

Phenomenological model of superconducting optoelectronic loop neurons

Jeffrey M. Shainline ^{1,*} Bryce A. Primavera, ^{1,2} and Saeed Khan ¹

¹*National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado, USA*

²*Department of Physics, University of Colorado Boulder, Boulder, Colorado, USA*



(Received 18 October 2022; accepted 20 January 2023; published 9 March 2023)

Superconducting optoelectronic loop neurons are a class of circuits potentially conducive to networks for large-scale artificial cognition. These circuits employ superconducting components including single-photon detectors, Josephson junctions, and transformers to achieve neuromorphic functions. To date, all simulations of loop neurons have used first-principles circuit analysis to model the behavior of synapses, dendrites, and neurons. These circuit models are computationally inefficient and leave opaque the relationship between loop neurons and other complex systems. Here we introduce a modeling framework that captures the behavior of the relevant synaptic, dendritic, and neuronal circuits at a phenomenological level without resorting to full circuit equations. Within this compact model, each dendrite is discovered to obey a single nonlinear leaky-integrator ordinary differential equation, while a neuron is modeled as a dendrite with a thresholding element and an additional feedback mechanism for establishing a refractory period. A synapse is modeled as a single-photon detector coupled to a dendrite, where the response of the single-photon detector follows a closed-form expression. We quantify the accuracy of the phenomenological model relative to circuit simulations and find that the approach reduces computational time by a factor of ten thousand while maintaining an accuracy of one part in ten thousand. We demonstrate the use of the model with several basic examples. The net increase in computational efficiency enables future simulation of large networks, while the formulation provides a connection to a large body of work in applied mathematics, computational neuroscience, and physical systems such as spin glasses.

DOI: 10.1103/PhysRevResearch.5.013164

I. INTRODUCTION

A technological platform capable of realizing networks at the scale and complexity of the brains of intelligent organisms would be a tool of supreme scientific utility. Neuromorphic hardware based on conventional silicon microelectronics has a great deal to offer in this regard [1–5]. Yet challenges remain, primarily concerning bottlenecks in the shared communication infrastructure that must be employed to emulate the connectivity of biological neurons. Alternative hardware may bring new benefits, and multiple techniques have been pursued over the years. The use of optics within neural systems has been pursued [6], beginning with free-space optics to realize Hopfield networks [7], integrated photonics to implement inference in a deep-learning framework [8,9], and waveguide approaches to three-dimensional optical interconnectivity [10,11]. Superconducting electronic circuits have also been pursued [12–14], with recent work targeting deep learning [15–17] and spiking neural networks [18,19]. We have argued elsewhere for the advantages of a superconducting optoelectronic approach [20,21]. In brief, light

for communication enables high fan-out with low latency across spatial scales from a chip to a many-wafer system. Superconducting electronics provide single-photon detection coupled to high-speed, low-energy analog neuromorphic computational primitives.

While the components of these superconducting optoelectronic networks (SOENs) have been demonstrated [11,22–24], full neurons have not. Prior to undertaking the effort and expense of realizing the requisite semiconductor-superconductor-photonic fabrication process, it is prudent to gain confidence that SOENs are indeed ripe for further investigation. This confidence can be gained through simulations of device, circuit, and system behavior using numerical simulations on digital computers. The constituent devices are most commonly modeled with circuit simulations carried out on a picosecond timescale to accurately capture the dynamics of Josephson junctions (JJs). Simulation of networks of large numbers of these neurons becomes computationally intensive. From the perspective of the neural system, the picosecond dynamics of the JJs are not of primary interest, and one would prefer to treat each synapse, dendrite, and neuron as an input-output device with a model that accurately captures the circuit dynamics on the nanosecond to microsecond timescales while not explicitly treating the picosecond behavior of the underlying circuit elements.

Here we introduce a phenomenological model of loop neurons and their constitutive elements that accurately captures the transfer characteristics of the circuits without solving the underlying circuit equations. Dendrites are revealed to be

*jeffrey.shainline@nist.gov

central to the system. Each dendrite is treated with a single, first-order ordinary differential equation (ODE) that describes the output of the element as a function of its instantaneous inputs and internal state. These equations take the form of a leaky integrator with a nonlinear driving term. The input to each dendrite is flux, and the output is an integrated current, which is coupled through a transformer into another dendrite. Synapses are dendrites with a closed-form expression for the input flux following each synapse event. The soma of a neuron is also modeled with the same equations as a dendrite with two modifications. First, when the output current reaches a specified level, threshold is reached, and an electrical signal is sent to a transmitter circuit, which produces a pulse of light. Second, a refractory dendrite triggered off this output is coupled back to the soma inhibitively to achieve a refractory period.

By working at the phenomenological level, the time to simulate single dendrites is decreased by a factor of ten thousand while maintaining an accuracy of one part in ten thousand. The speed advantage grows with the size of the system being simulated and the duration of the simulation. The functional form of the speedup with size and duration has not been fully investigated, as solving the full system of circuit equations becomes very time consuming for even small systems. In addition, the model makes transparent the qualitative behavior of all components of the system, including the similarities to biological neurons as well as other physical systems such as Ising models, spin glasses, and pulse-coupled oscillators.

We begin by motivating the form of the model based on circuit considerations. We then describe the means by which the form of the driving term in the leaky integrator is obtained. Error is quantified by comparison with full circuit simulations, and convergence is investigated as a function of time step size. Numerical examples of dendrites, synapses, and neurons are presented. An example of a neuron with a dendritic arbor performing an image-classification task is given to illustrate the utility of the model. Further extensions to enable theoretical treatment of very large systems are discussed.

II. OVERVIEW OF LOOP NEURONS

A schematic diagram of a loop neuron is shown in Fig. 1, depicting the complex dendritic tree that appears central to the computations of loop neurons [25,26]. In these neurons, integration, synaptic plasticity, and dendritic processing are implemented with inductively coupled loops of supercurrent. It is due to the prominent role of superconducting storage loops that we refer to devices of this type as loop neurons [20]. Operation of loop neurons is as follows. Photons from upstream neurons are received by a superconducting single-photon detector (SPD) at each synapse. Using a superconducting circuit comprising two JJs coupled to the SPD, synaptic detection events are converted into an integrated supercurrent which is stored in a superconducting loop. The amount of current added to the integration loop during a photon detection event is determined by the synaptic weight. The synaptic weight is dynamically adjusted by another circuit combining SPDs and JJs, and all involved circuits are analog. When the integrated current of a given neuron reaches a (dynamically variable) threshold, an amplification cascade

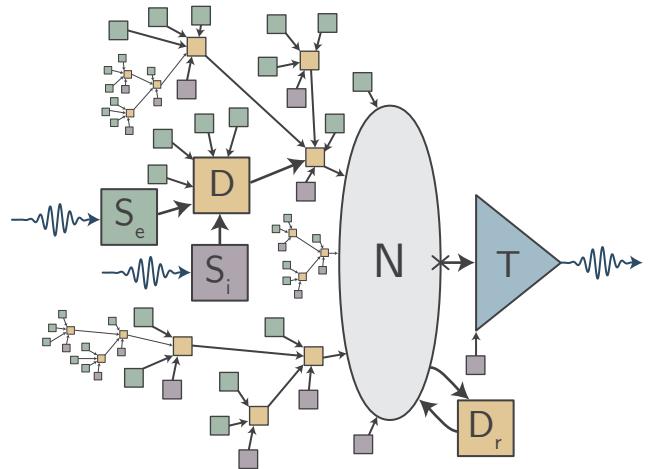


FIG. 1. Schematic of a loop neuron with an elaborate dendritic tree. The complex structure consists of excitatory and inhibitory synapses (S_e and S_i) that feed into dendrites (D). Each dendrite performs computations on the inputs and communicates the result to other dendrites for further processing or on to the cell body of the neuron (N). The neuron cell body 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 via photonic waveguides.

begins, resulting in the production of light from a waveguide-integrated semiconductor light emitter. The photons thus produced fan out through a network of dielectric waveguides and arrive at the synaptic terminals of other neurons where the process repeats.

The core active component of loop neurons is a circuit known as a superconducting quantum interference device (SQUID), which comprises two JJs in parallel. A circuit diagram is shown in Fig. 2(a). The SQUID is a three-terminal device with a bias, ground, and an input which couples flux into the loop formed by the two JJs and inductors. For the present purpose, the flux input is the active signal. When the SQUID is current-biased below the critical current of the two JJs and the flux input is below a bias-dependent threshold, the SQUID remains superconducting, and the voltage across the device is zero. When the applied flux exceeds the threshold at a given bias point, the JJs will begin to emit a series of voltage pulses known as fluxons, and the time-averaged voltage across the device will become nonzero.

