# Fluxonic Processing of Photonic Synapse Events

Jeffrey M. Shainline, *Member, IEEE*

*(Invited Paper)*

***Abstract*—Much of the information processing performed by a biological neuron occurs in the dendritic tree. For artificial neural systems using light for communication, it is advantageous to convert signals to the electronic domain at synaptic terminals, so dendritic computation can be performed with electrical circuits. Here, we present circuits based on Josephson junctions and mutual inductors that act as dendrites, processing signals from synapses receiving single-photon communication events with superconducting detectors. We show simulations of circuits performing basic temporal filtering, logical operations, and nonlinear transfer functions. We further show how the synaptic signal from a single photon can fan out locally in the electronic domain to enable the dendrites of the receiving neuron to process a photonic synapse event or pulse train in multiple different ways simultaneously. Such a technique makes efficient use of photons, energy, space, and information.**

***Index Terms*—Neural systems, superconducting electronics, integrated photonics.**

## I. Introduction

A BIOLOGICAL neuron is a complex information processing device [1], integrating signals from thousands of inputs and producing pulses when those signals reach threshold. These neuronal firing events consume the most energy of any operation performed by a neuron. To optimize spatial, temporal, and energy efficiency, the neurons receiving the signals must extract as much information as possible from each pulse [2]. Neurons accomplish this through processing occurring in synapses and dendrites. Because neural information is based on sequences of pulses, the relevant processing involves applying temporal and logical filters to extract relevant data. For example, synapses perform temporal filtering of pulse trains to identify rising edges and to identify pulse trains exceeding some duration or number of pulses [3]. Dendrites receive and further process synaptic signals. The operations performed by dendrites include leaky integration [4]; logical operations [5]; identification of coincidences [5] and sequences [6], [7] between synapses from different neurons; and nonlinear thresholding transfer functions on signals from groups of synapses [8]. Inhibitory neurons in the network can temporarily suppress the activity of a dendrite to dynamically direct attention to information of interest [9], thereby adapting the structural network into myriad functional networks [10].

Within a point-neuron model [4], each neuron performs leaky integration of the synaptic activities with a single decay time constant, $\tau$. Thus, a neuron is capable of answering the question, “Is the sum of activity across all synapses in the last $\tau$ seconds greater than threshold?” If the answer is yes, the neuron produces a pulse. While such a model may be useful for certain neuromorphic computations, the functionality of this neuron model is significantly reduced in comparison to its biological referent. Computations occurring at synapses and dendrites in biological neurons allow those neurons to answer subtle and varied questions such as, “How long has it been since neuron $i$ last produced a pulse?” “How many pulse trains have begun and then ceased on neuron $i$ in the last $\tau_i$ seconds?” “How many times have neurons $i$ and $j$ fired within $\tau_{ij}$ seconds of each other in the last $\tau_q$ seconds?” “Have five or more of the neurons in cluster $x$ fired in the last $\tau_x$ seconds?”

The present work is concerned with artificial hardware capable of neural information processing. Such hardware is anticipated to be used both in the scientific study of the mechanisms of neural information processing as well as in technological applications that benefit from neuromorphic computing. In previous studies, we have considered artificial hardware based on superconducting optoelectronic circuits to achieve point-neuron functionality [11]–[13]. The present work builds on those circuit concepts, introducing new superconducting electronic circuitry to perform functions associated with dendritic processing in biological systems. For hardware to be efficient for neural information processing, synaptic and dendritic operations must be efficiently manifest in constituent devices. We have argued elsewhere that light is promising for communication in neural systems because it enables the fan-out and energy efficiency necessary for large neural systems, and that utilization of superconducting single-photon detectors enables communication at the lowest possible light levels [11]. Subsequent work considered specific synaptic and neuronal circuits suitable for point-neuron behavior, introducing circuits capable of transducing single-photon communication events to the electronic domain for further information processing [12], [13]. References 12 and 13 discussed basic synaptic functionality, plasticity, neuronal integration, thresholding, and the production of light during a neuronal firing event—all functions necessary for point neurons implemented with superconducting optoelectronic hardware. Due to the prominent role of flux storage loops, these circuits are referred to as loop neurons. The significance of light and

Manuscript received April 3, 2019; revised June 28, 2019; accepted July 3, 2019. Date of publication July 9, 2019; date of current version August 7, 2019.

The author is with the National Institute of Standards and Technology, Boulder, CO 80305 USA (e-mail: jeffrey.shainline@nist.gov).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTQE.2019.2927473

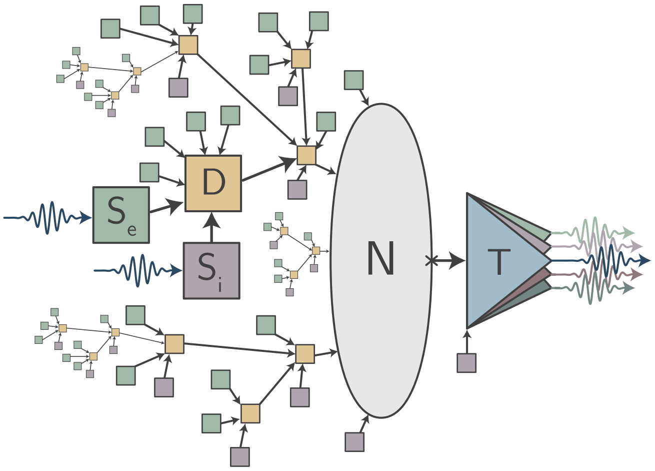

Fig. 1. Schematic of the neuron under consideration. Optical signals are represented by wavy, colored arrows, while electrical signals are represented by straight, black arrows. The complex structure consists of excitatory and inhibitory synapses ($\mathrm{S_e}$ and $\mathrm{S_i}$) that feed into dendrites (D). Each dendrite performs computations on the inputs and communicates the result to other dendrites for further processing or on to the cell body of the neuron (N). The neuron itself acts as the final thresholding stage, and when its threshold is reached, light is produced by the transmitter (T), which is routed to downstream synaptic connections. Multiple photons with different colors are shown to emanate from the transmitter, indicating the potential to use different frequencies of light for different operations [13].

superconductors for scaling was analyzed in [14]. Other work has investigated photonics [15]–[19] and superconducting electronics [20]–[25] for neuromorphic computation, but to our knowledge none of this work has pursued dendritic processing beyond point neurons or the integration of photonics with superconducting electronics to leverage their complementary strengths for communication and computation.

The purpose of this paper is to consider specific circuits implementing a more elaborate model for superconducting optoelectronic neural information processing in which the dendritic tree extracts significantly more information about synaptic activities than a simple sliding average. The model is illustrated schematically in Fig. 1, and three elemental circuits that serve as building blocks to perform many synaptic and dendritic functions are shown in Fig. 2. The model involves synapses, dendrites, and a neuron cell body. We define these terms below. These elements are envisioned to operate in the context of networks of superconducting optoelectronic loop neurons described elsewhere [12], [13].

In the present article, we work with the following component definitions. A *synapse* is a circuit that receives photons with a superconducting nanowire single-photon detector and transduces the signal to an electrical current circulating in a storage loop. Specifically, Fig. 2(a) is the circuit diagram of a synapse. A *dendrite* is a circuit that receives as input a signal proportional to the electrical output of one or more synapses and/or dendrites, performs a transfer function on the sum of the inputs, and produces an electrical current circulating in a storage loop as the output. Specifically, Fig. 2(b) is the circuit diagram of a dendrite. A *neuron cell body* receives as input a signal proportional to the electrical output of one or more synapses and/or dendrites, performs a threshold operation on the sum of the inputs, and produces as output a pulse of photons if the threshold is exceeded. Following the production of a pulse, the neuron cell body experiences a refractory period wherein threshold is temporarily significantly elevated, making subsequent pulses temporarily unlikely or impossible. The circuits that accomplish the thresholding and electrical-to-optical transduction are discussed in Refs. 12 and 13. Based on this defintion,

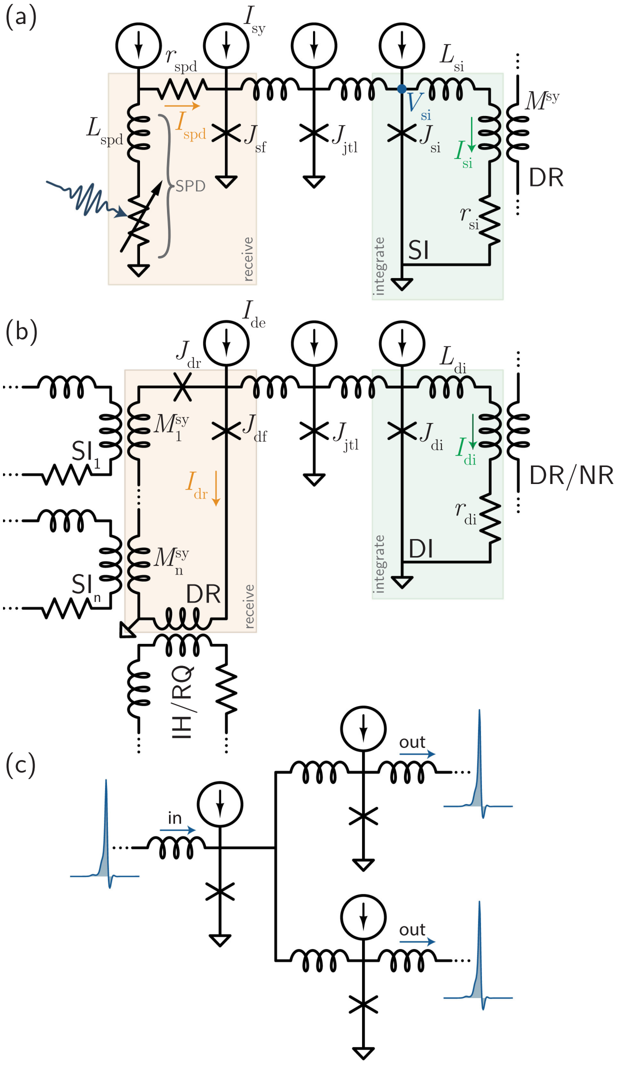

Fig. 2. Diagrams of the circuits under consideration. (a) Synaptic transducer. A photonic communication event with one or more photons diverts current from the single-photon detector (SPD) to the synaptic firing junction ($I_{\mathrm{spd}}$ to $J_{\mathrm{sf}}$). A series of fluxons is produced, and these fluxons traverse the Josephson transmission line ($J_{\mathrm{jtl}}$) and result in an integrated current in the synaptic integration loop ($I_{\mathrm{si}}$). This synaptic signal is communicated to a dendritic receiving (DR) loop through a mutual inductor. (b) Dendritic circuit. The dendritic receiver loop sums the signals from afferent synapses, and upon reaching the threshold established by the dendritic firing junction ($J_{\mathrm{df}}$), one or a series of fluxons is generated. Inhibitory (IH) or rapid query (RQ) synapses can also be established on the DR loop. The generated signal is communicated to other dendritic receiving loops or to the neuronal receiving (NR) loop of the neuron cell body. While drawn in the same place, either IH or RQ will be present on a given dendrite, and these loops require opposite signs of mutual inductance. (c) Fluxon pulse splitter. This circuit is used to make electronic copies of the information generated by a synapse or a dendrite, and the amplitudes of the current pulses at the outputs are restored to the input level.

the components N and T in the schematic of Fig. 1 comprise the analog to the cell body, although it is more biologically accurate to associate the transmitter T in this hardware with the axon hillock of a biological neuron.

A dendrite can behave similarly to a neuron cell body in that both can perform a nonlinear threshold function on their inputs. However, dendrites and neurons are very different with respect to both physics and functions. In the context of the hardware under consideration we make the distinction that a dendrite produces an electrical output that is to be communicated locally, while a neuron cell body produces an optical output that is to be communicated to synapses that may be spatially distant. Additionally, dendrites are more appropriately conceived as nonlinear filters, with different dendrites performing different transfer functions. In contrast, all neuron cell bodies are envisioned to perform only the thresholding function leading to spike production. It follows that the outputs from dendrites are functions with analog amplitude and a continuous temporal envelope, while the outputs from neuron cell bodies are stereotypical spike events wherein the amplitude is intended to be constant across spikes and the temporal envelope is intended to approximate a delta function. The amplitude of the output from a neuron cell body carries no information, and all information output from the neuron cell body is encoded in the timing of the spikes. The flexibility to implement nonlinear transformations in the electrical domain relatively easily in comparison to optical implementations motivates these hardware design choices.

