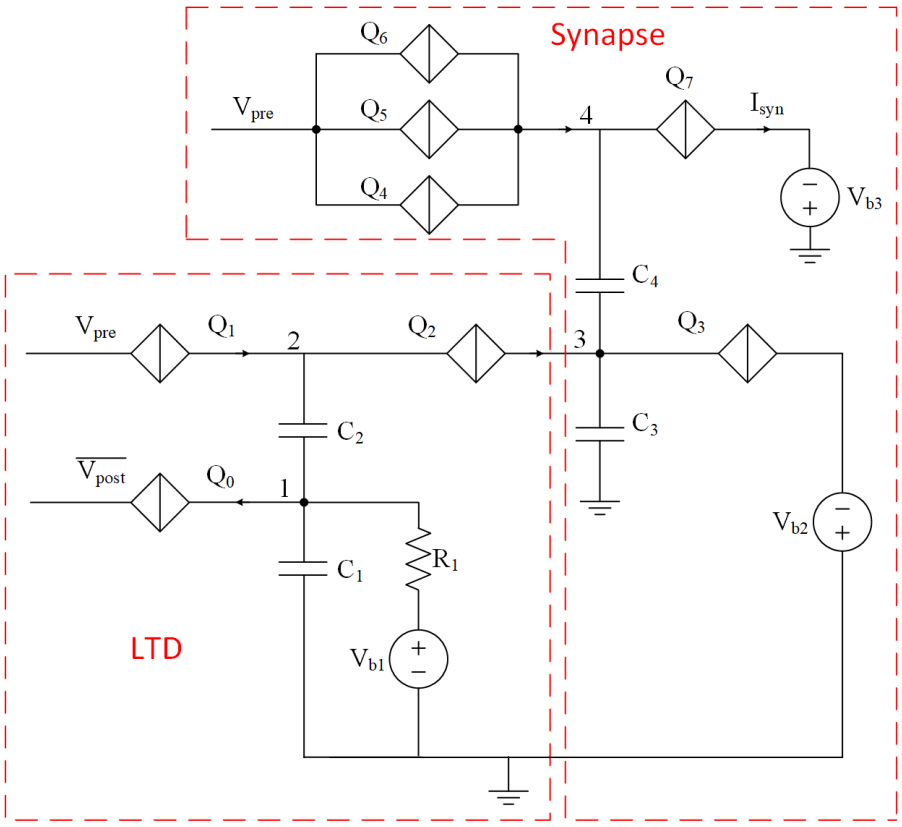


Figure 4.5: An LTD circuit with a multi-weight synaptic circuit, from [16]. The number of sequential pulses at I_{syn} can be reduced when the timings of pulses from V_{pre} and $\overline{V_{\text{post}}}$ trigger an LTD learning event.

Figure 4.1, the weight can be potentiated by applying negative pulses at V_{w}. Here we propose an LTP circuit that can generate negative current pulses to potentiate the synaptic weight according to the relative timing information between a presynaptic neuron and a postsynaptic neuron. The circuit shown in Figure 4.7 is an LTP circuit with a multi-weight synaptic circuit. Similar to the LTD circuit shown in Figure 4.3, the LTP circuit has two inputs $\overline{V_{\text{pre}}}$ and V_{post}. Q_2 and Q_3 are identical and biased by voltage V_{b3}. The initial voltage at node 1 is set by bias voltage V_{b1}. Voltage at node 2 (V_2) is set by bias voltage V_{b2} so that the voltage across Q_2 and Q_3 is near their critical voltages. When there are no inputs at $\overline{V_{\text{pre}}}$, V_{post} cannot switch Q_1. A negative voltage pulse from $\overline{V_{\text{pre}}}$ can switch Q_0, taking two electrons from capacitor C_1. The voltage drop at node 1 results in a voltage drop at node 2 as well. The slight voltage change at node 2 allows the upcoming voltage pulse from V_{post} to switch Q_1 and in turn switch Q_2 and Q_3, resulting in a current pulse that contains a charge of $4e$. Since there are only two electrons coming from V_{post}, the voltage drop at node 2 allows Q_4 to be switched and allows C_3 to provide another pair

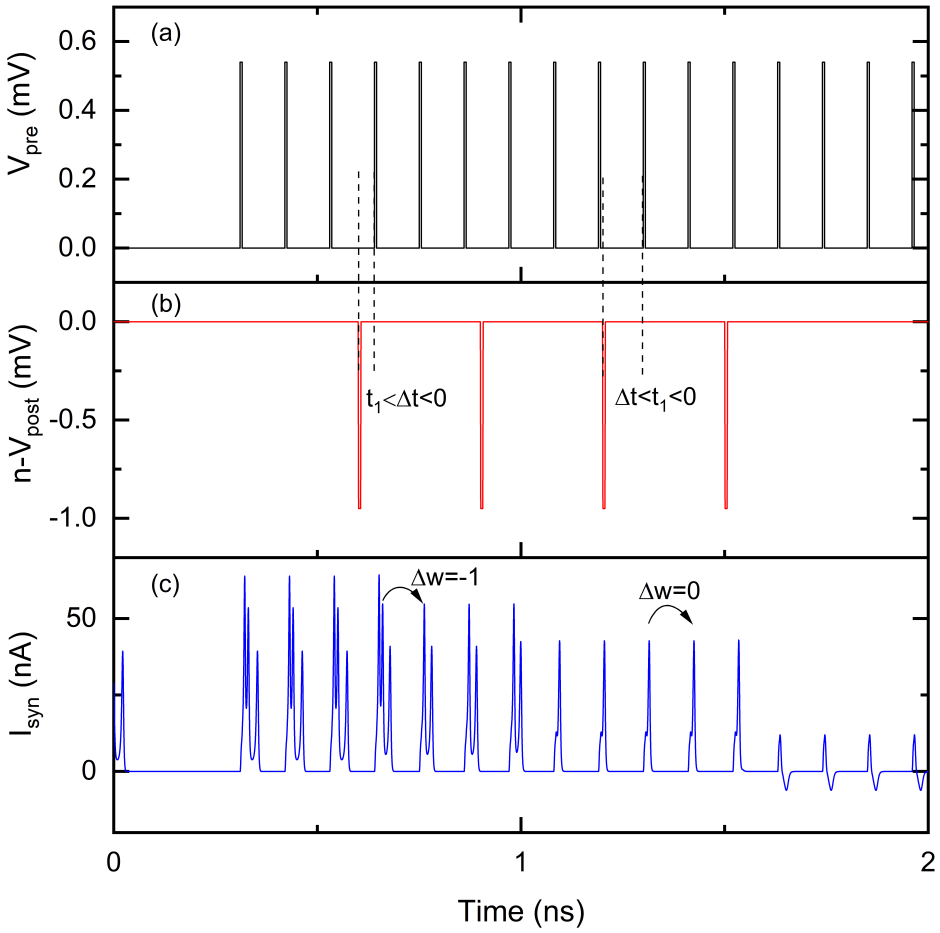


Figure 4.6: Simulation results of the circuit shown in Figure 4.5, from [16]. V_{pre} was 0.54 mV and $\overline{V_{\text{post}}}$ was 0.95 mV. The critical voltage values used for Q_0 to Q_7 were 0.75 mV, 0.55 mV, 0.3 mV, 2 mV, 0.54 mV, 0.52 mV, 0.5 mV and 0.34 mV, respectively. C_1 and C_3 were 9.2 fF, and C_2 and C_4 were 1.2 fF. V_{b1}, V_{b2} and V_{b3} were 0.03 mV, 0.77 mV and 0.53 mV, respectively. R_1 was 10 kΩ. (a) Input signal V_{pre}. (b) Input signal $\overline{V_{\text{post}}}$. (c) Output signal I_{syn}.

of electrons. This circuit behaves like an "inverter" circuit that can convert positive voltage (or current from an upstream neuron) pulses to negative current pulses. By choosing appropriate biasing conditions and critical voltage value of Q_4, we only allow Q_4 to be switched for a maximum of three times, which represents a maximum weight change of 3. Each time Q_4 is switched, a pair of electrons flow from C_3 to C_2 and voltage at node 3 drops by $2e/C_3$, which makes the synaptic weight increase by 1.

In Figure 4.8, we assume LTP is effective when $0 < \Delta t < t_2$, where $t_2 \approx 34$ ps is primarily determined by the resistance of R_1 in Figure 4.7. The width of LTP learning window

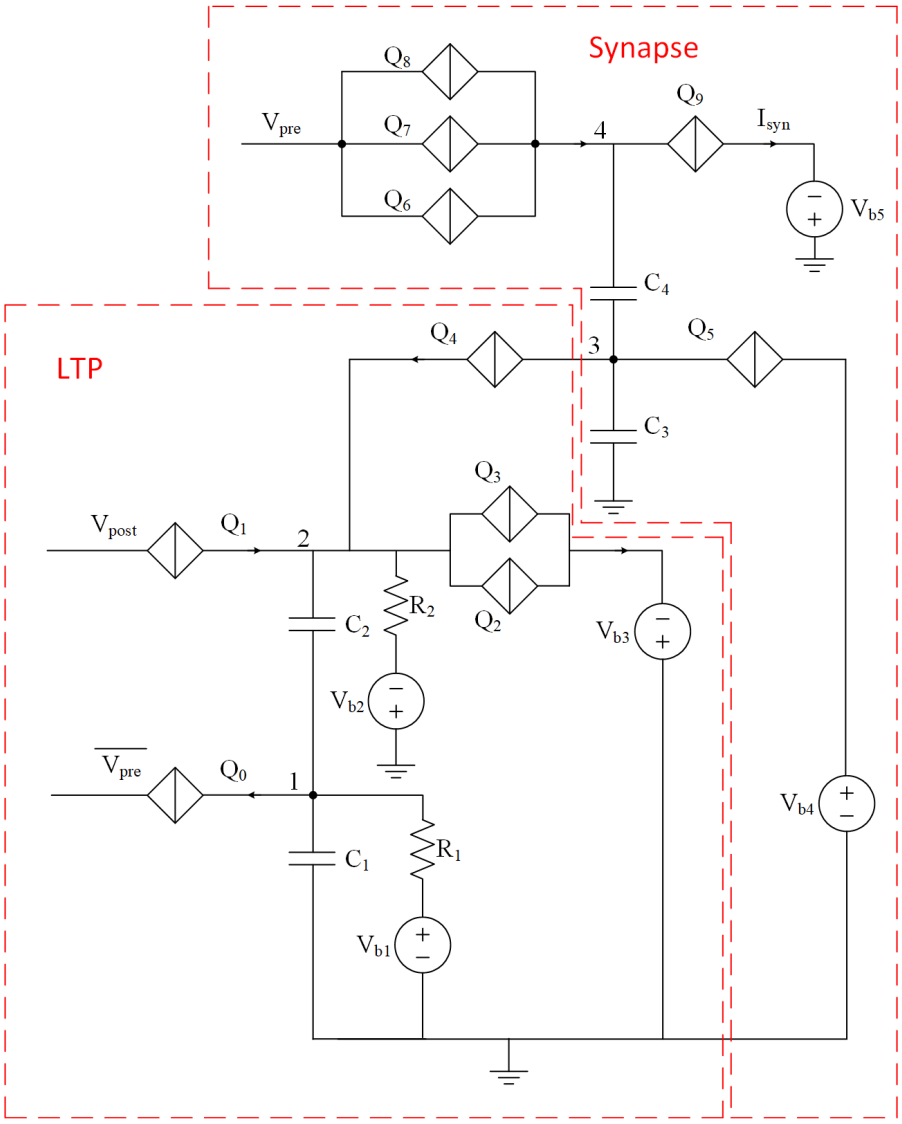


Figure 4.7: An LTP circuit with a multi-weight synaptic circuit, from [16]. The number of sequential pulses at I_{syn} can be increased when the timings of pulses from V_{pre} and V_{post} trigger an LTP learning event.

was not a concern during this simulation, as the circuit parameters were chosen to demonstrate LTP functions but not for a specific LTP learning window. The initial weight of the synapse was set to 0. Different periodic pulses were applied at V_{pre} and V_{post} in the simulation to demonstrate LTP learning. For example, the sixteenth pulse from V_{pre} is slightly ahead of the seventh pulse from V_{post}, which triggers LTP for the multi-weight synapse. As a result, the weight changes from 1 to 2. The upcoming pulse from V_{pre} can trigger two sequential current pulses at I_{syn}. However, the second pulse from V_{pre} has a relatively large time interval ($\sim$ 100 ps) with the first pulse from V_{post}, which does not trigger a weight change.

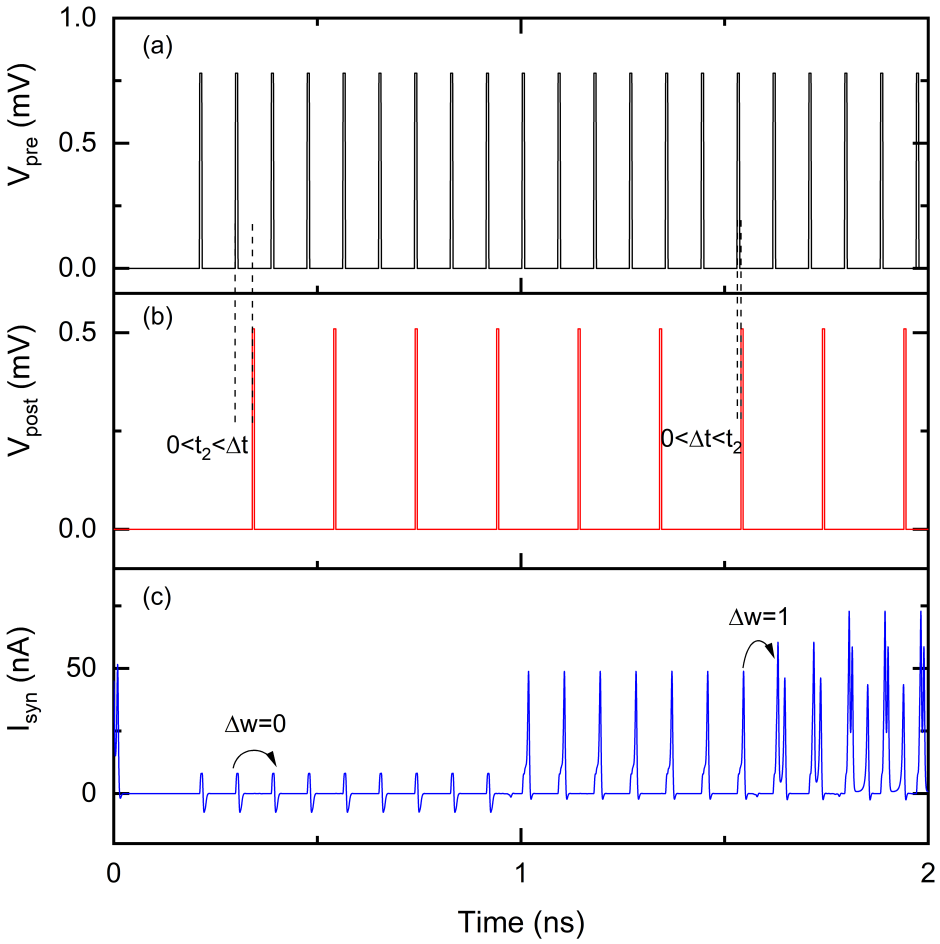


Figure 4.8: Simulation results of the circuit shown in Figure 4.7, from [16]. V_{pre} was 0.78 mV and $\overline{V_{\text{pre}}}$ was 0.54 mV. V_{post} was 0.51 mV. The critical voltage values used for Q_0 to Q_9 were 0.4 mV, 0.5 mV, 1 mV, 1 mV, 0.58 mV, 2 mV, 1.04 mV, 1.02 mV, 1 mV and 0.28 mV, respectively. C_1, C_2, C_3 and C_4 were 9 fF, 1 fF, 9.2 fF and 2 fF, respectively. V_{b1}, V_{b2}, V_{b3}, V_{b4} and V_{b5} were 0.05 mV, 0.2 mV, 1.1 mV, 1.01 mV and 0.6 mV, respectively. R_1 and R_2 were 10 kΩ. (a) Input signal V_{pre}. (b) Input signal V_{post}. (c) Output signal I_{syn}.

We replaced negative input voltage pulses from $\overline{V_{\text{pre}}}$ with positive input voltage pulses from V_{pre} in the circuit shown in Figure 4.9. This circuit contains another "inverter" circuit to convert positive voltage pulses from V_{pre} to negative current pulses. As we explained earlier, the "inverter" circuit can take electrons from capacitor C_2 to temporally reduce voltage at node 2. Like many other technologies, signals transmission and processing in QPSJ-based circuits exhibit delays. The extra "inverter" circuit in Figure 4.9 also adds extra delay. The learning window shifts by $t_0 \simeq 10$ ps and becomes $\sim t_0 < \Delta t < t_2 + t_0$, as shown in Figure 4.10.

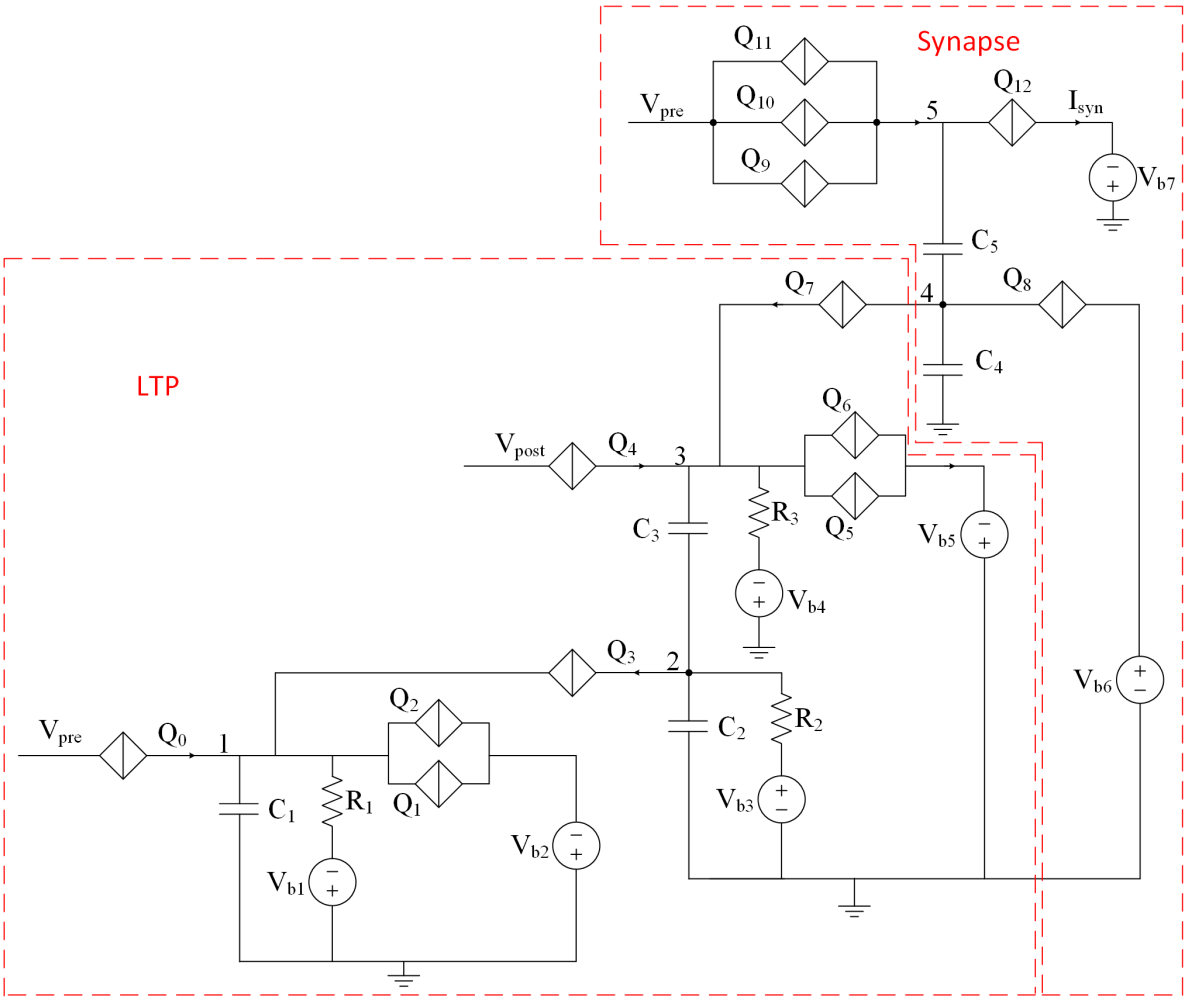


Figure 4.9: A modified LTP circuit with a multi-weight synaptic circuit, from [16]. The number of sequential current pulses at I_{syn} can be increased when the timings of pulses from V_{pre} and V_{post} trigger an LTP learning event.

Although the input signals are identical during the simulations, the output results of I_{syn} in Figure 4.10 are different from results in Figure 4.8. We observed that LTP occurs in Figure 4.10 where Δt is relatively large ($t_2 < \Delta t < t_2 + t_0$) but does not occur where Δt is very small ($0 < \Delta t < t_0$). Proper choice of resistance values and potentially adding a delay circuit (e.g., using a QPSJ transmission line circuit) for some of the input signals can adjust the LTP learning window to desired values.

4.5 A spike timing dependent plasticity circuit

A simplified STDP rule is illustrated in Figure 4.11. The synaptic weight is depressed if $t_1 < \Delta t < t_2$ while the synaptic weight is potentiated if $t_3 < \Delta t < t_4$. This simplified STDP rule

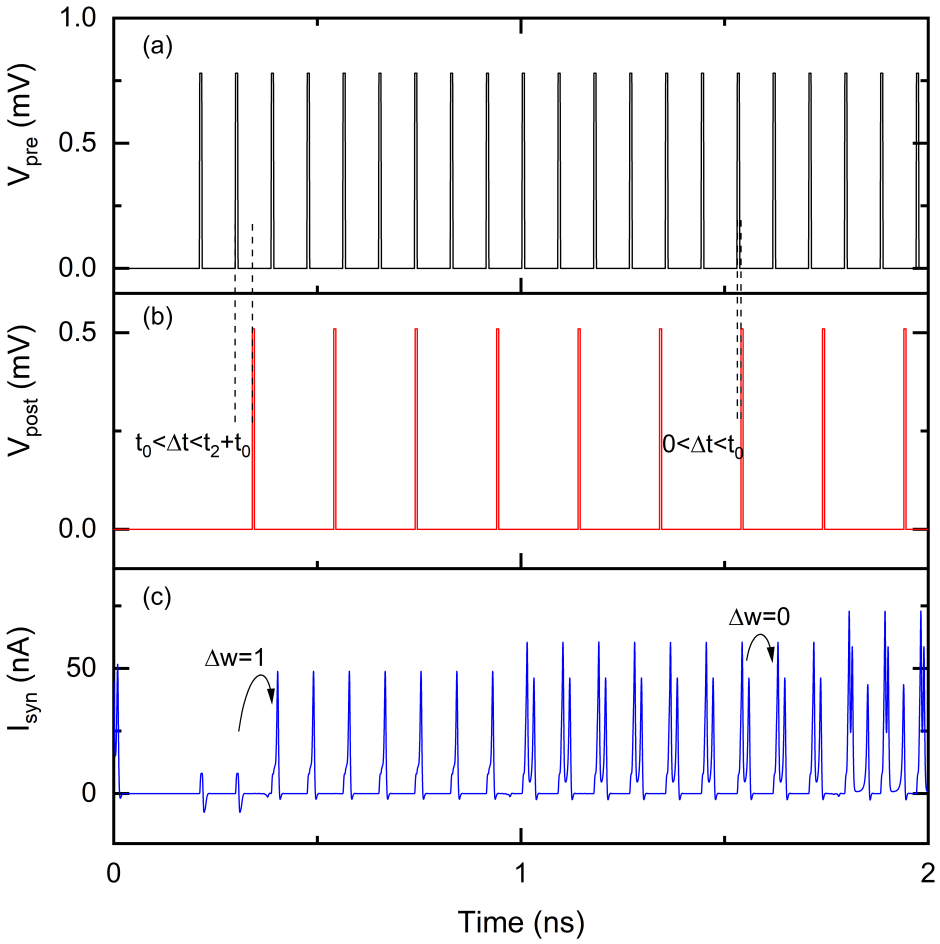


Figure 4.10: Simulation results of the circuit shown in Figure 4.9, from [16]. V_{pre} was 0.78 mV and V_{post} was 0.51 mV. The critical voltage values used for Q_0 to Q_{12} were 0.8 mV, 0.95 mV, 0.95 mV, 0.36 mV, 0.5 mV, 1 mV, 1 mV, 0.58 mV, 2 mV, 1.04 mV, 1.02 mV, 1 mV and 0.28 mV, respectively. C_1, C_2, C_3, C_4 and C_5 were 1 fF, 9 fF, 1 fF, 9.2 fF and 2 fF, respectively. V_{b1}, V_{b2}, V_{b3}, V_{b4}, V_{b5}, V_{b6} and V_{b7} were 0.2 mV, 1.1 mV, 0.05 mV, 0.2 mV, 1.1 mV, 1.01 mV and 0.6 mV, respectively. R_1, R_2 and R_3 were 10 kΩ. (a) Input signal V_{pre}. (b) Input signal V_{post}. (c) Output signal I_{syn}.

can be realized by combining the LTD and LTP circuit, as shown in Figure 4.12. Charge is injected onto or taken from capacitor C_6, resulting in a weight depression or potentiation for the multi-weight synapse. The LTD portion has an additional bias voltage V_{b8} and a resistor R_5 to provide voltage bias for Q_{10} and Q_{11}, which is different from the original LTD circuit shown in Figure 4.3. The LTD and LTP learning windows, defined by $t_1 < \Delta t < t_2$ and $t_3 < \Delta t < t_4$, can be adjusted by choosing appropriate circuit parameters or adding extra circuit components for signal delays. The simulation results of this circuit are shown in Figure 4.13.

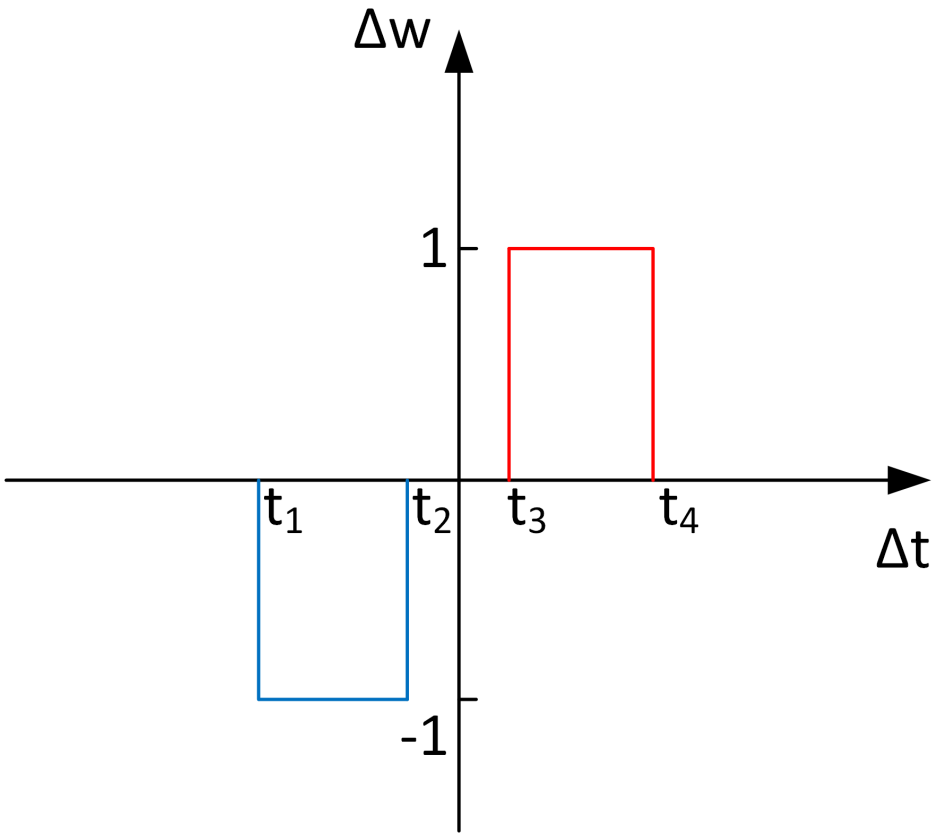


Figure 4.11: A simplified STDP learning rule. LTD occurs if $t_1 < \Delta t < t_2$ while LTP occurs if $t_3 < \Delta t < t_4$.