To form a dendrite from a SQUID, we first ensure that it is biased below the critical current of the JJs so it is quiescent when no flux is applied. The output of the SQUID is captured by an L - R loop that performs current integration with a leak [Fig. 2(b)]. When sufficient flux is input to the SQUID to drive the JJs to begin producing voltage pulses, the pulses drive current into the L - R loop, these pulses are summed in the inductor, and the accumulated signal leaks with the L/R time constant of the loop. This configuration of a SQUID coupled to an L - R loop is referred to as a dendrite, the SQUID that receives input flux is referred to as the receiving (R) loop, and the L - R component is referred to as the integration (I) loop. The current stored in the integration loop produces the signal that will be communicated to other dendrites. To

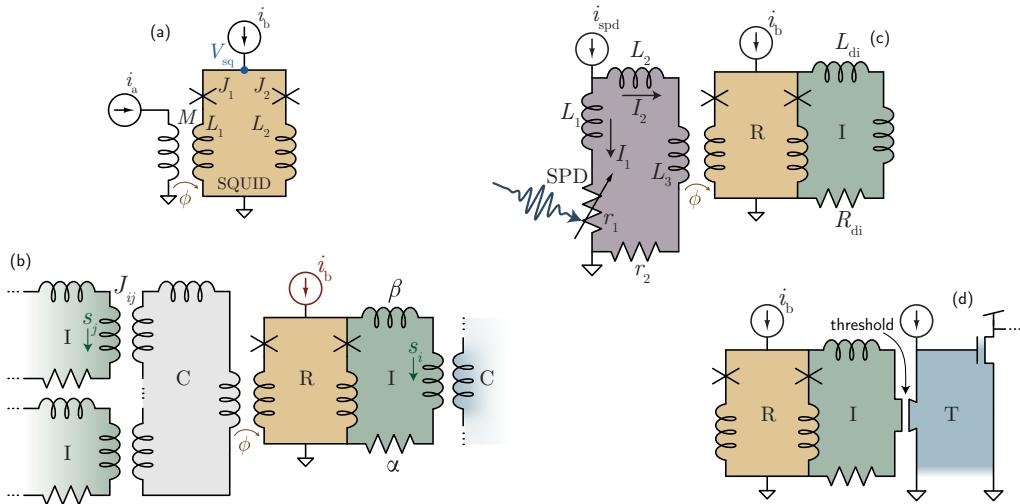


FIG. 2. Circuit diagrams. (a) A two-junction SQUID that forms the dendritic receiving loop. (b) A dendrite with receiving (R) and integrating (I) loop. The integrating loops of two other dendrites are input to a collection coil (C) that delivers flux to the receiving loop. (c) A synapse formed with an SPD input to a dendrite. (d) A soma realized as a dendrite with thresholding component in the integration loop and initial stage of the transmitter (T).

receive signal from multiple input dendrites, a passive collection coil (C) is used. A circuit diagram of a dendrite with two inputs, a collection coil, a receiving loop, and an integrating loop coupled to an output is shown in Fig. 2(b). All coupling between dendrites is through magnetic flux communicated through mutual inductors. The use of transformers for this purpose mitigates cross talk and enables high fan-in [26].

To form a synapse from a dendrite, we attach an SPD to the flux input to the receiving loop. This circuit is shown in Fig. 2(c). When an SPD detects a photon, it rapidly switches from zero resistance to a large resistance, diverting the bias current to the other branch of the circuit. This current is coupled into the receiving loop of the dendrite as flux, driving the dendrite above threshold and adding current to the dendrite's integration loop.

To form a neuron from a dendrite, two modifications are required. First, the integration loop must be equipped with a thresholding element that drives a transmitter circuit to produce light when the integrated signal reaches this threshold. This thresholding element coupled to the transmitter is shown in Fig. 2(d); it is referred to as a tron and is described in more detail in Sec. V. Second, an additional dendrite is attached to the neuron that provides negative feedback to achieve a refractory period (not shown in the circuit diagram for simplicity). The refractory period is a brief period of quiescence following a neuronal spike event. When the neuron's integration loop reaches threshold, this refractory dendrite is driven, accumulates signal in its integration loop, and this signal suppresses the state of the neuron's receiving loop. The time constant of the refractory dendrite establishes the refractory period of the neuron.

To summarize, a dendrite is a SQUID with a flux input adding current to a leaky integration loop. A synapse is a single-photon detector coupled to a dendrite. The soma of a neuron is a dendrite with a thresholding element in the integration loop as well as a second dendrite that provides feedback for refraction. To construct full loop neurons, many

synapses are coupled into a dendritic arbor which feeds forward into the soma. The output of the soma is an optical pulse that couples light to a network of waveguides and delivers faint photonic signals to downstream synapses where they are received with single-photon detectors. This construction is illustrated schematically in Fig. 1, and the basic circuits are shown in Fig. 2. The current in a dendritic integration loop is analogous to the membrane potential of a biological dendrite [27,28], and these signals are the principal dynamical variables of the system. Inhibitory synapses can be achieved through mutual inductors with the opposite sign of coupling. Complex arbors with multiple levels of dendritic hierarchy can be implemented to perform various computations [25] as well as to facilitate a high degree of fan-in [26].

We see that dendrites are central to loop neurons. The states of current in all dendritic integration loops specifies the state of the system. A phenomenological model of loop neurons must therefore capture the temporal evolution of these currents as well as the coupling between dendrites. With these concepts in mind we proceed to construct the model.

III. DENDRITE MODEL

As mentioned in Sec. II, a SQUID is the primary active element of a dendrite. To motivate the phenomenological dendrite model we require a quantitative understanding of SQUID operation. The two-junction SQUID of Fig. 2(a) with symmetrical inductances was modeled using a first-principles circuit model [29], and the results are shown in Fig. 3. Figure 3(a) shows a time trace of the voltage across the SQUID when it is in the voltage state. The peaks corresponding to fluxon production are evident, and the time-averaged voltage is also shown. While the voltage is a rapidly varying function of time on the picosecond scale, the time-averaged voltage is steady.

This behavior is analyzed systematically in Fig. 3(b). The time-averaged voltage of a symmetric SQUID is plotted as a function of the applied flux for several values of the bias

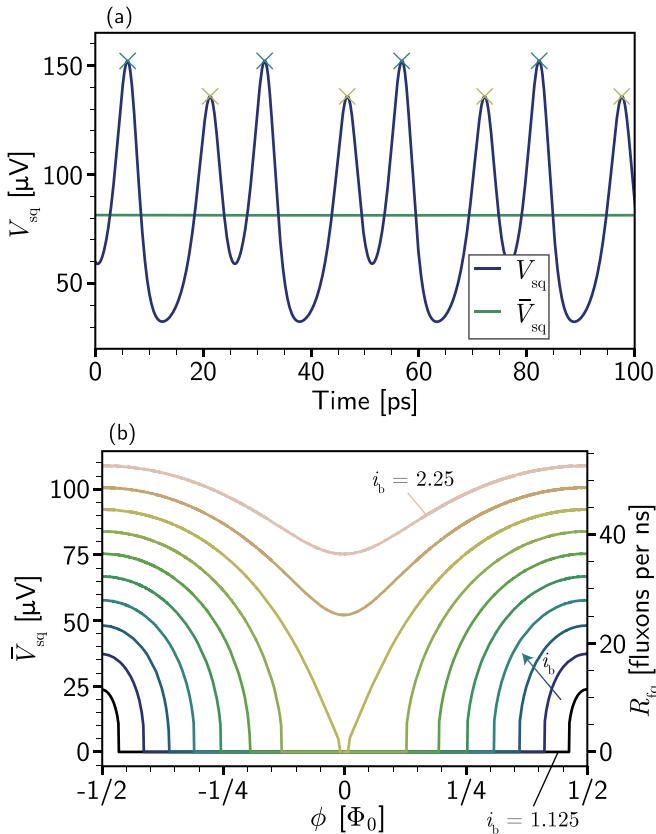


FIG. 3. The response of a SQUID. (a) Time trace of a SQUID biased in the voltage state. Fluxon peaks are marked with crosses of different colors for the two JJs. The time average is also shown as calculated by taking the average of the time trace between two fluxons produced by the same junction. (b) The time-averaged voltage across the SQUID as a function of applied flux, ϕ , normalized to the magnetic flux quantum, Φ_0 , for several values of normalized bias current, i_b .

current, which has been normalized to the critical current of a single junction ($i_b = I_b/I_c$). The time-averaged voltage, plotted on the left y axis, is proportional to the rate of flux-quantum production, plotted on the right y axis. The relationship results from the single-valued nature of the superconducting wave function around the closed SQUID loop, which requires that

$$\int_0^{t_{\text{fq}}} V_{\text{sq}}(t) dt = \Phi_0. \quad (1)$$

Equation (1) informs us that the time required to produce a single fluxon, t_{fq} , is related to the voltage across the squid, V_{sq} . For constant voltage, we obtain $r_{\text{fq}} = 1/t_{\text{fq}} = V_{\text{sq}}/\Phi_0$. When viewed over timescales appreciably longer than t_{fq} , it makes sense to speak of a rate of flux-quantum production, r_{fq} , and this is the first element of our model: when driven to the active state, a dendrite will begin to produce fluxons, which carry current, and we can track this current by monitoring the rate of fluxon production while ignoring the picosecond dynamics by which the JJs produce the fluxons.