The term *dendritic tree* refers collectively to all the synapses and dendrites that feed into a neuron cell body. Figure 1 is intended to illustrate the potential complexity and diversity of the dendritic tree. The output optical signals from a neuron cell body reach downstream synapses through a network of dielectric waveguides, optical fibers, and free-space interconnects, as described in [13] and [26]–[28]. These optical paths are collectively referred to as the *axonal tree* and are not shown in the schematic diagram. We define a *neuron* to be a system comprising a dendritic tree, a neuron cell body, and an axonal tree.

This article is focused on the dendritic tree, the circuits that comprise it, and some of the functions these circuits can perform. The functions considered here are leaky integration, temporal filtering of afferent pulse trains, logical operations, detection of coincidences between activities of input neurons, inhibition, and power-law memory retention of synaptic activity. In biological systems, these functions occur through nonlinearities resulting from dendritic conductances and arbor morphology [5], [6]. The Josephson circuits presented here are not intended to quantitatively reproduce biological behaviors, but rather to perform logical, temporal, and nonlinear functions in the spirit of synaptic and dendritic processing. Josephson circuits are remarkably capable of these operations due to the nonlinearity established by the existence of a critical current; the avoidance of cross talk and current leakage pathways enabled by coupling through mutual inductors; and the ability to establish essentially arbitrary time constants across many orders of magnitude by choosing the inductance and resistance of current storage loops.

This work is based on time-domain circuit simulations of the three elemental circuits shown in Fig. 2 when combined in various configurations. In Section II we review the basic operations of a synapse that transduces a single-photon communication event to the superconducting electronic domain for information processing, and in Section III we consider operations performed on pulse trains at a single synapse, usually associated with short-term plasticity and synaptic computation. In Section IV we consider the detection of coincidences between two or more synapses, and we show how the same circuits can be used with broken temporal symmetry to identify sequences of activity. For these various fragments of information to be utilized only when relevant, inhibition can be used to silence specific dendrites at appropriate times, as discussed in Section V. A central premise of the work in Refs. [11]–[14], [28] is that scalable neural systems will benefit from the fan-out and efficiency of few-photon communication. Yet when superconducting electronic circuits are employed for computation, even few-photon communication events represent a significant energy expense relative to the extremely low energy per operation of the superconducting circuits. In Section VI we discuss the use of superconducting splitters to make copies of photonic synapse events so that answers to all of the questions listed above can be simultaneously present in the dendritic tree through processing of the signal from a single photon. Section VII contains a discussion of the results.

## II. Photon-to-Fluxon Transduction at a Synapse

Analysis of fluxonic processing of photonic synapse events begins with consideration of the circuit that transduces a single-photon detection event to the superconducting electronic domain in the form of a series of fluxons. The circuit that accomplishes this is shown in Fig. 2(a). This circuit was first introduced in [12] and described in more detail in [13]. The circuit comprises an initial receiver/transducer section, consisting of a superconducting-nanowire single-photon detector (SPD) [29]–[32] in parallel with a Josephson junction (JJ) [33]–[35]. In the steady state, the SPD (drawn as a variable resistor in series with an inductor) has zero resistance, and thus its entire bias current flows directly through it to ground. The synaptic firing junction, $J_{\mathrm{sf}}$, is biased below its critical current ($I_{\mathrm{c}}$) by the synaptic bias current, $I_{\mathrm{sy}}$. Upon absorption of a photon, the variable resistor of the SPD switches temporarily to a high-resistance state (5 k$\Omega$) for a short duration (200 ps) [36]. The current through the SPD is diverted across a resistor ($I_{\mathrm{spd}}$ across $r_{\mathrm{spd}}$ in Fig. 2(a)) and to $J_{\mathrm{sf}}$. At this point, the sum of the currents across $J_{\mathrm{sf}}$ exceeds $I_{\mathrm{c}}$, and the junction produces a series of fluxons [33]–[35]. These fluxons propagate along the Josephson transmission line [34], [35], and are stored in the synaptic integration (SI) loop. The Josephson transmission line serves to isolate the activity of the receiver portion of the circuit from the integration loop, allowing their circuit parameters to be optimized independently. After the 200 ps photon detection event, the bias current returns to the SPD with the time constant of $\tau_{\mathrm{spd}} = L_{\mathrm{spd}}/r_{\mathrm{spd}}$. This time constant has a minimum functional value determined by the electro-thermal properties of the nanowire [36], and throughout this work this time constant is fixed at $\tau_{\mathrm{si}} = 10$ ns, and the bias

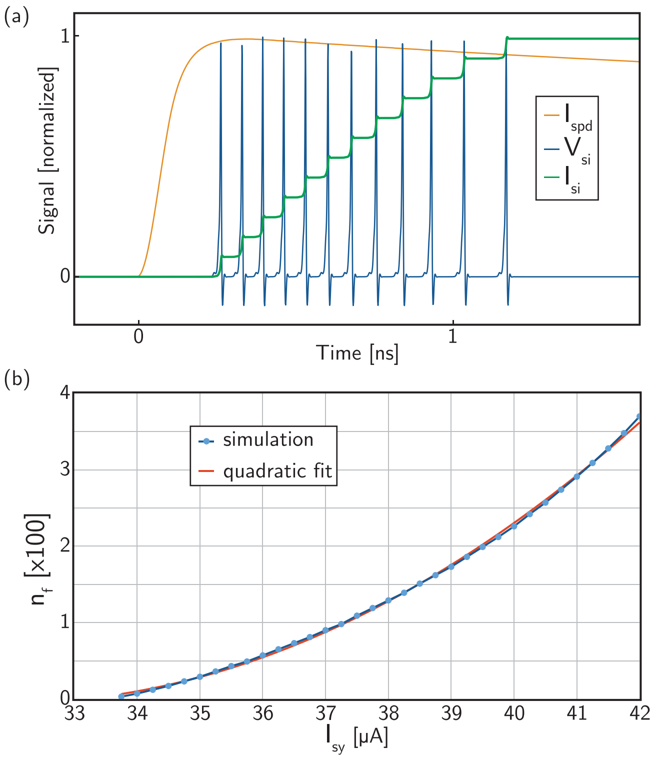

Fig. 3. Basic operation of the photon-to-fluxon synaptic transducer. (a) Temporal activity of the circuit in Fig. 2(a) during a synaptic firing event. The traces are color-coded with the currents and voltages labeled in Fig. 2(a), and all traces have been independently normalized for display on the same plot. (b) Demonstration of variable synaptic weight. The number of fluxons generated during a synaptic firing event ($n_{\mathrm{f}}$) is plotted as a function of the synaptic bias current ($I_{\mathrm{sy}}$). A fit to a second-order polynomial is also shown.

to the SPD is fixed at 10 $\mu$A. The number of fluxons created during a synaptic firing event depends on the net current across $J_{\mathrm{sf}}$ as well as the duration during which $J_{\mathrm{sf}}$ is biased above $I_{\mathrm{c}}$. With $\tau_{\mathrm{si}}$ and the bias to the SPD fixed, the number of fluxons, and thus the synaptic weight, are dynamically adaptable by changing the synaptic bias current, $I_{\mathrm{sy}}$. More details regarding $I_{\mathrm{sy}}$ and the associated plasticity mechanisms are given in [13].

The temporal activity of the circuit in Fig. 2(a) during a synaptic firing event is shown in Fig. 3(a). Throughout this work, WRSpice [37] has been used to simulate all circuits. All JJs have $I_{\mathrm{c}} = 40\ \mu$A and $\beta_{\mathrm{c}} = 0.95$. The yellow trace in Fig. 3(a) shows the current diverted from the SPD after a photon has been received. The blue trace shows the voltage pulses as the fluxons enter the SI loop. As each fluxon enters the loop, it introduces a discrete, fixed value of current given by $I_\phi = \Phi_0/L_{\mathrm{si}}$, where $\Phi_0 \approx 2 \times 10^{-15}$ Wb is the magnetic flux quantum, and $L_{\mathrm{si}}$ is the inductance of the synaptic integration loop. We assume the value of $L_{\mathrm{si}}$ is chosen by design independently for each synapse and set in hardware at the time of fabrication. The green trace in Fig. 3(a) shows the increase in current as the fluxons enter the SI loop during a synaptic firing event. The discrete steps with each fluxon are evident, and the total amount of current added to the SI loop during a synaptic firing event depends on both the number of fluxons generated during the firing event (controlled dynamically by $I_{\mathrm{sy}}$) and the inductance of the SI loop (set in hardware as $L_{\mathrm{si}}$).

The role of $I_{\mathrm{sy}}$ is to adapt the synaptic weight by changing the number of fluxons generated during a synaptic firing event. In Fig. 3(b) we show the number of fluxons generated during a synaptic firing event as a function of $I_{\mathrm{sy}}$. The fit shows close agreement with a quadratic function. This method of establishing and adapting the synaptic weight has several important properties. First, it is slowly varying, so small changes in $I_{\mathrm{sy}}$ result in small changes in the synaptic weight. Second, the function is monotonic, so increases in $I_{\mathrm{sy}}$ always result in increased synaptic efficacy, while decreases in $I_{\mathrm{sy}}$ always result in decreases in synaptic efficacy. This is necessary to enable activity-based plasticity mechanisms [38], [39], which have been explored in the context of these circuits in [13]. Third, the bias $I_{\mathrm{sy}}$ can be bounded so synaptic strength never exceeds a certain limit, and runaway activity is not possible. Finally, the integer number of fluxons generated can be made to cover a broad range so that analog synapses of relatively high bit depth can be achieved. Figure 3(b) shows that over eight bits (256 levels) can be utilized, and throughout this work we find the range of eight to 10 bits to be a comfortable working range for the circuits under consideration. This is much lower than the 64-bit processors used for high-arithmetic-depth numerical calculations. Yet neural computation benefits from performing lower-resolution operations with high efficiency; accuracy is achieved through redundancy and parallelism. Additionally, effects of noise due to operation at finite temperature will further reduce the ability to resolve distinct synaptic weight values. The effects of this noise will be investigated in future work.

From Fig. 3, we can also gain some insight into the manufacturability of these circuits. With this technology, we aspire to achieve large-scale systems capable of advanced cognitive computing. Such systems will potentially comprise billions of synapses and 10 times as many JJs. These JJs will have a statistical distribution of critical currents due to fabrication variations. During operation, the biases delivered to the junctions will also have a statistical distribution. The data in Fig. 3(b) inform us that synapses with a broad range of bias conditions will contribute signal upon receiving synaptic events. Here we show that synapses will be operational if the bias current varies by 8 $\mu$A around 38 $\mu$A, giving a margin of 20%. When the system is initially fabricated and turned on, variations in junction critical currents and biases will result in a statistical distribution of synaptic weights. Over time, as the system operates and learns, these bias currents will be finely adjusted based on the activity-dependent plasticity mechanisms described in [13], mitigating any deleterious effects of fabrication variations.

After a photonic communication event has been detected, the synaptic weight has been set as the number of fluxons created, and current has been added to the SI loop, further processing ensues. The electrical current generated by the synapse event can be stored for a chosen amount of time. This is determined by the leak rate of the SI loop, selected by design and set in hardware with the time constant $\tau_{\mathrm{si}} = L_{\mathrm{si}}/r_{\mathrm{si}}$. Note that $\tau_{\mathrm{si}}$ is entirely independent of $\tau_{\mathrm{spd}}$, and because we consider superconducting circuits, memory of a synaptic event can persist indefinitely.

Also note that while the amount of current added to the SI loop during a synaptic firing event depends on $L_{\mathrm{si}}$, $r_{\mathrm{si}}$ can be chosen independently from $L_{\mathrm{si}}$, thereby enabling the amount of current and its storage time to be separately selected. The current can be released quickly, on the order of the SPD reset time of 10 ns, or it can be stored 10 or 100 times longer to retain a memory of the event for as long as required. In this work we mainly consider decay times spanning two orders of magnitude, from $\tau_{\mathrm{si}} = 10$ ns to $\tau_{\mathrm{si}} = 1$ $\mu$s.