We use a customized spike train applied at V_{pre} and a periodic spike train applied at V_{post} and $\overline{V_{\text{post}}}$ to demonstrate STDP learning functionality. The initial weight of the synapse is set to 0. At the beginning of this simulation, no current pulses are presented at I_{syn} when applying voltage pulses at V_{pre}. As synaptic weight changes according to the relative timings of presynaptic pulses and postsynaptic pulses, the output current pulses at I_{syn} also change over time. Specifically, both presynaptic and postsynaptic voltage pulses are transmitted to the LTD and LTP units. However, using the specific device parameter values during this simulation, the LTD unit only generates depression pulses to the synaptic circuit if -10 ps $< \Delta t < -2$ ps. The LTP unit only generates potentiation pulses to the synaptic circuit if 16 ps $< \Delta t <$ 41 ps. These results demonstrate the simplified learning rule realized by this STDP circuit.

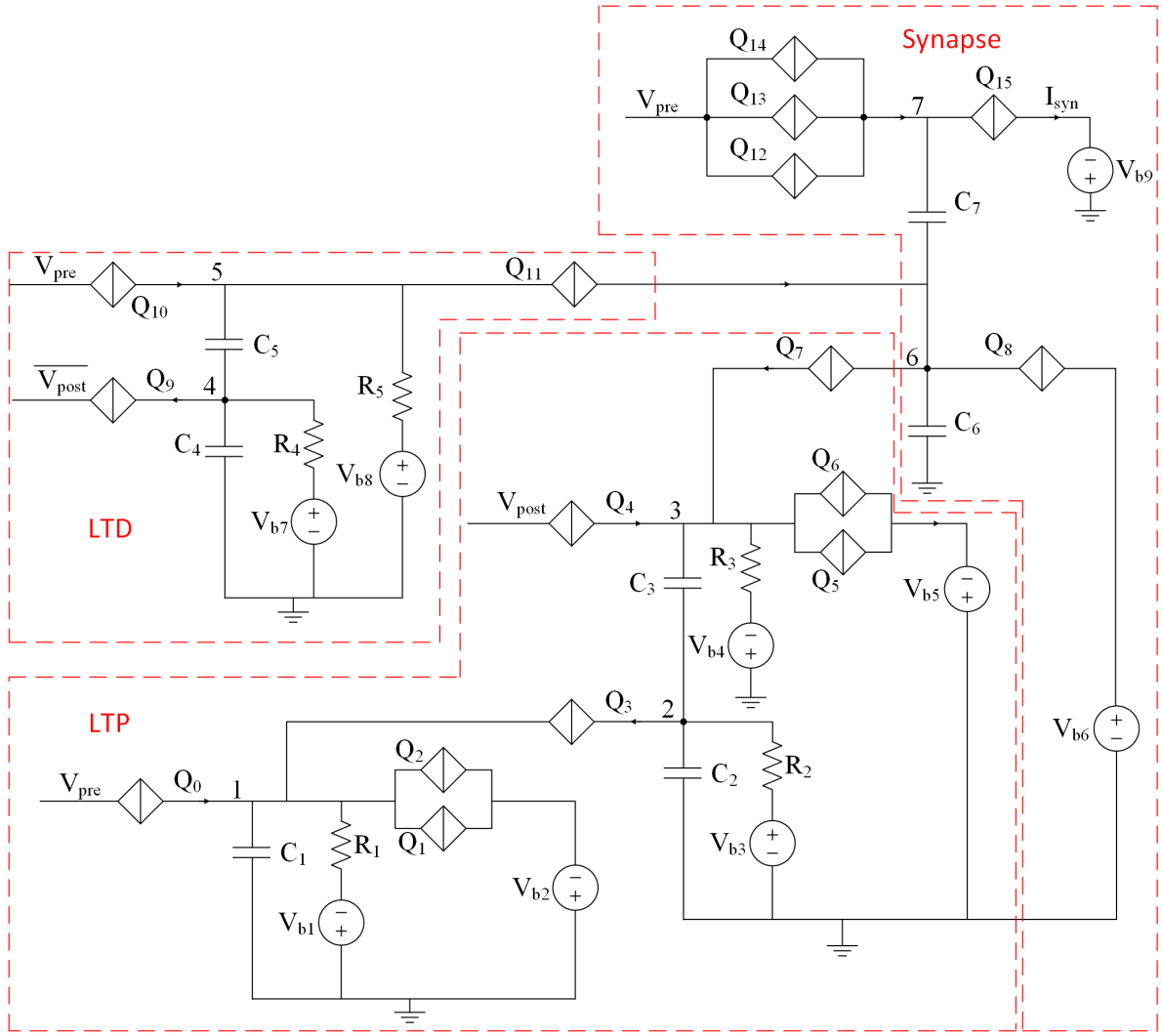


Figure 4.12: An STDP circuit with a multi-weight synaptic circuit, from [16]. The STDP circuit is comprised with an LTD circuit and an LTP circuit. The number of sequential current pulses at I_{syn} can be updated according to the timings of pulses from V_{pre} and V_{post}.

4.6 A modified LTD circuit towards a more complex STDP learning rule

The proposed STDP learning circuit has the ability to update synaptic weight by 1 or -1 during each learning event. However, for more complex computation tasks, more synaptic weight states are desired. In this case, the weight change per learning event is expected to be more than 1 unit. We designed a modified LTD circuit towards a more complex STDP learning rule while the modified STDP circuit is yet to be designed at this point. The modified LTD circuit can update weight up to 3 per learning event, as shown in Figure 4.14. The initial voltage at node 1 is set by the bias voltage V_{b1} when there are no inputs at $\overline{V_{\text{post}}}$. The critical voltages of Q_3, Q_4 and Q_5 are V_{C3}, V_{C4} and V_{C5} respectively while $V_{\text{C3}} < V_{\text{C4}} < V_{\text{C5}}$. QPSJ Q_0, Q_1

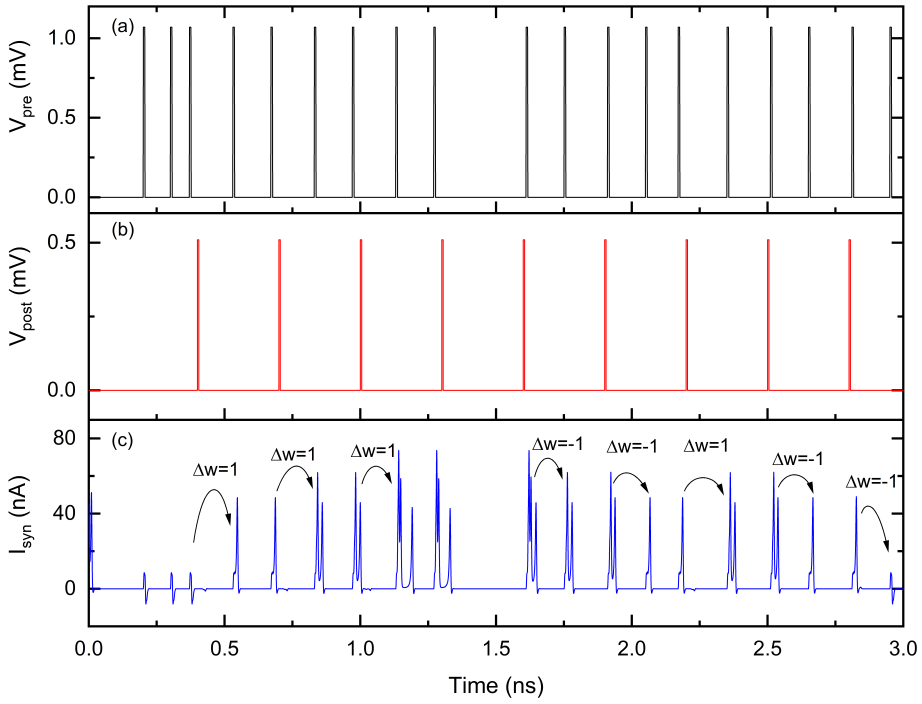


Figure 4.13: Simulation results of the circuit shown in Figure 4.12, from [16]. V_{pre} was 1.07 mV. V_{post} was 0.51 mV and $\overline{V_{\text{post}}}$ was 0.51 mV. The critical voltage values used for Q_0 to Q_{15} were 0.8 mV, 0.95 mV, 0.95 mV, 0.36 mV, 0.5 mV, 1 mV, 1 mV, 0.46 mV, 2 mV, 0.75 mV, 0.55 mV, 0.3 mV, 1.37 mV, 1.35 mV, 1.33 mV and 0.28 mV, respectively. C_1, C_2, C_3, C_4, C_5, C_6 and C_7 were 1 fF, 9 fF, 1 fF, 9.2 fF, 1.2 fF, 9.2 fF and 2 fF, respectively. V_{b1}, V_{b2}, V_{b3}, V_{b4}, V_{b5}, V_{b6}, V_{b7}, V_{b8} and V_{b9} were 0.2 mV, 1.1 mV, 0.05 mV, 0.2 mV, 1.1 mV, 0.89 mV, 0.46 mV, 0.3 mV and 0.6 mV, respectively. R_1, R_2, R_3, R_4 and R_5 were 10 kΩ, 10 kΩ, 20 kΩ, 10 kΩ and 20 kΩ, respectively. (a) Input signal V_{pre}. (b) Input signal V_{post}. (c) Output signal I_{syn}.

and Q_2 are identical. In this design, if a voltage pulse from V_{pre} arrives first, none of Q_3 to Q_5 can be switched by V_{pre}. Therefore, no current pulses are generated at I_{learning}. However, if a negative voltage pulse from $\overline{V_{\text{post}}}$ arrives first, Q_0 to Q_2 are switched and three pairs of electrons are taken from capacitor C_1. The voltages at node 1 and node 2 drop by $6e/C_1$, where C_1 is the capacitance of capacitor C_1. The voltage change at node 2 allows the upcoming voltage pulse from V_{pre} to switch Q_3, Q_4 and Q_5 and in turn switch Q_6, resulting in three sequential current pulses at I_{learning}. The voltage at node 1 and node 2 will recover to their initial states gradually. If the relative timings of V_{pre} and $\overline{V_{\text{post}}}$ are very small (-34 ps $< \Delta t < -10$ ps), all three QPSJs Q_3, Q_4 and Q_5 can be switched upon the arrival of a voltage pulse at V_{pre}. However, if the voltage pulse arrives at V_{pre} later, only Q_3 and Q_4 can be switched. There are only two sequential current pulses at I_{learning}. Depending on the relative timings of V_{pre} and

$\overline{V_{\text{post}}}$, different numbers of sequential current pulses can be generated at I_{learning}. The resulting learning window resembles a quantized STDP curve and also can be adjusted by the resistance value of R_1.

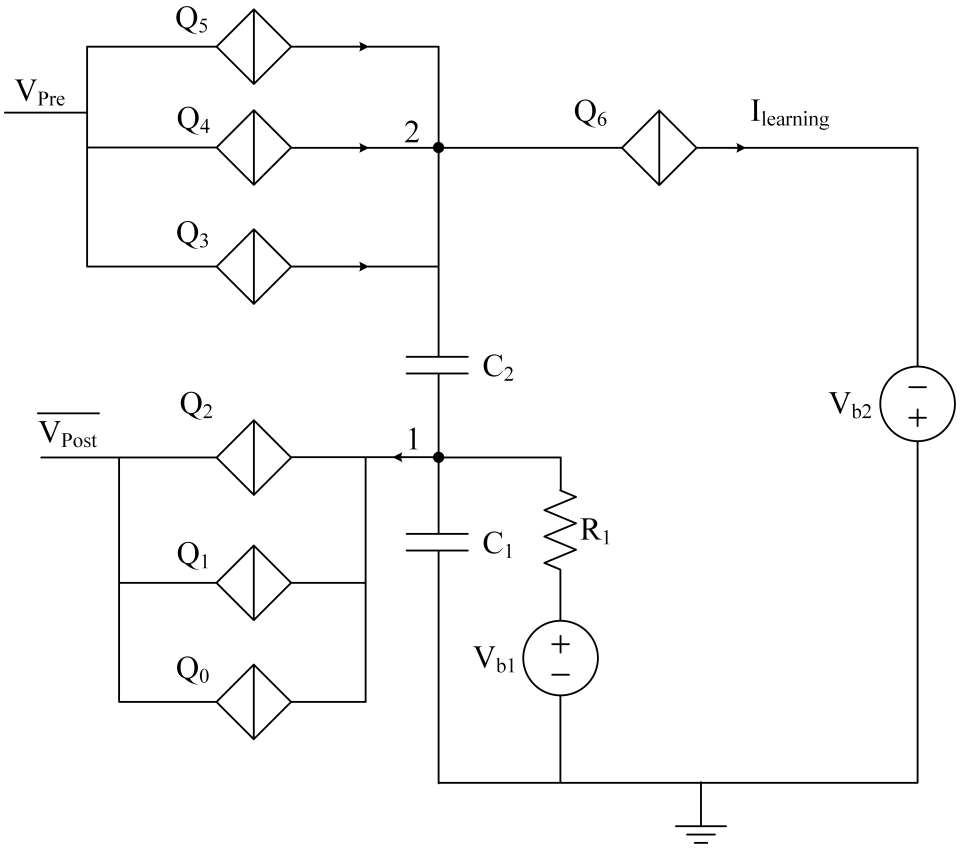


Figure 4.14: A modified LTD circuit that generates depression pulses to a synapse. The number of sequential current pulses generated at I_{learning} depends on the the timing difference $\Delta t = t_{\text{post}} - t_{\text{pre}}$.

In Figure 4.15, we show the comparison of different STDP learning rules. The complex STDP learning rule is based on the modified LTD circuit. We investigated the number of sequential current pulses at I_{learning} by setting different Δt in the simulation with a step size of 1 ps. The results at $\Delta t > 0$ are a mirror of the simulation results at $\Delta t < 0$ while no actual LTP circuit was simulated. The resistance value of R_1 is chosen to be 10 kΩ, which allows voltages at node 1 and 2 to recover fast. Therefore, the resulting learning curve resembles a quantized STDP curve compared to the biological STDP shown in Figure 1.8. Based on this design scheme, a complex STDP circuit is expected to be designed and simulated to realize the learning curve shown in Figure 4.15.

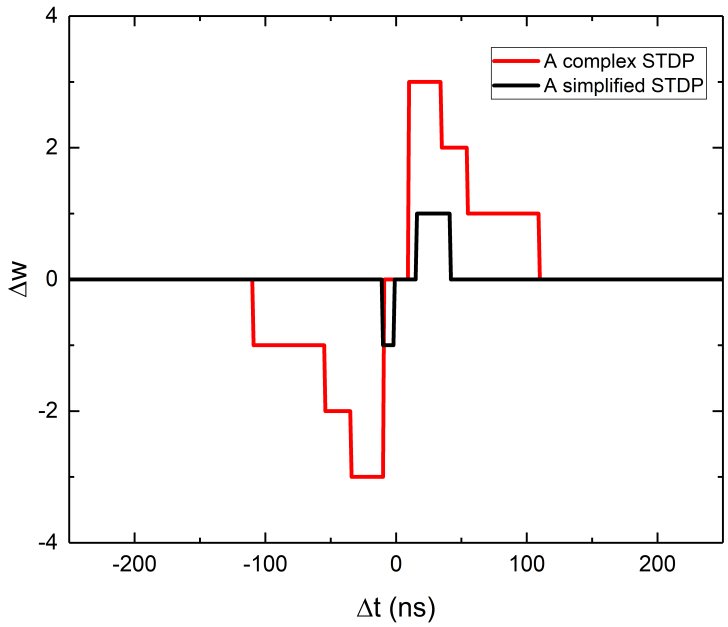


Figure 4.15: A comparison of different STDP learning rules. The complex STDP learning rule is based on the modified LTD circuit shown in Figure 4.14. The results at $\Delta t > 0$ are a mirror of the simulation results at $\Delta t < 0$ while no actual LTP circuit was simulated.

4.7 Summary

In this chapter, we proposed a simplified STDP learning circuit based on QPSJs. This circuit is designed to update synaptic weight automatically according to the timings of presynaptic and postsynaptic spikes. The weight change per learning event is either 1 or -1, which is compatible with previously designed multi-weight synaptic circuit. The simplified STDP learning circuit is consisting of an LTD circuit and an LTP circuit, which contributes to weight depression and weight potentiation, respectively. The functionality of the STDP circuit, incorporating with a multi-weight synaptic circuit, has been demonstrated in WRspice simulations. We also presented a modified LTD learning circuit. This circuit is more flexible since it can update the synaptic weight up to 3 per learning event. A quantized STDP learning curve has been generated based on this modified LTD circuit.

The simplified learning rule presented in this paper aims to provide a simple learning method to update synaptic weights according to relative timings of presynaptic and postsynaptic

pulses, but has interesting differences compared to its biological counterpart. One aspect is that the superconducting circuit processes information for signals with pulse rates in the tens of GHz scale, which is many orders of magnitude faster than a human brain that typically operates at tens of Hz. Another aspect is the effective learning window for a circuit in Figure 4.12 is -10 ps to -2 ps for LTD and 16 ps to 41 ps for LTP using the specific parameters in this simulation. Though this learning window does not have the exact shape of a more realistic STDP, it may still be useful for implementation to solve practical problems. We also note that the learning window can be adjusted by slightly modifying the STDP circuit in Figure 4.12, in addition to what we mentioned earlier to fix delay issues noted in this dissertation. For example, adding QPSJs in parallel with Q_9 and increasing the resistance of R_4 could extend the effective learning window for LTD. We have not yet combined input and output neuron circuits, synaptic circuit, fan-out circuit and STDP circuit to demonstrate a large network application. While voltage biasing in QPSJ-based circuits has advantages, as circuit sizes grow and become more complex, challenges related to biasing and impedance matching will likely become more critical [15]. We believe that these challenges, which are also found as challenges in other technologies (e.g., current distribution in large JJ-based circuits), do have engineering solutions and require additional work. We also note that these solutions may exist as trade-offs with circuit operation speed and may impact the overall power or energy efficiency. Circuit modifications and new circuit configurations to realize interconnection circuits for synapse feedback loops may also be needed. These aspects are expected to be the focus of potential improvements in future studies. In the next chapter, we will discuss the advantages of QPSJ-based circuits in terms of power dissipation and processing speed.

Chapter 5

Power Dissipation, Processing Speed and Experimental Challenges

5.1 Power dissipation and processing speed

Power dissipation and switching speed are important considerations for densely integrated electronics. As superconducting charge-based circuits, QPSJ neuron circuits present the opportunity for very low power dissipation per switching event and with similar switching speed compared to Josephson junction-based circuits. An additional advantage of these circuits is nearly zero static power dissipation, as the junctions will be in a Coulomb blockade condition when they are biased below their critical voltage. Given the relatively new and unexplored situation of both QPSJ-based and JJ-based neuron circuits, a complete characterization of their power dissipation and switching speeds is challenging to perform at this time. Nonetheless, in this section, we present estimations based on device switching to compare the power dissipation and switching speed of single-neuron cells for both JJ and QPSJ-based circuits. The power per switching event for a JJ circuit is given by [132]:

$$P = \frac{I_{\mathrm{b}}\Phi_0}{t_1 + t_2} \tag{5.1}$$

$$t_1 = RC \tag{5.2}$$

$$t_2 = \frac{L_{\mathrm{k}}}{R} \tag{5.3}$$

where t_1 is the charging time of intrinsic capacitor, R is normal resistance, C is intrinsic capacitance, t_2 is discharging time of intrinsic capacitor, L_{k} is kinetic inductance, P is power, I_{b} is bias current and R_{p} is parallel resistance.

For a practical QPSJ using InO_x material, the power per switching event and switching speed for given dimensions of a nanowire can be estimated using the phase-slip energy model developed by Mooij et al. [119], which can be summarized as:

$$P = \frac{V_{\mathrm{b}} 2e}{t_1 + t_2} \tag{5.4}$$

$$t_1 = \frac{L}{R} \tag{5.5}$$

$$t_2 = RC_{\mathrm{k}} \tag{5.6}$$

where t_1 is the charging time of kinetic capacitor, R is normal resistance, L is intrinsic inductance, t_2 is discharging time of kinetic capacitor, C_{k} is kinetic capacitance, P is power and V_{b} is bias voltage. For both technologies, the switching delay is determined by the charging or discharging time of the capacitor, whichever is larger. Results from this analysis are shown in Figure 5.1 and Figure 5.1. A QPSJ that has a V_{C} of 0.7 mV has a switching energy of $\sim 2eV_{\mathrm{C}}$, i.e., $\sim$ 0.224 zJ. Therefore, the energy per spike is approximately 0.45 zJ (two switching events) for a QPSJ-based neuron shown in Figure 3.1, compared with approximately 0.33 aJ (one switching event) for a typical JJ-based neuron [44].

Some additional comments about this analysis and associated assumptions are warranted. The parameters chosen for QPSJ and JJ technologies are typical values from published experimental results for devices at their nominal operating temperatures. The presented power dissipation numbers do not take into account the device operating temperature. Classical JJ-based circuits usually operate at a temperature of 4.2 K. Quantum phase-slip events have been experimentally observed at temperatures up to 700 mK [133] and recent experiments report QPS phenomena in NbN nanowires at a temperature of 1.92 K [56]. These results support the possibility of realizing useful QPS events at higher temperatures, perhaps up to and beyond 4.2 K, with proper choice of materials and refined fabrication processes. Furthermore, comparison to CMOS-based implementations provides additional insight. In CMOS neuromorphic systems, IBM's digital neuron consumes 45 pJ per spike [134] and even the most recent neuron circuits can currently only reach several fJ/spike [135]. While these results are estimates based

on assumed material properties, they show promise for QPSJ-based neuromorphic circuits and logic that may exhibit competitive power-delay properties compared to other solid-state technologies.

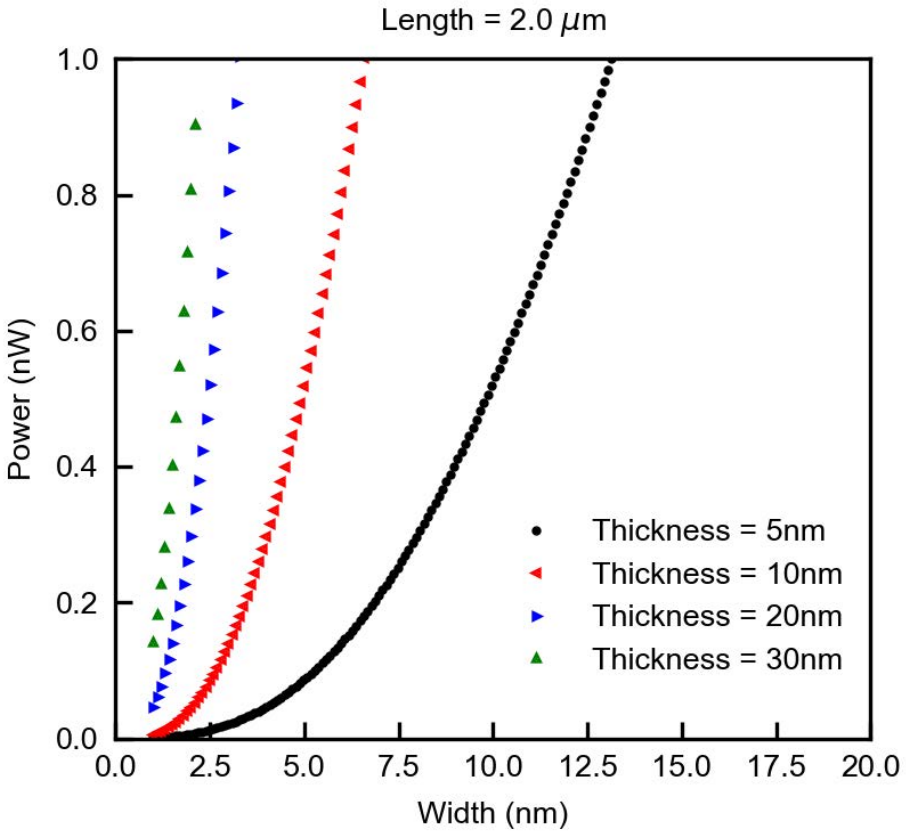


Figure 5.1: Power dissipated per switching event in an InO_x QPSJ, from [13]. Values were calculated based on a model by Mooij et al. for a nanowire of length 2 μm.

5.2 Power dissipation analysis

For QPSJ circuits that do not have dissipative components (e.g., resistors), the main energy dissipation is due to switching energies of multiple QPSJs, which are extremely small. Some of the learning circuits we presented in the previous chapter have resistors in series with bias voltage, also consuming low energy. We performed simulations to find out the energy dissipation among different voltage sources under different operating conditions.

We consider the power dissipation under three different conditions for the circuit shown in Figure 4.3. In the case of no inputs from V_{pre} or $\overline{V_{\text{post}}}$, this circuit will adapt to a resting state

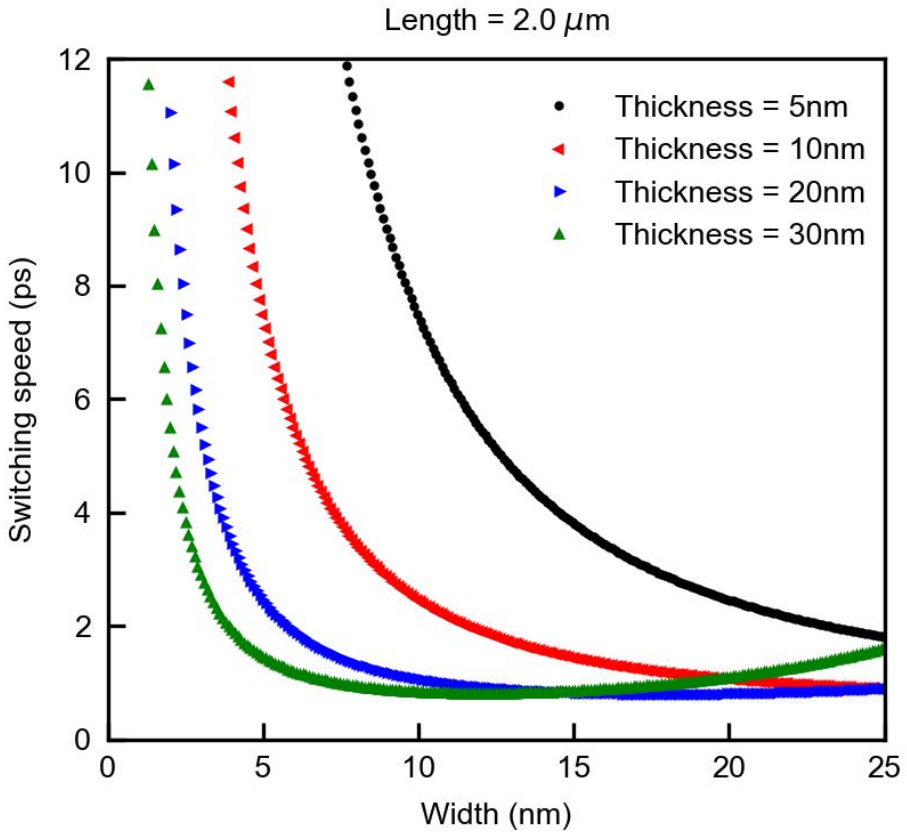


Figure 5.2: Switching speed in an InO_x QPSJ, from [13]. Values were calculated based on a model by Mooij et al. for a nanowire of length 2 μm.

after initialization. No power is generated from all the voltage sources, as shown in Figure 5.3. If there is a voltage pulse from $\overline{V_{\text{post}}}$, we can observe dynamic power consumption at P_{post} and P_{vb1} in Figure 5.4, which corresponds to the energy dedicated to Q_0 and energy consumed at resistor R_1, respectively. The learning happens when a pulse from $\overline{V_{\text{post}}}$ is slight ahead of a pulse from V_{pre}. In this case, the power dissipation is presumed to be the maximum for this circuit operation. We can observe dynamic power dissipation at each input and output voltage source, as shown in Figure 5.5. The total energies dissipated are calculated by the integral of dynamic powers over an active period. We estimate that the energies dissipated at V_{pre}, $\overline{V_{\text{post}}}$, V_{b1} and V_{b2} are 79.8 yJ, 134 yJ, 9.45 yJ, and 159 yJ, respectively. These simulation results suggest that the power dissipation in QPSJ-based circuits is expected to be extremely low, and QPSJ technology could be a good candidate for low-power neuromorphic computing.