Several features of Fig. 3(b) are pertinent to the present study. First, for a given value of i_b , a finite value of flux is required before the SQUID enters the voltage state. This

provides a nonzero threshold that is relevant to dendritic computation. This threshold can be adjusted with the bias current. Second, the response is periodic in applied flux, with the period being $\Phi_0/2$, where $\Phi_0 = h/2e \approx 2 \times 10^{-15} \text{ V s} = 2 \text{ mV ps}$ is the magnetic flux quantum. To maintain a monotonic response, the applied flux must be limited to this value [26]. Third, if the inductors L_1 and L_2 are equal, the response of the SQUID is symmetric about $\Phi = 0$. These features will be discussed further as the study proceeds.

The second element of the model captures the integration and leak of the current generated when the SQUID is driven above threshold. These behaviors are accomplished by adding an L - R loop to the output of the SQUID, as shown by the integration loop labeled I in Fig. 2(b). The current integrated in this branch of the circuit is the quantity of interest for the dendrite. It is this quantity that will couple to other dendrites or the neuron cell body, and it is this quantity we wish to track with our phenomenological model. Dendrites comprising a receiving loop coupled to an integrating loop [Fig. 2(b)] are referred to as RI dendrites. We know from elementary circuit theory that the L - R loop will result in exponential decay of signal with a time constant of the dendritic integration loop given by $\tau_{\text{di}} = L_{\text{di}}/R_{\text{di}}$.

We can now write down a postulated expression for the signal s stored in the integration loop of a dendrite:

$$\beta \frac{ds}{d\tau} = r(\phi, s; i_b) - \alpha s. \quad (2)$$

Equation (2) states that the signal s grows in time due to the driving term, which is the rate of flux quantum production, denoted by r . This function r depends on the applied flux to the receiving loop of the dendrite, $\phi = \Phi/\Phi_0$, as well as of the signal present in the integration loop, s . The rate also depends on the bias current, i_b , as a parameter that throughout this work is assumed to be held constant over times much longer than the inter-fluxon interval. In Eq. (2), we have formulated the model in dimensionless units, where $s = I_{\text{di}}/I_c$, and I_{di} is the current present in the dendritic integration loop. $\beta = 2\pi I_c L_{\text{di}}/\Phi_0$ is a dimensionless parameter that quantifies the inductance of the loop, and $\alpha = R_{\text{di}}/R_j$, where R_j is the shunt resistance of each JJ in the resistively and capacitively shunted junction model [29–32]. The signal s decays at a rate related to α and β . Specifically, the time constant for decay is given by $\tau_{\text{di}} = L_{\text{di}}/R_{\text{di}} = \beta/\omega_c \alpha$, where ω_c is the Josephson characteristic frequency discussed in Appendix A, and we include the subscript on τ_{di} to refer to the dendritic integration loop and distinguish that quantity from the dimensionless time variable entering Eq. (2). Equation (2) is a leaky integrator ODE. The drive term is the rate of flux quantum production, and the leak term gives simple exponential decay, as expected from an L - R circuit.

Dendrites are coupled to each other through flux. The coupling flux from dendrites indexed by j to dendrite i is given by

$$\phi_i = \sum_{j=1}^n J_{ij} s_j, \quad (3)$$

where J_{ij} is a coupling term proportional to the mutual inductance that includes contributions from all the transformers present on the collection coil in Fig. 2(b). Equation (3) shows

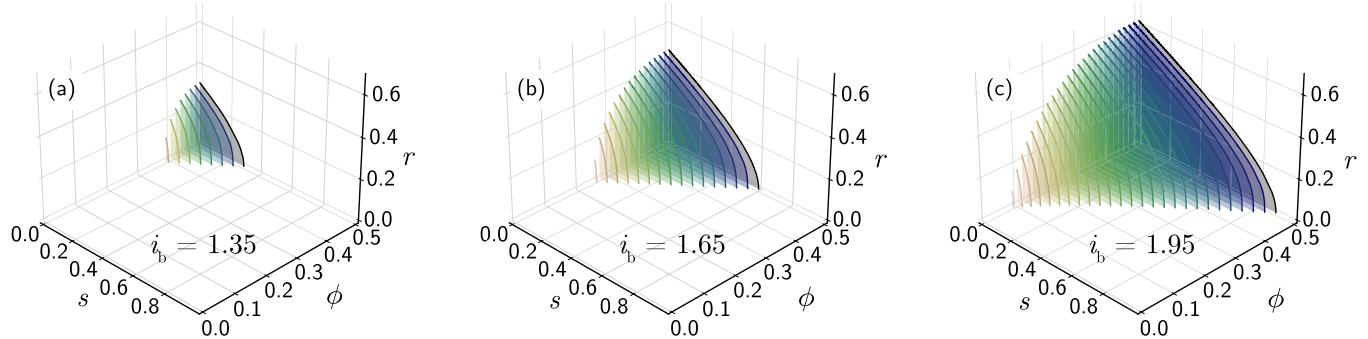


FIG. 4. Rate arrays for the dendrite with a receiving and integrating loop for three values of i_b . The normalized bias current (a) $i_b = 1.35$, (b) 1.65, and (c) 1.95.

that coupling between dendrites is due to the signal in the integration loop of one dendrite being communicated as flux into the receiving loop of a subsequent dendrite. The signal in the subsequent dendrite is then obtained through the evolution of Eq. (2) with the flux from the first dendrite providing the flux ϕ in the function r and the signal from the second dendrite providing the s term. The i_b term entering r refers to the bias on the second dendrite and is treated here as a parameter rather than a dynamical variable. Further details regarding the derivation of these expressions is given in Appendix C.

Equations (2) and (3) constitute the phenomenological model of a dendrite. A neuron or network can be simulated by solving these coupled equations for all dendrites in the system. However, we have not specified the form for the rate function, $r(\phi, s; i_b)$, which is central to the model.

We have arrived at Eq. (2) as a postulate; this expression is not directly obtained from the underlying circuit equations. The postulate is that there will be a function $r(\phi, s; i_b)$ such that Eq. (2) provides an accurate account of the signal present in a dendrite's integration loop under the circumstances of interest for loop neurons, provided we only inquire about the signal over timescales appreciably longer than the interfluxon interval, which is on the order of 10 ps. We aim to interrogate dendrites on timescales of 100 ps or longer, with neuron and network activity of interest on timescales from nanoseconds to the longest timescales that can be simulated under the limits of computational resources.

Our procedure for obtaining $r(\phi, s; i_b)$ is as follows. A dendrite with a SQUID as a receiving loop and an L - R branch as an integrating loop is numerically modeled with the circuit equations given in Appendix A. A constant value of flux is applied to the receiving loop, and the rate of flux-quantum production is monitored as a function of time while current accumulates in the integration loop. For these simulations, the resistance of the integration loop is set to zero. This procedure is repeated for many values of ϕ and i_b to obtain what we refer to as “rate arrays,” which are shown in Fig. 4, where $r(\phi, s; i_b)$ is plotted as a function of ϕ and s for three values of i_b . Here we work in dimensionless units, so the units of r are fluxons generated per unit of dimensionless time τ , which is related to the JJ characteristic frequency (Appendix A). It can be seen that the value of $r(\phi, s; i_b)$ is monotonically increasing with ϕ over the range considered here, while accumulation of s decreases the rate of flux-quantum production. This decrease is because addition of current to the integration loop diverts

the bias away from the SQUID, so the voltage is decreased, and the rate is reduced in accordance with Eq. (1). When sufficient signal is accumulated in the integration loop, the rate of flux quantum production drops to zero, and we say the loop is saturated.

For a small value of i_b [Fig. 4(a)] a large amount of flux is required to drive the dendrite above threshold to the active state, and a small signal s is present at saturation. As i_b is increased [Figs. 4(b) and 4(c)] the threshold is reduced, and the saturation level is increased. The qualitative shape of r for different values of i_b is similar, and the surfaces for smaller i_b are seen to fit inside those for larger i_b . For this reason, we refer to these surfaces as r -shells, with each shell corresponding to a fixed value of i_b . Network adaptation, training, and learning consist in finding the values of i_b that achieve the desired input-output relationships for a given computational or control task.

While Fig. 4 shows r for $0 \leq \phi \leq 1/2$, these surfaces are symmetric about zero for the circuit of Fig. 2(b) with symmetric inductances in the SQUID [as is also the case in the SQUID response of Fig. 3(b)], and they are periodic in normalized flux ϕ with period of unity (period of Φ_0 in SI units). One consequence of this response is that it may be necessary to restrict the applied flux to $\phi \leq 1/2$ to retain a monotonic response [26], and if inhibition is applied to a dendrite ($\phi < 0$) it may be necessary to limit this flux below the threshold for activity so that inhibition does not drive the dendrite to the active state. These concepts are discussed further in Sec. VI.

From the rate arrays, we obtain several quantities of interest that facilitate dendrite design. In particular, the value of ϕ for which r becomes nonzero as a function of i_b determines the dendrite threshold. This function is shown in Fig. 5(a) for both positive and negative flux. These traces can be leveraged to anticipate when a dendrite will become active as well as the maximum value of inhibitory flux that can be applied before activity results, which is necessary in design of the refractory dendrite that quenches activity of a soma. The saturation values of s as a function of applied flux and bias current are also important for anticipating the maximum value of signal that a dendrite will accumulate, and therefore the maximum value of flux it will couple into another dendrite. We refer to this value as s_{\max} , and Fig. 5(b) shows its behavior as a function of both positive and negative applied flux ($s_{\max}^{+/-}$). The inset shows the value versus i_b for the maximum absolute value of applied flux, $|\phi| = 1/2$.