The reason for focusing on these time scales is as follows. In biological neural systems, processing among local clusters of neurons occurs primarily through fast activity in the range of gamma frequencies (30 Hz–80 Hz) [40], [41]. This frequency range emerges because it reaches the upper limit of speed for the excitatory pyramidal neurons participating in the activity. In the superconducting optoelectronic hardware under consideration, this upper speed limit is in the tens of megahertz, limited by the reset time of the SPDs in the synapses and of the transmitter circuits that generate neuronal firing events [13]. Here we take the upper firing rate to be 100 MHz for numerical simplicity. Therefore, we expect the neurons under consideration to demonstrate behavior similar to gamma oscillations, bursting with inter-spike intervals on the order of 10 ns. Similarly, biological neural systems process information across the network as a whole through slower activity at theta frequencies (4 Hz–8 Hz) [40], [41]. Mapping this scaling onto the system under consideration, we pay particular attention to gamma oscillations occurring at 100 MHz as well as theta oscillations occurring at 10 MHz. It is for this reason that we consider $\tau_{\mathrm{si}}$ ranging from 10 ns to 1 $\mu$s and spike trains in the 50 MHz to 100 MHz range.

In addition to signal decay from a synaptic integration loop, we must also consider saturation, as shown in Fig. 4. As stated above, the current associated with a fluxon being generated in a loop of inductance $L$ is $I_\phi = \Phi_0/L$. This current circulates in the direction opposing the applied bias to the JJ. The number of fluxons that can enter the loop before the cumulative opposing bias equals $I_{\mathrm{c}}$ is given by $I_{\mathrm{c}}/I_\phi = LI_{\mathrm{c}}/\Phi_0 = \beta_{\mathrm{L}}/2\pi$, where $\beta_{\mathrm{L}}$ is a common parameter quantifying the flux storage capacity of a superconducting loop. $\beta_{\mathrm{L}}/2\pi$ gives an estimate for how many fluxons a given SI loop will be able to store before saturation, and the exact number also depends on the applied bias. In Fig. 4 we show the integrated current in an SI loop as a function of time in response to a periodic train of pulses with 20 ns inter-spike interval. Here we fix $\tau_{\mathrm{si}} = \infty$ and vary the inductance of the loop. In these simulations, the value of $I_{\mathrm{sy}}$ was fixed at 38 $\mu$A, so 129 flux quanta ($> 2^7$) are generated during each synaptic firing event until the loop nears saturation, at which point the effective synaptic weight is suppressed, demonstrating a simple form of short-term plasticity. With a small value of $L_{\mathrm{si}}$, the quantity $\beta_{\mathrm{L}}/2\pi = L_{\mathrm{si}}I_{\mathrm{c}}/\Phi_0 = 150$, and the loop saturates after a single synaptic firing event. With an intermediate value of $L_{\mathrm{si}} = 77.5$ nH, $\beta_{\mathrm{L}}/2\pi = 1.5 \times 10^3$, and seven synaptic firing events fill the loop. With a large value of $L_{\mathrm{si}} = 775$ nH, $\beta_{\mathrm{L}}/2\pi = 1.5 \times 10^4$, and the loop can hold the activity from nearly 100 synaptic firing events with this value of $I_{\mathrm{sy}}$. All these values of inductance are straightforward to achieve with high-kinetic-inductance materials. Note that in digital superconducting electronics $\beta_{\mathrm{L}}/2\pi = 1.5$, so a loop can hold a single fluxon to represent a bit. Figure 4 shows the control one has in design over the capacity of the SI loop. The loop can operate as a binary device switching from a low to high state with each synapse event, or it can act as an analog device capable of representing many synapse events with distinct values of current. This saturation is a simple form of nonlinearity present in the synapse.

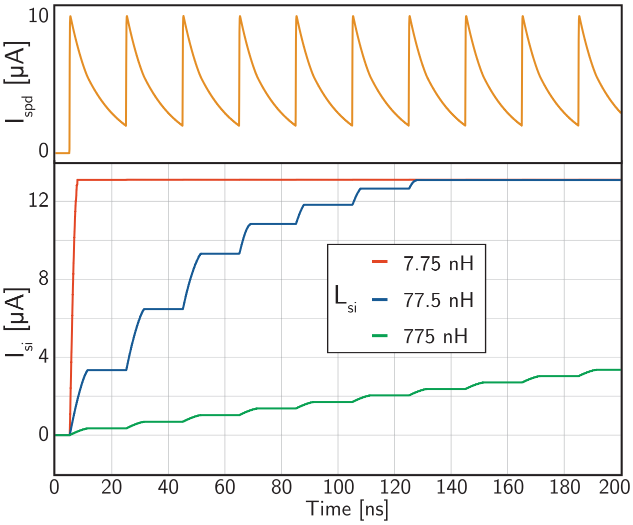

Fig. 4. Filling of the SI loop in response to a pulse train. The upper panel shows the current pulses generated by detection events at the SPD. The lower panel shows time traces of $I_{\mathrm{si}}$. Here $\tau_{\mathrm{si}} = \infty$ to focus attention on the manner in which the SI loop fills with current, rather than how it decays. A loop with small inductance ($L_{\mathrm{si}} = 7.75$ nH) will saturate after a single photon detection event, while a loop with large inductance ($L_{\mathrm{si}} = 775$ nH) can store the signals from many synaptic firing events.

As we have described, the two basic degrees of freedom of the SI loop are the signal storage time and storage capacity. We now proceed to explore the use of such synapses to extract information from pulse trains.

## III. Operations on Pulse Trains at a Single Synapse

As an example of one form of processing that can be performed using the synaptic circuit of Fig. 2(a), Fig. 5 considers the operation of rate-to-current conversion. The first term of the Volterra expansion of a spike train corresponds to the time-averaged spike rate [4], so a neuron must be able to decode this information. This can be accomplished with the synaptic transducer of Fig. 2(a) when the SI loop is given a leak rate, as discussed above. The circuit behaves as a standard leaky integrator modeled as $\dot{I}_{\mathrm{si}} = \alpha - I_{\mathrm{si}}/\tau_{\mathrm{si}}$, where $\alpha$ is the rate of current added to the SI loop by synaptic firing events. The leaky integrator model has the steady-state solution $I_{\mathrm{si}} = \alpha\tau_{\mathrm{si}}$, indicating that the current in the loop is proportional to the rate of input spikes. In Fig. 5(a) we show temporal traces of the current $I_{\mathrm{si}}$ in the presence of afferent activity at various rates for a loop with $\tau_{\mathrm{si}} = 100$ ns and $L_{\mathrm{si}} = 77.5$ nH, and it can be seen that the time-averaged value of $I_{\mathrm{si}}$ reaches steady-state. In Fig. 5(b) we show the time-averaged current, $\bar{I}_{\mathrm{si}}$, as a function

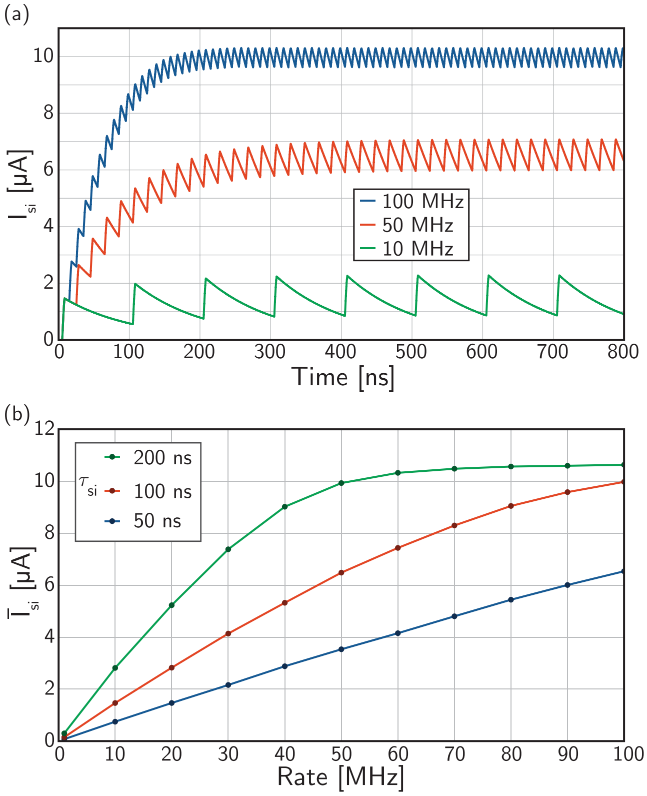

Fig. 5. Rate-to-current conversion at a synaptic transducer. (a) $I_{\mathrm{si}}$ as a function of time as pulse trains of various frequencies are incident upon the synapse. (b) Systematic analysis of rate-to-current mapping for SI loops of three decay time constants. To obtain these curves, temporal traces like those of (a) have been analyzed once the steady state has been reached. Each data point in (b) results from time-averaging a trace such as those in (a) over a single interspike interval: $\bar{I}_{\mathrm{si}} = (t_2 - t_1)^{-1} \int_{t_1}^{t_2} I_{\mathrm{si}}(t) dt$, where $t_1$ and $t_2$ are arrival times of consecutive photons at the synapse after the steady state has been reached.

of the synaptic firing rate for three values of $\tau_{\mathrm{si}}$. With the value $\tau_{\mathrm{si}} = 50$ ns, the response is linear across the entire range of gamma and theta frequencies. Linear rate-to-current conversion holds as long as the integration time of the loop is short enough to avoid saturation, that is, $\alpha\tau_{\mathrm{si}} < I_{\mathrm{si}}^{\mathrm{sat}}$. With $\tau_{\mathrm{si}} = 200$ ns, the loop reaches saturation, and higher input frequencies do not code unique information. If linear operation is desired, one must choose the time constant of the loop to be commensurate with the frequencies to be detected, or if nonlinear saturation is desired, longer integration times can be utilized. If increased dynamic range is advantageous, one can utilize the splitter of Fig. 2(c) to activate multiple SI loops with different time constants from the same photonic synapse, as described in Section VI.

The synaptic transducer and SI loop of Fig. 2(a) on its own can achieve straightforward rate-to-current conversion to make use of rate-coded neuronal information. Yet when $I_{\mathrm{si}}$ is coupled to the circuit of Fig. 2(b) through a mutual inductor, significantly more functionality can be achieved, as we will discuss shortly. Let us first describe the basic operation of the circuit in Fig. 2(b), which we refer to as a dendritic processing circuit or dendrite. The dendritic processing circuit of Fig. 2(b) is similar to the synaptic transducer circuit of Fig. 2(a). Unlike the synapse, which receives photonic input, the dendrite receives input as flux coupled through mutual inductors. In the steady state, all junctions are biased below $I_{\mathrm{c}}$. Afferent input to the dendritic receiving (DR) loop from one or more SI loops during a time window established by the synaptic time constants, $\tau_{\mathrm{si}}$, increases the bias to the dendritic firing junction ($J_{\mathrm{df}}$). When the net bias to $J_{\mathrm{df}}$ exceeds $I_{\mathrm{c}}$, one or more fluxons will be produced, they will traverse the JTL, and they will add flux to the dendritic integration (DI) loop, just as in the case of the synapse. The role of the dendritic reset junction ($J_{\mathrm{dr}}$) is to release the flux generated by $J_{\mathrm{df}}$ from the DR loop, thereby resetting the loop to the state prior to firing. The signal integrated in the DI loop is coupled either to the DR loop of another dendrite or the neuronal receiving (NR) loop of the neuron cell body.