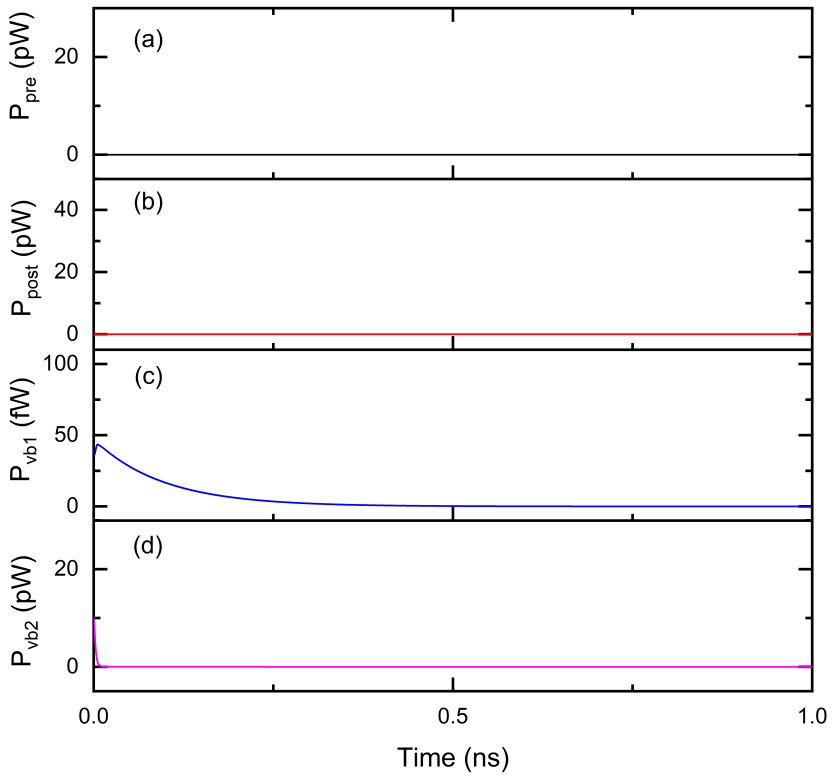


Figure 5.3: Power consumption of voltage sources in the circuit shown in Figure 4.3 when no inputs are presented. (a) Power from V_{pre}. (b) Power from $\overline{V_{\text{post}}}$. (c) Power from V_{b1}. (d) Power from V_{b2}.

5.3 Experimental challenges and scalability

We reiterate that the work described here is simulation based, with goals of exploring neuromorphic circuits based on QPSJs and motivating future experimental work. In this section, we first discuss experimental challenges. Then we discuss scalability and tolerance of these QPSJ-based circuits. Finally, we propose potential improvement for QPSJ-based neuromorphic circuits in our future design and simulation.

The circuits presented here have been designed and analyzed based on the ideal model. As stated previously, the experimental realization of QPSJ-based circuits is in its early stages and presents multiple challenges. Charge based circuits are subject to interactions with stray charges and charge fluctuations in the local environment. Random offset charge may affect the operation of Coulomb blockade devices such as single electron transistors, quantum dots,

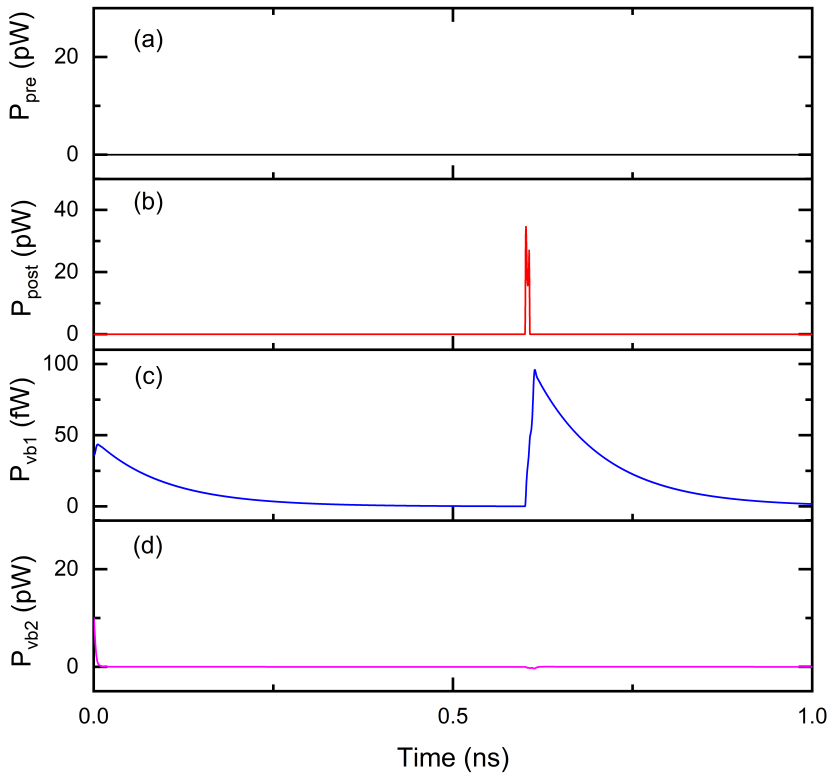


Figure 5.4: Power consumption of voltage sources in the circuit shown in Figure 4.3 when there is an input pulse at $\overline{V_{\text{post}}}$. (a) Power from V_{pre}. (b) Power from $\overline{V_{\text{post}}}$. (c) Power from V_{b1}. (d) Power from V_{b2}.

etc. [136]. Similar issues are expected to impact operation of QPSJ-based circuits. This issue has been considered previously and one of the possible methods to solve the offset charge problem is to add a tunable gate voltage on the island for charge modulation [137, 136, 103], though this results in a more complex circuit and requires more control signals. Furthermore, superconductive QPSJ-based technology may experience fan-in and fan-out challenges when the circuits are scaled up. This is evident in JJ-based technologies, as well [94]. In particular, for QPSJ-based neuron circuits, fan-in is highly dependent on the integration (i.e., membrane) capacitance, while fan-out is determined by the number of QPSJ splitter circuits that can be driven [57]. The challenge is how to guarantee that each QPSJ in a complex circuit has appropriate but generally not identical bias voltage. Additionally, there is a possible need for isolation of these circuits with high impedances as suggested in references [56] and [138].

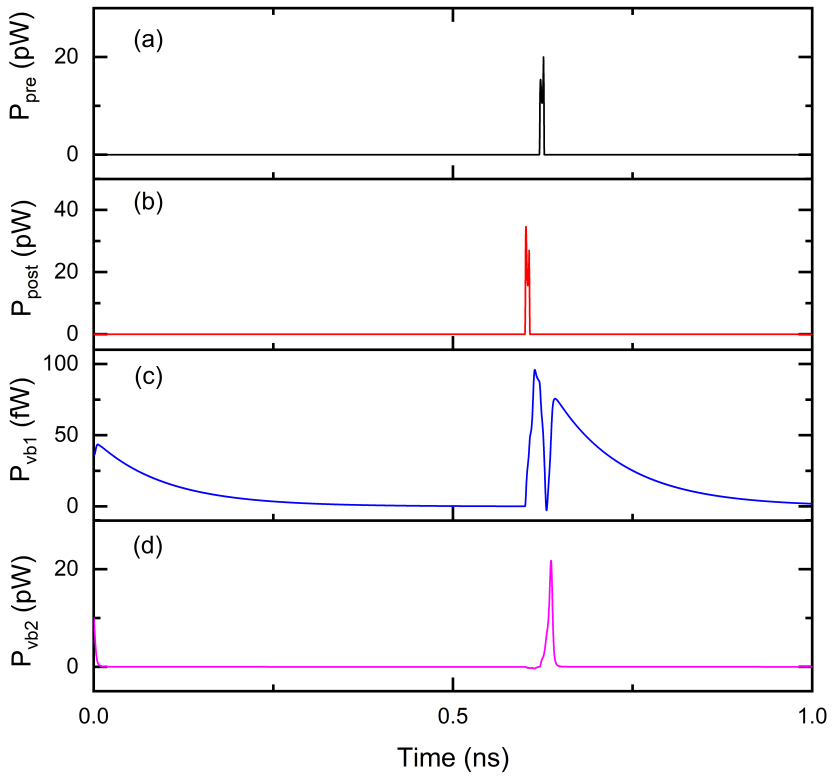


Figure 5.5: Power consumption of voltage sources in the circuit shown in Figure 4.3 when there are input pulses at V_{pre} and $\overline{V_{\text{post}}}$. (a) Power from V_{pre}. (b) Power from $\overline{V_{\text{post}}}$. (c) Power from V_{b1}. (d) Power from V_{b2}.

These impedances are neglected in this study, as their effects on operation of the circuits are not expected to change significantly. It is interesting to note that this challenge is somewhat similar to JJ current biasing in single flux quantum circuit technologies [93].

A growing number of research is being performed in this area to discover appropriate materials and device dimensions for a functional coherent quantum phase-slip. However, the fabrication of reproducible and controllable QPSJs is still at an early stage where the conditions to create and control a coherent quantum phase-slip are still under investigation. Although the DC I-V characteristics of a QPSJ have been experimentally observed as exact duals to a JJ, the AC characteristics of a QPSJ (i.e., dual Shapiro steps) have not been experimentally demonstrated. Charge fluctuation and charge noise are believed to be a great challenge for charge-based circuits.

The neuromorphic systems we present here have the ability to scale up without limitations in theory. For area scalability, each nanowire is expected to be only a few nanometers wide and a few micrometers long (and potentially in a meandered configuration). A large number of QPSJs can fit into a small chip area. Therefore, scalability is currently assumed to be a benefit of using QPSJ-based circuits. Biasing is a concern of QPSJ-based circuits, like many other superconducting circuits. Both voltage-biasing and current-biasing have advantages and disadvantages in the context of superconducting circuits. Supplying current and distributing it properly are known issues for JJ-based, SFQ-like circuitry, and are still an active area of research and development. The voltage bias aspect of QPSJ-based circuits has a sense of similarity to CMOS. We envision voltage biasing for QPSJ-based circuits that can be achieved using voltage regulators and/or resistor networks (perhaps at higher temperature stages in a cryostat). The actual number of required voltages is dependent on the circuit and complexity of the network and is an aspect that we are currently studying. Furthermore, using one voltage source to bias multiple QPSJs or logic blocks in series can potentially reduce the bias complexity. On the other hand, using one bias voltage for multiple QPSJs may require very finely tuned parameters for each QPSJ and may potentially reduce circuit tolerance. To enhance scalability, a large network may be broken into multiple smaller blocks where one or a small number of bias voltages can be used for each block. For example, we can bias one neuron and multiple synapses with only one bias voltage and the downstream neuron and synapses, separated by a fan-out circuit, use another bias voltage. One interesting aspect that is implied here is potentially being able to use voltage bias for some portions of the network (QPSJ-based pieces) and current bias for other portions (JJ-based pieces).

The operating margin of QPSJ-based circuits in this work is not only determined by the design itself but also the practical device performance. For example, the practical critical voltage was found to be smaller than 100 μV in [139], but up to several mV in [56]. We take the circuit in Figure 3.17 as an example to analyze each device tolerance in this circuit. The tolerance of the circuit shown in Figure 3.17 highly depends on the critical voltage variation of each QPSJ. The critical voltage variation of parallel QPSJs Q_5 to Q_8 in Figure 3.17 should not be greater than $2e/C_1$ in order to guarantee parallel QPSJs are switched at the same time. If capacitance

C_1 is 10 fF, $2e/C_1$ is $\sim$ 32 μV, resulting in a maximum tolerance of 16% for QPSJs with 100 μV critical voltage. However, it only allows 1.6% maximum tolerance for QPSJs with 1 mV critical voltage. Ideally, the critical voltage difference between Q_1 and Q_2 or Q_2 and Q_3 should be the same as the voltage change on node 1 after an input pulse from V_{w}, which is $\sim 2e/C_1$. The critical voltage variation of Q_4 should not be greater than $2e/C_2 + 2e/C_1$. The critical voltage of Q_0 has more tolerance as long as it can be switched by the input pulses from V_{w}. In general, in the simulation shown in Figure 3.18, we have found that a tolerance of less than 1% is required for critical voltage for parallel QPSJs (Q_1 to Q_3, Q_5 to Q_8) in the synaptic circuits. The tolerance of Q_4 and Q_0 are 7% and 40%, respectively. We also found that the tolerances of different QPSJs in this circuit are relatively independent. Varying the critical voltage value of one QPSJ does not significantly affect the tolerances of the other QPSJs. We note that this level of tolerance is also dependent on the nominal value of the varied parameter and is specific to the particular circuit. Developing a better understanding of the impact of device parameter mismatch and non-uniformity is critical to scaling this technology and is an ongoing effort.

The issues mentioned here are the challenges that researchers in this area are currently facing and we look forward to solutions through future research.

5.4 Summary

In this chapter, we introduced the equations used to estimate the power dissipation and processing speed of a single QPSJ. A QPSJ has nearly zero static power dissipation due to Coulomb blockade and has extremely low switching energy compared to other technologies. Having dissipative components in a QPSJ circuit does not significantly increase power dissipation, as demonstrated in the simulation of a typical learning circuit. This is because the current flows in QPSJ circuits are extremely low and are at a level of nA. The switching delay of a typical QPSJ ranges from a few ps to tens of ps, which suggests a processing speed up to hundreds of GHz. In theory, the QPSJ technology has a similar processing speed to the JJ technology but consumes less power and is promising for neuromorphic computing. To further explore the applications of QPSJ-based neuromorphic circuits, we addressed the experimental and design

challenges related to QPSJ technology. Another concern is the scalability of QPSJ-based neuromorphic circuits. The circuits, while configured properly, have the ability to scale up without any limitations. This requires an appropriately designed voltage biasing scheme for QPSJs along with improved device tolerance. In conclusion, the QPSJ technology is a promising candidate for high-speed and low-power applications, while many unsolved experimental challenges will continue to be a highlighted research topic for this technology. In the next chapter, we will introduce the method used to fabricate ultra narrow superconducting NbN nanowires in search of QPS phenomenon and show the experimental results and mathematical analysis of selected NbN nanowires.

Chapter 6

NbN Nanowire Fabrication and Characterization

6.1 Materials and methods

6.1.1 Introduction

The physical implementation of QPSJs requires the fabrication of ultra-narrow superconducting nanowires that have cross-section dimensions approaching the materials' coherence length. In this work, we focus on exploring the underlying conditions for ultra-narrow superconducting nanowires to show evidence of QPS. We considered InO_x, NbTiN, and NbN as highly disordered superconducting materials that could be good candidates for QPSJs. Further analysis of these materials regarding phase-slip energy for a QPSJ will be discussed. We chose NbN and performed experiments to optimize NbN sputtering process for the fabrication of NbN nanowires. A lift-off method is introduced to fabricate NbN nanowires using electron beam lithography (EBL). Selected NbN nanowires were characterized at temperatures down to 1.5 K.

6.1.2 Materials and device dimensions

We estimated that InO_x, NbN, and NbTiN are some of the suitable materials for QPSJs, due to their high normal resistivity (at low temperatures) and high phase-slip potential energy. We use equations provided in [119] to estimate the phase-slip energies E_S and the appropriate range of device dimensions.

$$E_S = a\frac{A}{\xi}k_B T_C \frac{R_q}{R_\xi}\exp(-b\frac{R_q}{R_\xi}) \tag{6.1}$$

where A is the length of the wire, ξ is the coherence length, k_{B} is the Boltzmann constant, T_{C} is the critical temperature, R_{q} is the quantum resistance and R_{ξ} is the resistance per coherence length, a and b are unknown constants. Parameters ξ, T_{C} and R_{ξ} are material dependent, while R_{ξ} is also dependent on device dimensions. The range of suitable physical dimensions for a particular material was discussed in [119]. The phase slip rate is determined by the ratio of $E_{\mathrm{S}}/E_{\mathrm{L}}$, where E_{L} is the inductive energy. The ratio of $E_{\mathrm{S}}/E_{\mathrm{L}}$ between the range 0.1 and 1 is put forth as a good estimate for crossover between superconducting and phase-slip behavior, which is desired for QPSJs. The typical values for these materials are shown in Table 6.1. Using practical values extracted from experiments for parameter a and b could predict the superconducting-insulating transition. The range of desired device dimensions for QPS are thus estimated.

Material	Resistivity ($\mu\Omega$·cm)	Critical temperature (K)	Coherence length (nm)
InO_x	14×10^{-5} [11]	2.7	10.0
NbN	400×10^{-6} [140]	13.0	4.0
NbTiN	93×10^{-6} [141]	15.0	4.3

Table 6.1: Typical parameter values for materials that are potential candidates for QPS.

6.1.3 Substrate selection and preparation

The choice of substrates has a significant impact on the success of subsequent fabrication stages. The adhesion between the substrate and deposited materials is a major consideration since the NbN film is amorphous, and the lattice matching between the substrate and the NbN film is not a concern. We used silicon substrates that have a surface roughness less than 5 Å to allow uniform NbN structures to be formed. Two types of silicon substrates were used for this experiment: doped p-type silicon and intrinsic silicon. The doped p-type silicon has a resistivity measured at 16 $\Omega\cdot$ cm. No oxide was grown on these substrates, but a native oxide was presumed present. All substrates were cleaned in an acetone sonication bath for at least 5 minutes, followed by isopropanol and DI water rinse, and were blown dry by nitrogen.

6.1.4 DC magnetron sputtering

The main deposition method to grow NbN thin film in this work is DC magnetron sputtering. Sputtering is a physical vapor deposition (PVD) coating technique where a target material used as the coating is bombarded with ionized gas molecules, and atoms from the source target are sputtered off into the plasma. The vaporized atoms are deposited onto the substrate as they condense as a thin film.

The target material is placed in a vacuum chamber parallel to the substrate to be coated. The vacuum chamber is pumped down to a base pressure and filled with a high purity inert process gas (e.g., Ar). A DC voltage is applied between the anode (the substrate) and the cathode (the magnetron gun), generating a plasma of energetic species from atoms in the process gas.

Ar has a relatively high mass and has the ability to convey kinetic energy upon impact during high energy molecular collisions in the plasma, which creates the gas ions that are the primary driving force of sputter thin film deposition. The electrically neutral Ar gas atoms are first ionized due to the forceful collision of these gas atoms onto the surface of the negatively charged source material, which ejects atoms off into the plasma. The ionized Ar gas atoms are then driven to the positively charged substrate, where the vaporized target coating atoms condense and form a thin film coating on the substrate to be coated. DC magnetron sputtering uses magnets behind the cathode to trap electrons over the negatively charged target material, providing a faster deposition rate. The deposition rate depends on the energy of the accelerated ions, which in turn depends on the potential difference between the two electrodes.

The high process pressure involved in sputtering can cause sputtered atoms collisions before reaching the substrate. These collisions change the angle of incidence at which the sputtered atoms arrive at the surface of the substrate. The result is that sputtering coating is more conformal and ideal for the deposition of noncrystalline films and structures. It is also suitable for applications that require uniform coverage of stepped features, such as the creation of interconnects and multi-layer device fabrication. DC magnetron sputtering is a cost-effective way of applying metal target coatings that are electrically conducting. However, DC Sputtering is

less effective for non-conducting insulating materials that can take on a polarized charge, such as dielectric coating materials.

6.1.5 Deposition in CHA Mark 50

Deposition of NbN was carried out in a CHA Mark 50 system manufactured by CHA Industries. CHA Mark 50 is a high vacuum deposition system that has sputtering and evaporation operation modes. Samples were mounted onto a 4" wafer that was loaded in the chamber before an overnight pump-down. A Ti gettering was performed to further reduce the system vacuum. No materials were deposited on the substrates during this period since the samples were shielded by a shutter. The base pressure was in the range of 1.5×10^{-7} Torr before NbN deposition. The power used in sputtering was 1 kW for all the depositions. The Ar pressure and the ratio of N_2/Ar flow rate were varied to obtain the highest T_C value. Before the deposition, a two-minute ion milling was performed to clean the substrate and remove any impurities. A five-minute pre-sputtering was performed immediately after that, using the same sputtering parameters. This was used to remove impurities from the surface of the target and achieve a constant deposition rate. During the deposition, the substrates placed in a disk were rotating at a constant speed to ensure film uniformity. Details of deposition parameters and measurement results will be given in the following subsections. Once the deposition was done, the chamber was pumped up, and the samples were unloaded.

6.1.6 Electron beam lithography

Electron beam lithography (EBL) is a lithography technique that uses a focused beam of electrons to draw custom shapes on a substrate covered with electron beam resist [142]. The electron beam changes the solubility of the resist, enabling selective removal of either the exposed or non-exposed areas of the resist by immersing it in the developer.

Unlike photolithography that uses a photomask to expose large areas indiscriminately, EBL can be used to write patterns into a resist layer without a photomask directly. Another feature of EBL is that dedicated doses can be assigned to different areas during the exposure while patterns of different geometries are exposed appropriately. In terms of resolution, EBL is

much better than photolithography, which is limited by the wavelength of the visible light used (e.g., a typical UV light has a 365 nm wavelength range). Electrons have a wavelength down to the picometer range, which can overcome the diffraction limits experienced with optical lithography. The reported EBL resolution can be less than 10 nm for isolated features and for features that have a pitch of 30 nm, leading to a density close to 700 Gbit/in^2 [143]. The feature size is not determined by beam size but is limited by forward scattering in the resist, while the pitch size is limited by secondary electron travel in the resist [144, 145].

In addition to secondary electrons, primary electrons from the electron gun with sufficient energy to penetrate the resist can be scattered from the resist and/or the substrate. This effect leads to a larger exposure area compared to the defined exposure area. The forward scattering is that the electron beam scatters laterally from the beam-defined area as it moves forward into the resist. The primary electrons are typically deflected by a small angle, and thus statistically broaden the beam in the resist. The back scattering is due to a collision with a heavy particle and leads to wide-angle scattering in the substrate. A large enough dose of backscattered electrons can lead to complete exposure of resist over an area much larger than defined by the beam spot. Due to the interactions of the primary beam electrons with the resist and substrate, the developed pattern is wider than the scanned pattern, which is known as the proximity effect.

During exposure, a beam of electrons is generated by an electron gun and accelerated towards the sample, similar to a scanning electron microscope (SEM). The acceleration voltage, also known as electron high tension (EHT) voltage, plays an important role in the volume of interaction between the primary beam and the used material. Increasing EHT voltage causes the electron beam to penetrate further into the resist/substrate with less forward scattering, which significantly affects the resolution. Other beam parameters include aperture size, stigmation, and shift. The magnet array used to deflect the beam has a limited range known as write field (WF). A typical write field ranges from 100 μm to 1 mm. Larger patterns require stage movement between different write fields, and write field alignment is the critical step to avoid stitching problems.

6.1.7 Electron beam resists and developers

Choosing appropriate electron beam resists is critical to obtain desired patterns and achieve the desired resolution. Most electron beam resists are polymeric materials whose solubility in a developer is proportional to their molecular weight. Like photoresists, electron beam resists are also classified as positive tone resists and negative tone resists. Exposure to electron beams will cause crosslinking of polymeric chains and an increase of molecular weight in a negative tone resist while causing chain scission and a decrease of molecular weight in a positive tone resist. The most commonly used electron beam resists we introduce here are polymethyl methacrylate (PMMA), copolymer (MMA (8.5) MMA), Zeon electron beam positive-tone (ZEP), and hydrogen silsesquioxane (HSQ).

PMMA and copolymer

PMMA resist is most commonly used as a high-resolution positive resist for direct-write electron beam and is also used as a protective coating for wafer thinning, as a bonding adhesive, and as a sacrificial layer. When exposed at higher dose levels, PMMA becomes a negative resist, as reported in [146]. Copolymer resist is based on a mixture of PMMA and $\sim 8.5\%$ methacrylic acid, formulated in the safer solvent ethyl lactate. Due to the chemical properties of methacrylic acid during the bake step at 200 °C in which dehydration leads to ring formation, copolymer has higher sensitivity and higher contrast than PMMA. Copolymer is commonly used in combination with PMMA in the bi-layer lift-off resist process.

Standard PMMA resists are formulated with 495,000 or 950,000 molecular weight in either chlorobenzene or the safer solvent anisole. Depending on the concentration, these products cover a wide range of film thicknesses, ranging from tens of nanometers to a few microns.

Typically, PMMA is developed in a 1:3 methyl isobutyl ketone (MIBK): isopropanol (IPA) solution to provide a relatively high resolution. However, higher concentrations of MIBK in IPA, such as 1:1 and 1:2, are also available if the resolution is not the primary concern. In addition to MIBK/IPA, IPA mixed with a small proportion of water was also reported as being an effective developer for PMMA [147]. Recent research has demonstrated an improvement of

PMMA resolution by using a 3:7 water/IPA developer, which helped them successfully fabricate gratings of minimum size 16 nm equal line spacing [148].

ZEP

ZEP resists are also high-resolution positive tone resists manufactured by the Zeon Corporation. ZEP resists generally require a lower dose than PMMA, and thus, write time is shorter. Another feature of ZEP resists is that they have excellent etch resistance, which is better for dry/wet etching process when etch selectivity is a concern. However, ZEP resists are much expensive than PMMA resists. Developers for ZEP resist family are ZED-N50, ZED-N60 and ZEP-A. ZED-N50, and ZED-N60 are for high sensitivity and high resolution, respectively, while ZEP-A is for thinner resists.

HSQ

HSQ is an inorganic material used as a negative tone resist for EBL. HSQ has excellent plasma etch resistance and selectivity to silicon. This resist has a better resolution compared to PMMA, and it has a reported minimum feature size of 2 nm [149]. However, HSQ generally needs a higher dose ($\sim 3\times$ higher than PMMA), resulting in a longer write time. In addition to that, it has higher line edge roughness than positive tone resists and potentially has more stitching errors. HSQ is typically developed in tetramethylammonium hydroxide (TMAH). After HSQ is exposed, it turns into silicon dioxide. Unlike organic resists that are soluble in solvents, the typical chemical used to remove HSQ is hydrogen fluoride (HF) or buffered oxide etchant (BOE).

6.1.8 Nanowire pattern definition in Raith e-LINE system

The Raith e-LINE EBL system utilizes thermal field emission filament technology and a laser-interferometer controlled stage. The system is featured by an accelerating voltage ranging from 0.5 to 30 kV with a minimum linewidth of 8 nm. The standard write field size is 100 μm, and the electron beam sequentially exposes each pixel (2 nm) before stitching the write fields together to generate the final patterns.

Substrates were prepared as illustrated in Section 6.1.3 before spinning PMMA resists. Bi-layer resist stacks PMMA 495K/PMMA 950 K were spun onto the substrate to achieve the best lift-off results, as shown in Figure 6.1. Since the molecular weight of PMMA 495K is smaller than that of PMMA 950K, PMMA 495K is more sensitive to electron beams. After the resist stacks are exposed to electron beams, the PMMA 495K layer is slightly over-exposed and has a bigger opening area after development. This provides an undercut profile that is ideal for the lift-off process and minimizes the possibility of "lily-pad" folded edges, especially after directional deposition. The "lily-pad" folded edge, illustrated in Figure 6.2, leading to non-ideal lift-off results and unexpected impacts on nanowire uniformity and potential device performance. Although this process has been optimized for lift-off results, the "lily-pad" issue is still inevitable due to the conformal deposition property of DC magnetron sputtering and the extremely thin resist thicknesses.

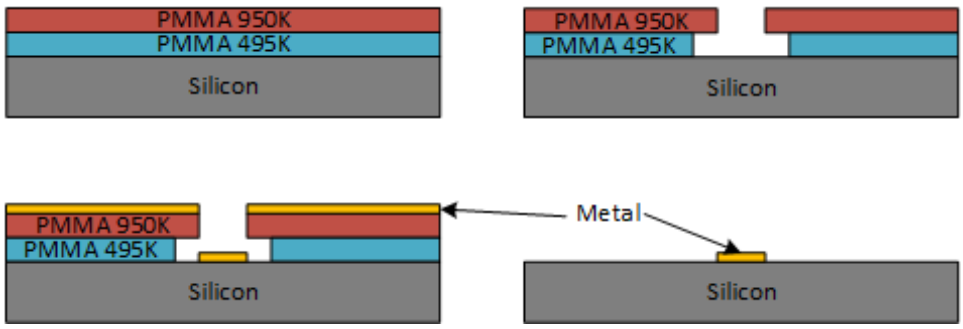


Figure 6.1: Bi-layer resist stacks for lift-off.