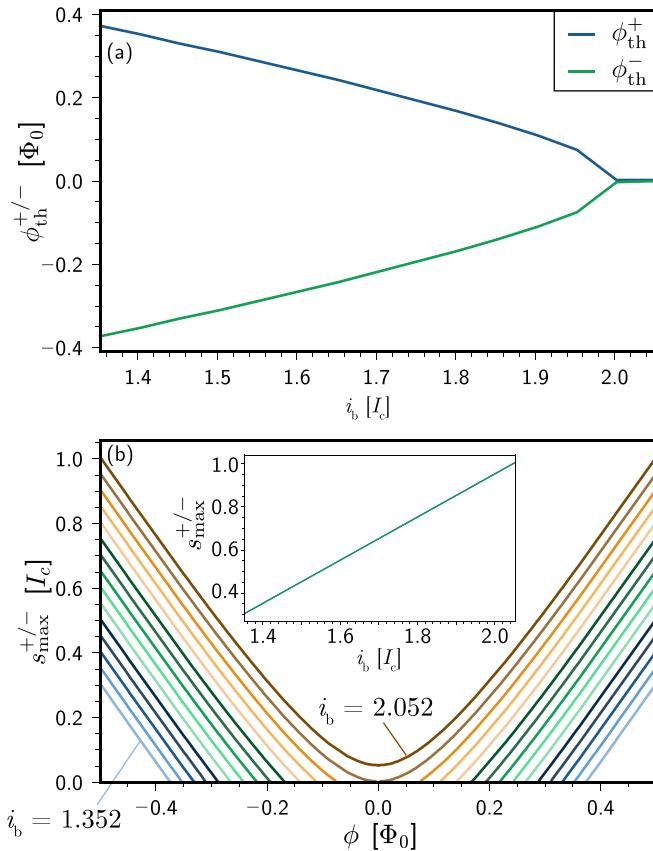


FIG. 5. Response parameters of the RI dendrite extracted from the rate arrays. (a) The thresholds to positive and negative applied flux as a function of the dendritic bias current. (b) The value of saturation signal as a function of the normalized flux applied to the receiving loop for several values of the normalized bias current, i_b .

Obtaining the rate arrays is somewhat computationally intensive, requiring a large number of numerical simulations of the dendrite circuit. However, these computations only need to be carried out once. Calculation of the rate arrays used in the present work took roughly 72 hours using a 3.7 GHz central processing unit with 10 cores and 20 logical processors. The rate arrays are then accessed by the phenomenological model as a look-up table, and the same arrays are used for all values of the integration loop inductance and leak rate. Using the same arrays for all values of inductance is possible because in all cases the inductance of the integration loop is much larger than that of the receiving loop, so the fraction of current bias that initially goes into the integration branch of the circuit is small in all cases of interest here. If properties of the receiving loop—the SQUID—are changed, new rate arrays will need to be calculated. Examples include changing the value of β_c of the JJs or changing the asymmetry of the inductances. However, there appears to be little advantage from employing SQUIDs with a variety of β_c values. Leveraging different asymmetry designs may be fruitful, but a small number of variants will likely suffice. We anticipate a symmetric SQUID and a highly asymmetric SQUID will be the two cases of primary interest.

Equipped with the phenomenological model [Eqs. (2) and (3)] as well as the rate arrays, we can compare the model

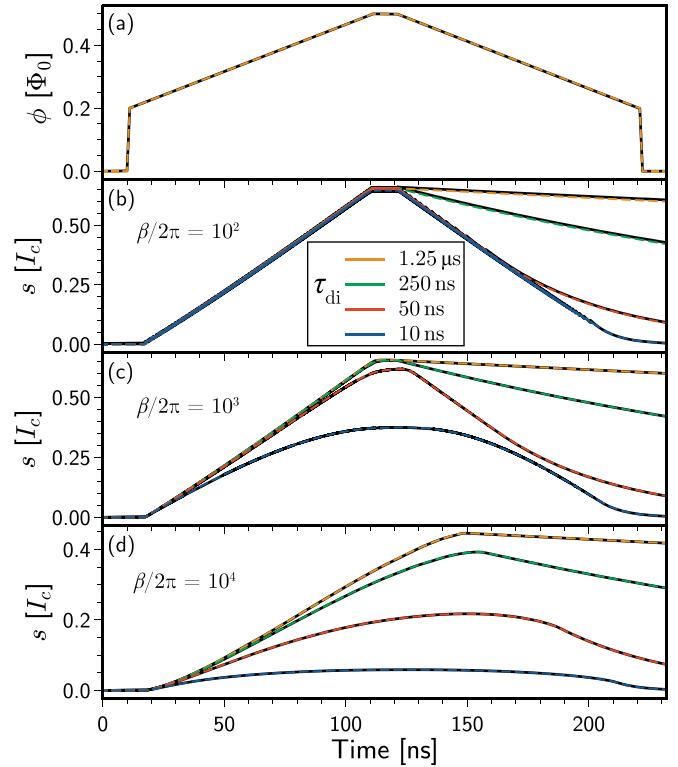


FIG. 6. Comparing the phenomenological model to circuit equations in the case of a linearly ramping drive. The time step for the phenomenological model was 100 ps. (a) The flux drive signal. (b) The smallest integration loop with $\beta/2\pi = 10^2$. (c) Integration loop with $\beta/2\pi = 10^3$. (d) $\beta/2\pi = 10^4$. In (a)–(d), the black traces show the values for the circuit simulations while the dashed colored traces show the values for the phenomenological model. The error of the drive signal was $\chi^2_{\text{drive}} < 5 \times 10^{-11}$ for all cases shown here.

performance to explicit solution of the circuit equations. Throughout this work, we numerically solve systems described by Eqs. (2) and (3) using a forward Euler method to facilitate computational speed. In discrete form, Eq. (2) becomes

$$s_{\tau+1} = s_\tau \left(1 - \Delta\tau \frac{\alpha}{\beta} \right) + \frac{\Delta\tau}{\beta} r(\phi_\tau, s_\tau; i_b). \quad (4)$$

We solve the circuit equations given in Appendix A using a Runge-Kutta method implemented with the PYTHON SCIPY function `solve_ivp` for initial-value problems, which uses an adaptive time mesh.

As a first test, Fig. 6 compares the phenomenological model to the circuit equations in the presence of a linear ramp function as the applied flux. The form of the applied flux into the receiving loop (ϕ) is shown in Fig. 6(a). The rate arrays are obtained only for discrete values of applied flux (200 values between 0 and 1/2), so this applied drive is relevant to quantify performance when the applied flux takes intermediate values between those explicitly present in the rate arrays. No interpolation was used. At each time step, the three inputs to the rate array (ϕ , s , and i_b) are rounded to the closest value for which the rate array has been evaluated, and that value of r is used in Eq. (4) at that time step. The dendrite signals are shown in Figs. 6(b)–6(d) for three values

of the inductance parameter, β , from $\beta/2\pi = 10^2$ – 10^4 . For each value of β , four time constants are considered from 10 ns to 1.25 μ s. The value of $\beta/2\pi$ quantifies the number of fluxons that can be accommodated in the storage loop, and the time constant is converted to the dimensionless parameter α for simulation in the dimensionless model. All data from the phenomenological model in Fig. 6 were obtained using a time step of 100 ps [converted to dimensionless time for Eq. (4)]. The circuit model time traces are shown as black lines, while the traces obtained with the phenomenological model are shown as various dashed colored lines. Accuracy of the phenomenological model is quantified with a χ^2 of the form

$$\chi^2 = \frac{\sum_{i=1}^{n_t^{\text{soen}}-1} |s_{\text{soen}}(t_i^{\text{soen}}) - s_{\text{ode}}^{\text{interp}}(t_i^{\text{soen}})|^2 \Delta t_i^{\text{soen}}}{\sum_{i=1}^{n_t^{\text{ode}}-1} |s_{\text{ode}}(t_i^{\text{ode}})|^2 \Delta t_i^{\text{ode}}}. \quad (5)$$

The subscript or superscript “soen” refers to the phenomenological model, while “ode” refers to the first-principles circuit model of Appendix A. In Eq. (5), n_t^{soen} is the number of time steps in a given SOEN simulation, s_{soen} is the state quantity obtained with the SOEN phenomenological model, t_i^{soen} is the time of the given simulation at time step i , $s_{\text{ode}}^{\text{interp}}$ is the state quantity obtained with the ODE circuit model solved with an adaptive time mesh and interpolated to the coarser phenomenological model time mesh, and $\Delta t_i^{\text{soen}} = t_{i+1}^{\text{soen}} - t_i^{\text{soen}}$. In the denominator, n_t^{ode} is the number of time steps in the ODE solution with the adaptive mesh, $s_{\text{ode}}(t_i^{\text{ode}})$ is the ODE solution on the nonuniform time grid without interpolation, and $\Delta t_i^{\text{ode}} = t_{i+1}^{\text{ode}} - t_i^{\text{ode}}$. In all cases shown in Fig. 6, the value of χ^2 is less than 7×10^{-5} , and the time required to run the ODE model exceeded that to run the SOEN model by a factor of more than one thousand. The accuracy is worst for the smallest storage loop ($\beta/2\pi = 10^2$) and the shortest time constant ($\tau_{\text{di}} = 10$ ns). All values of χ^2 and ratios of simulation times are given in Table I in Appendix D.