The use of mutual inductors is advantageous for coupling multiple synapses to a single dendrite because mutual inductors reduce cross talk between synapses to a very low level. In general, SI loops have a self-inductance of at least 1 nH, and possibly up to 10 $\mu$H. The mutual inductors considered here are asymmetric with the inductor in the SI loop being on the order of 100 pH and the coupled inductor in the DR loop being on the order of 10 pH. The total inductance of the DR loop is on the order of 100 pH. Thus, when current is circulating in one SI loop, appreciable current is coupled to the DR loop, while the parasitic current coupled into other SI loops is significantly smaller. Using typical numbers from the circuits studied in this work, the parasitic current coupled to an adjacent SI loop is roughly one thousandth the current induced in the DR loop, with $I_{\mathrm{dr}}$ being on the order of microamps. More generally, in the limit that $L_{\mathrm{si}} \gg M_{\mathrm{sy}}$, this induced current scales as $M_{\mathrm{sy}}/(N_{\mathrm{sy}}L_{\mathrm{si}})$, where $N_{\mathrm{sy}}$ is the number of synaptic loops coupled to the DR loop. For typical values of $M_{\mathrm{sy}}$ and $L_{\mathrm{si}}$, this quantity is on the order of $10^{-3}$ for $N_{\mathrm{sy}} = 1$ and decreases as synapses are added to the loop. The ratio of parasitic current induced in adjacent SI loops to the intended current induced in the DR loop is independent of $N_{\mathrm{sy}}$, and in the same limit of $L_{\mathrm{si}} \gg M_{\mathrm{sy}}$ we find this ratio is $M_{\mathrm{sy}}/L_{\mathrm{si}}$, which again is on the order of $10^{-3}$ for typical circuit parameters.

The dendritic circuit under consideration is reminiscent of a standard circuit from flux quantum logic that converts a DC pulse to a single flux quantum (DC-to-SFQ converter) [34], [35]. The circuit is also similar to the neuron circuit presented in [21]. The main computational attributes of the dendrite come from the biasing conditions and interplay between $J_{\mathrm{df}}$ and $J_{\mathrm{dr}}$. If the biases are established such that when $J_{\mathrm{df}}$ produces a fluxon, the current added to $J_{\mathrm{dr}}$ is insufficient to switch $J_{\mathrm{dr}}$ until the added biases from the SI loop(s) decay, the device acts like a DC-to-SFQ converter. $J_{\mathrm{df}}$ will produce exactly one fluxon, and the DR loop will then be inactivated until the counter bias across $J_{\mathrm{dr}}$ due to the SI loop(s) decays, at which point $J_{\mathrm{dr}}$ will produce a fluxon countering the one produced by $J_{\mathrm{df}}$, and the loop will be reset. In this configuration, the dendritic receiver has a binary character.

The circuit can also operate in an analog mode, wherein the dendrite can produce a continuous stream of fluxons, much like the synaptic transducer. To achieve this operation, $J_{\mathrm{dr}}$ is biased closer to $I_{\mathrm{c}}$ so that a fluxon generated by $J_{\mathrm{df}}$ is sufficient to switch $J_{\mathrm{dr}}$. Thus, each time $J_{\mathrm{df}}$ produces a fluxon, it is rapidly canceled by $J_{\mathrm{dr}}$, and the DR loop is reset with no net flux. $J_{\mathrm{df}}$ will continue to produce fluxons as long as it is held above $I_{\mathrm{c}}$, and in the presence of synaptic activation (current in one or more SI loops), a stream of fluxons will be generated by $J_{\mathrm{df}}$ and stored in the DI loop. This stream may contain a large number of fluxons until the DI loop saturates, so we consider this an analog mode of operation.

Whether operating in binary or analog, the effect of the dendrite is to perform a nonlinear transfer function on its inputs and provide the output signal to the DI loop in the form of supercurrent. Just as in the SI loop, the DI loop can be configured to saturate rapidly (small $\beta_{\mathrm{L}}$) or store the signal from many threshold events (large $\beta_{\mathrm{L}}$), and the loop can be configured with a decay time constant ($\tau_{\mathrm{di}} = L_{\mathrm{di}}/r_{\mathrm{di}}$) spanning a broad range, from time scales shorter than a gamma interspike interval to as long as superconductivity can be maintained. With these basic operating principles in mind, we proceed to consider examples of dendritic processing with this circuit.

We first consider operations usually associated with synaptic computation [3], namely short-term-facilitating and short-term-depressing plasticity. Some synapses are observed to provide no response or very weak response to the first pulse of a train, with the efficacy of the synapse increasing as the pulse train proceeds. This behavior is referred to as short-term-facilitating plasticity, and it can be due to dynamics within the synapse itself or to the conductance properties of a dendrite or series of dendritic compartments. Here we simulate analogous behavior with a single synaptic transducer (Fig. 2(a)) coupled to a single dendritic processing circuit (Fig. 2(b)).

To achieve short-term-facilitating plasticity, we design an SI loop that can store the signals from multiple synaptic firing events before saturation, and we bias $J_{\mathrm{df}}$ so that the additional current induced by the first few synaptic firing events does not push the junction over $I_{\mathrm{c}}$, but after multiple synaptic firing events, $I_{\mathrm{c}}$ is exceeded and flux is added to the DI loop. We design the dendrite in analog mode for this behavior. Circuit simulations of short-term-facilitating plasticity are shown in Fig. 6. Figure 6(a) shows the afferent pulse train. The first pulse occurs at 5 ns, and the interspike interval is 20 ns. Figures 6(b) and (c) show the accumulated current in the DI loop as a function of time. In Fig. 6(b) the effect of the synaptic bias current, $I_{\mathrm{sy}}$ is shown. The primary effect of the dynamically reconfigurable bias current is to shift the curve left or right. With a stronger synaptic weight, more current will be added to the SI loop with each synaptic firing event, and therefore more current will be induced by the mutual inductor into the DR loop. Thus, fewer synaptic firing events are required to reach threshold in the dendritic compartment. In this example, $I_{\mathrm{sy}}$ can shift the threshold from three to eight synaptic firing events. In Fig. 6(c), the synaptic bias current is fixed at 38 $\mu$A, while the dendritic bias current, $I_{\mathrm{de}}$, is varied. Change in $I_{\mathrm{de}}$ has less of an effect on the number of pulses required to reach threshold, but it significantly affects the number of fluxons generated by $J_{\mathrm{df}}$ each time a synaptic firing event occurs, which is related to the slope of the traces in Fig. 6(c). The effect of the dendritic bias current, $I_{\mathrm{de}}$, is therefore analogous to the effect of the synaptic bias current, $I_{\mathrm{sy}}$, during a synaptic firing event. We therefore anticipate that $I_{\mathrm{de}}$ will provide a dynamically reconfigurable circuit parameter that can be used to establish a "dendritic weight" and can be used for long-term plasticity and learning.

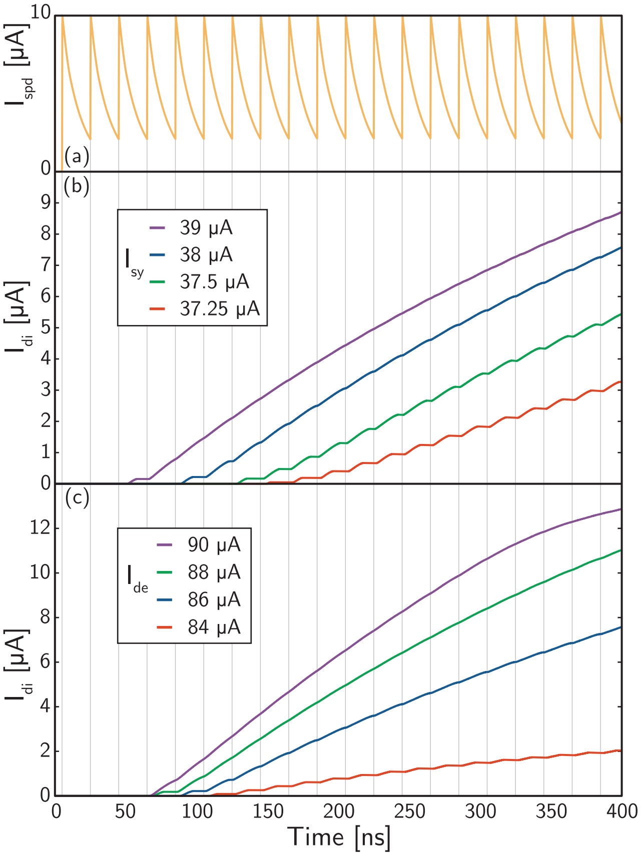

Fig. 6. Short-term-facilitating plasticity. With a single synapse coupled to a dendrite, a nonlinearity can be induced wherein multiple synaptic firing events are required to generate a signal. (a) The afferent pulse train. (b) The current in the DI loop, $I_{\mathrm{di}}$, as a function of time for several values of the synaptic bias current, $I_{\mathrm{sy}}$. (c) $I_{\mathrm{di}}$ for several values of the dendritic bias current, $I_{\mathrm{de}}$. The blue curve is the same in (b) and (c). The vertical lines spanning (a)–(c) represent the times of the synaptic firing events.

As mentioned above in the context of synapses, we wish to anticipate how fabrication imperfections in JJ critical currents as well as variations in bias conditions will affect circuit operation. The data in Fig. 6 show that the dendritic response is quantitatively sensitive to the value of the bias currents, or similarly the junction critical current. However, the qualitative nature of the response is consistent across a useful range of operating parameters. In large-scale systems, the intention is not to precisely control the response of each dendrite or synapse quantitatively at the time of fabrication, but rather to fabricate a complex network with a statistical distribution of device parameters and to employ adaptive plasticity functions that finely adjust biasing conditions through activity-dependent feedback, as discussed in [13], to adapt the circuits to operating points useful for network computation. Such adaptation over time through synaptic and dendritic plasticity are in the spirit of biological neural systems that cannot be constructed with specific values for each synaptic weight or precise dendritic morphology. Nevertheless, it remains to be seen if Josephson circuits can be manufactured with tight enough tolerances to enable the functions proposed here. This question is one of the most pertinent to determining the feasibility of large-scale superconducting optoelectronic networks.

While facilitating behavior effectively strengthens a synapse as a pulse train proceeds, short-term-depressing plasticity gives the opposite behavior. In an extreme form, this mechanism can be used to convey only the onset of a pulse train, while blocking subsequent spikes. To demonstrate this behavior, we consider the dendritic processing circuit in binary mode. Circuit simulations are shown in Fig. 7. Consider first the upper panel, Fig. 7(a–c). The current pulses from the SPD due to the afferent spike train are shown in Fig. 7(a), and the resulting current in the SI loop is shown in Fig. 7(b). The activity consists of two groups of three spikes. The current in the DI loop is shown in Fig. 7(c). A single pulse enters the DI loop at the onset of the first spike in the train. In Fig. 7(b), we have marked with a red line the value of $I_{\mathrm{si}}$ below which reset occurs in the DR loop. We see that the first spike of the second group of three occurs just before $I_{\mathrm{si}}$ drops below the reset value. The second group of pulses is not identified as a new spike train, so no additional signal is added, and $I_{\mathrm{di}}$ continues decaying with $\tau_{\mathrm{di}}$. By contrast, in the lower panel (Fig. 7(d–f)), the onset of the second group of pulses occurs 20 ns later than in the upper panel, giving the current in the SI loop (and therefore the DR loop) time to decay below the reset value. In this case, when the second group of pulses begins, it is identified as a new train, and additional signal is added to the DI loop, again in the form of a single fluxon. The reset delay can be established in hardware across a broad range of values through $\tau_{\mathrm{si}}$ and can be adjusted over a smaller range dynamically through $I_{\mathrm{de}}$. The dendritic receiving loop does not have any resistance of its own, so the current decay time constants in that loop are entirely determined by the SI loops.