Nanowire patterns were designed to fit onto a 5 mm $\times$ 5 mm die, which was wire-bonded to a leadless chip carrier (LCC). Each design contains eight nanowires for a four-point measurement setup. Each nanowire has two leads connected to one side (e.g., $V+$ and $I+$), and the other two leads (e.g., $V-$ and $I-$) are shared with three other nanowires to maximize lead utilization during cryogenic measurement. The nanowire width and length vary for each different design. Nanowires are typically positioned close to the center of a standard write field (100 μm $\times$ 100 μm) to avoid stitching errors. The patterns are split to two exposures to reduce exposure time. The interconnection lines were exposed first using a larger step size and smaller dose value, followed by the exposure of nanowires using a smaller step size and larger dose value.

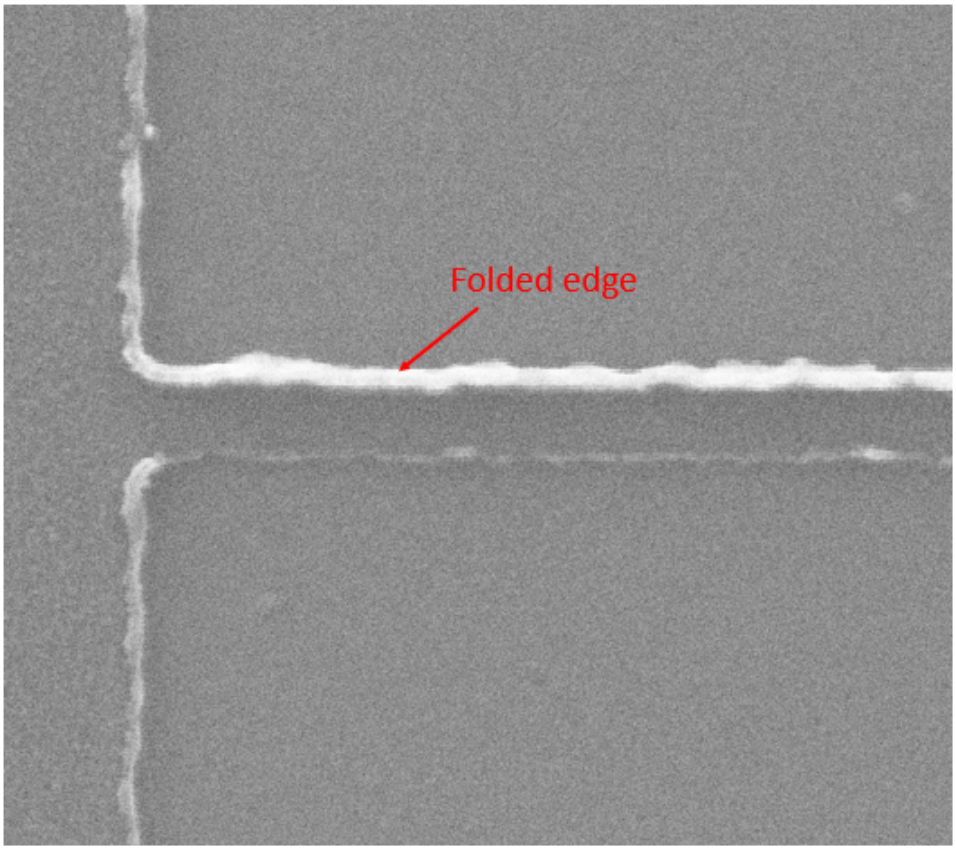


Figure 6.2: SEM image of a “lily-pad” folded edge after lift-off.

The dose values were optimized for the specific length and width of the nanowire. We implemented 1 μm offset at both sides of the nanowires to alleviate the alignment issue between two exposures, as shown in Figure 6.3. The 100 μm $\times$ 100 μm contact pads were written after the NbN layer was lifted off, using alignment marks for overlay exposure. After that, A layer of 250 nm thick Al was deposited onto the samples, followed by a lift-off process. The nanowire samples were then diced into 5 mm $\times$ 5 mm dies and were packaged for testing.

6.2 Characterization of NbN thin films

6.2.1 Measurement equipment

Cryogenic measurements were mainly carried out in a pulse tube-based cryostat (Cryo Industries of America), with temperature control from $\sim$ 1.5 K and up. Each sample was mounted onto a sample holder that has PCBs for different measurement configurations. An image of mounted samples is shown in Figure 6.4. Figure 6.4(a) shows that an NbN witness sample

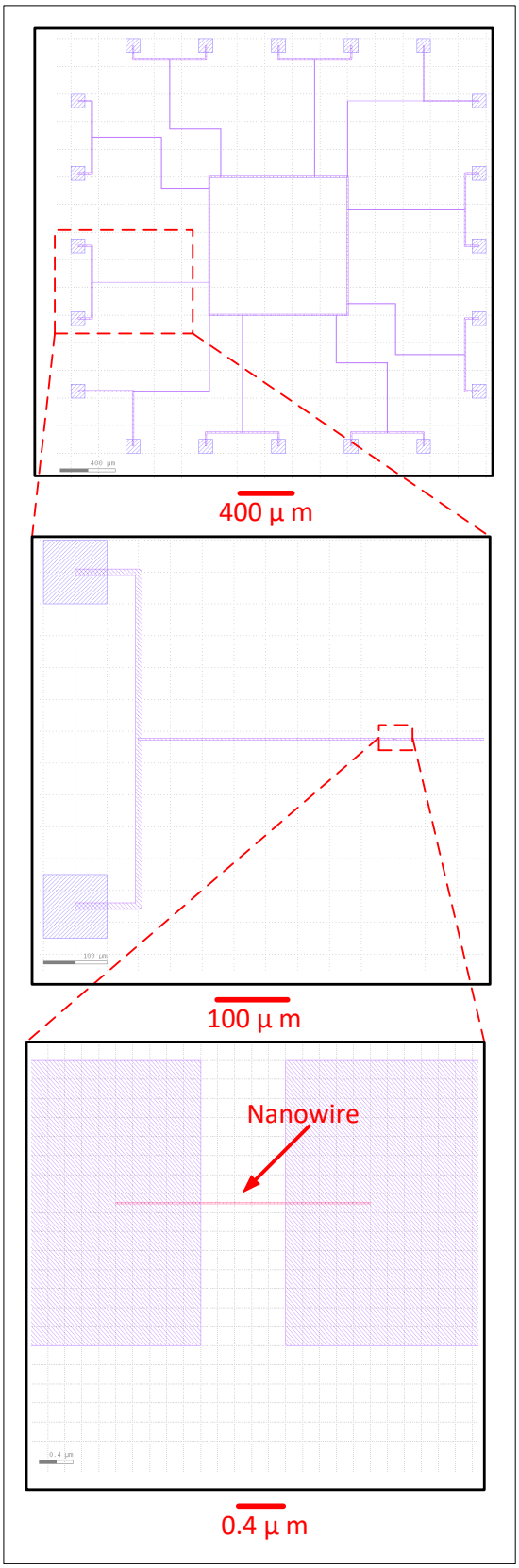


Figure 6.3: The full EBL mask design for nanowires. Each design contains eight nanowires for a four-point measurement setup. The figure shows the EBL pattern at different magnifications.

was mounted on the sample holder, using pogo pins, to perform a four-point resistance measurement. In Figure 6.4(b), a packaged 5 mm $\times$ 5 mm NbN nanowire sample was mounted on a PCB for various measurements. A micro-Ohm meter (Agilent 34420A) was used to measure DC resistance. Pulse generator, current amplifier, lock-in amplifier and etc., were used to conduct I-V characteristic, differential resistance measurements at cryogenic temperatures. In addition, a Physical Property Measurement System (PPMS) was used to conduct DC and AC resistance measurements with precise temperature control at temperatures down to 1.8 K.

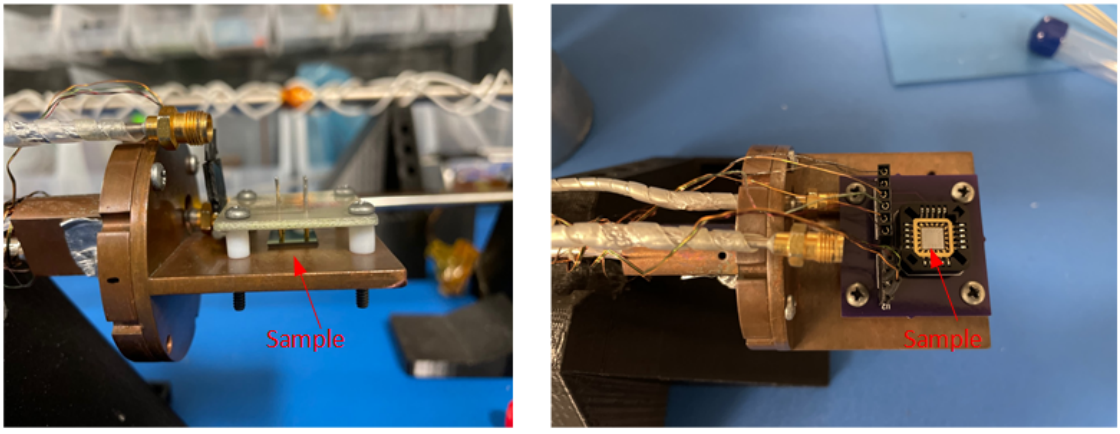


Figure 6.4: The pulse-tube based cryostat measurement setup. (a) An NbN witness sample is mounted to the sample holder. (b) A 5 mm $\times$ 5 mm NbN nanowire sample is wire-bonded to a leadless chip carrier on a PCB.

6.2.2 Experimental results

NbN was deposited onto an oxidized silicon substrate using different Ar pressures and ratios of N_2/Ar flow rate. The deposition time was 20 minutes, resulting in a thickness from 80 nm to 150 nm. The T_C values of different samples are summarized in Table 6.2. The $R(T)$ measurement results are shown in Figure 6.5. In each of the cases, the normal state resistance and room temperature resistance were inversely proportional to T_C, except for sample 10, which suggests that the N_2 percentage in sample 10 is too low. Therefore, the film has properties that are closer to Nb. We noticed that samples 1 and 2 were not superconducting, probably due to very low Ar pressure. X-ray diffraction (XRD) analysis was performed on these witness samples to

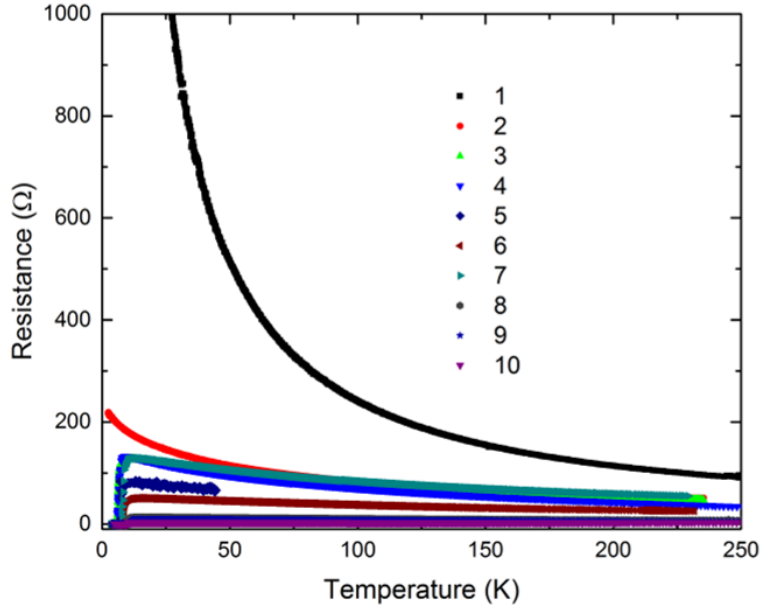
determine the crystallographic structure of deposited NbN, as shown in Figure 6.6. Sample 9 was measured at two different spots, labeled by S9-1 and S9-2.

Sample number	N_2/Ar ratio	Ar pressure (mTorr)	T_C (K)
1	0.200	1.5	N/A
2	0.100	1.5	N/A
3	0.060	4.0	6.2
4	0.043	4.5	6.0
5	0.040	4.5	6.8
6	0.040	4.0	8.5
7	0.033	4.5	7.7
8	0.035	4.0	10.0
9	0.027	4.0	11.0
10	0.021	4.0	9.8

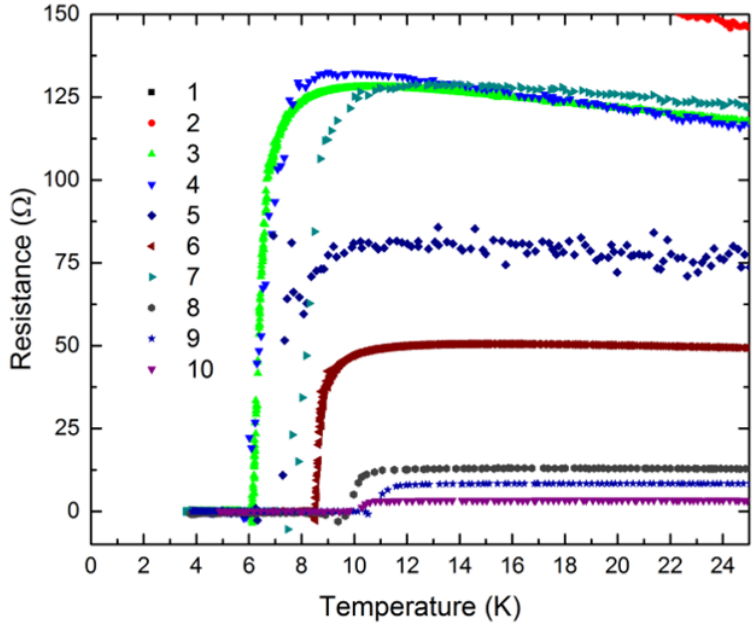
Table 6.2: T_C of NbN witness samples using different Ar pressures and ratios of N_2/Ar flow rate.

The XRD analysis results suggested that there were multiple NbN crystal structures existing on these samples. Samples that are superconducting have a peak at 35.5 degrees, while samples that are not superconducting do not have a peak at 35.5 degrees, which means δ- NbN might contribute to the superconductivity of these samples. We also noticed that there were impurities corresponding to a peak at 33 degrees.

The ratio of N_2/Ar flow rate was further narrowed down into an optimized region to explore the maximum T_C that could be achieved in this CHA deposition system. These experiments were performed after an upgrade of the mass flow controller, which resulted in a different ratio of N_2/Ar flow rate compared to the previous experiments. The power and Ar flow rate were fixed during these experiments. The results are summarized in Table 6.3. The resulting T_C is in a range from 10.3 K to 11.0 K, indicating a maximum T_C of $\sim$ 11 K, which was consistent with all the previous experiments. The following NbN depositions for nanowires were performed using a typical N_2/Ar ratio of 0.1, but a reduced deposition time for thinner films.



(a)



(b)

Figure 6.5: Resistance vs temperature measurement results for NbN witness samples using different Ar pressures and ratios of N_2/Ar flow rate. (a) Results at a temperature range from 0 K to 250 K. (b) Results at a temperature range from 0 K to 25 K.

6.3 Low temperature characteristics of NbN nanowires

6.3.1 Introduction

NbN nanowires were fabricated using the method introduced in Section 6.1. The nanowires and interconnection wires were first patterned by EBL. NbN thin film was then deposited onto the

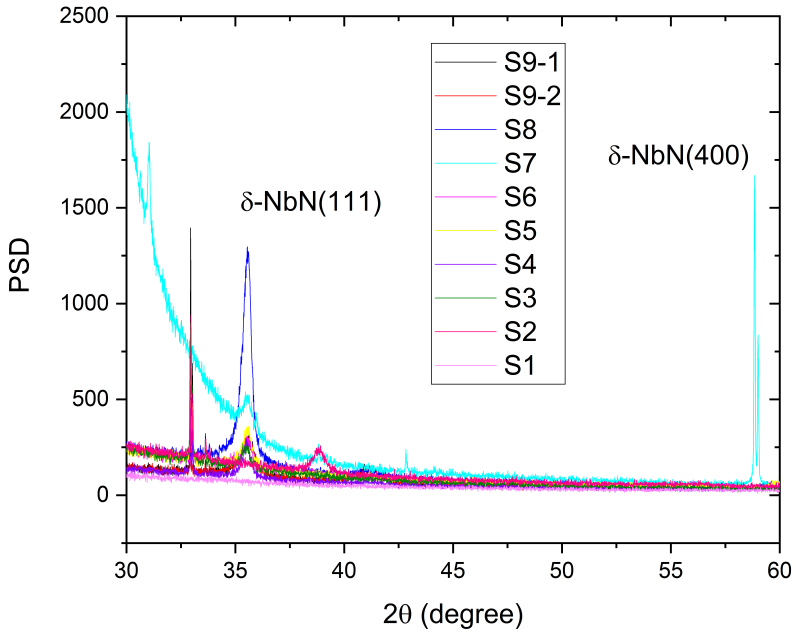
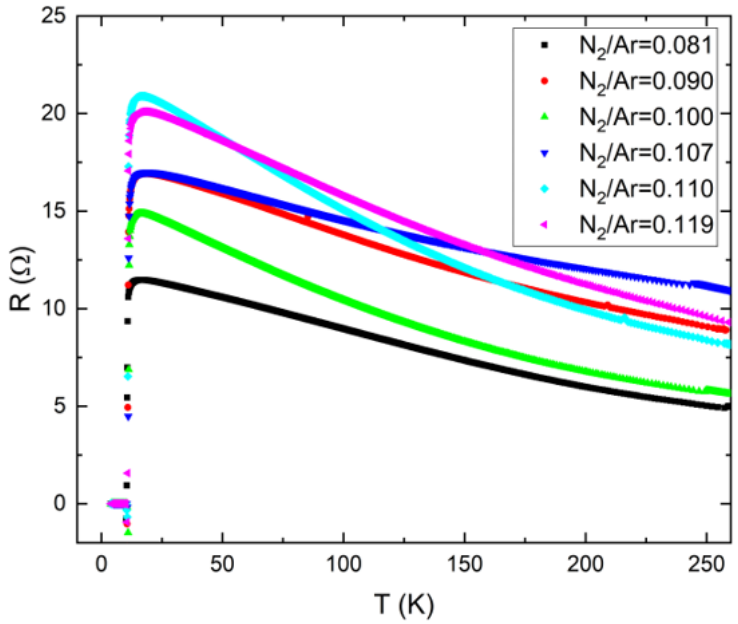


Figure 6.6: XRD analysis results for NbN witness samples using different Ar pressures and ratios of N_2/Ar flow rate.

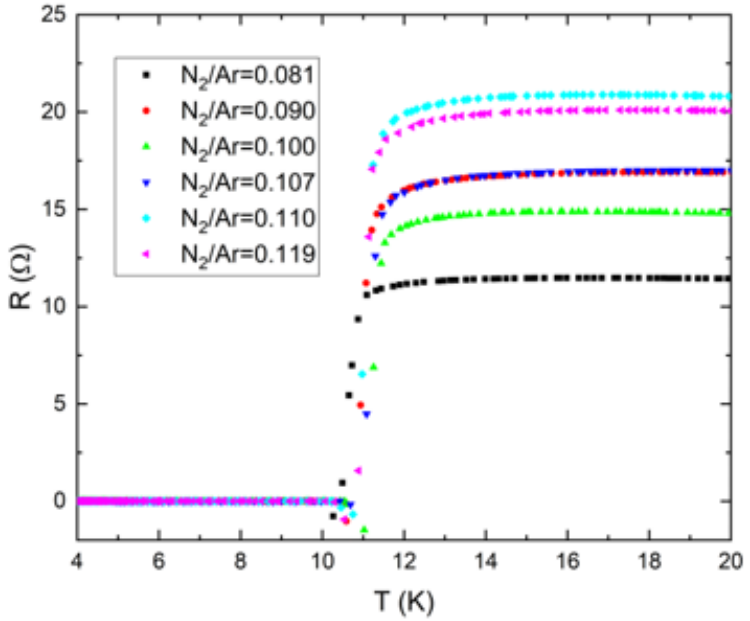
N_2/Ar ratio	Resistance at 293 K (Ω)	Resistance at 15 K (Ω)	RRR	T_C (K)
0.081	4.29	11.46	0.375	10.3
0.090	8.19	16.80	0.487	10.7
0.100	5.07	14.86	0.341	11.0
0.107	9.99	16.84	0.593	10.7
0.110	7.30	20.86	0.350	10.8
0.119	7.88	20.03	0.393	10.5

Table 6.3: T_C of NbN witness samples using different ratios of N_2/Ar flow rate.

samples using CHA Mark 50 at a deposition rate of $\sim$ 6.5 nm/min for three minutes, followed by a lift-off process. The contact pads were also defined by EBL. After that, a layer of 250 nm aluminum was deposited and lifted off. The nanowire sample was diced into a 5 mm $\times$ 5 mm die that can be wire-bonded to a leadless chip carrier and fit into a socket on the test PCB. Nanowires with different widths, lengths, and thicknesses were fabricated and tested to seek evidence of QPS. However, controlling the parameters in these narrow and thin nanowires in order to maintain homogeneous behavior across them is not trivial. Even geometrically identical nanowires fabricated on the same samples tend to exhibit slight variations in behavior.



(a)



(b)

Figure 6.7: Resistance vs temperature measurement results for NbN witness samples using different ratios of N_2/Ar flow rate. (a) Results at a temperature range from 0 K to 260 K. (b) Results at a temperature range from 4 K to 20 K.

In this section, we will show experimental results of selected NbN nanowires and analyze experimental data using mathematical models.

6.3.2 Temperature dependence of NbN nanowires

Superconducting nanowires of NbN were fabricated from thin films of thickness $\sim$ 20 nm, measured by Atomic force microscope (AFM). Prior to the sputtering run for nanowires, a test NbN run was conducted for a period of 30 minutes, giving a thickness of $\sim$ 140 nm and a T_C of 10 K. The T_C of the witness sample for the nanowire run was 7.9 K, as shown in Figure 6.8. The T_C values for both runs were slightly lower than what we had before, which were typically 11 K for a 30-minutes run and 9 K for a 3-minute run.

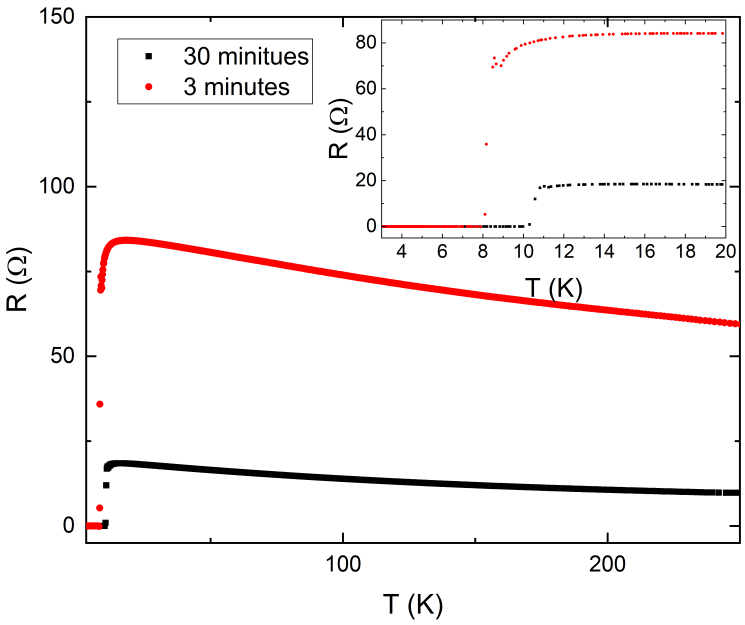


Figure 6.8: Resistance vs temperature measurement results for NbN witness samples of two different NbN runs.

We investigated the temperature dependence of NbN nanowires with different lengths and widths on different substrates. Figure 6.9 and 6.10 show the results for 5 μm and 10 μm long NbN nanowires on doped silicon substrates that has a resistivity of 16 $\Omega\cdot$ cm. The DC resistance of each nanowire was measured using a four-point probe method at temperatures down to 3.2

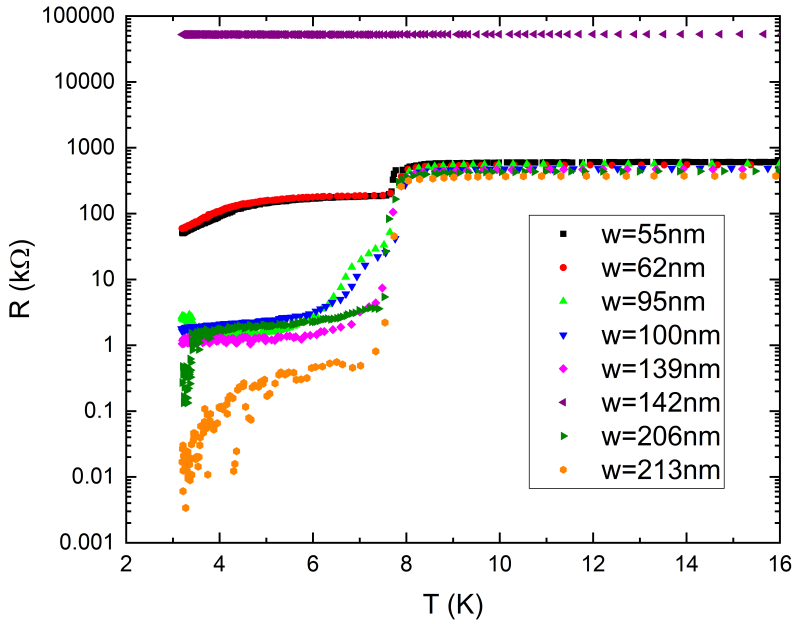


Figure 6.9: Resistance vs temperature measurement results for 5 μm long NbN nanowires on doped silicon substrates. Nanowire widths are from 62 nm to 213 nm.

K. DC current of 200 nA from Yokogawa GS200 source meter was injected into the nanowire, and DC voltage was measured by a nano-volt/micro-ohm meter (Agilent 34420A).

In Figure 6.9 and 6.10, all the nanowires showed a transition at a temperature of $\sim$ 8 K. This is due to the superconducting transition for wider NbN interconnection wires, which have widths up to 10 μm. There are also two nanowires (142 nm in Figure 6.9 or 148 nm in Figure 6.10) that show extremely high resistances, before and after the transition. This indicates that the high resistances of these two nanowires are probably from the nanowires themselves, not from interconnection parts. We will have a more detailed discussion regarding the extremely high residual resistance in the following sections.

Most of the other nanowires shown in Figure 6.9 and 6.10 have resistances less than 1 kΩ at temperatures below 3.5 K, which implies that they probably show zero resistance at lower temperatures. For example, this is evident in the 213 nm wide nanowire shown in Figure 6.9, which has a sharp transition at $\sim$ 3.2 K and its resistance drops to a few ohms. This phenomenon is in agreement with results shown in Figure 6.11 from [17], which was explained by the LAMH-TAPS model. However, the narrowest nanowires of either 5 μm or 10 μm length

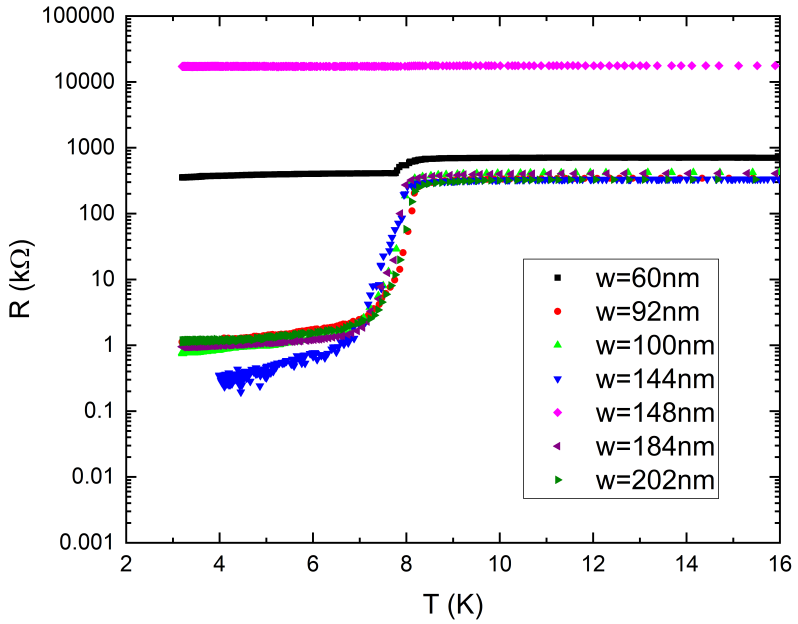


Figure 6.10: Resistance vs temperature measurement results for 10 μm long NbN nanowires on doped silicon substrates. Nanowire widths are from 60 nm to 202 nm.