As another test case, a series of square flux pulses was input to a dendrite for the same values of β and τ_{di} , as shown in Fig. 7. The heights of these pulses were drawn randomly from the interval between the flux threshold [Fig. 5(a)] and the maximum value of 1/2. The pulse durations were drawn randomly from the interval between 5 and 40 ns, while the pauses between pulses were drawn randomly from the interval between 10 and 100 ns. In Fig. 7, 10 such pulses were applied. The largest value of χ^2 was 3.32×10^{-4} , again in the case of smallest β and τ_{di} , while the time to complete the ODE simulations was longer than that for the SOEN simulations by at least 10³ in all cases. Complete data for the square pulses is given in Table II in Appendix D. To push the limits of speed, the same linear ramp and square pulse cases were carried out using a time step of 1 ns in the phenomenological model. The results for the square-pulse drive are shown in Fig. 8 for the same values of β and τ_{di} . In all cases, the agreement between the two models is good, except for the smallest inductance of $\beta/2\pi = 10^2$. Poor performance in these cases results from the fact that the term $\Delta\tau \frac{\alpha}{\beta}$ in Eq. (4) becomes large, leading to numerical instability. More detail is shown in Fig. 9 where a temporal zoom on just the first input square pulse is presented. The numerical performance is seen to be unacceptable

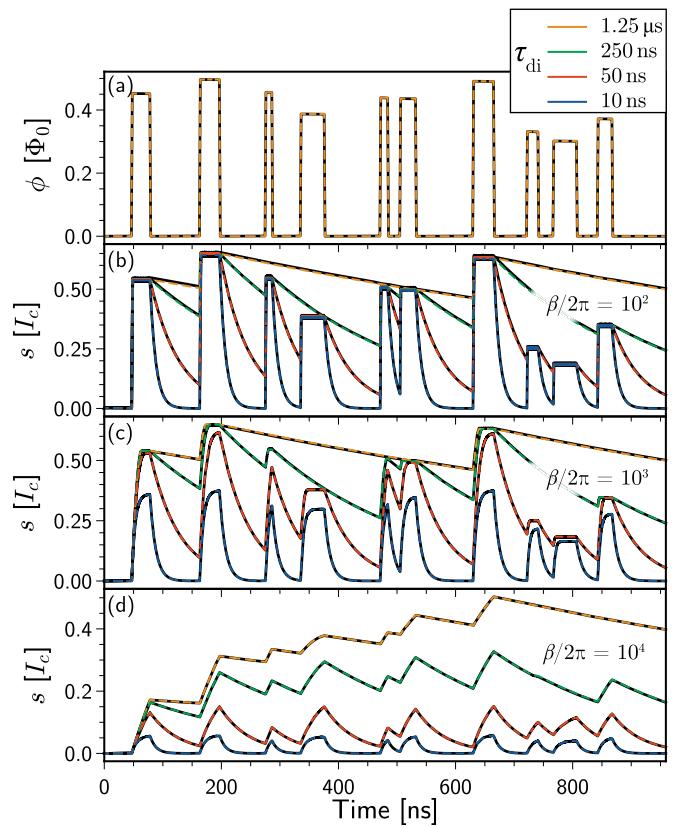


FIG. 7. Comparing the phenomenological model to circuit equations in the case of a random sequence of square pulses. The time step for the phenomenological model was 100 ps. (a) The flux drive signal. (b) The smallest integration loop with $\beta/2\pi = 10^2$. Integration loop with (c) $\beta/2\pi = 10^3$ and (d) 10^4 . The error of the drive signal was $\chi_{\text{drive}}^2 = 5.03 \times 10^{-5}$ for all cases shown here.

for the smallest value of β and only marginally acceptable for values above that. To determine the largest acceptable time step, square-pulse-input simulations were conducted for numbers of input pulses ranging from 10 to 160 with time steps ranging from 10 ps to 10 ns for the same values of β as above. To keep the computation time manageable, only $\tau_{\text{di}} = 250$ ns was considered, but individual test cases were conducted with other values to ensure the conclusions were insensitive to this number. The results of the χ^2 values for each simulation as well as the ratio of computational run times were stored, and the results for the case of 160 random square pulse inputs is shown in Fig. 10. Further data on the other cases is given in Fig. 27 in Appendix D. The χ^2 values again show that the lowest value of β is the most difficult to match, and the χ^2 converges close to 10^{-4} with a time step of 200 ps. All other β values approach $\chi^2 = 10^{-5}$ with this same time step. The simulated time interval in this case was 12 μ s, and with a time step of 200 ps it took over ten thousand times longer to solve the system of circuit ODEs as to step through the phenomenological model. Based on this analysis, a time step of 200 ps appears an optimal compromise between speed and accuracy. This time step is likely the convergence point in these simulations as 200 ps is the rise and fall time of the square pulses. To simulate 12 μ s of activity of a single

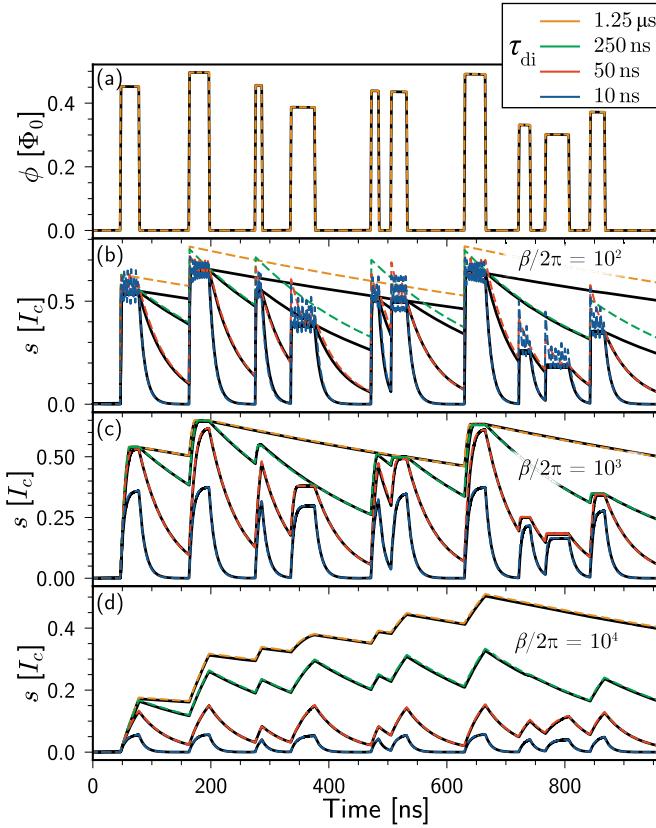


FIG. 8. Similar comparison as Fig. 7 except the time step was 1 ns. The error in reproducing the drive signal was $\chi^2_{\text{drive}} = 7.35 \times 10^{-3}$.

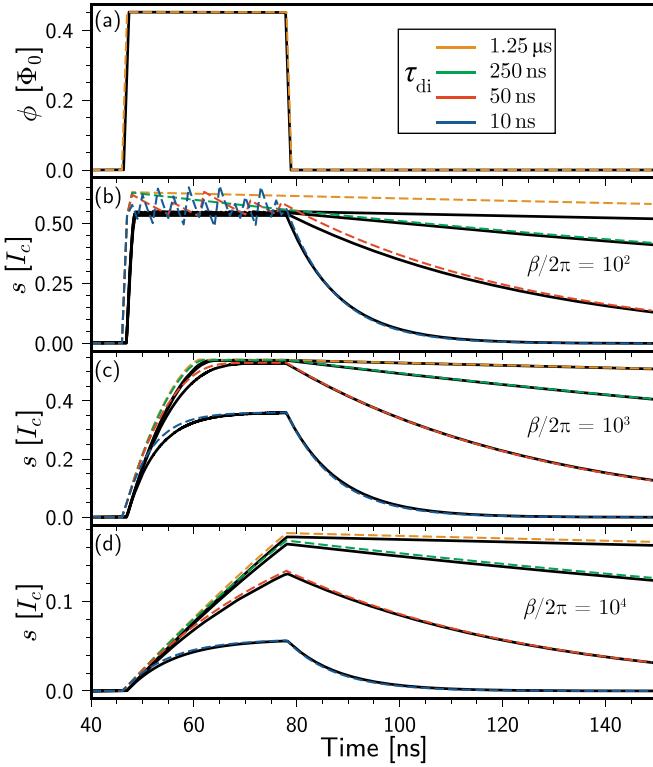


FIG. 9. The same simulations as Fig. 8 except zoomed on the time window around the first input pulse.