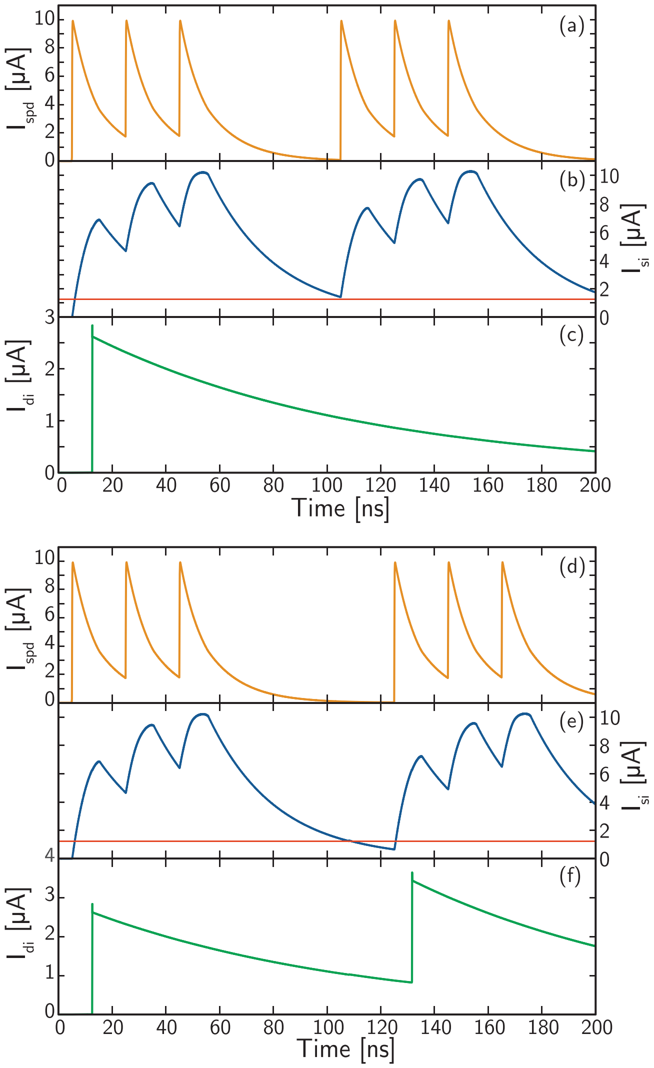

Fig. 7. Short-term-depressing plasticity. With a single synapse coupled to a dendrite, the first pulse of a train can generate signal in the DI loop, and the response of the synapse is depressed for subsequent pulses until the signal in the SI loop decays below a certain level, resulting in reset. (a) Afferent activity. (b) The resulting signal in the SI loop. The red line shows the reset level, which is not quite reached before the second series of three pulses occurs. (c) The signal in the DI loop, $I_{\mathrm{di}}$. Only a single pulse enters the DI loop because the break between the pulse trains was not long enough to achieve reset. (d) Afferent activity with a slightly longer delay between the two pulse trains. (e) The signal in the SI loop, dropping briefly below the reset threshold. (f) The resulting current, $I_{\mathrm{di}}$, showing two pulses generated as the dendrite recognized these as two separate pulse trains.

While we refer to this operation of the dendritic processing circuit as binary, this term refers to the all-or-nothing response of the DR loop. The DI loop may be independently configured to store anywhere from one to many fluxons, so the output of the circuit can be chosen independently to represent either a binary signal or to give an analog representation of the number of afferent pulse trains occurring within a time period set by $\tau_{\mathrm{di}}$. If the DI loop is configured with large $\beta_{\mathrm{L}}$ and $\tau_{\mathrm{di}}$ on the order of theta time scales, the dendrite will keep track of how many gamma-frequency pulse trains have occurred, thereby keeping track of oscillations on theta time scales. Because the maximum signal level in the DI loop can be made the same as in an SI or DI loop keeping track of gamma activity, such dendritic processing is capable of representing gamma and theta information with equal weight. Alternatively, using the same circuit configuration except employing an SI loop with a time constant close to $\tau_{\mathrm{spd}}$ will cause the DI loop to receive a single fluxon each time the synapse receives a photon. In this mode of operation, the circuit achieves single-photon-to-single-fluxon transduction, converting each photon detection event to an identical, binary signal. If synaptic weighting is not required, and dendritic weights alone can suffice, the signal from a photon-detection event can immediately be converted to a single fluxon, and energy efficiency can be gained.

To summarize the operations we have investigated so far, the synaptic firing circuit on its own can accomplish rate-to-current conversion, reporting a temporal average of recent activity. By coupling the synaptic firing circuit to a dendritic processing circuit, we can construct a dendrite that generates signal only when a pulse train persists for a certain duration. We can use the same circuits with slightly different biasing configuration to construct a dendrite that generates signal only when a pulse train begins after a certain period of rest. All of these operations correspond to temporal filters performed on spike trains occurring at a single synapse. Yet an important function of dendritic processing is to identify coincidences and sequences between the activities of multiple neurons. We now consider this task.

## IV. Detecting Coincidences Between Neurons

The second term in a Volterra expansion of the activities of two neurons corresponds to coincidences between the two neurons [4]. We can use the same dendritic processing circuit of Fig. 2(b) to detect coincidences, provided two SI loops are coupled to the DR loop through mutual inductors. In the simplest case, we wish to know whether two synapses have fired within a certain time period of each other. This can be achieved by giving both SI loops the same value of $\tau_{\mathrm{si}}$. The response of such a circuit is shown in Fig. 8(a), where the current induced in the DI loop is shown as a function of the time delay between the two synaptic firing events for several values of $I_{\mathrm{de}}$ with $\tau_{\mathrm{si}} = 100$ ns. For the two lower values of $I_{\mathrm{de}}$, the circuit can be thought of as an AND gate with an analog extension to the time domain: if synapse $i$ AND synapse $j$ fire within a time period set by $\tau_{\mathrm{si}}$, a signal dependent on the time difference is added to the DI loop. For larger values of $I_{\mathrm{de}}$, the circuit performs an OR operation, because for arbitrarily large $\Delta t$, the current in one SI loop alone is sufficient to switch $J_{\mathrm{df}}$ and generate some signal in the DI loop. A similar coincidence detection circuit was proposed in [13] based on two SPDs. The advantage of the circuit presented here is that the computation occurs in the electronic domain, bringing the advantage of energy efficiency as well as the ability to perform multiple dendritic operations simultaneously through the use of fluxonic pulse splitters (Section VI).

The dendritic tree may benefit from the ability to detect not just coincidences, but also the specific sequence in which synapse events occurred [7]. This can be achieved by breaking the symmetry between the two synapses with $\tau_{\mathrm{si1}} \gg \tau_{\mathrm{si2}}$. We consider this scenario in Fig. 8(b). Here, $\tau_{\mathrm{si1}}$ is still 100 ns, but $\tau_{\mathrm{si2}}$ is much shorter (2.5 ns–10 ns), and we again plot the current added to the DI loop as a function of $\Delta t = t_2 - t_1$, where $t_i$ is the time of a synapse event on synapse $i$. In this case, the response function is skewed toward $\Delta t > 0$. It is probable that any current induced in the DI loop is due to an event on synapse one followed by an event on synapse two. Yet with this design, the contribution from $\Delta t < 0$ does not vanish completely. We have plotted the response for three values of $\tau_{\mathrm{si2}}$. We see that as we decrease $\tau_{\mathrm{si2}}$, the error due to current added when $\Delta t < 0$ decreases as $\tau_{\mathrm{si2}}$ decreases. Thus, we can tighten the timing tolerance by decreasing $\tau_{\mathrm{si2}}$. With $\tau_{\mathrm{si2}} = 2.5$ ns, errors do not occur if $t_2$ is prior to $t_1$ by 8 ns, less than the interspike interval of a gamma sequence, rendering this circuit capable of providing reliable information regarding the temporal order of activity between two synapses.

The coincidence and sequence operations of the dendritic processing circuit provide information regarding activity at two synapses. We would like to extend this to perform nonlinear operations on groups of multiple synapses. This can be straightforwardly achieved by coupling multiple synapses to a single dendrite, using the same circuits we have been discussing so far. In Fig. 8(c) we show the value of $I_{\mathrm{de}}$ resulting from a variable number of synapses firing simultaneously, with 10 total synapses coupled to a DR loop. We have chosen the circuit parameters so the bias added to $J_{\mathrm{df}}$ by a single synapse event is insufficient to exceed $I_{\mathrm{c}}$. The transfer function of the circuit is highly nonlinear, approximating a sigmoidal activation function. Thus, the current generated in the DI loop is not the sum of independent SI currents (see [4, p. 101]). The threshold number of active synapses can be set in design across a broad range, and as the three traces reveal, this number can be dynamically adjusted with $I_{\mathrm{de}}$. In both Fig. 8(a) and Fig. 8(c) we see that $I_{\mathrm{de}}$ can be used in a

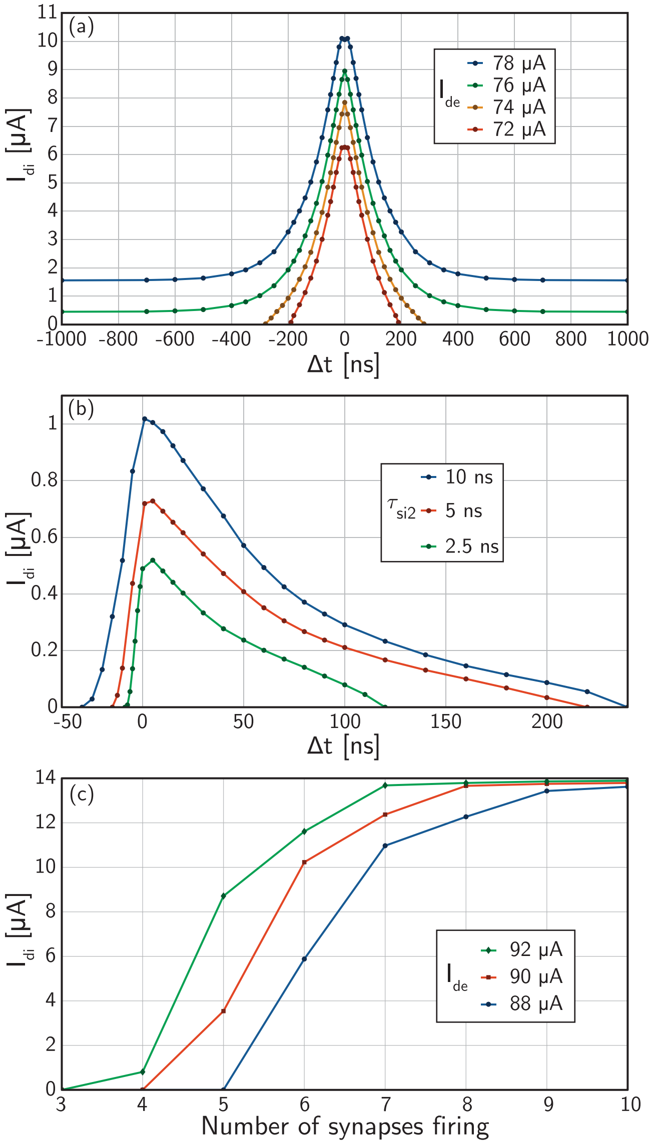

Fig. 8. Multiple synapses from different neurons coupled to a single DI loop. (a) Two synapses with the same time constant. The current induced in the DI loop ($I_{\mathrm{di}}$) is plotted as a function of the time between the two synapse events ($\Delta t$) for four values of the dendritic bias current ($I_{\mathrm{de}}$). (b) Two synapses with significantly different time constants. $I_{\mathrm{di}}$ is plotted versus $\Delta t$ for three values of the fast synaptic time constant ($\tau_{\mathrm{si2}}$). (c) Ten synapses from different neurons coupled to a single DI loop. The current generated in the DI loop is plotted as a function of the number of synapses firing simultaneously for three values of the dendritic bias current, $I_{\mathrm{de}}$.

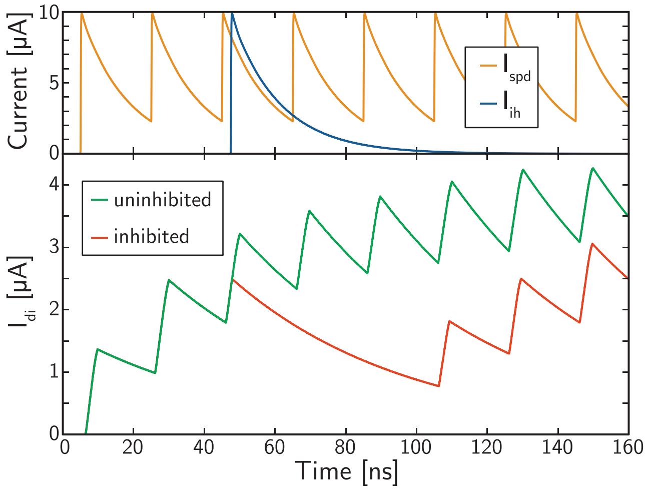

Fig. 9. The effect of inhibition. A single excitatory synapse and a single inhibitory synapse are coupled to a DR loop. The upper panel shows the signal from an afferent pulse train as well as the single inhibitory pulse. The lower panel shows the current generated in the DI loop ($I_{\mathrm{di}}$) with and without the inhibitory pulse. In this example, the circuit has been configured so that following the inhibitory pulse, no amount of activity on the excitatory synapse can drive $J_{\mathrm{df}}$ above $I_{\mathrm{c}}$, and the dendrite is completely suppressed until the signal in the IH loop has decayed.

manner analogous to the synaptic bias current, pointing again to the potential for reconfigurable efficacy and learning. While Fig. 8(c) only considers simultaneous synaptic activity, the true response of the dendrite would convolve the temporal responses of the constituent synapses. Similar principles to those demonstrated in Figs. 8(a) and (b) shape the net dendritic contribution.