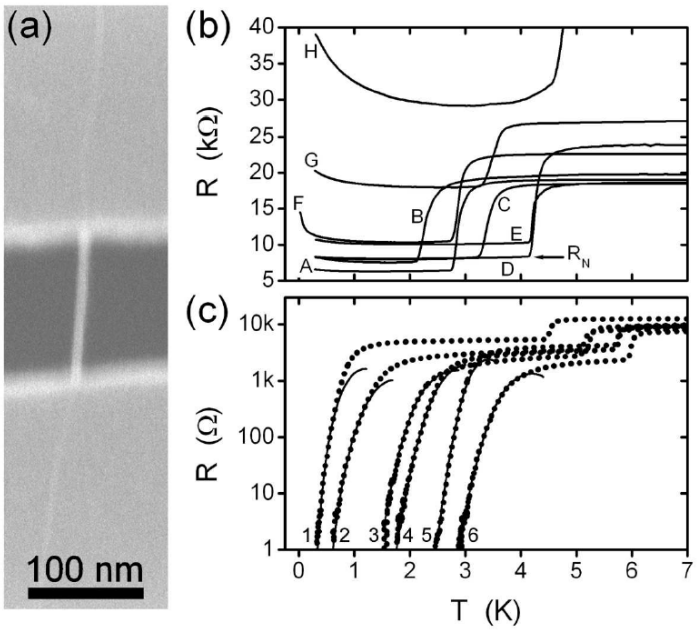


Figure 6.11: Experimental results of MoGe nanowires from [17]. (a) SEM image of an 8 nm nanowire over the trench in SiN. (b) R(T) curves for insulating nanowires. (c) R(T) curves for superconducting nanowires. Solid curves suggest fits to the LAMH-TAPS theory.

show very high resistive tails below 8 K. In particular, the 60 nm wide nanowire in Figure 6.10 does not show significant resistance change at temperatures between 3.2 K and 7.8 K, which

could be a sign of phase-slip phenomena. Further analysis of this nanowire will be presented in Section 6.3.5.

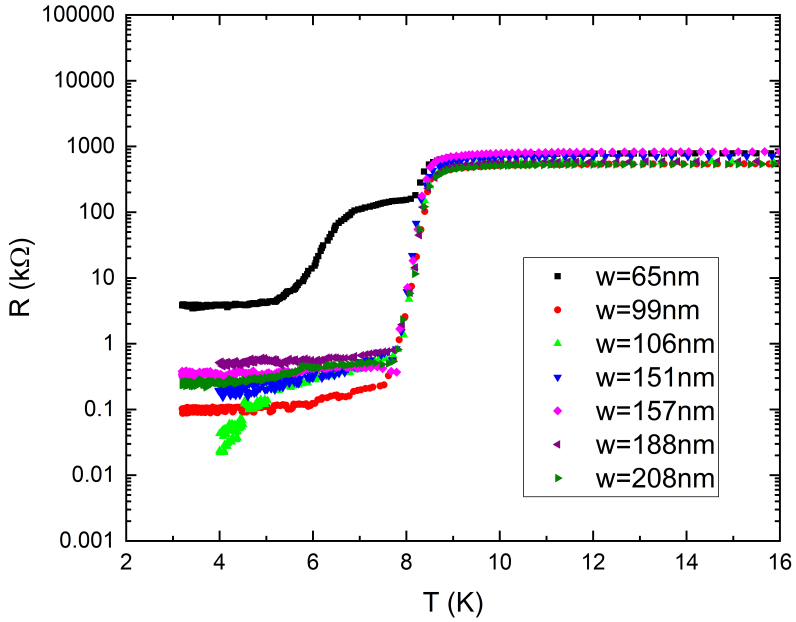


Figure 6.12: Resistance vs temperature measurement results for 10 μm long NbN nanowires on intrinsic silicon substrates. Nanowire widths are from 65 nm to 208 nm.

The resistance-temperature results for 10 μm long nanowires on intrinsic silicon substrates are shown in Figure 6.12. At the same length, the resistances of nanowires on intrinsic silicon substrates are typically smaller than the resistances of nanowires on doped silicon substrates at temperatures below T_C. Similarly, most of the wide nanowires exhibit low resistance at temperatures below T_C. The 65 nm wide nanowire, however, shows a second transition at a temperature of 7 K, and the resistance further reduces to a few $k\Omega$. This nanowire probably behaves like TAPS at temperatures around T_C. We classified all these nanowires into three categories based on their experimental results. The results and analytical results will be presented in the following subsections.

6.3.3 I-V characteristics of selected wide nanowires

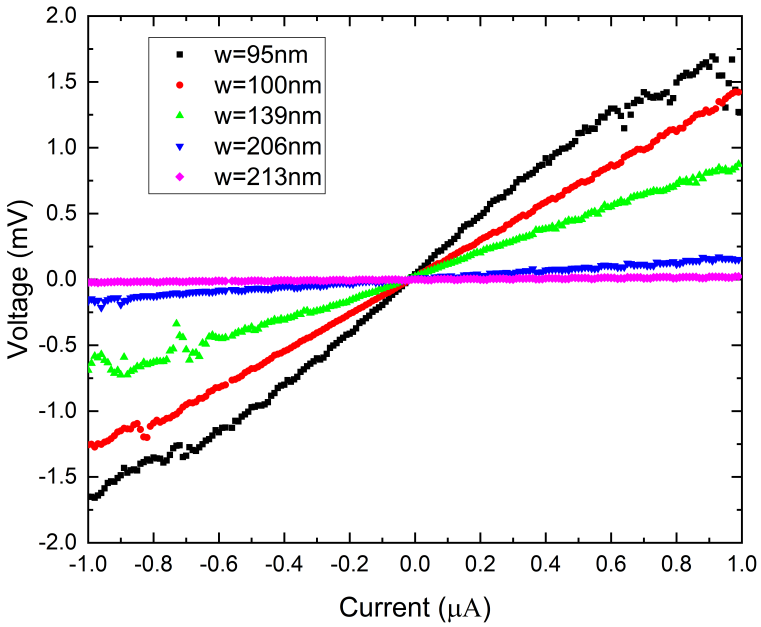
Most of the wide NbN nanowires were expected to be superconducting at temperatures below 7.9 K, according to the $R(T)$ measurement results shown in Section 6.3.2. To further explore

the nonlinearity of these nanowires, I-V measurements were conducted at a temperature of $\sim$ 3.2 K. Sweeping DC current was injected into the nanowires by a source meter, and the resulting DC voltage was measured by a nano-volt/micro-ohm meter. We measured each nanowire at a low current regime from -1 μA to 1 μA with 10 nA current step size and at a high current regime from -10 μA to 10 μA with 50 nA current step size.

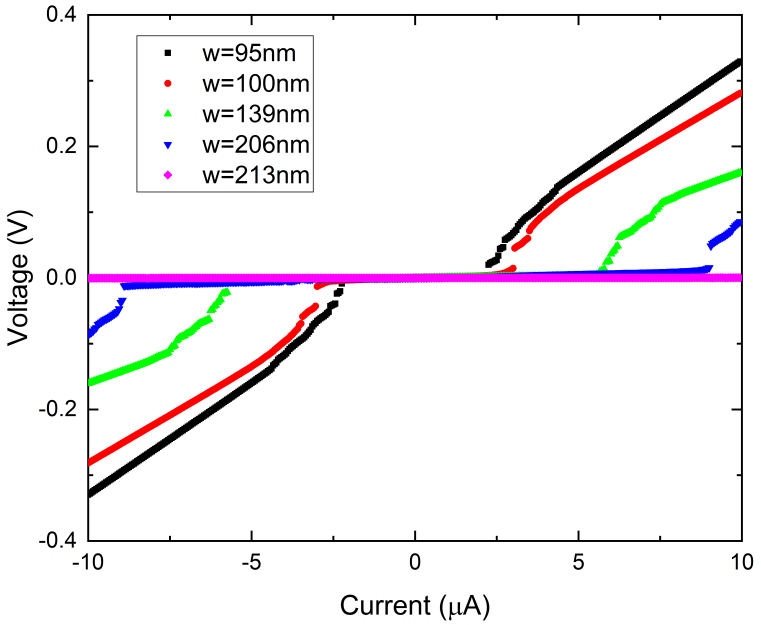
In Figure 6.13(a), all the nanowires show linear I-V characteristics at currents below 1 μA. In general, the residual resistance is inversely proportional to the width w of the nanowire. At higher current biases shown in Figure 6.13(b), a transition is observed for most of the nanowires. This transition is similar to a superconducting-normal transition. The current at which the transition occurs is similar to the critical current I_C of a superconductor. This current is also proportional to the width of the nanowire. It is also interesting to see non-zero resistances existing in wide NbN nanowires, which may be due to TAPS. These results support our hypothesis that superconducting properties of NbN dominate in wider NbN nanowires.

We also find similar results in Figure 6.14 for longer nanowires. They all have residual resistances less than 1 kΩ at bias currents below 1 μA. Only one nanowire shows a transition at approximately 7.5 μA. The transition in this plot is too fast due to the relatively large step size. The other nanowires are expected to show transitions at current biases larger than 10 μA. Nanowires from this sample do not follow the rule of resistance-width dependency, which is probably because of sample-to-sample variations and inhomogeneity of some nanowires (e.g., effects of "lily-paddings") during the fabrication process. The nanowires shown in Figure 6.15 exhibit even smaller resistances compared to their counterparts shown in Figure 6.14, although they all have the same length and thickness. For example, the 188 nm wide nanowire in Figure 6.15 has a resistance of $\sim 500\,\Omega$ while the 184 nm wide nanowire in Figure 6.14 has a resistance of $\sim 800\,\Omega$. This implies that the substrate may affect the low temperature performance of NbN superconducting nanowires. The critical currents of the three nanowires in Figure 6.15 are also expected to be greater than 10 μA.

Based on these experimental results and results of other nanowire samples we do not show in this dissertation, we can imply that the width of NbN superconducting nanowires is a critical condition for phase-slip phenomena. Specifically, nanowires that are $\sim$ 20 nm thick are

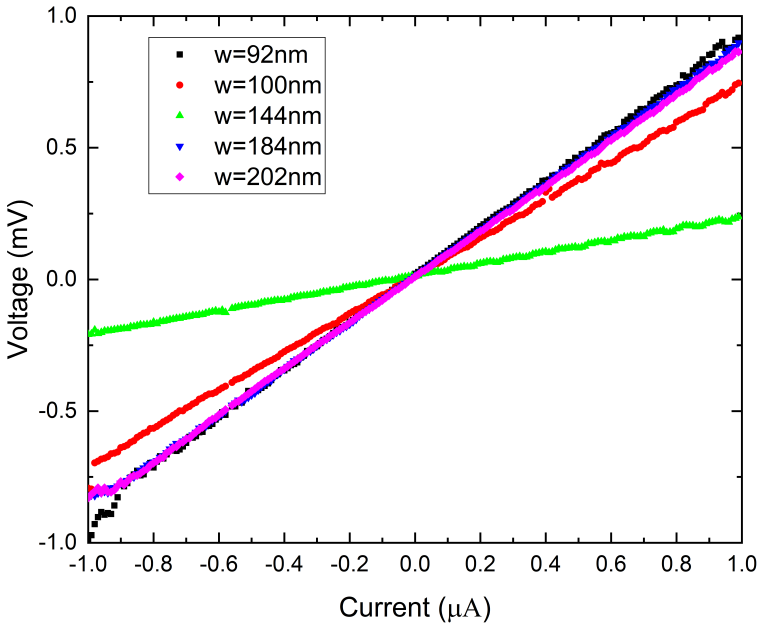


(a)

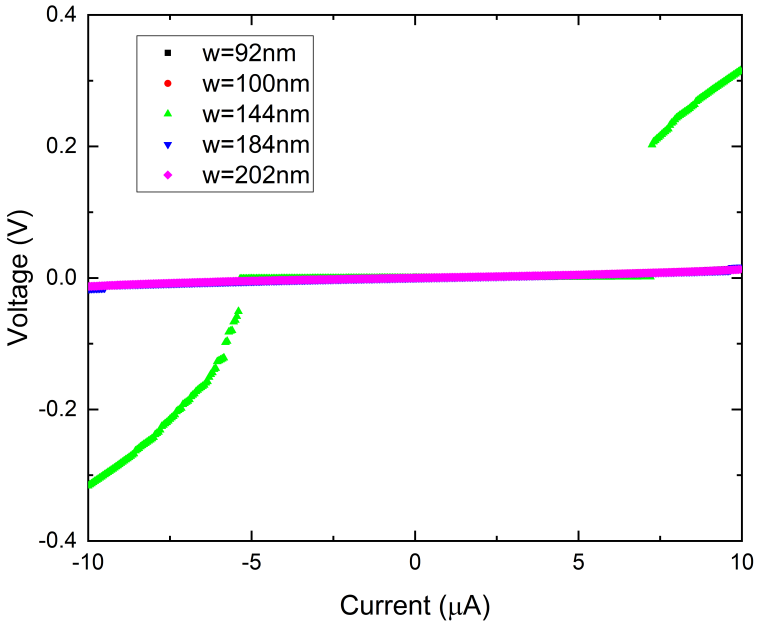


(b)

Figure 6.13: I-V characteristics of 5 μm long NbN nanowires on doped silicon substrates. (a) Bias current is from -1 μA to 1 μA with 10 nA current step size. (b) Bias current is from -10 μA to 10 μA with 50 nA current step size.

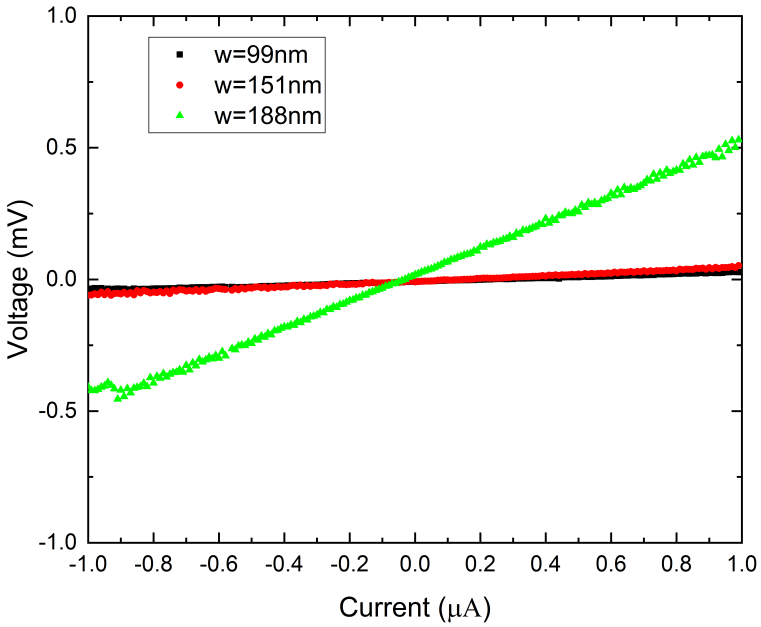


(a)

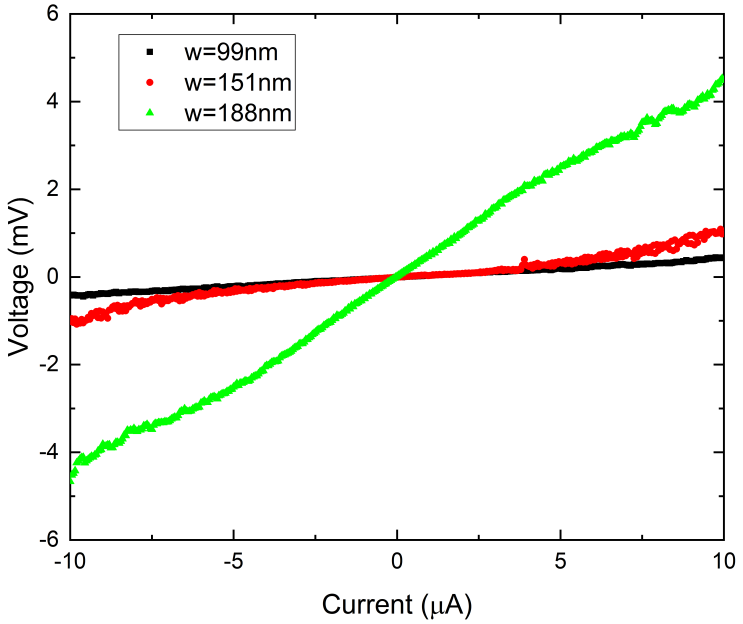


(b)

Figure 6.14: I-V characteristics of 10 μm long NbN nanowires on doped silicon substrates. (a) Bias current is from -1 μA to 1 μA with 10 nA current step size. (b) Bias current is from -10 μA to 10 μA with 50 nA current step size.



(a)



(b)

Figure 6.15: I-V characteristics of 10 μm long NbN nanowires on intrinsic silicon substrates. (a) Bias current is from -1 μA to 1 μA with 10 nA current step size. (b) Bias current is from -10 μA to 10 μA with 50 nA current step size.

expected to show relatively large residual resistance at temperatures below T_C if their widths are below $\sim$ 70 nm. This could be a result of TAPS or QPS. Wider nanowires of the same thickness may still show small residual resistances at temperatures below T_C, probably due to TAPS. In the following sections, we will focus on the narrowest nanowires we fabricated and show experimental results and analysis regarding the existence of phase-slip phenomena in ultra-narrow NbN nanowires.

6.3.4 Characterization of selected nanowires that showed extremely high resistance

The nanowires of interest in this work were the narrowest ones that may exhibit phase-slips, as suggested in [55]. Selected nanowires that had widths below 60 nm have been tested in a pulse tube system and a PPMS system. An SEM image of a typical narrow nanowire is shown in Figure 6.16. This nanowire was fabricated on an intrinsic silicon substrate, with 20 μm length, 57 nm width, and 20 nm thickness. We can see some brighter areas along the edge in Figure 6.16(b), which are presumed to be "lily-pad" folded edges. In Figure 6.16(a), we can see that the nanowire is a little away from the centers of two interconnection wires because of the alignment error between two exposures in EBL. This problem was potentially fixed by adding a constant offset during the second exposure.

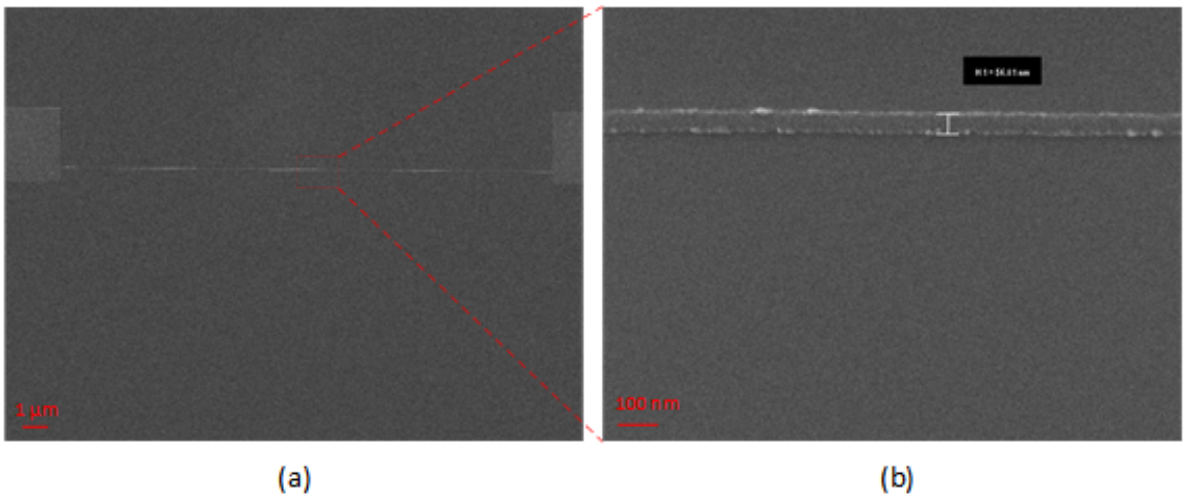


Figure 6.16: SEM images of NbN nanowire with 20 μm length, 57 nm width and 20 nm thickness. (a) Image of the nanowire. (b) Zoom-in view of the central area of this nanowire.

We started by studying the AC resistance of this nanowire at different temperatures. Using the Electrical Transport Option (ETO) module of the PPMS system, we measured the AC resistance of the nanowire from 14 K down to 1.8 K. The AC testing current has an amplitude of 10 nA and a frequency of 1.5 Hz. In Figure 6.17, the AC resistance increases as the temperature drops from 14 K, and it starts to decrease at $\sim$ 10 K, which is slightly higher than T_{C}. The resistance drop at temperatures around T_{C} is due to the superconducting transition of wider interconnection NbN wires. The resistance has the lowest value at $\sim$ 8.7 K and shows an upturn below that temperature. We noticed that the AC resistance of this nanowire was extremely high compared to the other nanowires we showed earlier in this chapter. The resistance is more than 10 MΩ even after the transition, which suggests a superconducting-insulating transition in this nanowire, similar to the results shown in [17].

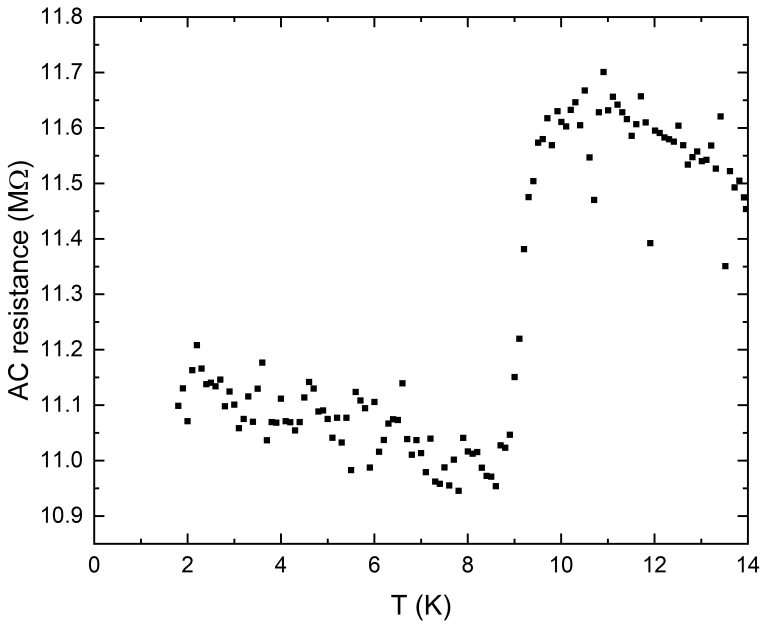


Figure 6.17: Experimental results of AC resistance vs temperature for a 57 nm wide nanowire. The AC testing current has an amplitude of 10 nA and a frequency of 1.5 Hz.

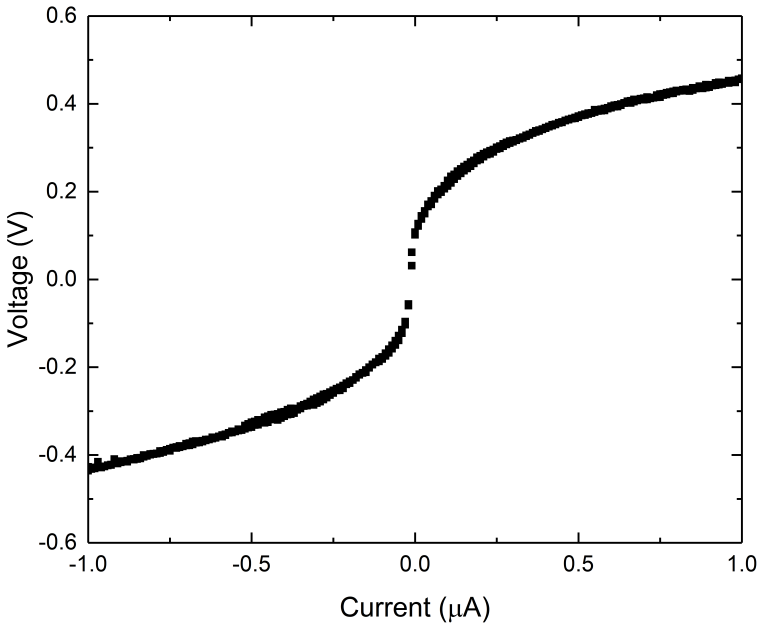
The I-V measurement was carried out in the pulse tube system at a base temperature of $\sim$ 3.2 K. The results shown in Figure 6.18(a) suggest a critical voltage of $\sim$ 100 mV at zero bias current, which is presumed to be the Coulomb blockade phenomenon. As the sweeping bias current increases to $\sim$ 2.3 μA, a superconducting-normal transition happens, indicating that the

critical current of this nanowire is close to 2.3 μA. These results show some similarities to the results of coherent quantum phase-slip presented in [56], which were from an NbN nanowire with a length of 9 μm, a width of 60 nm, and a thickness of 18 nm.

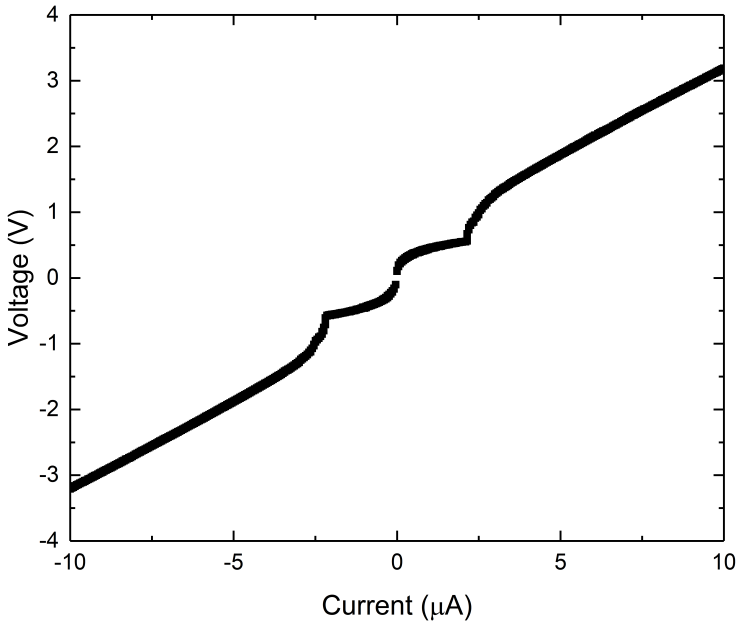
We used two different measurement setups to explore the underlying nonlinearity of this nanowire. The first setup is shown in Figure 6.19, which was for the measurements in the pulse tube system at a fixed temperature. These experiments were carried out at a temperature of 3.2 K, which is the lowest stable temperature in this system. Slowly ramped DC bias current was modulated by a small AC current. The resulting AC voltage, measured by the SR830 lock-in amplifier, is proportional to the AC resistance (differential resistance dV/dI). The same measurement was also carried out in the PPMS system, which provided a more precise temperature control and a more flexible configuration using the build-in ETO module. The experimental results are shown in Figure 6.20.

The AC testing current has an amplitude of 10 nA and a frequency of 1.5 Hz. The bias current ranges from -1 μA to 1 μA with 50 nA step size. At all different temperatures shown in Figure 6.20, the AC resistance increases significantly as bias current drops at the low bias regime. This is presumed to be the Coulomb blockade, which is consistent with the results shown in Figure 6.18(a). We found that it was hard to capture the details within the Coulomb blockade area due to the relatively large step size. The peak value of AC resistance at each temperature does not follow a trend of either increasing or decreasing. We tested this sample again in the PPMS system at higher temperatures using a step size of 2 nA.