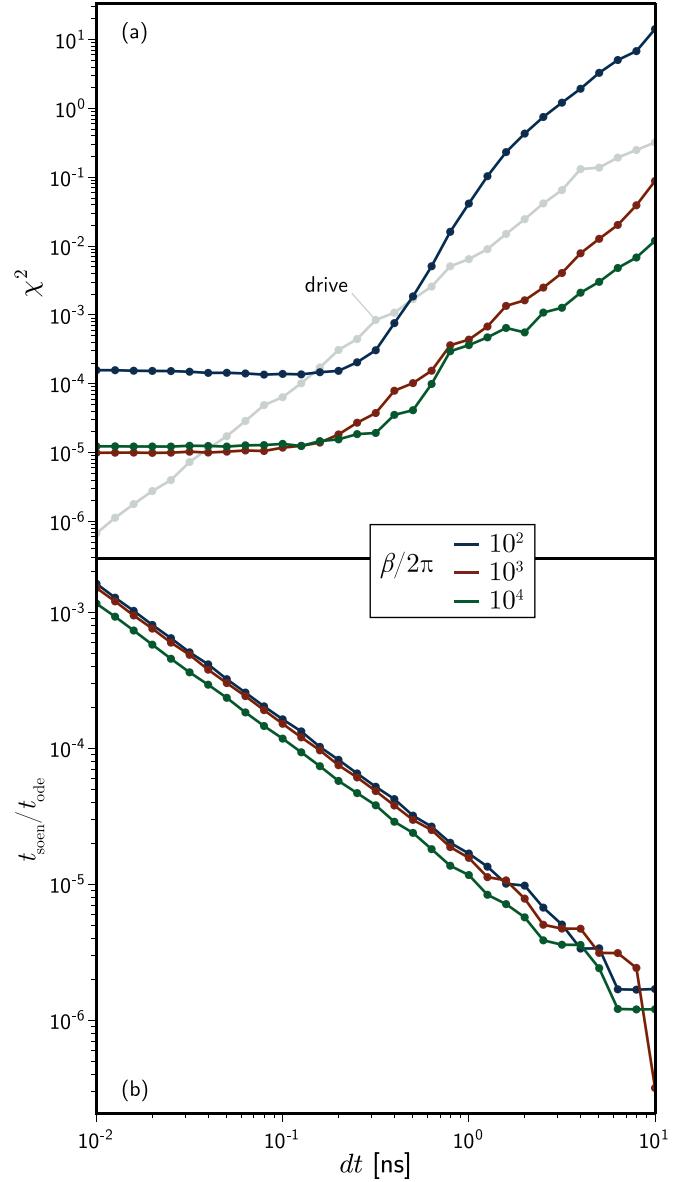


FIG. 10. Quantification of the phenomenological model vs the circuit model. For the dendrite considered here, $\tau_{di} = 250$ ns and $i_b = 1.702$. These simulations included 160 random square pulses, and the circuit model was set to converge to a relative and absolute accuracy of 10^{-5} in the `solve_ivp` function. (a) The values of χ^2 as a function of time step dt for three values of β . (b) The ratio of the time required to run the simulations for the two numerical approaches as a function of time step dt .

dendrite with the phenomenological model required 1.02 s, while simulating the same time interval with the ODE model required 2.78 hours. Attempting to simulate longer time intervals with the circuit model becomes impractical. Before proceeding to treat synaptic inputs to dendrites, we note an extension of the model to dendrites with additional circuit complexity.

The response function of the two-JJ dendrite is useful for many operations, but in certain cases it is desirable for the value of the saturation current to be less sensitive to the applied flux. This can be accomplished by adding a Josephson transmission line between the receiving and integrating

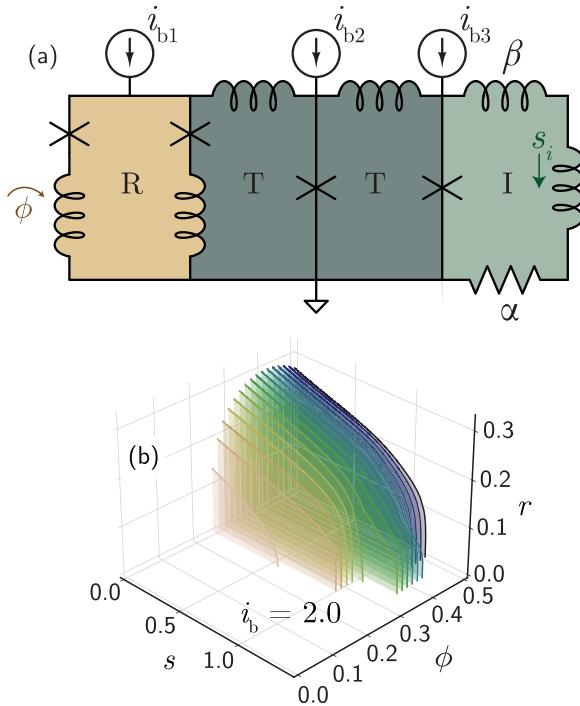


FIG. 11. The RTTI dendrite. (a) Circuit diagram showing the addition of two transmission loops. (b) An example r -shell for the RTTI dendrite.

loops of the dendrite, as shown in Fig. 11(a). Due to the presence of two transfer loops between the receiving and integrating loops, we refer to this as the RTTI dendrite (receive-transfer-transfer-integrate). An example rate array for the RTTI dendrite is shown in Fig. 11(b). While it is similar to the RI dendrite in that it has a threshold of applied flux and decreases with integrated s , the overall shape is more complicated, with appreciable structure close to saturation. The details of this structure can be used for various computations. In particular, the fact that the response is relatively flat for certain ranges of ϕ can be used to achieve a digital response in which the output is insensitive to the exact input value of ϕ . This attribute will be exploited below to achieve the OR gate. As for the case of the RI dendrite, we extract the thresholds and saturation values from the rate arrays. These functions are shown in Fig. 12. In Fig. 12(a), it can be observed that the thresholds for positive and negative applied flux are now slightly asymmetric due to the breaking of the symmetry of the circuit with the low-inductance transfer loop to the right of the receiving loop. More importantly, the values of $s_{\max}^{+/-}$ shown in Fig. 12(b) are seen to have several regions of nearly flat response where the saturation level of the dendritic integration loop does not depend on applied flux. Such a response is useful for obtaining digital behaviors, as we show below when discussing logic gates in Sec. IV. The peak value of s_{\max}^+ as a function of i_b is shown in the inset to Fig. 12(a).

To simulate the RTTI dendrite with a conventional model, nine coupled ODEs are used, making the simulations even slower than the RI dendrite with its five coupled equations. With the phenomenological model, the RTTI dendrite reduces to a single ODE, just as the RI dendrite. The only difference is the form of the r -shells that provide the driving term.

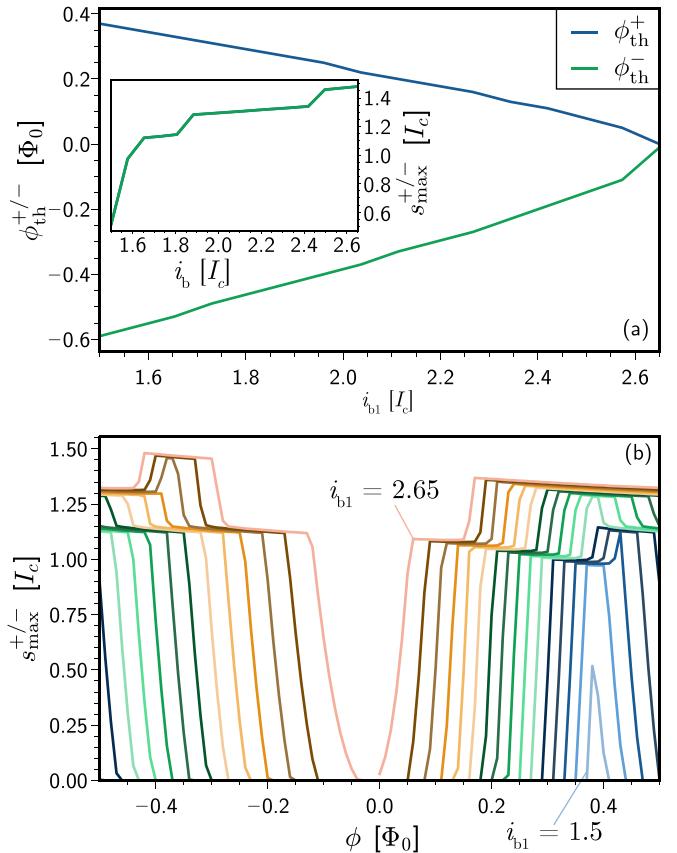


FIG. 12. The values of flux thresholds (a) and saturation levels (b) for the RTTI dendrite extracted from the rate arrays, analogous to these quantities shown in Fig. 5 for the RI dendrite.