All operations discussed thus far are excitatory. We now turn our attention to inhibition of the dendritic response.

## V. Inhibition and Rapid Query

The dendritic tree offers the most information to the neuron when it can be dynamically adapted into diverse functional networks. Inhibition can enable such adaptation (as well as many additional functions [42]) by temporarily silencing specific dendrites or entire branches of the dendritic tree. To accomplish this with the dendritic processing circuit under consideration, we couple an additional loop to the DR, except with mutual inductor of reverse coupling to oppose the bias to $J_{\mathrm{df}}$. We refer to this as an inhibitory (IH) loop, as shown in Fig. 2(b). The circuit parameters can be chosen so that following a synaptic event on the inhibitory synapse no amount of activity on the excitatory synapses can drive $J_{\mathrm{df}}$ above $I_{\mathrm{c}}$. As discussed above, the AND/OR logical operations become coincidence detections when extended to the time domain [5], and when the previously considered AND circuit is augmented with an inhibitory input, the logical operation becomes AND-NOT [5].

Simulated operation of a dendrite with a single excitatory and single inhibitory synapse is shown in Fig. 9. The upper panel shows a temporal trace of excitatory activity, which consists of a pulse train at 50 MHz. A single inhibitory synapse event occurs shortly after the third pulse of the excitatory train. The lower panel shows the current circulating in the DI loop as a function of time for cases with and without the inhibitory synapse event. Without inhibition, current is added to and decays from the DI loop, as expected. When inhibition occurs, the effect of excitation is immediately quenched. Following the inhibitory synapse event, $I_{\mathrm{di}}$ begins decaying with time constant $\tau_{\mathrm{di}}$. Inhibition decays with a completely independent time constant, $\tau_{\mathrm{ih}} = L_{\mathrm{ih}}/r_{\mathrm{ih}}$, just as all other loops discussed thus far. When the inhibitory current has decayed sufficiently, the effect of the excitatory pulse train resumes.

The duration over which the dendrite is inhibited is controlled by $\tau_{\mathrm{ih}}$, and for the network to be rapidly adaptable under the influence of inhibition, this time constant will be as short as a gamma-range interspike interval. If inhibition is required over theta time scales, repeated activity on the inhibitory neuron can keep the dendrite suppressed. However, this may not be the most energy-efficient mode of operation. Given the circuits under consideration, we can utilize a mode of operation complimentary to inhibition. In this configuration, the mutual inductors and bias to the DR loop are chosen so that even with all afferent SI loops saturated, the current across $J_{\mathrm{df}}$ cannot exceed $I_{\mathrm{c}}$. Only when an additional, unique synapse fires does the current exceed $I_{\mathrm{c}}$.

We refer to these unique synapses as *rapid query* synapses, and we explain their function as follows. The role of a rapid query synapse is complimentary to the role of an inhibitory synapse. In the typical operation of a dendrite, the response of the dendrite depends on the activities of the input excitatory synapses, and the role of inhibition is to effectively cancel the excitatory inputs. In one sense, the function of a rapid query synapse is the opposite of the function of an inhibitory synapse. With rapid query, a dendrite is designed so that the sum of all excitatory synaptic responses is insufficient to evoke a dendritic response. However, the function of the rapid query synapse is to drive the dendrite right up to its threshold, and therefore any excitatory input present at the time the rapid query synapse fires evokes a dendritic response. The circuit implementing this unique synapse is identical to all synapses considered thus far (Fig. 2(a)), but the function is unique. This synapse is designed to saturate with each synapse event and to decay rapidly, providing an identical response to each synapse event and no synaptic weight variation. The action of this synapse is to allow $J_{\mathrm{df}}$ to quickly sample the value of $I_{\mathrm{dr}}$ at a given instant in time. A dendrite with one or more excitatory synapses and a single rapid query synapse behaves as if it were always under the influence of inhibition until the rapid query synapse fires, briefly releasing it from inhibition. When the rapid query synapse fires, the current generated in the DI loop provides an answer to the question, "How much current is in the DR loop?"

As stated above, the function of a rapid query synapse is the opposite of the function of an inhibitory synapse. However, in another sense the objective is the same. A primary function of inhibition in neural systems is to dynamically adapt a given structural network into multiple functional networks. When inhibition is applied to a dendrite, the dendrite is functionally disconnected from the neuron cell body. Similarly, with rapid query, a dendrite is functionally disconnected from the neuron cell body at all times except when a rapid query synapse has fired. Rapid query operation provides another means to rapidly adapt the structural network into myriad functional networks, and rapid query is likely to be more energy efficient than inhibition when information stored in certain dendrites need not be accessed frequently. In biological neural systems, a given neuron either makes inhibitory connections or excitatory connections, but not both. This is referred to as Dale's law. It may be a consequence of physiological limitations, or it may be due to an information-processing advantage resulting from differentiating the responsibility of excitatory neurons that spread information and inhibitory neurons that adapt the functional network. We anticipate that such differentiation will be advantageous in superconducting optoelectronic networks as well, in which case neurons dedicated to inhibition will make only inhibitory synaptic connections, and neurons dedicated to rapid query will form only rapid query synapses. We refer to these neurons as rapid query neurons. The role of a rapid query neuron is to quickly cause the information stored in a collection of dendrites to be communicated from those dendrites further along the respective dendritic trees toward the neuron cell bodies, thereby rapidly functionally connecting those dendrites to the active network.

Figure 10 considers rapid query operation. The circuit under consideration comprises a single excitatory synapse and a single rapid query synapse coupled to a DR loop in the configuration of Fig. 2(b). In the present example, three excitatory synapse events occur, as seen in the upper panel of Fig. 10(a). Two rapid query synapse events are also shown in that panel. The first rapid query event follows the first excitatory pulse by 30 ns, and with $\tau_{\mathrm{si}} = 20$ ns, only a small amount of current is added to the DI loop. The second excitatory event is not followed by a rapid query event, and no current is added to DI. The third excitatory event is followed by a rapid query event with 10 ns delay, and significantly more current is induced in the DI loop.

The behavior of this circuit is summarized more systematically in Fig. 10(b). Here we plot the current induced in the DI loop as a function of the time delay between the rapid query and excitatory events for two values of $\tau_{\mathrm{si}}$. We see that the signal generated by rapid query follows the exponential decay of the SI loop, thus providing an accurate mapping of $I_{\mathrm{si}}$ to $I_{\mathrm{di}}$ at the time rapid query was performed.

We plot the exponential functions of Fig. 10(b) on a log-log graph to emphasize that each SI loop provides information over a single time scale determined by $\tau_{\mathrm{si}}$. It would be desirable to find a means by which a memory trace may be extended across multiple time scales from a single photonic synapse event. This increased temporal dynamic range is one example of what can be achieved if electronic copies of photonic synapse events are produced. This fluxonic fan-out is the subject of the next section.

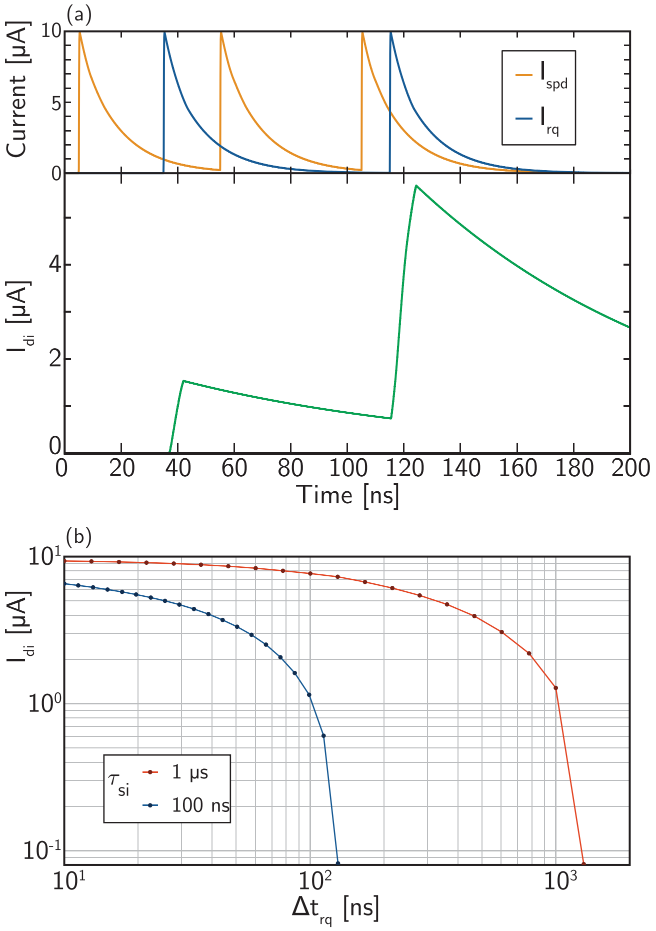

Fig. 10. Rapid query. (a) Temporal response of a single excitatory synapse and a single rapid query synapse coupled to a DR loop. The upper panel shows the pulses resulting from photon detection events on the excitatory and rapid query synapses, while the lower panel shows the the current generated in the DI loop by the two rapid query events. (b) Systematic quantification of the result of rapid query activity following a single excitatory pulse. The current generated in the DI loop, $I_{\mathrm{di}}$, is plotted as a function of the time delay between the excitatory synapse event and the rapid query synapse event. All simulations in this figure were conducted with $\tau_{\mathrm{rq}} = 10$ ns.

## VI. Fluxonic Fan-Out From Photonic Synapses

In neural systems using light for communication, generation and detection of photons are likely to consume the most energy. We have described several example operations that can be performed in the electronic domain to extract information from photonic pulse trains and their synaptic reception. We would like to perform them all simultaneously without requiring an additional photonic synapse for each. We can straightforwardly copy fluxons with a pulse splitter, a common means of achieving fan-out of flux-quantum signals [43]. We can therefore simply copy the output signals from a single photonic synapse to multiple independent SI loops that can each perform different temporal filters and feed into different dendrites. We refer to these as electronic synapses, and we anticipate that each photonic synapse will feed multiple electronic synapses.

The circuit for splitting pulses is shown in Fig. 2(c). A fluxon enters from the left, and when it switches the initial junction, the current of the resulting fluxon is split to two subsequent junctions. These junctions are biased such that the amount of current is sufficient to exceed $I_c$, thus producing fluxons at both junctions with restored signal level. For the application at hand, the splitter of Fig. 2(c) can be placed following $J_{\mathrm{jt1}}$ in Fig. 2(a) or Fig. 2(b). Thus, signals produced by synapses or dendrites can be copied and processed independently to extract distinct information through multiple temporal filters and logical operations. The circuit of Fig. 2(c) achieves direct one-to-two fan-out. If a greater number of copies is desired, the same circuit can be repeated in a tree. The limits of this fan out will depend on one's tolerance for circuit complexity. We speculate that in mature systems, a given photonic synapse may split to as many as 10 electronic synapses. The axonal arbor could implement one-to-one-thousand fan-out with branching photonic waveguides across a broad spatial range and an additional one-to-ten fan-out from each photon detector to electronic synapses across a much more limited spatial range. Total fan-out would then be one-to-ten-thousand, comparable to biological neurons in many regions of the brain [44].