Figure 6.21 shows the experimental results of AC resistance vs. DC bias current at 10 K and 11 K. To our surprise, at temperatures higher than T_{C}, the nanowire also exhibits Coulomb blockade at low bias regime. Therefore, the Coulomb blockade is probably not due to any coherent quantum phenomena, particularly quantum phase-slip phenomena. To demonstrate the repeatability of these results, we performed more measurements using different methods and techniques. For example, we added low pass filters at input lines to block any high-frequency noise. In addition to that, a different experimental setup was used to explore the I-V characteristics. Instead of the current biasing scheme, we used slowly sweeping DC voltage to bias the nanowire and used SR570 low noise current preamplifier to sense the small DC current and



(a)



(b)

Figure 6.18: I-V characteristics of a 20 μm long and 57 nm wide NbN nanowire on an intrinsic silicon substrate. (a) Bias current is from -1 μA to 1 μA with 10 nA current step size. (b) Bias current is from -10 μA to 10 μA with 50 nA current step size.

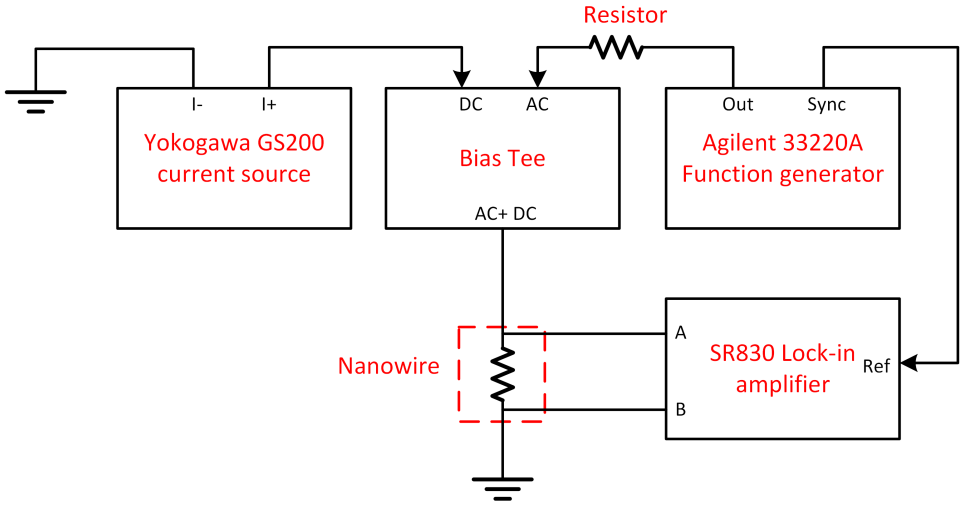


Figure 6.19: Block diagram of measurement setup for differential resistance measurement in the pulse tube system.

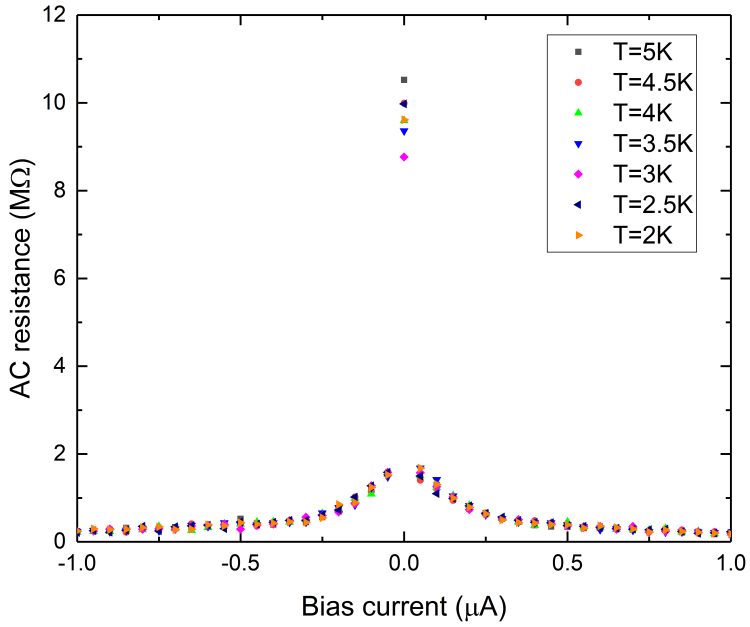


Figure 6.20: Experimental results of AC resistance vs DC bias current as a function of temperature. The AC testing current has an amplitude of 10 nA, a frequency of 1.5 Hz and a step size of 50 nA.

measured the output of SR570 by a nano-volt/micro-ohm meter. The resulting I-V characteristic curves were similar to the curves shown in Figure 6.18. The AC differential resistance was also re-measured by using AC currents with different amplitudes and frequencies, which did not provide any significantly different experimental results. We found that the Coulomb blockade

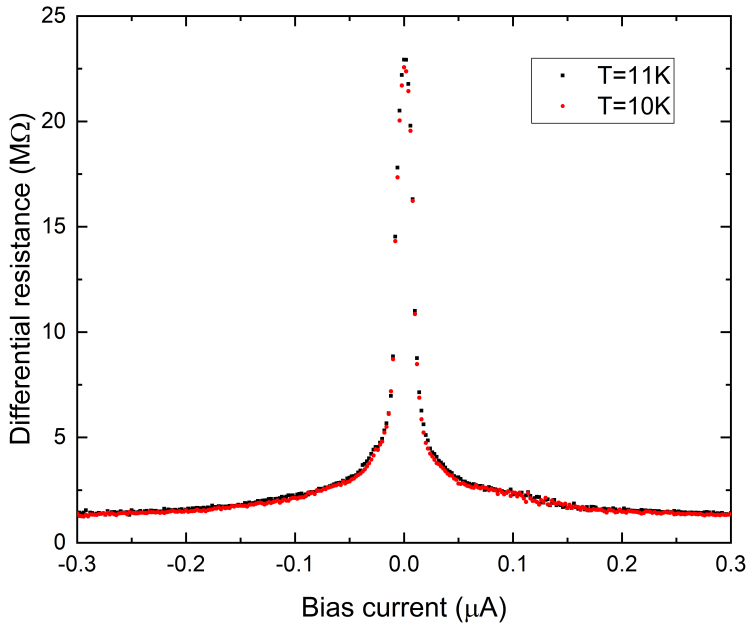


Figure 6.21: Experimental results of AC resistance vs DC bias current as a function of temperature. The AC testing current has an amplitude of 10 nA, a frequency of 1.5 Hz and a step size of 2 nA.

also existed in some other nanowires (e.g., the 142 nm wide nanowire in Figure 6.9) regardless of their geometries and substrate types, as shown in Figure 6.22. This phenomenon usually did not exist at room temperature but became more pronounced as temperature decreased. Therefore, the results shown in this section are repeatable and not due to QPS. One hypothesis is that the Coulomb blockade exhibited by the nanowire could be a result of unfiltered noise inside or outside the testing equipment and systems. Another possible explanation is that a low capacitive tunnel junction may exist in this nanowire, which significantly reduces electrical conductance under a certain bias threshold[120]. Because the capacitance is extremely small, the voltage build-up on the capacitor is relatively large and can prevent other electrons from tunneling. That is the reason why the conductivity was suppressed at low current bias.

6.3.5 Mathematical analysis of a selected narrow nanowire that showed high resistive tail

In Section 6.3.2, we show that some narrow nanowires have resistive tails at temperatures below T_C in the $R(T)$ measurement results, which could be a sign of phase-slips. In this section, we

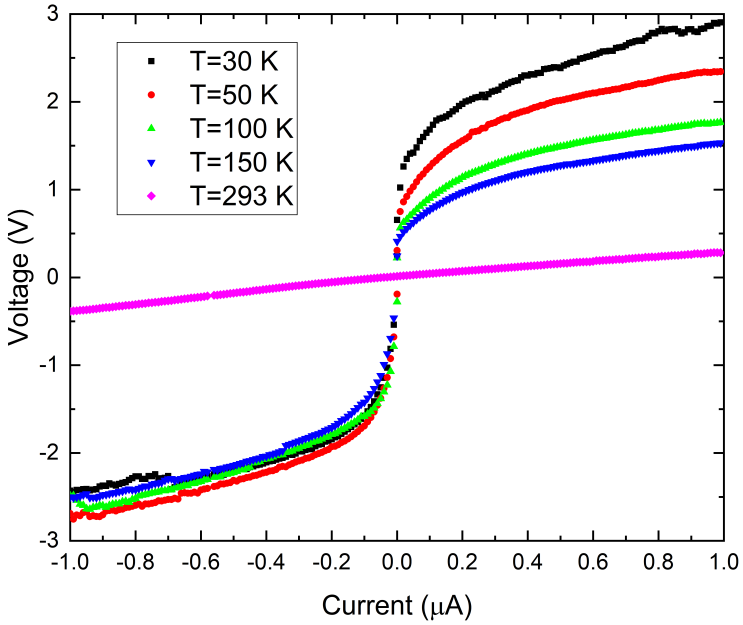
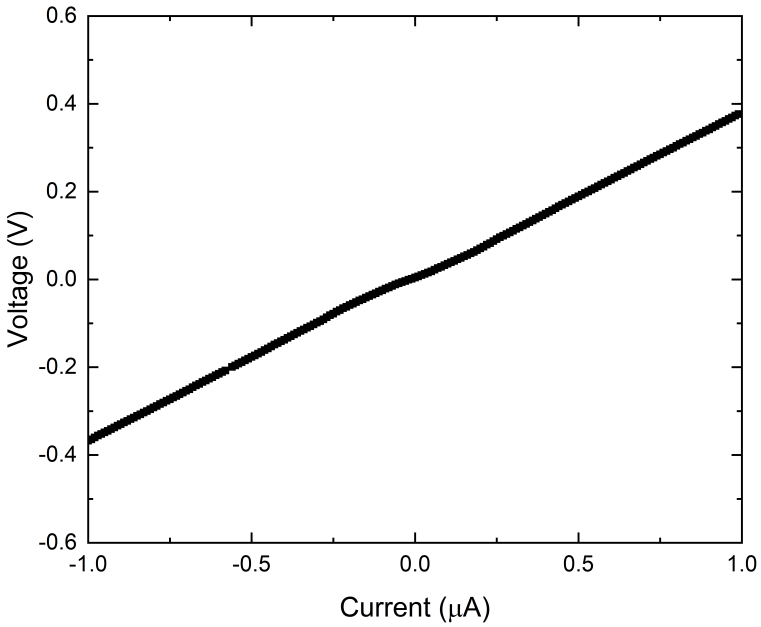
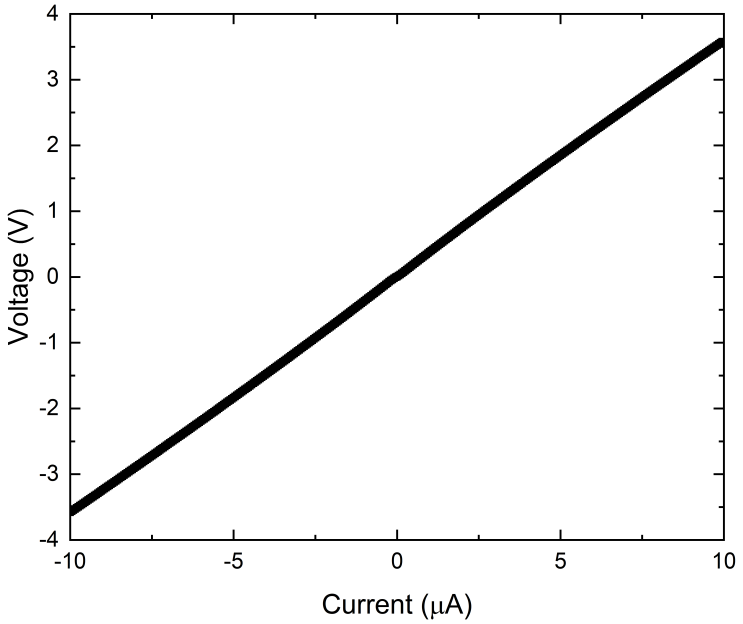


Figure 6.22: Temperature dependent I-V characteristics of a 142 nm wide nanowire on doped silicon substrate. The Coulomb blockade phenomenon did not exist at room temperature, but became more pronounced as temperature decreased.

use mathematical phase-slip models (both TAPS and QPS) to fit the experimental data in search of evidence of phase-slip phenomena. An example is a nanowire that has a width of 60 nm, a length of 10 μm, and a thickness of 20 nm on the doped silicon substrate. In addition to the $R(T)$ results shown in Figure 6.10, we performed I-V measurements for this nanowire at a temperature of 3.2 K using the current biasing method. The I-V experimental results are shown in Figure 6.23. According to these results, this nanowire shows resistive behavior at 3.2 K when the bias current is below 10 μA. The resistance is over 300 kΩ, which is consistent with the results shown in Figure 6.10. Furthermore, the temperature-dependent I-V curves shown in Figure 6.24 were obtained by using the voltage biasing method we introduced in the previous section. Slowly ramped DC voltage was applied on the nanowire, and the resulting DC current was amplified by the SR570 low noise current preamplifier and measured by the nano-volt/micro-ohm meter. At all three different temperatures, this nanowire shows resistive behaviors, and the resistance decreases slightly as temperature decreases from 7.2 K to 3.2 K. These results are also consistent with the results shown in Figure 6.23.



(a)



(b)

Figure 6.23: I-V characteristics of a 60 nm wide, 10 μm long and 20 nm thick NbN nanowire on the doped silicon substrate. (a) Bias current is from -1 μA to 1 μA with 10 nA current step size. (b) Bias current is from -10 μA to 10 μA with 50 nA current step size.

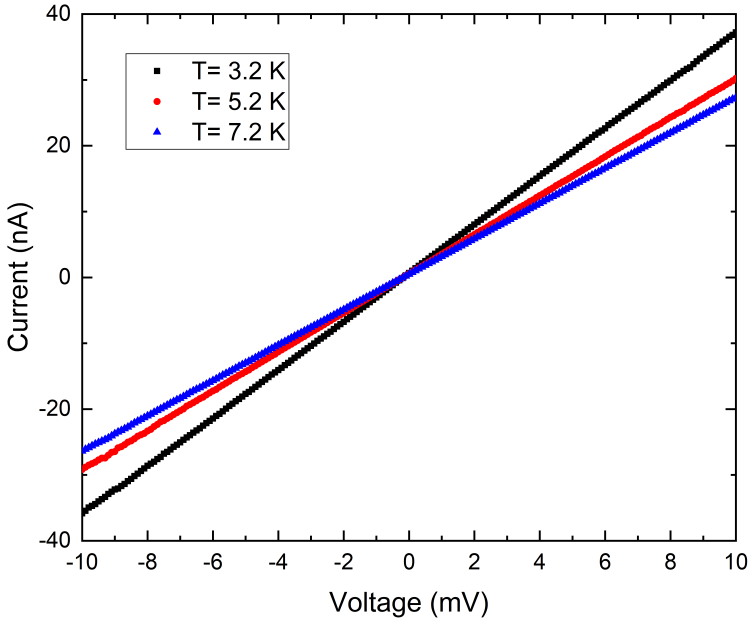


Figure 6.24: Temperature dependent I-V characteristics of a 60 nm wide, 10 μm long and 20 nm thick NbN nanowire on the doped silicon substrate. The nanowire shows resistive behaviors at temperatures of 3.2 K, 5.2 K and 7.2 K.

We used a TAPS model and a QPS model to fit $R(T)$ data of this nanowire. The TAPS model is described in [150], and defined by the following equations:

$$R_{\mathrm{TAPS}}(T) = R_{\mathrm{n}} e^{-\Delta F/k_{\mathrm{B}}T} \tag{6.2}$$

$$\Delta F(T) \approx 0.83[L/\xi(0)](R_{\mathrm{q}}/R_{\mathrm{n}})k_{\mathrm{B}}T_{\mathrm{C}}(1-T/T_{\mathrm{C}})^{3/2} \tag{6.3}$$

where $\xi(0)$ is the GL coherence length at 0 K. The normal state resistance is estimated to be 410 kΩ, which is the resistance right after superconducting transition. This resistance does not include the resistance of wider interconnection wires, which is nearly zero after the transition. The GL coherence length we used in this fitting is 5 nm, an estimate from [151]. The QPS model was proposed by Golubev et al. [152, 153], given by the following equations:

$$R_{\mathrm{QPS}}(T) = BR_{\mathrm{q}}S_{\mathrm{QPS}}\frac{L}{\xi(T)}e^{-S_{\mathrm{QPS}}} \tag{6.4}$$

$$S_{\mathrm{QPS}} = C\frac{R_{\mathrm{q}}}{R_{\mathrm{n}}}\frac{L}{\xi(T)} \tag{6.5}$$

$$\xi(T) = 0.907\xi(0)(1 + (1 - 0.25t)\xi(0)/t)^{-1/2}(1 - t^2)^{-1/2} \tag{6.6}$$

where $\xi(T)$ is the GL coherence length with $t = T/T_{\mathrm{C}}$, and B and C are fitting parameters. The values for fitting parameters were chosen to be B = 0.15 and C = 0.085 in order to fit the QPS model.

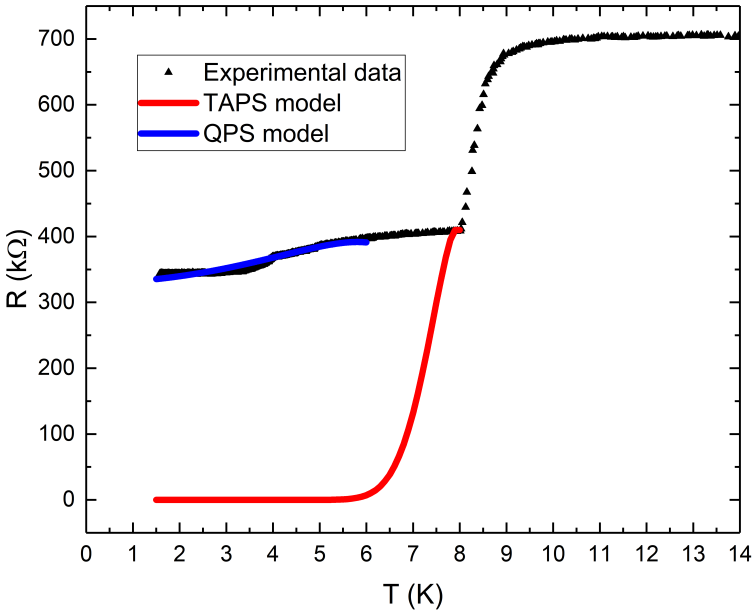


Figure 6.25: TAPS model and QPS model are used to fit experimental data. The black curve shows the experimental data of $R(T)$ results for a 60 nm wide, 10 μm long and 20 nm thick nanowire. The red curve and blue curve show the fits for TAPS model and QPS model, respectively.

In Figure 6.25, we observe that the resistance in a TAPS model drops significantly at temperatures below T_{C} due to the suppression of thermal fluctuation at very low temperatures. The QPS model is used to predict $R(T)$ dependence at temperatures below T_{C}. In this model, the resistance approaches zero at a temperature of T_{C} and has an upturn below T_{C}. Therefore, this QPS model is not appropriate for fitting data at temperatures around T_{C}, and we do not include data of the QPS model at temperatures higher than 6 K in this fitting. Using the appropriate fitting parameters, the QPS model has a good fit to the experimental data at temperatures between

1.5 K and 6 K. According to these results, the nanowire discussed here does not show any similarities to TAPS. The $R(T)$ measurement results of this nanowire may potentially show a sign of incoherent QPS. Further experiments at lower temperatures down to mK will help confirm the QPS phenomenon in this nanowire. It is also interesting to see whether this nanowire shows coherent QPS behaviors (e.g., Coulomb blockade phenomenon) at lower temperatures.

6.4 Summary

In this chapter, we provided an overview of the material selection, fabrication techniques and methods, design and measurement configurations used in this work. We discussed the development of our NbN thin film compositions and introduced the methods to fabricate NbN nanowires. We also presented characterization results of NbN thin films and low temperature characteristics of NbN nanowires with different geometries on different substrates. We showed various of experimental results of selected nanowires that showed different electrical behaviors and also implemented mathematical TAPS and QPS models to fit experimental data.

We grew NbN thin films on an oxidized silicon substrate via DC magnetron sputtering technique. We optimized the ratio of N_2/Ar flow rate and Ar pressure to obtain the possible highest T_C of NbN in CHA Mark 50 deposition system. We used EBL to define nanowire patterns on bi-layer PMMA resist stacks, followed by deposition of NbN thin film and lift-off. The NbN nanowires were thus formed on substrates of interest, which were low-conductivity p-type silicon substrate and the intrinsic silicon substrate. The T_C of our NbN was 11 K for 150 nm thick films, and we expected to see a reduction of T_C as thickness reduced to $\sim$ 20 nm. The T_C of 20 nm thick NbN was 8 - 9 K. As the coherence length is inversely proportional to T_C, there is always a trade-off between coherence length and critical temperature. Thus, we believe that T_C from a range of 8 - 9 K might be good enough for nanowires to exhibit QPS.

We presented the low temperature characteristics of nanowires fabricated from thin films of NbN. The temperature dependence of resistance for most of the nanowires showed an increasingly insulating behavior as dimensions were confined through the reduction of widths. This is due to the disorder in the electrical characteristics as the current path in the material is reduced. Nanowires with widths greater than 65 nm were demonstrated to be superconducting

at temperatures below 7.9 K through measurement results of resistance-temperature dependency and I-V characteristic curves. The critical current I_C of most of the wider nanowires was proportional to the width of the nanowire.

We also observed different electrical characteristics in some of the narrow nanowires. We used small AC testing current to obtain temperature dependence of AC resistance, which showed a very high resistive tail at temperatures below T_C. The I-V characteristic curve indicated that a potential Coulomb blockade existed in one narrow nanowire. At low current bias, the current was nearly zero before the voltage reached a critical voltage. At high current bias, the nanowire showed a critical current value of ~ 2.3 μA. These results indicate that the nanowire has a nonlinear I-V characteristic, which behaves like an insulator at a low bias regime while behaves like a superconductor at high a bias regime. The DC bias current dependence of differential resistance measurements was carried out in both the pulse tube system and the PPMS system. The results were consistent with the previous results. However, we also observed Coulomb blockade behavior at temperatures higher than T_C, which suggested that the observed phenomenon in this nanowire was not due to QPS. More experiments were performed for this nanowire, and repeatable results were observed. These unexplained phenomena were expected to be a result of unfiltered noise or low-capacitance tunnel junctions formed in this device.

More efforts have been put into this work in search of evidence of phase-slips, in addition to the experiments we discussed earlier. We used the TAPS model and QPS model to fit experimental data of a narrow nanowire that showed a very high resistive tail at low temperatures. The TAPS model suggested a significant reduction of resistance as temperature decreased, which was not consistent with experimental data. However, using appropriate fitting parameters for a QPS model, the mathematical model could fit the experimental data. Therefore, this nanowire was expected to show some behaviors of QPS. We can use lower temperatures measurements, for example, in a dilution refrigerator, to help conclude that the observed phenomenon is indeed QPS. The remaining high resistance at even lower temperatures (e.g., mK) could be evidence

for QPS. The demonstration of incoherent QPS through $R(T)$ measurements in a dilution refrigerator, and more interestingly, the demonstration of coherent QPS are the goals of future research on this topic.

In the next chapter, we will summarize all the work presented in this dissertation, including the design and simulation of QPSJ-based neuromorphic systems and the design, fabrication, and testing of superconducting NbN nanowires. We will discuss the future improvement of QPSJ-based neuromorphic circuits and potential applications of QPSJ-based neuromorphic systems. New methods to fabricate nanowires, design of nanowires, and future experiments to demonstrate QPS will also be presented.

Chapter 7

Conclusion

7.1 Conclusions

Neuromorphic computing, as a non-von Neumann computational architecture, has received more and more attention recently. Inspired by human brains, neuromorphic computing has brought together electronics, biology, materials, and algorithm. The massive interconnections in CMOS-based neuromorphic hardware (e.g., VLSI neuromorphic systems) require significant power, which is several orders of magnitude higher than a human brain, to perform the same functionality. However, superconductors that provide loss-less interconnections could be an ideal candidate for large-scale neuromorphic computing. In this dissertation, we contribute to the development of superconducting neuromorphic hardware systems using an emerging QPSJ technology with innovative circuit design.

We presented circuit design to emulate multiple functional components existing in a biological nervous system, such as neuron (soma), synapse, axon, and dendrite. A simplified STDP learning circuit was also proposed to perform potential learning functions for unsupervised learning. These circuits can be combined to form a spiking neural network with an embedded learning algorithm. The functionalities of proposed neuromorphic circuits were demonstrated through circuit simulation in WRspice using a QPSJ SPICE model.

Circuit performance with respect to power dissipation and processing speed has been discussed in detail. The equations used to estimate energy and delay per spiking event are provided. The calculation results show that the energy per spiking event for a QPSJ neuron circuit is typically several orders of magnitude less than the energy dissipation from many competing

technologies. The processing speed of QPSJ technology is expected to be at the same level as JJ technology, which is as fast as hundreds of GHz or even faster. We also showed the simulation results of power dissipation in a simple learning circuit. The experimental challenges of QPSJ technology, and circuit design challenges such as scalability and tolerance, have been addressed. The promising simulation and calculation results also motivated us to fabricate reproducible and controllable QPSJs.

We showed the efforts that have been made to develop NbN sputtering process and fabricate ultra narrow superconducting nanowires in search of quantum phase-slips. We compared different superconducting materials that are suitable for QPS based on the ratio of phase-slip energy and inductive energy. The methods used to deposit NbN thin films and fabricate NbN nanowires from NbN thin film are introduced. We used EBL to define nanowire patterns and the lift-off method to transfer NbN nanowires onto silicon substrates. The narrowest nanowire we have successfully fabricated is $\sim$ 50 nm wide. We characterized NbN thin films using $R(T)$ measurement results and XRD analysis results, based on which we developed NbN sputtering process. The highest T_{C} of deposited NbN was $\sim$ 11 K at a thickness of 140 nm and T_{C} reduced to 8 - 9 K at a thickness of 20 nm.

We compared experimental results of nanowires with different lengths and widths on different substrates. The wider nanowires (widths from 92 nm to 208 nm) typically showed superconductivity, as demonstrated by $R(T)$ and I-V measurement results. We also noted that the nanowires on the doped silicon substrate had more residual resistance compared to the nanowires on the intrinsic silicon substrate, even they had similar geometries. Most of these nanowires exhibited a critical current I_{C} with a value below 10 μA.

One of the narrowest nanowires appeared to show some QPS characteristics. We observed very high residual resistance (more than 11 MΩ) below T_{C} in this nanowire. I-V measurement results showed a critical voltage of $\sim$ 100 mV at zero bias current and a critical current of 2.3 μA. A Coulomb blockade was formed at low bias current, as seen from differential resistance (dV/dI) measurement results. This phenomenon did not only exist at temperatures below T_{C}, but also at temperatures much higher than T_{C}. We have demonstrated that these results were repeatable through experiments using different measurement setups. This phenomenon was

expected to be a result of noise or an unexpected low-capacitance tunnel junction formed in this nanowire.

Phase-slip models, including the TAPS model and QPS model, were used to fit the $R(T)$ measurement results of a narrow nanowire. We observed that in a TAPS model, the residual resistance dropped fast and approached zero at $\sim$ 6 K as temperature decreased. The TAPS model is not consistent with our experimental data. The QPS model, when using appropriate fitting parameters, has a good fit for the experimental data. The very high residual resistance that appeared at temperatures down to 1.5 K is expected to remain when the temperature keeps decreasing. We look forward to performing more experiments to demonstrate QPS in future work, as discussed in the following section.

7.2 Future work

Design of superconducting neuromorphic circuits and systems based on QPSJs, and fabrication of reproducible and controllable QPSJs are ongoing research topics. The work we presented here is only a small step towards the successful applications of QPSJ technology in the neuromorphic computing area.

A human brain has billions of neurons and trillions of synaptic connections. To emulate a human brain, we need a very large neural network consisting of functional neurons, synapses, etc. This requires future work on combining different circuits, tuning parameters, dealing with fan-in and fan-out challenges, and applying appropriate learning algorithms. Improving circuit tolerance is another big concern. We mentioned that some of the QPSJ-based circuits had very small operating margins. This could be an issue when a neural network is scaled up. Therefore, optimizing circuit design to improve circuit tolerance is a necessary step towards a large-scale neuromorphic system. Although much superconducting neuromorphic hardware has been proposed since the last few decades, we could hardly witness any physical implementations and practical applications of these advanced technologies. To advance QPSJ-based neuromorphic hardware to a useful point, we envision many useful applications of this hardware platform, such as pattern recognition and character recognition. This requires ongoing

efforts on the aforementioned integration of different circuits to build a large network, development of encoding and decoding methods for high-speed signals, training with a large dataset, and performing real-time learning.