IV. SYNAPSES

As described in Secs. II and III, the dendritic receiving loop is a SQUID. In many applications in science and technology, a SQUID is used as a measurement device with unmatched sensitivity for detecting magnetic flux. Based on the response curves of Fig. 3(b), it can be seen that a SQUID can be used as a flux-to-voltage transducer. Such a device can also be used to measure very low current levels when the current is coupled into the SQUID as flux using a transformer. It is in this mode of operation that we convert a dendrite into a synapse.

When an SPD detects a photon, a current pulse is diverted out of the circuit, and that current returns with an L/R time of around 35 ns, which sets the recovery time and maximum rate of synapse events (≈ 30 MHz). To couple an SPD into a dendrite to form a synapse, this current output can be coupled into the SQUID that forms the receiving loop through a transformer, as shown in Fig. 2(c). The equations of motion for the SPD can be solved exactly under a simple model [33] to obtain the currents as a function of time. In this model, the resistance of the SPD is zero until a photon is detected, at which point it switches to a finite resistance for a finite duration. We are interested in the flux applied to a dendrite by the SPD receiver circuit; this flux is the product of the current I_2 with the mutual inductance of the transformer, which is a quantity we can select in design. Therefore obtaining the currents is sufficient

to obtain the flux, which we write as

$$\phi_r = \phi_{\text{peak}} \left(1 - \frac{\tau_{\text{rise}}}{\tau_{\text{fall}}} \right) \times \begin{cases} [1 - e^{-t/\tau_{\text{rise}}}], & \text{for } 0 \leq t \leq t_0, \\ [1 - e^{-t_0/\tau_{\text{rise}}}] e^{-(t-t_0)/\tau_{\text{fall}}}, & \text{for } t > t_0. \end{cases} \quad (6)$$

In Eq. (6), it is assumed a synapse event occurred at time $t = 0$, which is the time at which the SPD entered the resistive state, and t_0 is the duration for which the SPD stays in the resistive state following the absorption of a photon, which we take to be 200 ps, following the model of Ref. [33]. This duration is commensurate with the 200 ps time step that we found to be optimal in Sec. III. The resistance r_1 [Fig. 2(c)] is the resistance of the SPD in the time window t_0 after absorbing a photon, while r_2 is a fixed resistance chosen to obtain a sufficiently long L/R recovery time to prevent the SPD from latching [34]. The total inductance of the circuit is $L_{\text{tot}} = L_1 + L_2 + L_3$, which includes the kinetic inductance of the SPD itself (L_1), the inductance of the transformer input coil (L_3), and any additional parasitic inductance in the circuit L_2 . In terms of these parameters, $\tau_{\text{rise}} = L_{\text{tot}}/(r_1 + r_2)$ and $\tau_{\text{fall}} = L_{\text{tot}}/r_2$. Typically τ_{rise} is a few tens of picoseconds and τ_{fall} is 30–50 ns [24]. The quantity ϕ_{peak} can be chosen in design and set by the transformer between the SPD and the dendrite. Throughout this work we take $\phi_{\text{peak}} = 0.5$ in units of Φ_0 .

We now consider several examples of the phenomenological model with the output of an SPD coupled as input to a dendrite. In Fig. 13, we show the response of the dendrite to a single synaptic pulse. The time course of the total applied flux to the receiving loop, ϕ , is shown in Fig. 13(a), and four values of β are shown in Figs. 13(b)–13(e), each for four values of τ_{di} . In each case, the dashed line corresponds to using a time step of 1 ns in the phenomenological model, while the solid line corresponds to 100 ps. Again we see that the worst performance occurs with small β and small τ_{di} . Similar data is shown in Fig. 14 for an arbitrary sequence of five synapse events. Dendritic integration is evident, and the effects of β and τ_{di} can be seen: for smaller β and larger τ_{di} , the integration loop reaches saturation after a small number of input pulses, while for larger β and/or smaller τ_{di} , the integrated signal does not reach saturation even by the end of the train of five pulses.

Using synapses and dendrites to perform various temporal extensions of basic logic operations is a primary function of biological neural computation. Here we consider four basic logic operations: AND, OR, AND-NOT, and XOR. Each of these operations is accomplished with a single dendrite receiving input from two synapses. The AND operation is shown in Fig. 15. The flux from the two synapses independently is shown in Fig. 15(a), their sum is shown in Fig. 15(b), and the signal in the dendritic integration loop is shown in Fig. 15(c). In Figs. 15–18, the flux from a single SPD to the dendrite is referred to as ϕ_{spd} , while the combined flux from both SPDs to the receiving loop of the dendrite is labeled ϕ . To achieve the AND response, the dendrite is configured such that the flux from a single synapse is insufficient to drive the dendrite above threshold, but the flux from two synapses together is adequate, and signal is added to the integration loop only when the pulses are coincident.

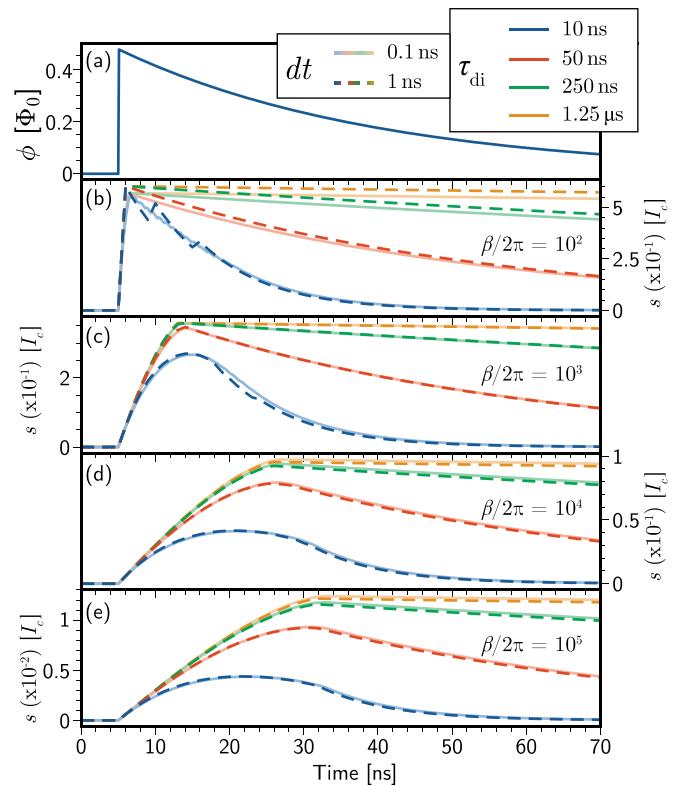


FIG. 13. Output of the phenomenological model for the case of a synapse. (a) Flux output from the SPD. (b) Signal s in the dendritic integration loop for the case of $\beta/2\pi = 10^2$. The solid lines result from a time step of 100 ps, while the dashed lines used 1 ns. [(c)–(e)] $\beta/2\pi = 10^3, 10^4$, and 10^5 , respectively.

Because the synaptic flux and the dendritic signal are both functions of time, it is important to know how the accumulated signal will vary with the delay between the arrival of the two synapse events. These data are shown in Fig. 15(d) for multiple values of the bias to the dendrite, i_b . Such a circuit can be used as a temporal coincidence detector, as described in Ref. [25]. The OR operation is shown in Fig. 16. The challenge with OR is that we would like the amplitude of the signal in the dendritic integration loop to be identical whether one or both synapses are active. Because both the rate of fluxon production and the value of saturation depend on the applied flux for the RI dendrite, this is difficult to accomplish. However, the flat response of the RTI dendrite can generate this input-output relationship, and that type of dendrite has been used to generate Fig. 16. In Fig. 16(b), we can see that when both synapses are active the applied flux to the receiving loop has twice the amplitude as when only a single synapse is active, yet the integrated signal in Fig. 16(c) is nearly identical whether one or both synapses are active. The response is possible due to the flat-top behavior of the RTI dendrite. We have biased the dendrite at a point such that when the flux from one synapse is present the dendrite just enters the plateau of the response, and when the flux from both is present it remains on the plateau.