As a simple example of the utility of pulse splitting, we consider one photonic synapse feeding into two electronic synapses with different time constants. Figure 11(a) summarizes the motivation for employing two different time constants. Instead of retaining a memory trace of a synapse event over only a single temporal scale, as occurs in a single SI loop with exponential decay with one time constant, we would prefer a signal with a power-law decay, so that information across temporal scales can be accessed. In Fig. 11(a) we compare $f(t) \propto t^{-q}$ to $g(t) \propto e^{-t/\tau}$ for three values of $\tau$, taking the power-law exponent to be $q = 1$. This figure illustrates the principle that a power-law temporal decay represents information across multiple orders of magnitude in time, while an exponential function only has a single time constant, and therefore only represents information across roughly one order of magnitude. For example, in Fig. 11(a), the power-law decay has an appreciable signal spanning two orders of magnitude in time. By contrast, the smallest value of $\tau$ provides no information past its cutoff, and the signal from the largest value of $\tau$ is nearly constant initially across more than an order of magnitude. The middle value gives a poor representation at the start and the end. However, we can obtain a suitable approximation to the power law function by superposing a small number of exponentials [45], as shown in Fig. 11(b). Here we represent a power law with unity exponent ($q = 1$), mapping two orders of magnitude in time to two orders of magnitude in signal. Convergence is shown in the inset. The error is improved by an order of magnitude when using two exponentials instead of one, and there is little advantage to using more than three for this task.

We implement this principle with the circuits under consideration by copying the signal from a photonic synapse to two electronic synapses coupled to a common passive superconducting loop via mutual inductors. We choose the time constants and couplings of the two SI loops to approximate the fitting technique employed to produce Fig. 11(b). Figure 11(c) shows the current in each of the SI loops as well as the common output loop. A power law with $q = 1.1$ is shown for comparison. This power-law temporal extension can be used in conjunction with many of the other operations discussed thus far, with the objective to use cheap fluxonic operations to extend the memory trace of expensive photonic activity across extra orders of magnitude in time. Such an operation performs a power-law mapping of a temporal signal to the dynamic range of the firing junction and allows a single dendrite to retain and access information regarding both gamma and theta frequencies.

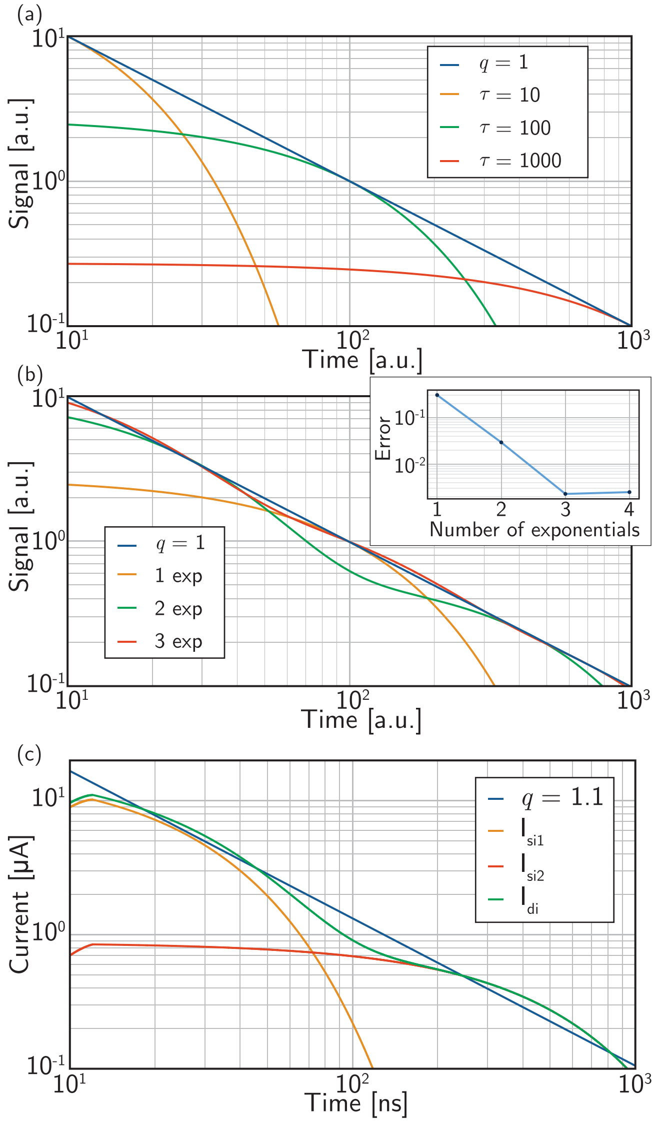

Fig. 11. Approximation of power-law temporal decay. (a) Illustration of the differences between power-law and exponential decay on a log-log plot. The functions plotted are $f(t) \propto t^{-q}$ and $g(t) \propto e^{-t/\tau}$, referred to generally as "Signal". (b) Approximating power-law decay through the superposition of multiple exponentials. A power law function with unity exponent is shown, as are approximations composed of one to three exponentials. Amplitudes and time constants were adjusted for best fit, and convergence is shown in the inset. (c) Approximating power-law decay with two SI loops coupled to a common DR loop. The time constants and mutual inductances have been chosen to approximate a power law using the same fitting algorithm that generated (b).

This example of using pulse splitting to access broader time spans is a straightforward extension of the behavior of a single SI loop. Additional functionality can be envisioned by combining pulse splitting with many of the functions discussed in this paper. Most importantly, by copying the output from a photonic synapse, each of the operations discussed here can be performed concurrently. With a single photon, the dendritic tree can be provided with information regarding the synapse's average firing rate across multiple temporal scales; the time since the last synaptic firing; various quantities regarding initiation and duration of pulse trains; coincidences and sequences with synapses from multiple other neurons; and inhibition and rapid query applied independently to each of these pieces of information.

## VII. Summary and Discussion

We have described several synaptic and dendritic operations achieved with Josephson junctions and mutual inductors. These include various logical operations, temporal filters, and nonlinear transfer functions applied to one or more synapses. The operations performed here are all accomplished with configurations of the building blocks shown in Fig. 2. We envision the dendritic tree to be comprised of a complex network of synapses and dendrites performing a multitude of computations on signals that fan in from photonic synapses, traverse the dendritic tree, and feed the neuron's final thresholding compartment, which triggers the production of light. A network will comprise many neurons, and each neuron is itself a network. We have described the dynamic functional adaptation of the dendritic network through inhibition and rapid query. Inhibitory activity nullifies targeted portions of the tree, while rapid query obtains local fragments of information and passes them along the tree. We have also described how electronic copies of photonic synapse events can enable several of these operations to be performed with the information from the detection of a single photon.

This work provides additional support for the hypothesis that superconducting computation is complimentary to photonic communication for achieving large-scale neural systems. While photons can achieve fan-out, they lack the required nonlinearities required for computation, especially at the low light levels required for energy efficiency. Further, photon motion cannot be halted to enable memory retention. Additionally, generating photons is more expensive than generating fluxons, and therefore only the minimum number of photons required for communication should be generated. Superconducting circuits are complimentary to photonic circuits in these regards. The proposed hardware may achieve greater than one-to-one-thousand fan-out in the photonic domain from each neuron to its connections [14], and subsequently, at each neuronal terminal an additional factor of roughly one-to-ten fan-out in the electronic domain, providing each receiving neuron with the capability of analyzing much more information about synaptic activity than would be available from a single synapse alone. Fan-in is envisioned to occur in the electronic domain as the dendritic tree computes and feeds its signals into the neuron cell body, ultimately resulting in a binary decision of whether or not to fire. Superconducting-nanowire single-photon detectors enable binary communication in that the response is nearly identical whether one or more photons are detected, and all computations—including synaptic weighting, nonlinear processing, and temporal integration—occur in superconducting electronic circuits with sub-nanosecond response times, native nonlinearities, and the potential for signal retention with no dissipation.

In mature superconducting optoelectronic circuits, system-level considerations will inform decisions regarding trade-offs between energy consumption and performance. One could reduce energy and area consumption by omitting dendritic processing entirely, but this would leave out important information processing. We do not attempt to fully address these trade-offs here, but we briefly consider the energy expended on synaptic, dendritic, and neuronal operations. At the physical level, these operations require light production during neuronal firing, photon detection during synapse events, and fluxonic processing in the dendritic tree. We would like the energy expended on light production, photon detection, and fluxonic processing to be roughly equal. Such an operating point is appealing because it indicates a global optimum wherein improvements to any one aspect of the system provide little added benefit, as other contributions become limiting factors. We can estimate where this operating point may reside by considering the three primary contributions to energy consumption. Similar analysis can be conducted regarding area. Production of a fluxon requires $E_{\mathrm{j}} = I_{\mathrm{c}}\Phi_0/2\pi = 1.3 \times 10^{-20}$ J for the junctions considered here, while production of a photon requires $E_{\mathrm{p}} = h\nu/\eta = \frac{1.6\times 10^{-19}}{\eta}$ J, where $\eta$ is the photon production efficiency, and we consider operation at $\lambda = 1.22$ $\mu$m [46]. Light generation is expensive because $\eta$ is unlikely to ever exceed 0.1 and may be limited to 0.01 or worse. Likewise, photon detection requires $E_{\mathrm{d}} = L_{\mathrm{spd}} I_{\mathrm{spd}}^2/2 = 1.3 \times 10^{-17}$ J for the superconducting-nanowire single-photon detector designs presented here. Due to the requirement of engineering reset dynamics in the detector, $L_{\mathrm{spd}}$ cannot be reduced below a certain value without decreasing the normal-state resistance of $J_{\mathrm{sf}}$, which requires increasing $I_{\mathrm{c}}$, which increases $E_{\mathrm{j}}$. Similarly, $I_{\mathrm{spd}}$ cannot decrease without using either junctions with smaller $I_{\mathrm{c}}$ or operating them in a noisy regime with bias close to $I_{\mathrm{c}}$, and would result in reduction of the dynamic range of the synaptic weight. This space of trade-offs is complex, and we make no attempt to identify the optimum in this work. We simply note that if $\eta = 0.01$, $I_{\mathrm{spd}} = 10$ $\mu$A, $L_{\mathrm{spd}} = 250$ nH, $I_{\mathrm{c}} = 40$ $\mu$A, and full analog processing of each synapse event generates $10^3$ fluxons on average, then light generation, detection, and fluxonic processing each contribute roughly equally to energy consumption. Full optoelectronic integration with few-photon binary communication and superconducting electronic analog computation offers a route to balance the energy budget while enabling the requisite communication and repertoire of computational functions for large-scale artificial cognitive systems.

One emphasis in this work has been on the interaction of inhibitory and rapid query neurons with dendrites to enable diverse functional networks. Inhibition is central to neural computation [42], with a key role being the formation and synchronization of adaptive neuronal modules that operate as task-specific processors [47], [48]. With inhibition, branches of the dendritic tree are functionally responsive by default and are selectively silenced by inhibitory synapse events. Inhibition can lead to synchronization by opening brief temporal windows when groups of neurons can fire [41]. With rapid query, branches of the dendritic tree are silent by default and are only functionally connected if rapid query synapse events occur. If the information in a given dendrite need not be accessed regularly, rapid query will be more energy efficient than continually performing inhibition. Like inhibition, rapid query may be useful for inducing synchronization. We do not propose rapid query instead of inhibition, but rather in addition. Both inhibition and rapid query may be leveraged to enable sub-threshold oscillations to be sampled only when required by the network, as occurs in biological neural systems to direct attention and amplify relevant information [9]. We posit the utility of a dedicated class of rapid query neurons in superconducting optoelectronic networks even though, to our knowledge, there is no such class of neurons in the biological domain. This may be due to a computational inadequacy of rapid query that we have overlooked, or it may be that the circuits under consideration are more amenable to such a mode of operation, which requires a degree of control over competing circuit parameters. There are dozens of different, specialized neurons in the mammalian brain, with multiple types of inhibitiory neurons playing specific roles [41], [42]. Superconducting optoelectronic networks take significant inspiration from the brain, but hardware discrepancies will inevitably lead to deviations in computation. Perhaps rapid query neurons are one such departure.

There are multiple possible extensions of the functions considered here as well as further details to be considered. XOR may be achieved with pulse splitting and lateral inhibition between dendrites. We have only considered binary inhibition, but weaker or multiple IH loops could be coupled to a DR loop to achieve partial inhibition. The neural operations considered here tend toward analog operation of the superconducting circuits, and we have presented circuits capable of representing signals with eight to 10 bits of resolution based on the $\beta_{\mathrm{L}}$ values chosen for the integration loops. However, this resolution is only available if noise is sufficiently low, so further investigation is required to determine a suitable tradeoff between loop inductance, signal resolution, and operating temperature. Future work may find different optimal values for different operations, and an improved balance between information capacity and hardware demands might be discovered. We have primarily considered signal storage loops with retention times on the order of what we suspect will be the gamma and theta frequencies of the system, but further research may find advantages of retaining fading memories for much longer than this or may reveal that even theta retention times are gratuitous.