The experimental results of NbN nanowires presented in this dissertation are a starting point for exploring fabrication and characterization methods for QPSJs. To obtain reproducible QPSJs, using different materials and fabrication methods are necessary. Other superconducting materials such as NbTiN, InO_x and NbSi can also be used for the fabrication of superconducting nanowires. We only used PMMA resist and lift-off method to fabricate nanowire in this work. There are more potential fabrication methods, such as dry etching method and ion milling method, which could help make high-quality nanowires and further reduce the width, as mentioned in [56]. Dry etching (e.g., reactive ion etching) is a more general approach to fabricating nano-scale devices. Using negative tone resist or positive tone resist along with stencil patterns, the defined nanowire patterns are protected by resist during etching. An advantage of this technique is that any over-etching will only result in narrower nanowires, which are desired in this work. In addition to that, nanowires made by dry etching do not exhibit "lily-paddings", which can be easily found in the lift-off process. As we discussed in Section 6.1.2, QPS is expected to exist in nanowires that have certain geometries, ratios of QPS energy, and inductive energy, and are embedded in high impedance environment [138]. This may require fabrication of additional circuit elements, such as high-resistance resistors and inductors of high kinetic inductance in series with the nanowire, as seen in [56, 51]. One ongoing experiment is using a focused ion beam method, proposed in [154], to irradiate NbN and potentially make the superconducting-insulating transition in NbN.

In this dissertation, we showed experimental results of DC characteristics. The future work also includes experiments on AC characteristics of a QPSJ (i.e., dual Shapiro steps), which have not been experimentally demonstrated. Another interesting point is the modulation of critical voltage V_C of QPSJs. Experimental results have shown that the critical voltage of a TiN nanowire can be tuned in a magnetic field [139]. The critical voltage of a single-charge transistor made of two QPSJs and a capacitor island can be tuned by modulating the charge on the capacitor island [103]. Inspired by these promising results, we expect to see that the critical

voltage of a single QPSJ can be modulated by the charge on a shunt capacitor in future studies. Most of the experimental results I presented in this dissertation were carried out at temperatures above 1.5 K. We can perform further measurements in a dilution refrigerator to confirm the QPS behaviors in some of these nanowires, as we work towards repeatable and controllable QPS devices at temperatures near and above 4 K. Once reproducible and controllable QPSJs are successfully fabricated, the next step is to fabricate and test simple neuromorphic circuits, such as the IFN circuit.

In conclusion, we presented an emerging superconducting technology based on quantum phase-slip junctions (QPSJs) along with a corresponding QPSJ SPICE model used to explore applications in neuromorphic circuit design. We showed a family of neuromorphic circuits to emulate their biological counterparts and demonstrated individual and combined circuit functions through simulations in WRspice. Simplified and complex STDP learning circuits were also proposed to potentially support the realization of unsupervised learning functions in QPSJ-based superconducting neuromorphic systems. In addition, we performed experiments to explore important fabrication conditions for QPSJs by fabricating ultra-narrow NbN nanowires and characterizing them at temperatures down to $\sim$ 1.5 K. All the work presented in this dissertation is towards the fabrication and testing of functional neuromorphic circuits using emerging QPSJ and QPSJ+JJ technologies, which are expected to advance superconducting neuromorphic hardware to a high performance level that is competitive with corresponding semiconducting approaches.

References

[1] K. Academy, "Anatomy of a neuron," accessed: 2021-09-22. [Online]. Available: https://www.khanacademy.org/science/biology/human-biology/neuron-nervous-system/a/overview-of-neuron-structure-and-function

[2] M. Devices, "Action potential," accessed: 2021-10-12. [Online]. Available: https://www.moleculardevices.com/applications/patch-clamp-electrophysiology/what-action-potential#gref

[3] S. Haykin, *Neural networks and learning machines, 3/E*. Pearson Education India, 2010.

[4] G.-q. Bi and M.-m. Poo, "Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type," *Journal of neuroscience*, vol. 18, no. 24, pp. 10 464–10 472, 1998.

[5] P. Crotty, D. Schult, and K. Segall, "Josephson junction simulation of neurons," *Physical Review E*, vol. 82, no. 1, p. 011914, 2010.

[6] M. L. Schneider, C. A. Donnelly, and S. E. Russek, "Tutorial: High-speed low-power neuromorphic systems based on magnetic Josephson junctions," *Journal of Applied Physics*, vol. 124, no. 16, p. 161102, 2018.

[7] J. M. Shainline, S. M. Buckley, A. N. McCaughan, J. T. Chiles, A. Jafari Salim, M. Castellanos-Beltran, C. A. Donnelly, M. L. Schneider, R. P. Mirin, and S. W. Nam, "Superconducting optoelectronic loop neurons," *Journal of Applied Physics*, vol. 126, no. 4, p. 044902, 2019.

[8] E. Toomey, K. Segall, and K. K. Berggren, "Design of a power efficient artificial neuron using superconducting nanowires," *Frontiers in neuroscience*, vol. 13, p. 933, 2019.

[9] N. G. N. Constantino, "Disorder in superconductors in reduced dimensions," Ph.D. dissertation, UCL (University College London), 2016.

[10] A. J. Kerman, "Flux–charge duality and topological quantum phase fluctuations in quasi-one-dimensional superconductors," *New Journal of Physics*, vol. 15, no. 10, p. 105017, 2013.

[11] O. Astafiev, L. Ioffe, S. Kafanov, Y. A. Pashkin, K. Y. Arutyunov, D. Shahar, O. Cohen, and J. Tsai, "Coherent quantum phase slip," *Nature*, vol. 484, no. 7394, p. 355, 2012.

[12] J. Mooij and Y. V. Nazarov, "Superconducting nanowires as quantum phase-slip junctions," *Nature Physics*, vol. 2, no. 3, p. 169, 2006.

[13] R. Cheng, U. S. Goteti, and M. C. Hamilton, "Spiking neuron circuits using superconducting quantum phase-slip junctions," *Journal of Applied Physics*, vol. 124, no. 15, p. 152126, 2018.

[14] ——, "Superconducting neuromorphic computing using quantum phase-slip junctions," *IEEE Transactions on Applied Superconductivity*, vol. 29, no. 5, pp. 1–5, 2019.

[15] ——, "High-speed and low-power superconducting neuromorphic circuits based on quantum phase-slip junctions," *IEEE Transactions on Applied Superconductivity*, vol. 31, no. 5, pp. 1–8, 2021.

[16] R. Cheng, U. S. Goteti, H. Walker, K. M. Krause, L. Oeding, and M. C. Hamilton, "Toward learning in neuromorphic circuits based on quantum phase slip junctions," *Frontiers in Neuroscience*, vol. 15, p. 1470, 2021.

[17] A. T. Bollinger, A. Rogachev, and A. Bezryadin, "Dichotomy in short superconducting nanowires: Thermal phase slippage vs. Coulomb blockade," *EPL (Europhysics Letters)*, vol. 76, no. 3, p. 505, 2006.

[18] B. Rajendran and F. Alibart, "Neuromorphic computing based on emerging memory technologies," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 6, no. 2, pp. 198–211, 2016.

[19] G. E. Moore, "Cramming more components onto integrated circuits," *Proceedings of the IEEE*, vol. 86, no. 1, pp. 82–85, 1998.

[20] J. D. Owens, M. Houston, D. Luebke, S. Green, J. E. Stone, and J. C. Phillips, "Gpu computing," *Proceedings of the IEEE*, vol. 96, no. 5, pp. 879–899, 2008.

[21] A. Steane, "Quantum computing," *Reports on Progress in Physics*, vol. 61, no. 2, p. 117, 1998.

[22] C. Mead, "Neuromorphic electronic systems," *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1629–1636, 1990.

[23] B. Linares-Barranco, T. Serrano-Gotarredona, L. A. Camuñas-Mesa, J. A. Perez-Carrasco, C. Zamarreño-Ramos, and T. Masquelier, "On spike-timing-dependent-plasticity, memristive devices, and building a self-learning visual cortex," *Frontiers in Neuroscience*, vol. 5, p. 26, 2011.

[24] J.-s. Seo, B. Brezzo, Y. Liu, B. D. Parker, S. K. Esser, R. K. Montoye, B. Rajendran, J. A. Tierno, L. Chang, D. S. Modha *et al.*, "A 45nm CMOS neuromorphic chip with a scalable architecture for learning in networks of spiking neurons," in *Custom Integrated Circuits Conference (CICC), 2011 IEEE*. IEEE, 2011, pp. 1–4.

[25] M. Davies, N. Srinivasa, T.-H. Lin, G. Chinya, Y. Cao, S. H. Choday, G. Dimou, P. Joshi, N. Imam, S. Jain *et al.*, "Loihi: A neuromorphic manycore processor with on-chip learning," *IEEE Micro*, vol. 38, no. 1, pp. 82–99, 2018.

[26] D. Pelvig, H. Pakkenberg, A. Stark, and B. Pakkenberg, "Neocortical glial cell numbers in human brains," *Neurobiology of aging*, vol. 29, no. 11, pp. 1754–1762, 2008.

[27] S. Ivan, K. and S. Rick, "Neuromorphic computing: from materials to systems architecture," U.S. Department of Energy, Office of Science, Tech. Rep., 2015.

[28] J. A. Renteria-Cedano, L. M. Aguilar-Lobo, J. R. Loo-Yau, and S. Ortega-Cisneros, "Implementation of a NARX neural network in a FPGA for modeling the inverse characteristics of power amplifiers," in *2014 IEEE 57th International Midwest Symposium on Circuits and Systems (MWSCAS)*, Aug. 2014, pp. 209–212.

[29] G. Lecerf, J. Tomas, S. Boyn, S. Girod, A. Mangalore, J. Grollier, and S. Saighi, "Silicon neuron dedicated to memristive spiking neural networks," in *Circuits and Systems, 2014 IEEE International Symposium*. IEEE, 2014, pp. 1568–1571.

[30] W. A. Borders, H. Akima, S. Fukami, S. Moriya, S. Kurihara, Y. Horio, S. Sato, and H. Ohno, "Analogue spin–orbit torque device for artificial-neural-network-based associative memory operation," *Applied Physics Express*, vol. 10, no. 1, p. 013007, 2016.

[31] S. Sidler, A. Pantazi, S. Woźniak, Y. Leblebici, and E. Eleftheriou, "Unsupervised learning using phase-change synapses and complementary patterns," in *International Conference on Artificial Neural Networks*. Springer, 2017, pp. 281–288.

[32] S. Sonde, B. Chakrabarti, Y. Liu, K. Sasikumar, J. Lin, L. Stan, R. Divan, L. E. Ocola, D. Rosenmann, P. Choudhury, K. Ni, S. K. R. S. Sankaranarayanan, S. Datta, and S. Guha, "Silicon compatible sn-based resistive switching memory," *Nanoscale*, vol. 10, pp. 9441–9449, 2018.

[33] P. A. Merolla, J. V. Arthur, R. Alvarez-Icaza, A. S. Cassidy, J. Sawada, F. Akopyan, B. L. Jackson, N. Imam, C. Guo, Y. Nakamura *et al.*, "A million spiking-neuron integrated circuit with a scalable communication network and interface," *Science*, vol. 345, no. 6197, pp. 668–673, 2014.

[34] E. Chicca, F. Stefanini, C. Bartolozzi, and G. Indiveri, "Neuromorphic electronic circuits for building autonomous cognitive systems," *Proceedings of the IEEE*, vol. 102, no. 9, pp. 1367–1388, 2014.

[35] A. Calimera, E. Macii, and M. Poncino, "The human brain project and neuromorphic computing," *Functional neurology*, vol. 28, no. 3, p. 191, 2013.

[36] E. Painkras, L. A. Plana, J. Garside, S. Temple, F. Galluppi, C. Patterson, D. R. Lester, A. D. Brown, and S. B. Furber, "Spinnaker: A 1-w 18-core system-on-chip for massively-parallel neural network simulation," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 8, pp. 1943–1953, 2013.

[37] D. P. Pappas, D. E. David, R. E. Lake, M. Bal, R. B. Goldfarb, D. A. Hite, E. Kim, H.-S. Ku, J. Long, C. McRae *et al.*, "Enhanced superconducting transition temperature in electroplated rhenium," *Applied Physics Letters*, vol. 112, no. 18, p. 182601, 2018.

[38] Y. Mizugaki, K. Nakajima, Y. Sawada, and T. Yamashita, "Superconducting neural circuits using SQUIDs," *IEEE Transactions on Applied Superconductivity*, vol. 5, no. 2, pp. 3168–3171, 1995.

[39] T. Onomi, Y. Maenami, and K. Nakajima, "Superconducting neural network for solving a combinatorial optimization problem," *IEEE Transactions on Applied Superconductivity*, vol. 21, no. 3, pp. 701–704, 2011.

[40] T. Hirose, T. Asai, and Y. Amemiya, "Spiking neuron devices consisting of single-flux-quantum circuits," *Physica C: Superconductivity and Its Applications*, vol. 445, pp. 1020–1023, 2006.

[41] ——, "Pulsed neural networks consisting of single-flux-quantum spiking neurons," *Physica C: Superconductivity and its applications*, vol. 463, pp. 1072–1075, 2007.

[42] S. E. Russek, C. A. Donnelly, M. L. Schneider, B. Baek, M. R. Pufall, W. H. Rippard, P. F. Hopkins, P. D. Dresselhaus, and S. P. Benz, "Stochastic single flux quantum neuromorphic computing using magnetically tunable josephson junctions," in *Rebooting Computing (ICRC), IEEE International Conference on*. IEEE, 2016, pp. 1–5.

[43] M. L. Schneider, C. A. Donnelly, S. E. Russek, B. Baek, M. R. Pufall, P. F. Hopkins, and W. H. Rippard, "Energy-efficient single-flux-quantum based neuromorphic computing," in *Rebooting Computing (ICRC), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1–4.

[44] M. L. Schneider, C. A. Donnelly, S. E. Russek, B. Baek, M. R. Pufall, P. F. Hopkins, P. D. Dresselhaus, S. P. Benz, and W. H. Rippard, "Ultralow power artificial synapses using nanotextured magnetic josephson junctions," *Science advances*, vol. 4, no. 1, p. e1701329, 2018.

[45] C. Bell, G. Burnell, C. Leung, E. Tarte, D.-J. Kang, and M. Blamire, "Controllable Josephson current through a pseudospin-valve structure," *Applied Physics Letters*, vol. 84, no. 7, pp. 1153–1155, 2004.

[46] V. V. Ryazanov, V. V. Bol'ginov, D. S. Sobanin, I. V. Vernik, S. K. Tolpygo, A. M. Kadin, and O. A. Mukhanov, "Magnetic Josephson junction technology for digital and memory applications," *Physics Procedia*, vol. 36, pp. 35–41, 2012.

[47] J. M. Shainline, S. M. Buckley, R. P. Mirin, and S. W. Nam, "Superconducting optoelectronic circuits for neuromorphic computing," *Physical Review Applied*, vol. 7, no. 3, p. 034013, 2017.

[48] E. Toomey, Q.-Y. Zhao, A. N. McCaughan, and K. K. Berggren, "Frequency pulling and mixing of relaxation oscillations in superconducting nanowires," *Physical Review Applied*, vol. 9, no. 6, p. 064021, 2018.

[49] B. A. Primavera and J. M. Shainline, "Considerations for neuromorphic supercomputing in semiconducting and superconducting optoelectronic hardware," *Frontiers in Neuroscience*, vol. 15, 2021.

[50] J. Mooij and C. Harmans, "Phase-slip flux qubits," *New Journal of Physics*, vol. 7, no. 1, p. 219, 2005.

[51] C. Webster, J. Fenton, T. Hongisto, S. Giblin, A. Zorin, and P. Warburton, "NbSi nanowire quantum phase-slip circuits: DC supercurrent blockade, microwave measurements, and thermal analysis," *Physical Review B*, vol. 87, no. 14, p. 144510, 2013.

[52] U. S. Goteti and M. Hamilton, "SPICE model implementation of quantum phase-slip junctions," *Electronics Letters*, vol. 51, no. 13, pp. 979–981, 2015.

[53] U. S. Goteti and M. C. Hamilton, "Charge-based superconducting digital logic family using quantum phase-slip junctions," *IEEE Transactions on Applied Superconductivity*, vol. 28, no. 4, pp. 1–4, June 2018.

[54] ——, "Complementary quantum logic family using Josephson junctions and quantum phase-slip junctions," *IEEE Transactions on Applied Superconductivity*, vol. 29, no. 5, pp. 1–6, 2019.

[55] N. Giordano, "Evidence for macroscopic quantum tunneling in one-dimensional superconductors," *Physical Review Letters*, vol. 61, no. 18, p. 2137, 1988.

[56] N. G. Constantino, M. S. Anwar, O. W. Kennedy, M. Dang, P. A. Warburton, and J. C. Fenton, "Emergence of quantum phase-slip behaviour in superconducting NbN nanowires: DC electrical transport and fabrication technologies," *arXiv preprint arXiv:1806.07708*, 2018.

[57] M. C. Hamilton and U. S. Goteti, "Superconducting quantum logic and applications of same," Jun. 12 2018, US Patent 9,998,122.

[58] A. K. Jain, J. Mao, and K. M. Mohiuddin, "Artificial neural networks: A tutorial," *Computer*, vol. 29, no. 3, pp. 31–44, 1996.

[59] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *The Journal of physiology*, vol. 117, no. 4, p. 500, 1952.

[60] K. Gerrow and A. Triller, "Synaptic stability and plasticity in a floating world," *Current opinion in neurobiology*, vol. 20, no. 5, pp. 631–639, 2010.

[61] J. R. Hughes, "Post-tetanic potentiation," *Physiological reviews*, vol. 38, no. 1, pp. 91–113, 1958.

[62] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943.

[63] S. Ghosh-Dastidar and H. Adeli, "Third generation neural networks: Spiking neural networks," in *Advances in Computational Intelligence*. Springer, 2009, pp. 167–178.

[64] E. D. Adrian and Y. Zotterman, "The impulses produced by sensory nerve-endings: Part ii. the response of a single end-organ," *The Journal of physiology*, vol. 61, no. 2, pp. 151–171, 1926.

[65] P. Dayan and L. F. Abbott, *Theoretical neuroscience: computational and mathematical modeling of neural systems*. MIT press, 2001.

[66] S. Wu, S.-i. Amari, and H. Nakahara, "Population coding and decoding in a neural field: a computational study," *Neural Computation*, vol. 14, no. 5, pp. 999–1026, 2002.

[67] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no. 6583, p. 607, 1996.

[68] Y. Dan and M.-m. Poo, "Hebbian depression of isolated neuromuscular synapses in vitro," *Science*, vol. 256, no. 5063, pp. 1570–1573, 1992.

[69] S. Cooke and T. Bliss, "Plasticity in the human central nervous system," *Brain*, vol. 129, no. 7, pp. 1659–1673, 2006.

[70] D. O. Hebb *et al.*, "The organization of behavior," 1949.

[71] E. L. Bienenstock, L. N. Cooper, and P. W. Munro, "Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex." DTIC Document, Tech. Rep., 1981.

[72] G. Rachmuth, H. Z. Shouval, M. F. Bear, and C.-S. Poon, "A biophysically-based neuromorphic model of spike rate-and timing-dependent plasticity," *Proceedings of the National Academy of Sciences*, vol. 108, no. 49, pp. E1266–E1274, 2011.

[73] S. Li, F. Zeng, C. Chen, H. Liu, G. Tang, S. Gao, C. Song, Y. Lin, F. Pan, and D. Guo, "Synaptic plasticity and learning behaviours mimicked through ag interface movement

in an ag/conducting polymer/ta memristive system," *Journal of Materials Chemistry C*, vol. 1, no. 34, pp. 5292–5298, 2013.

[74] J. M. Brader, W. Senn, and S. Fusi, "Learning real-world stimuli in a neural network with spike-driven synaptic dynamics," *Neural computation*, vol. 19, no. 11, pp. 2881–2912, 2007.

[75] W. Levy and O. Steward, "Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus," *Neuroscience*, vol. 8, no. 4, pp. 791–797, 1983.

[76] D. Debanne, B. Gähwiler, and S. M. Thompson, "Asynchronous pre-and postsynaptic activity induces associative long-term depression in area ca1 of the rat hippocampus in vitro." *Proceedings of the National Academy of Sciences*, vol. 91, no. 3, pp. 1148–1152, 1994.

[77] H. Markram, J. Lübke, M. Frotscher, and B. Sakmann, "Regulation of synaptic efficacy by coincidence of postsynaptic aps and epsps," *Science*, vol. 275, no. 5297, pp. 213–215, 1997.

[78] Y. Dan and M.-m. Poo, "Spike timing-dependent plasticity of neural circuits," *Neuron*, vol. 44, no. 1, pp. 23–30, 2004.

[79] J. M. Young, W. J. Waleszczyk, C. Wang, M. B. Calford, B. Dreher, and K. Obermayer, "Cortical reorganization consistent with spike timing–but not correlation-dependent plasticity," *Nature neuroscience*, vol. 10, no. 7, pp. 887–895, 2007.

[80] Z. Q. Wang, H. Y. Xu, X. H. Li, H. Yu, Y. C. Liu, and X. J. Zhu, "Synaptic learning and memory functions achieved using oxygen ion migration/diffusion in an amorphous ingazno memristor," *Advanced Functional Materials*, vol. 22, no. 13, pp. 2759–2765, 2012.

[81] K. D. Cantley, A. Subramaniam, H. J. Stiegler, R. A. Chapman, and E. M. Vogel, "Neural learning circuits utilizing nano-crystalline silicon transistors and memristors," *IEEE transactions on neural networks and learning systems*, vol. 23, no. 4, pp. 565–573, 2012.

[82] Y. Zhang, Z. Zeng, and S. Wen, "Implementation of memristive neural networks with spike-rate-dependent plasticity synapses," in *Neural Networks (IJCNN), 2014 International Joint Conference on*. IEEE, 2014, pp. 2226–2233.

[83] J. Lisman and N. Spruston, "Questions about stdp as a general model of synaptic plasticity," *Spike-timing dependent plasticity*, vol. 26, p. 53, 2010.

[84] S. Mitra, S. Fusi, and G. Indiveri, "Real-time classification of complex patterns using spike-based learning in neuromorphic VLSI," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 3, no. 1, pp. 32–42, 2009.

[85] N. Qiao, H. Mostafa, F. Corradi, M. Osswald, F. Stefanini, D. Sumislawska, and G. Indiveri, "A reconfigurable on-line learning spiking neuromorphic processor comprising 256 neurons and 128k synapses," *Frontiers in neuroscience*, vol. 9, p. 141, 2015.

[86] G. G. Turrigiano and S. B. Nelson, "Homeostatic plasticity in the developing nervous system," *Nature Reviews Neuroscience*, vol. 5, no. 2, pp. 97–107, 2004.

[87] P. U. Diehl and M. Cook, "Unsupervised learning of digit recognition using spike-timing-dependent plasticity," *Frontiers in computational neuroscience*, vol. 9, 2015.

[88] D. Querlioz, O. Bichler, P. Dollfus, and C. Gamrat, "Immunity to device variations in a spiking neural network with memristive nanodevices," *IEEE Transactions on Nanotechnology*, vol. 12, no. 3, pp. 288–295, 2013.

[89] R. Brette and W. Gerstner, "Adaptive exponential integrate-and-fire model as an effective description of neuronal activity," *Journal of neurophysiology*, vol. 94, no. 5, pp. 3637–3642, 2005.

[90] Y. Mizugaki, K. Nakajima, Y. Sawada, and T. Yamashita, "Superconducting neural circuits using fluxon pulses," *Applied physics letters*, vol. 62, no. 7, pp. 762–764, 1993.

[91] ——, "Implementation of new superconducting neural circuits using coupled squids," *IEEE transactions on applied superconductivity*, vol. 4, no. 1, pp. 1–8, 1994.

[92] A. E. Schegolev, N. V. Klenov, I. I. Soloviev, and M. V. Tereshonok, "Adiabatic superconducting cells for ultra-low-power artificial neural networks," *Beilstein journal of nanotechnology*, vol. 7, no. 1, pp. 1397–1403, 2016.

[93] K. K. Likharev and V. K. Semenov, "RSFQ logic/memory family: A new Josephson-junction technology for sub-terahertz-clock-frequency digital systems," *IEEE Transactions on Applied Superconductivity*, vol. 1, no. 1, pp. 3–28, 1991.

[94] H. Suzuki, N. Fujimaki, H. Tamura, T. Imamura, and S. Hasuo, "A 4K Josephson memory," *IEEE Transactions on Magnetics*, vol. 25, no. 2, pp. 783–788, 1989.

[95] N. V. Klenov, A. E. Schegolev, I. I. Soloviev, S. V. Bakurskiy, and M. V. Tereshonok, "Energy efficient superconducting neural networks for high-speed intellectual data processing systems," *IEEE Transactions on Applied Superconductivity*, vol. 28, no. 7, pp. 1–6, 2018.

[96] K. Berggren, Q. Xia, K. K. Likharev, D. B. Strukov, H. Jiang, T. Mikolajick, D. Querlioz, M. Salinga, J. R. Erickson, S. Pi *et al.*, "Roadmap on emerging hardware and technology for machine learning," *Nanotechnology*, vol. 32, no. 1, p. 012002, 2020.

[97] E. Varoquaux and O. Avenel, "Phase slip phenomena in superfluid helium," *Physica B: Condensed Matter*, vol. 197, no. 1-4, pp. 306–314, 1994.

[98] K. C. Wright, R. Blakestad, C. Lobb, W. Phillips, and G. Campbell, "Driving phase slips in a superfluid atom circuit with a rotating weak link," *Physical review letters*, vol. 110, no. 2, p. 025302, 2013.

[99] N. D. Mermin and H. Wagner, "Absence of ferromagnetism or antiferromagnetism in one-or two-dimensional isotropic heisenberg models," *Physical Review Letters*, vol. 17, no. 22, p. 1133, 1966.

[100] J. Lukens, R. Warburton, and W. Webb, "Onset of quantized thermal fluctuations in" one-dimensional" superconductors," *Physical Review Letters*, vol. 25, no. 17, p. 1180, 1970.

[101] R. Newbower, M. Beasley, and M. Tinkham, "Fluctuation effects on the superconducting transition of tin whisker crystals," *Physical Review B*, vol. 5, no. 3, p. 864, 1972.

[102] C. N. Lau, N. Markovic, M. Bockrath, A. Bezryadin, and M. Tinkham, "Quantum phase slips in superconducting nanowires," *Physical review letters*, vol. 87, no. 21, p. 217003, 2001.

[103] T. Hongisto and A. Zorin, "Single-charge transistor based on the charge-phase duality of a superconducting nanowire circuit," *Physical Review Letters*, vol. 108, no. 9, p. 097001, 2012.

[104] K. Y. Arutyunov, T. T. Hongisto, J. S. Lehtinen, L. I. Leino, and A. L. Vasiliev, "Quantum phase slip phenomenon in ultra-narrow superconducting nanorings," *Scientific reports*, vol. 2, no. 1, pp. 1–7, 2012.

[105] J. S. Langer and V. Ambegaokar, "Intrinsic resistive transition in narrow superconducting channels," *Physical Review*, vol. 164, no. 2, p. 498, 1967.

[106] V. L. Ginzburg and L. D. Landau, "On the theory of superconductivity," in *On Superconductivity and Superfluidity*. Springer, 2009, pp. 113–137.

[107] A. J. van Run, J. Romijn, and J. E. Mooij, "Superconduction phase coherence in very weak aluminium strips," *Japanese Journal of Applied Physics*, vol. 26, no. S3-2, p. 1765, jan 1987. [Online]. Available: https://doi.org/10.7567/jjaps.26s3.1765

[108] R. F. Voss and R. A. Webb, "Macroscopic quantum tunneling in 1-μm nb josephson junctions," *Physical Review Letters*, vol. 47, no. 4, p. 265, 1981.

[109] L. Jackel, J. Gordon, E. Hu, R. Howard, L. Fetter, D. Tennant, R. Epworth, and J. Kurkijärvi, "Decay of the zero-voltage state in small-area, high-current-density josephson junctions," *Physical Review Letters*, vol. 47, no. 9, p. 697, 1981.

[110] J. M. Martinis, M. H. Devoret, and J. Clarke, "Experimental tests for the quantum behavior of a macroscopic degree of freedom: The phase difference across a josephson junction," *Physical Review B*, vol. 35, no. 10, p. 4682, 1987.