In several instances we have indicated that the dendritic bias current $I_{\mathrm{de}}$ can be used to adjust circuit operation, pointing to a means of achieving learning and plasticity between synapses and dendrites [7] or between two dendrites. This subject deserves further investigation, but at present we simply note that similar circuits used for spike-timing-dependent plasticity in [13] can be used to implement such activity-based weight update functions.

We have only considered the first layer of dendritic hierarchy, but the same dendritic building block of Fig. 2(b) can be tiled essentially arbitrarily. The depth of this tree is enabled by the logic-level restoration occurring in the basic circuit. Design of the DI loop is independent of the DR loop, and regardless of the configuration of the inputs to the DR loop, as long as threshold can be reached, flux can be added to the DR loop, and a restored current level can be attained with as few as one fluxon. In this work, that current level is around 10 $\mu$A, but it could be designed to be higher or lower as needed. This logic-level restoration enables a many-compartment dendritic tree to be as deep as needed for the desired information processing, pointing to numerous theoretical questions. At the base of the tree is the soma, or cell body. The soma receives signals just as any of the other dendrites, but its output feeds into an amplifier chain that leads to the production of light [13]. Because nanowire single-photon detectors have a binary response, each neuron-to-synapse communication event also results in logic-level restoration, but between neurons and synapses rather than dendrites.

Beyond specifics related to the superconducting optoelectronic hardware implementation, this work touches on important theoretical questions regarding neural information. We have based circuit designs around the hypothesis that incorporating significant dendritic structure beyond the point-neuron model is important for neural processing. Quantification of dendritic information processing is difficult in biological experiments due to the length scales involved, the sensitivity of the neurons and dendrites under study, and the inability to design or control the circuits being investigated. The circuits presented here can be precisely designed, fabricated, manipulated, and measured, potentially leading to traction on theoretical models of dendritic processing. The goal of the dendritic tree is to provide as much information as possible about the temporal activity on a neuron's afferent synapses. Proper design will maximize knowledge in the dendritic tree and the arbor's ability to communicate that information to the cell body. Versatile hardware implementations of neurons with various dendritic processing capabilities may serve to elucidate the important functions of dendrites in biological and artificial neural systems.

## Acknowledgment

This is a contribution of NIST, an agency of the US government, not subject to copyright.

## References

[1] C. Koch, "Computation and the single neuron," *Nature*, vol. 385, pp. 207–210, 1997.

[2] S. Laughlin and T. Sejnowski, "Communication in neuronal networks," *Science*, vol. 301, pp. 1870–1874, 2003.

[3] L. Abbott and W. Regehr, "Synaptic computation," *Nature Rev.*, vol. 431, pp. 796–803, 2004.

[4] W. Gerstner and W. Kistler, *Spiking Neuron Models*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 2002.

[5] G. Stuart and N. Spruston, "Dendritic integration: 60 years of progress," *Nature Neurosci.*, vol. 18, pp. 1713–1721, 2015.

[6] K. Stiefel and T. Sejnowski, "Mapping function on neuronal morphology," *J. Neurophysiol.*, vol. 98, pp. 513–526, 2007.

[7] J. Hawkins and S. Ahmad, "Why neurons have thousands of synapses, a theory of sequence memory in neocortex," *Frontiers Neural Circuits*, vol. 10, p. 23, 2016.

[8] S. Sardi, R. Vardi, A. Sheinin, A. Goldental, and I. Kanter, "New types of experiments reveal that a neuron functions as multiple independent threshold units," *Scientific Rep.*, vol. 7, 2017, Art. no. 18036.

[9] A. Engel, P. Fries, and W. Singer, "Dynamic predictions: Oscillations and synchrony in top-down processing," *Nature Rev. Neurosci.*, vol. 2, pp. 704–716, 2001.

[10] S. Bressler and V. Menon, "Large-scale brain networks in cognition: Emerging methods and principles," *Trends Cogn. Sci.*, vol. 14, pp. 277–290, 2010.

[11] J. Shainline, S. Buckley, R. Mirin, and S. Nam, "Superconducting optoelectronic circuits for neuromorphic computing," *Phys. Rev. Appl.*, vol. 7, 2017, Art. no. 034013.

[12] J. Shainline *et al.*, "Circuit designs for superconducting optoelectronic loop neurons," *J. Appl. Phys.*, vol. 124, 2018, Art. no. 152130.

[13] J. Shainline *et al.*, "Superconducting optoelectronic loop neurons," *J. Appl. Phys.*, vol. 126, 2019, Art. no. 044902.

[14] J. Shainline, "The largest cognitive systems will be optoelectronic," presented at the IEEE Int. Conf. Rebooting Comput., McLean, VA, USA, Nov. 2018.

[15] M. Nahmias, B. Shastri, A. Tait, and P. Prucnal, "A leaky integrate-and-fire laser neuron for ultrafast cognitive computing," *IEEE J. Sel. Topics Quantum Electron.*, vol. 19, no. 5, Sep./Oct. 2013, Art. no. 1800212.

[16] A. Tait *et al.*, "Neuromorphic photonic networks using silicon photonic weight banks," *Nature Sci. Rep.*, vol. 7, 2017, Art. no. 7430.

[17] P. Prucnal and B. Shastri, *Neuromorphic Photonics*, 1st ed. New York, NY, USA: CRC Press, 2017.

[18] Y. Shen *et al.*, "Deep learning with coherent nanophotonic circuits," *Nature Photon.*, vol. 11, pp. 441–446, 2016.

[19] I. Chakraborty, G. Saha, A. Sengupta, and K. Roy, "Toward fast neural computing using all-photonic phase change spiking neurons," *Scientific Rep.*, vol. 8, 2018, Art. no. 12980.

[20] T. Hirose, T. Asai, and Y. Amemiya, "Pulsed neural networks consisting of single-flux-quantum spiking neurons," *Physica C*, vol. 463, pp. 1072–1075, 2007.

[21] P. Crotty, D. Schult, and K. Segall, "Josephson junction simulation of neurons," *Phys. Rev. E*, vol. 82, 2010, Art. no. 011914.

[22] S. Russek *et al.*, "Stochastic single flux quantum neuromorphic computing using magnetically tunable Josephson junctions," presented at the IEEE Int. Conf. Rebooting Comput., San Diego, CA, USA, Oct. 2016.

[23] K. Segall *et al.*, "Synchronization dynamics on the picosecond time scale in coupled Josephson junction networks," *Phys. Rev. E*, vol. 95, 2017, Art. no. 032220.

[24] M. Schneider *et al.*, "Ultralow power artificial synapses using nanotextured magnetic Josephson junctions," *Sci. Adv.*, vol. 4, 2018, Art. no. 1701329.

[25] H. Katayama, T. Fujii, and N. Hatakenaka, "Theoretical basis of squid-based artificial neurons," *J. Appl. Phys.*, vol. 124, 2018, Art. no. 152106.

[26] J. Chiles *et al.*, "Multi-planar amorphous silicon photonics with compact interplanar couplers, cross talk mitigation, and low crossing loss," *APL Photon.*, vol. 2, 2017, Art. no. 116101.

[27] J. Chiles, S. Buckley, S. Nam, R. Mirin, and J. Shainline, "Design, fabrication, and metrology of 10 x 100 multi-planar integrated photonic routing manifolds for neural networks," *APL Photon.*, vol. 3, 2018, Art. no. 106101.

[28] J. Shainline, "Optoelectronic intelligence," unpublished.

[29] G. Gol'tsman *et al.*, "Picosecond superconducting single-photon optical detector," *Appl. Phys. Lett.*, vol. 79, p. 705, 2001.

[30] C. Natarajan, M. Tanner, and R. Hadfield, "Superconducting nanowire single-photon detectors: Physics and applications," *Supercond. Sci. Tech.*, vol. 25, 2012, Art. no. 063001.

[31] D. Liu *et al.*, "Electrical characteristics of superconducting nanowire single photon detector," *IEEE Trans. Appl. Supercond.*, vol. 23, no. 3, Jun. 2013, Art. no. 2200804.

[32] F. Marsili *et al.*, "Detecting single infrared photons with 93% system efficiency," *Nature Photon.*, vol. 7, pp. 210–214, 2013.

[33] M. Tinkham, *Introduction to Superconductivity*, 2nd ed. New York, NY, USA: Dover, 1996.

[34] T. V. Duzer and C. Turner, *Principles of Superconductive Devices and Circuits*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.

[35] A. M. Kadin, *Introduction to Superconducting Circuits*, 1st ed. Hoboken, NJ, USA: Wiley, 1999.

[36] J. Yang *et al.*, "Modeling the electrical and thermal response of superconducting nanowire single-photon detectors," *IEEE Trans. Appl. Supercond.*, vol. 17, no. 2, pp. 581–585, Jun. 2007.

[37] S. Whiteley, "Josephson junctions in SPICE3," *IEEE Trans. Mag.*, vol. 27, no. 2, pp. 2902–2905, Mar. 1991.

[38] S. Song, K. Miller, and L. Abbott, "Competitive Hebbian learning through spike-timing-dependent synaptic plasticity," *Nature Neurosci.*, vol. 3, pp. 919–926, 2000.

[39] H. Markram, W. Gerstner, and P. Sjostrom, "Spike-timing-dependent plasticity: A comprehensive overview," *Frontiers Synaptic Neurosci.*, vol. 4, p. 2, 2012.

[40] G. Buzsaki and A. Draguhn, "Neuronal oscillations in cortical networks," *Science*, vol. 304, pp. 1926–1929, 2004.

[41] G. Buzsaki, *Rhythms of the Brain*. London, U.K.: Oxford Univ. Press, 2006.

[42] L. Roux and G. Buzsaki, "Tasks for inhibitory interneurons in intact brain circuits," *Neuropharmacology*, vol. 88, pp. 10–23, 2015.

[43] K. Likharev and V. Semenov, "RSFQ logic/memory family: A new Josephson-junction technology for sub-terahertz-clock-frequency digital systems," *IEEE Trans. Appl. Supercond.*, vol. 1, no. 1, pp. 3–28, Mar. 1991.

[44] V. Braitenberg and A. Schuz, *Cortex: Statistics and Geometry of Neuronal Connectivity*. Berlin, Germany: Springer, 1998.

[45] J. Beggs, "The criticality hypothesis: How local cortical networks might optimize information processing," *Philos. Trans. Roy. Soc. A*, vol. 366, pp. 329–343, 2007.

[46] S. Buckley *et al.*, "All-silicon light-emitting diodes waveguide-integrated with superconducting single-photon detectors," *Appl. Phys. Lett.*, vol. 111, 2017, Art. no. 141101.

[47] F. Varela, J.-P. Lachaux, E. Rodriguez, and J. Martinerie, "The brainweb: Phase synchronization and large-scale integration," *Nature Rev. Neurosci.*, vol. 2, pp. 229–239, 2001.

[48] E. Salinas and T. Sejnowski, "Correlated neuronal activity and the flow of neural information," *Nature Rev. Neurosci.*, vol. 2, pp. 539–550, 2001.

**Jeffrey M. Shainline** (M'16) was born in Albuquerque, NM, USA, in 1982. He received the B.S. degree in physics from the University of Colorado, Boulder, CO, USA, in 2005, and the Ph.D. degree in physics from Brown University, Providence, RI, USA, in 2010. From 2010 to 2013, he was a Postdoctoral Researcher with the laboratory of Dr. Miloš Popović investigating the use of optical interconnects in CMOS electronics before joining the National Institute of Standards and Technology (NIST), Boulder, in 2013 as a National Research Council Postdoctoral Fellow. In 2017, he became a member of the scientific staff at NIST, where he works in the group of Dr. Richard Mirin and Dr. Sae Woo Nam and leads the Physics and Hardware for Information Project. His current research interests include superconducting optoelectronic hardware for cognitive computing as well as more general aspects of artificial intelligence. He is a member of the Optical Society of America and the American Physical Society.