[111] A. Davidson and M. Beasley, "Duality between superconducting and semiconducting electronics," *IEEE Journal of Solid-State Circuits*, vol. 14, no. 4, pp. 758–762, 1979.

[112] K. Likharev, "Single-electron transistors: Electrostatic analogs of the dc squids," *IEEE transactions on magnetics*, vol. 23, no. 2, pp. 1142–1145, 1987.

[113] A. Kadin, "Duality and fluxonics in superconducting devices," *Journal of applied physics*, vol. 68, no. 11, pp. 5741–5749, 1990.

[114] A. Larkin, K. Likharev, and Y. N. Ovchinnikov, "Secondary quantum macrpscopic effects in weak superconductivity," *Physica B+ C*, vol. 126, no. 1-3, pp. 414–422, 1984.

[115] A. Zaikin and S. Panyukov, "Dynamics of a quantum dissipative system: Duality between coordinate and quasimomentum spaces," *Physics Letters A*, vol. 120, no. 6, pp. 306–311, 1987.

[116] K. Likharev and A. Zorin, "Theory of the bloch-wave oscillations in small josephson junctions," *Journal of low temperature physics*, vol. 59, no. 3, pp. 347–382, 1985.

[117] I. S. Beloborodov, F. Hekking, and F. Pistolesi, "Influence of thermal fluctuations on an underdamped josephson tunnel junction," in *New Directions in Mesoscopic Physics (Towards Nanoscience)*. Springer, 2003, pp. 339–349.

[118] W. Guichard and F. W. Hekking, "Phase-charge duality in josephson junction circuits: Role of inertia and effect of microwave irradiation," *Physical Review B*, vol. 81, no. 6, p. 064508, 2010.

[119] J. Mooij, G. Schön, A. Shnirman, T. Fuse, C. Harmans, H. Rotzinger, and A. Verbruggen, "Superconductor–insulator transition in nanowires and nanowire arrays," *New Journal of Physics*, vol. 17, no. 3, p. 033006, 2015.

[120] D. Averin and K. Likharev, "Coulomb blockade of single-electron tunneling, and coherent oscillations in small tunnel junctions," *Journal of low temperature physics*, vol. 62, no. 3, pp. 345–373, 1986.

[121] J. B. Reece, L. A. Urry, M. L. Cain, S. A. Wasserman, P. V. Minorsky, R. B. Jackson *et al.*, *Campbell biology*. Pearson Boston, 2011, vol. 9.

[122] G. Indiveri, B. Linares-Barranco, T. J. Hamilton, A. Van Schaik, R. Etienne-Cummings, T. Delbruck, S.-C. Liu, P. Dudek, P. Häfliger, S. Renaud *et al.*, "Neuromorphic silicon neuron circuits," *Frontiers in neuroscience*, vol. 5, p. 73, 2011.

[123] E. M. Izhikevich, "Which model to use for cortical spiking neurons?" *IEEE transactions on neural networks*, vol. 15, no. 5, pp. 1063–1070, 2004.

[124] L. F. Abbott, "Lapicque's introduction of the integrate-and-fire model neuron (1907)," *Brain research bulletin*, vol. 50, no. 5, pp. 303–304, 1999.

[125] D. S. Jeong and C. S. Hwang, "Nonvolatile memory materials for neuromorphic intelligent machines," *Advanced Materials*, vol. 30, no. 42, p. 1704729, 2018.

[126] M. Abd El Qader, R. Singh, S. N. Galvin, L. Yu, J. Rowell, and N. Newman, "Switching at small magnetic fields in Josephson junctions fabricated with ferromagnetic barrier layers," *Applied Physics Letters*, vol. 104, no. 2, p. 022602, 2014.

[127] I. P. Nevirkovets, S. E. Shafraniuk, and O. A. Mukhanov, "Multiterminal superconducting-ferromagnetic device with magnetically tunable supercurrent for memory application," *IEEE Transactions on Applied Superconductivity*, vol. 28, no. 7, pp. 1–4, 2018.

[128] D. Purves, G. J. Augustine, D. Fitzpatrick, W. C. Hall, A.-S. LaMantia, J. O. McNamara, and L. E. White, "Neuroscience. 4th," *Sunderland, Mass.: Sinauer. xvii*, vol. 857, p. 944, 2008.

[129] C. S. von Bartheld, J. Bahney, and S. Herculano-Houzel, "The search for true numbers of neurons and glial cells in the human brain: a review of 150 years of cell counting," *Journal of Comparative Neurology*, vol. 524, no. 18, pp. 3865–3895, 2016.

[130] M. Ito and M. Kano, "Long-lasting depression of parallel fiber-Purkinje cell transmission induced by conjunctive stimulation of parallel fibers and climbing fibers in the cerebellar cortex," *Neuroscience letters*, vol. 33, no. 3, pp. 253–258, 1982.

[131] T. V. Bliss and T. Lømo, "Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path," *The Journal of physiology*, vol. 232, no. 2, pp. 331–356, 1973.

[132] A. M. Kadin, *Introduction to superconducting circuits*. Wiley New York, 1999.

[133] T. Aref, A. Levchenko, V. Vakaryuk, and A. Bezryadin, "Quantitative analysis of quantum phase slips in superconducting $Mo_{76}Ge_{24}$ nanowires revealed by switching-current statistics," *Physical Review B*, vol. 86, no. 2, p. 024507, 2012.

[134] P. Merolla, J. Arthur, F. Akopyan, N. Imam, R. Manohar, and D. S. Modha, "A digital neurosynaptic core using embedded crossbar memory with 45pJ per spike in 45nm," in *Custom Integrated Circuits Conference*. IEEE, 2011, pp. 1–4.

[135] I. Sourikopoulos, S. Hedayat, C. Loyez, F. Danneville, V. Hoel, E. Mercier, and A. Cappy, "A 4-fJ/spike artificial neuron in 65 nm CMOS technology," *Frontiers in Neuroscience*, vol. 11, p. 123, 2017.

[136] A. M. Hriscu, "Theoretical proposals of quantum phase-slip devices," Ph.D. dissertation, TU Delft, Delft University of Technology, 2012.

[137] A. Hriscu and Y. V. Nazarov, "Coulomb blockade due to quantum phase slips illustrated with devices," *Physical Review B*, vol. 83, no. 17, p. 174511, 2011.

[138] K. Y. Arutyunov, J. Lehtinen, and T. Rantala, "The quantum phase slip phenomenon in superconducting nanowires with high-impedance environment," *Journal of Superconductivity and Novel Magnetism*, vol. 29, no. 3, pp. 569–572, 2016.

[139] I. Schneider, K. Kronfeldner, T. Baturina, and C. Strunk, "Quantum phase slips and number-phase duality in disordered TiN nanostrips," *Physical Review B*, vol. 99, no. 9, p. 094522, 2019.

[140] J. Peltonen, O. V. Astafiev, Y. P. Korneeva, B. Voronov, A. Korneev, I. Charaev, A. Semenov, G. Golt'Sman, L. B. Ioffe, T. Klapwijk *et al.*, "Coherent flux tunneling through NbN nanowires," *Physical Review B*, vol. 88, no. 22, p. 220506, 2013.

[141] L. Zhang, W. Peng, L. You, and Z. Wang, "Superconducting properties and chemical composition of NbTiN thin films with different thickness," *Applied Physics Letters*, vol. 107, no. 12, p. 122603, 2015.

[142] R. Pease, "Electron beam lithography," *Contemporary Physics*, vol. 22, no. 3, pp. 265–290, 1981.

[143] C. Vieu, F. Carcenac, A. Pepin, Y. Chen, M. Mejias, A. Lebib, L. Manin-Ferlazzo, L. Couraud, and H. Launois, "Electron beam lithography: resolution limits and applications," *Applied surface science*, vol. 164, no. 1-4, pp. 111–117, 2000.

[144] A. N. Broers, A. Hoole, and J. M. Ryan, "Electron beam lithography—Resolution limits," *Microelectronic engineering*, vol. 32, no. 1-4, pp. 131–142, 1996.

[145] K. W. Lee, S. Yoon, S. Lee, W. Lee, I. Kim, C. E. Lee, and D. Kim, "Secondary electron generation in electron-beam-irradiated solids: resolution limits to nanolithography," *Journal of the Korean Physical Society*, vol. 55, no. 4, pp. 1720–1723, 2009.

[146] A. Hoole, M. Welland, and A. Broers, "Negative PMMA as a high-resolution resist-the limits and possibilities," *Semiconductor science and technology*, vol. 12, no. 9, p. 1166, 1997.

[147] D. Brambley and R. Bennett, "Electron-beam resist technology for GaAs microwave device fabrication," *GEC journal of research*, vol. 13, no. 1, pp. 42–53, 1996.

[148] S. Yasin, D. Hasko, and H. Ahmed, "Comparison of MIBK/IPA and water/IPA as PMMA developers for electron beam nanolithography," *Microelectronic engineering*, vol. 61, pp. 745–753, 2002.

[149] V. R. Manfrinato, L. Zhang, D. Su, H. Duan, R. G. Hobbs, E. A. Stach, and K. K. Berggren, "Resolution limits of electron-beam lithography toward the atomic scale," *Nano letters*, vol. 13, no. 4, pp. 1555–1558, 2013.

[150] A. Rogachev, A. Bollinger, and A. Bezryadin, "Influence of high magnetic fields on the superconducting transition of one-dimensional Nb and MoGe nanowires," *Physical review letters*, vol. 94, no. 1, p. 017004, 2005.

[151] K. Masuda, S. Moriyama, Y. Morita, K. Komatsu, T. Takagi, T. Hashimoto, N. Miki, T. Tanabe, and H. Maki, "Thermal and quantum phase slips in niobium-nitride nanowires based on suspended carbon nanotubes," *Applied Physics Letters*, vol. 108, no. 22, p. 222601, 2016.

[152] A. D. Zaikin, D. S. Golubev, A. van Otterlo, and G. T. Zimányi, "Quantum phase slips and transport in ultrathin superconducting wires," *Physical review letters*, vol. 78, no. 8, p. 1552, 1997.

[153] D. S. Golubev and A. D. Zaikin, "Quantum tunneling of the order parameter in superconducting nanowires," *Physical Review B*, vol. 64, no. 1, p. 014504, 2001.

[154] S. A. Cybart, E. Cho, T. Wong, B. H. Wehlin, M. K. Ma, C. Huynh, and R. Dynes, "Nano josephson superconducting tunnel junctions in YBa2Cu3O7 7–δ directly patterned with a focused helium ion beam," *Nature nanotechnology*, vol. 10, no. 7, pp. 598–602, 2015.

[155] U. S. Goteti *et al.*, "Superconducting digital logic families using quantum phase-slip junctions," 2019.

Appendices

Appendix A

Device Parameters in WRspice Simulation

Circuit simulation in this dissertation has been performed in WRspice using a QPSJ SPICE model described in [155]. In this chapter, the device parameters for circuit simulation in WRspice are presented as follows.

Component	Value	Comment
V_{in}	1 mV	Input voltage
Q_1, Q_2	0.7 mV	Critical voltage
C	0.23 fF	Capacitance
V_{b}	1 mV	Bias voltage

Table A.1: Device parameters for the simulation shown in Figure 3.2.

Component	Value	Comment
V_{in}	1 mV	Input voltage
Q_1, Q_2	0.7 mV	Critical voltage
C	0.5 fF	Capacitance
V_{b}	1 mV	Bias voltage

Table A.2: Device parameters for the simulation shown in Figure 3.4.

Component	Value	Comment
V_{in}	1 mV	Input voltage
Q_1, Q_2	0.7 mV	Critical voltage
C	0.5 fF	Capacitance
V_{b}	1 mV	Bias voltage

Table A.3: Device parameters for the simulation shown in Figure 3.6.

Component	Value	Comment
V_{in}	0.8 mV	Input voltage
Q_1-Q_{10}	0.7 mV	Critical voltage
C	15 fF	Capacitance
V_{b}	1 mV	Bias voltage
R_{b}	9 kΩ	Resistance

Table A.4: Device parameters for the simulation shown in Figure 3.8.

Component	Value	Comment
V_{in}	1 mV	Input voltage
Q_1-Q_N	0.7 mV	Critical voltage
C	0.3/2/5 fF	Capacitance
V_{b}	1 mV	Bias voltage
R	0/5/10 kΩ	Resistance

Table A.5: Device parameters for the simulation shown in Figure 3.9.

Component	Value	Comment
V_{in}	1.4 mV	Input voltage
Q_1	0.7 mV	Critical voltage
L_1	4 pH	Inductance
V_{b}	0.5 mV	Bias voltage
I_{b}	140 μA	Bias current
J_1	200/300 μA	Critical current
R	12 Ω	Resistance

Table A.6: Device parameters for the simulation shown in Figure 3.11.

Component	Value	Comment
V_{in}	1 mV	Input voltage
Q_1	0.7 mV	Critical voltage
L_1	2 pH	Inductance
L_2	4 pH	Inductance
V_{b}	0.5 mV	Bias voltage
I_{b}	160 μA	Bias current
J_1	10/50/350/400 μA	Critical current
J_2	200 μA	Critical current
R	3 Ω	Resistance

Table A.7: Device parameters for the simulation shown in Figure 3.13.

Component	Value	Comment
V_r	0.6 mV	Read voltage
V_w	0.6 mV	Write voltage
Q_0, Q_1, Q_4	0.7 mV	Critical voltage
Q_2, Q_3	0.4 mV	Critical voltage
C_1, C_2	3 fF	Capacitance
V_{b1}	0.8 mV	Bias voltage
V_{b2}	1 mV	Bias voltage

Table A.8: Device parameters for the simulation shown in Figure 3.15.

Component	Value	Comment
V_r	0.6 mV	Read voltage
V_w	0.6 mV	Write voltage
Q_0, Q_1, Q_4	0.7 mV	Critical voltage
Q_2, Q_3	0.4 mV	Critical voltage
C_1, C_2	3 fF	Capacitance
V_{b1}	0.8 mV	Bias voltage
V_{b2}	1 mV	Bias voltage

Table A.9: Device parameters for the simulation shown in Figure 3.16.

Component	Value	Comment
V_r	0.54 mV	Read voltage
V_w	1.2 mV	Write voltage
Q_0	0.6 mV	Critical voltage
Q_1	0.54 mV	Critical voltage
Q_2	0.52 mV	Critical voltage
Q_3	0.51 mV	Critical voltage
Q_4	0.3 mV	Critical voltage
Q_5-Q_8	0.09 mV	Critical voltage
C_1	7.8 fF	Capacitance
C_2	1.2 fF	Capacitance
V_{b1}	0.2 mV	Bias voltage
V_{b2}	0.5 mV	Bias voltage

Table A.10: Device parameters for the simulation shown in Figure 3.18.

Component	Value	Comment
V_r	0.54 mV	Read voltage
V_w	0 mV	Write voltage
Q_1	0.75 mV	Critical voltage
Q_2	0.65 mV	Critical voltage
Q_3	0.42 mV	Critical voltage
Q_4, Q_5	0.6 mV	Critical voltage
Q_6	0.58 mV	Critical voltage
Q_7	2 mV	Critical voltage
C_1	9 fF	Capacitance
C_2	1 fF	Capacitance
C_3	1.2 fF	Capacitance
C_4	9.2 fF	Capacitance
V_{b1}	0.5 mV	Bias voltage
V_{b2}	1.07 mV	Bias voltage
V_{b3}	0.58 mV	Bias voltage
R_1	10 kΩ	Resistance

Table A.11: Device parameters for the simulation shown in Figure 3.20.

Component	Value	Comment
V_r	0.54 mV	Read voltage
V_w	0.9 mV	Write voltage
Q_1	0.75 mV	Critical voltage
Q_2	0.65 mV	Critical voltage
Q_3	0.42 mV	Critical voltage
Q_4, Q_5	0.6 mV	Critical voltage
Q_6	0.58 mV	Critical voltage
Q_7	2 mV	Critical voltage
C_1	9 fF	Capacitance
C_2	1 fF	Capacitance
C_3	1.2 fF	Capacitance
C_4	9.2 fF	Capacitance
V_{b1}	0.5 mV	Bias voltage
V_{b2}	1.07 mV	Bias voltage
V_{b3}	0.58 mV	Bias voltage
R_1	10 kΩ	Resistance

Table A.12: Device parameters for the simulation shown in Figure 3.21.

Component	Value	Comment
Q_1-Q_{10}	0.5 mV	Critical voltage
L_1, L_2	0.1 nH	Inductance
L_3	0.01 nH	Inductance
V_b	0.5 mV	Bias voltage
I_b	70 μA	Bias current
J_1	40 μA	Critical current
J_2	50 μA	Critical current

Table A.13: Device parameters for the simulation shown in Figure 3.23.

Component	Value	Comment
V_in	3 mV	Input voltage
Q_1-Q_{50}	0.7 mV	Critical voltage
C_1-C_{50}	0.23 fF	Capacitance
V_b	20 mV	Bias voltage

Table A.14: Device parameters for the simulation shown in Figure 3.25.

Component	Value	Comment
V_{r1}, V_{r2}	0.8 mV	Read voltage
V_w1, V_w2	0.8 mV	Write voltage
Q_0, Q_1, Q_5, Q_6	0.7 mV	Critical voltage
Q_2, Q_3, Q_7, Q_8	0.41 mV	Critical voltage
Q_4, Q_9	0.68 mV	Critical voltage
Q_{10}-Q_{12}	0.38 mV	Critical voltage
C_1-C_4	3 fF	Capacitance
C_5	13 fF	Capacitance
V_b1, V_b2	0.8 mV	Bias voltage
V_b3	1.34 mV	Bias voltage

Table A.15: Device parameters for the simulation shown in Figure 3.30.

Component	Value	Comment
V_{r1}, V_{r2}	0.5 mV	Read voltage
V_{w1}, V_{w2}	1.2 mV	Write voltage
Q_0, Q_9	0.6 mV	Critical voltage
Q_1, Q_{10}	0.51 mV	Critical voltage
Q_2, Q_{11}	0.52 mV	Critical voltage
Q_3, Q_{12}	0.54 mV	Critical voltage
Q_4, Q_{13}	0.36 mV	Critical voltage
Q_5-Q_8, Q_{14}-Q_{17}	0.15 mV	Critical voltage
Q_{18}-Q_{27}	0.075 mV	Critical voltage
C_1, C_3	5 fF	Capacitance
C_2, C_4	1 fF	Capacitance
C_5	330 fF	Capacitance
V_{b1}, V_{b2}	0.15 mV	Bias voltage
V_{b3}	0.63 mV	Bias voltage

Table A.16: Device parameters for the simulation shown in Figure 3.33.

Component	Value	Comment
V_r	0.54 mV	Read voltage
V_w	0.7 mV	Write voltage
Q_0	0.3 mV	Critical voltage
Q_1	0.5 mV	Critical voltage
Q_2	0.52 mV	Critical voltage
Q_3	0.54 mV	Critical voltage
Q_4	0.31 mV	Critical voltage
C_1	9.2 fF	Capacitance
C_2	1.2 fF	Capacitance
V_{b1}	0.5 mV	Bias voltage

Table A.17: Device parameters for the simulation shown in Figure 4.2.

Component	Value	Comment
V_{pre}	0.54 mV	presynaptic voltage
$\overline{V_{\text{post}}}$	0.95 mV	postsynaptic voltage
Q_0	0.75 mV	Critical voltage
Q_1	0.56 mV	Critical voltage
Q_2	0.31 mV	Critical voltage
C_1	9.2 fF	Capacitance
C_2	1.2 fF	Capacitance
R_1	10/20/30/40 kΩ	Resistance
V_{b1}	0.03 mV	Bias voltage
V_{b2}	0.5 mV	Bias voltage

Table A.18: Device parameters for the simulation shown in Figure 4.4.

Component	Value	Comment
V_{pre}	0.54 mV	presynaptic voltage
$\overline{V_{\text{post}}}$	0.95 mV	postsynaptic voltage
Q_0	0.75 mV	Critical voltage
Q_1	0.55 mV	Critical voltage
Q_2	0.3 mV	Critical voltage
Q_3	2 mV	Critical voltage
Q_4	0.54 mV	Critical voltage
Q_5	0.52 mV	Critical voltage
Q_6	0.5 mV	Critical voltage
Q_7	0.34 mV	Critical voltage
C_1	9.2 fF	Capacitance
C_2	1.2 fF	Capacitance
C_3	9.2 fF	Capacitance
C_4	1.2 fF	Capacitance
R_1	10 kΩ	Resistance
V_{b1}	0.03 mV	Bias voltage
V_{b2}	0.77 mV	Bias voltage
V_{b3}	0.53 mV	Bias voltage

Table A.19: Device parameters for the simulation shown in Figure 4.6.

Component	Value	Comment
V_{pre}	0.78 mV	presynaptic voltage
$\overline{V_{\text{pre}}}$	0.54 mV	presynaptic voltage
V_{post}	0.51 mV	postsynaptic voltage
Q_0	0.4 mV	Critical voltage
Q_1	0.5 mV	Critical voltage
Q_2	1 mV	Critical voltage
Q_3	1 mV	Critical voltage
Q_4	0.58 mV	Critical voltage
Q_5	2 mV	Critical voltage
Q_6	1.04 mV	Critical voltage
Q_7	1.02 mV	Critical voltage
Q_8	1 mV	Critical voltage
Q_9	0.28 mV	Critical voltage
C_1	9 fF	Capacitance
C_2	1 fF	Capacitance
C_3	9.2 fF	Capacitance
C_4	2 fF	Capacitance
R_1	10 kΩ	Resistance
R_2	10 kΩ	Resistance
V_{b1}	0.05 mV	Bias voltage
V_{b2}	0.2 mV	Bias voltage
V_{b3}	1.1 mV	Bias voltage
V_{b4}	1.01 mV	Bias voltage
V_{b5}	0.6 mV	Bias voltage

Table A.20: Device parameters for the simulation shown in Figure 4.8.

Component	Value	Comment
V_{pre}	0.78 mV	presynaptic voltage
V_{post}	0.51 mV	postsynaptic voltage
Q_0	0.8 mV	Critical voltage
Q_1	0.95 mV	Critical voltage
Q_2	0.95 mV	Critical voltage
Q_3	0.36 mV	Critical voltage
Q_4	0.5 mV	Critical voltage
Q_5	1 mV	Critical voltage
Q_6	1 mV	Critical voltage
Q_7	0.58 mV	Critical voltage
Q_8	2 mV	Critical voltage
Q_9	1.04 mV	Critical voltage
Q_{10}	1.02 mV	Critical voltage
Q_{11}	1 mV	Critical voltage
Q_{12}	0.28 mV	Critical voltage
C_1	1 fF	Capacitance
C_2	9 fF	Capacitance
C_3	1 fF	Capacitance
C_4	9.2 fF	Capacitance
C_5	2 fF	Capacitance
R_1	10 kΩ	Resistance
R_2	10 kΩ	Resistance
R_3	10 kΩ	Resistance
V_{b1}	0.2 mV	Bias voltage
V_{b2}	1.1 mV	Bias voltage
V_{b3}	0.05 mV	Bias voltage
V_{b4}	0.2 mV	Bias voltage
V_{b5}	1.1 mV	Bias voltage
V_{b6}	1.01 mV	Bias voltage
V_{b7}	0.6 mV	Bias voltage

Table A.21: Device parameters for the simulation shown in Figure 4.10.

Component	Value	Comment
V_{pre}	1.07 mV	presynaptic voltage
V_{post}	0.51 mV	postsynaptic voltage
$\overline{V_{\text{post}}}$	0.51 mV	postsynaptic voltage
Q_0	0.8 mV	Critical voltage
Q_1	0.95 mV	Critical voltage
Q_2	0.95 mV	Critical voltage
Q_3	0.36 mV	Critical voltage
Q_4	0.5 mV	Critical voltage
Q_5	1 mV	Critical voltage
Q_6	1 mV	Critical voltage
Q_7	0.46 mV	Critical voltage
Q_8	2 mV	Critical voltage
Q_9	0.75 mV	Critical voltage
Q_{10}	0.55 mV	Critical voltage
Q_{11}	0.3 mV	Critical voltage
Q_{12}	1.37 mV	Critical voltage
Q_{13}	1.35 mV	Critical voltage
Q_{14}	1.33 mV	Critical voltage
Q_{15}	0.28 mV	Critical voltage
C_1	1 fF	Capacitance
C_2	9 fF	Capacitance
C_3	1 fF	Capacitance
C_4	9.2 fF	Capacitance
C_5	1.2 fF	Capacitance
C_6	9.2 fF	Capacitance
C_7	2 fF	Capacitance
R_1	10 kΩ	Resistance
R_2	10 kΩ	Resistance
R_3	20 kΩ	Resistance
R_4	10 kΩ	Resistance
R_5	20 kΩ	Resistance
V_{b1}	0.2 mV	Bias voltage
V_{b2}	1.1 mV	Bias voltage
V_{b3}	0.05 mV	Bias voltage
V_{b4}	0.2 mV	Bias voltage
V_{b5}	1.1 mV	Bias voltage
V_{b6}	0.89 mV	Bias voltage
V_{b7}	0.46 mV	Bias voltage
V_{b8}	0.3 mV	Bias voltage
V_{b9}	0.6 mV	Bias voltage

Table A.22: Device parameters for the simulation shown in Figure 4.13.

Component	Value	Comment
V_{pre}	0.54 mV	presynaptic voltage
$\overline{V_{\text{post}}}$	0.95 mV	postsynaptic voltage
Q_0-Q_2	0.7 mV	Critical voltage
Q_3	0.5 mV	Critical voltage
Q_4	0.52 mV	Critical voltage
Q_5	0.54 mV	Critical voltage
Q_6	0.3 mV	Critical voltage
C_1	3 fF	Capacitance
C_2	1.2 fF	Capacitance
R_1	10 kΩ	Resistance
V_{b1}	72.6 μV	Bias voltage
V_{b2}	0.5 mV	Bias voltage

Table A.23: Device parameters for the simulation shown in Figure 4.15.

Appendix B

Nanowire Fabrication Traveler

Step #	Process	Parameters
Nanowire definition and deposition		
1	Substrate clean	(1) Acetone sonication bath for 5 min. (2) Rinse in IPA/DI water. (3) N_2 dry.
2	Dehydration bake	110 °C on hotplate for 10 min.
3	Spin PMMA 495K A2	(1) 500 rpm, Acc. = 500 rpm/s, time = 5 s. (2) 2000 rpm, Acc. = 1000 rpm/s, time = 60 s.
4	Soft bake	180 °C on hotplate for 3 min.
5	Spin PMMA 950K A2	(1) 500 rpm, Acc. = 500 rpm/s, time = 5 s. (2) 4000 rpm, Acc. = 1000 rpm/s, time = 60 s.
6	Soft bake	180 °C on hotplate for 5 min.
7	EBL for interconnection patterns	Area dose = 280 μC/cm^2, stepsize = 20 nm.
8	EBL for nanowire patterns	Area dose = 250-440 μC/cm^2, stepsize = 4 nm.
9	Development	(1) 1 min in MIBK:IPA (1:3). (2) 20 s in IPA.
10	NbN deposition	(1) 2 min of ion milling. (2) Power = 1kW, N_2/Ar = 0.1, Ar pressure = 4 mTorr. (3) 5 min pre-sputtering. (4) 3 min sputtering.
11	Lift-off	(1) Acetone sonication bath for 20 min. (2) Rinse in IPA/DI water. (3) N_2 dry.
Contact pads definition and deposition		
12	Dehydration bake	110 °C on hotplate for 10 min.
13	Spin PMMA 950K A7	(1) 500 rpm, Acc. = 500 rpm/s, time = 5 s. (2) 2000 rpm, Acc. = 1000 rpm/s, time = 60 s.
14	Soft bake	180 °C on hotplate for 2 min.
15	EBL for contact pads	Area dose = 250 μC/cm^2, stepsize = 20 nm.
16	Development	(1) 2.5 min in MIBK:IPA (1:3). (2) 20 s in IPA.
17	Al deposition	(1) 2 min of ion milling. (2) Thickness = 250 nm.
18	Lift-off	(1) Acetone sonication bath for 5 min. (2) Rinse in IPA/DI water. (3) N_2 dry.
Dicing and packaging		
19	Dicing	Dice samples into 5 mm $\times$ 5 mm dies.
20	Wire-bonding	Wire-bond dies to LCC packages.

Table B.1: Traveler of NbN nanowire fabrication.