The response as a function of the delay between the two synapse events is shown in Fig. 16(d). For lower bias points the flat-top response only occurs for a relatively short

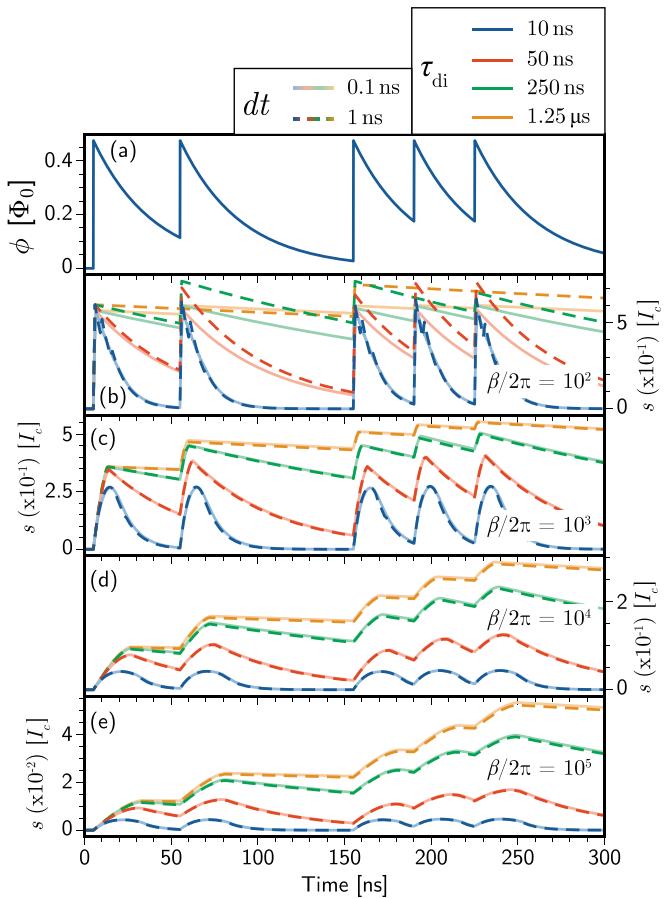


FIG. 14. Similar data to Fig. 13 except signals are in response to a train of five input synapse events.

coincidence window, providing a response more like the AND gate. Yet for higher bias points the response is very flat, independent of temporal delay between the synapse events, just as we desire for the OR gate. We will see in Sec. V that this response is useful in achieving the desired classification task. The AND-NOT operation is shown in Fig. 17. This logic gate produces a logical one at the output if and only if the first input is a logical one and the second input is a logical zero. This is distinct from the NAND gate. For AND-NOT, we return to the RI dendrite. In this case, synapse one is excitatory, while synapse two is inhibitory with the opposite sign of flux applied as well as a reduced amplitude. Figure 17(c) shows that the dendrite only becomes active when synapse one receives a pulse in the absence of synapse two receiving a pulse. Extension to various temporal delays in Fig. 17(d) shows the effect of the asymmetric coupling of the two synapses. The XOR operation is shown in Fig. 18. XOR is accomplished in much the same manner as AND-NOT, except the coupling between the two synapses is equal. Thus, when either synapse is active in the absence of the other the dendrite becomes active. In this instance, we are making use of the fact that the RI dendrite response is symmetric with respect to positive or negative flux. The response as a function of relative delay between the synapse events is shown in Fig. 18(d), and it is evident that cancellation of the two pulses requires relatively high timing precision. The response

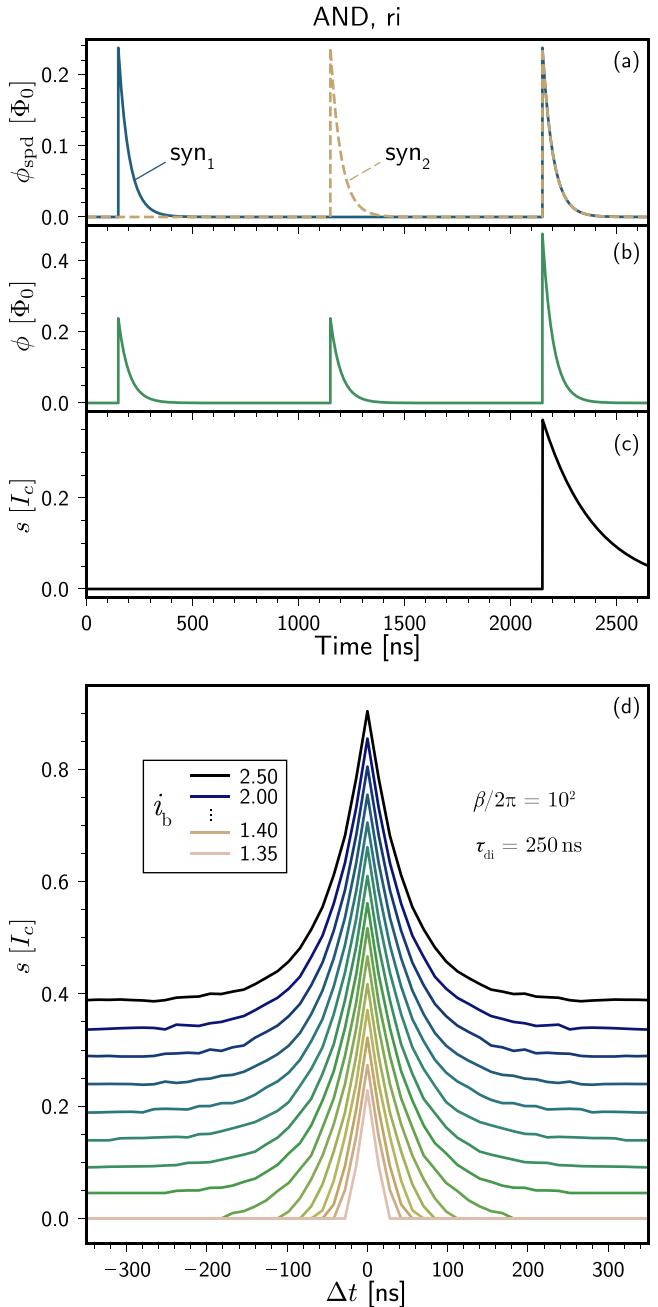


FIG. 15. Operation of the AND gate. (a) Input flux from the two synapses separately. (b) Combined input flux. (c) Dendrite output signal. (d) Analysis of the gate response as a function of time delay between the two synapse events for several values of i_b .

over longer delays is shown as an inset. The requisite timing precision can be adjusted with the L/R time of the SPD. This implementation of XOR in a single dendrite is analogous to that observed in human pyramidal neurons in cortical layers two and three [35]. In addition to discrete responses to few synapse events, the phenomenological model can help us understand the transfer functions performed by synapses coupled to dendrites when many pulses are incident.

In Fig. 19, we analyze these responses for various bias conditions, revealing the role of i_b for sculpting dendritic behavior. Figure 19(a) shows the peak of the accumulated

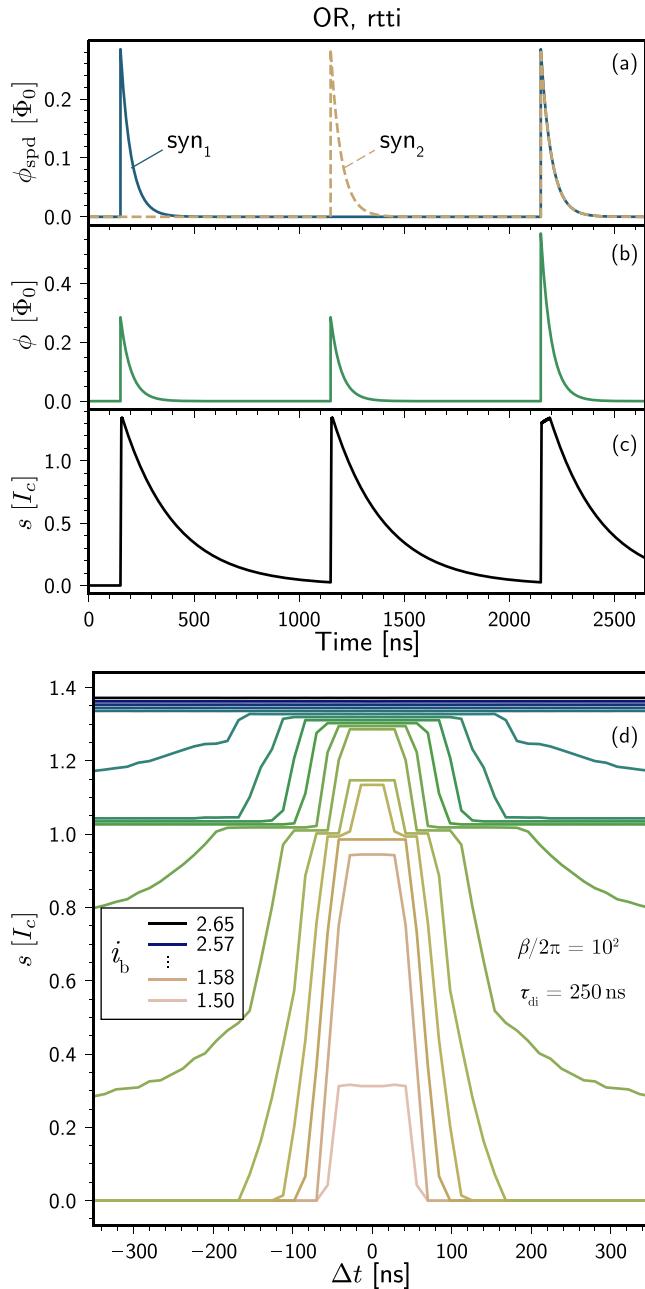


FIG. 16. Operation of the OR gate. (a) Input flux from the two synapses separately. (b) Combined input flux. (c) Dendrite output signal. (d) Analysis of the gate response as a function of time delay between the two synapse events for several values of i_b .

signal in the dendritic integration loop as a function of the number of synapse events input in a burst. Traces from the same color family correspond to the same value of i_b . Lighter traces correspond to pulses input at a slower rate relative to the time constant of the dendrite, while darker traces show the response for higher-frequency inputs. The response is unsurprisingly larger for higher input frequencies and can be broadly adjusted with the bias current. The sigmoidlike response with threshold and saturation has been useful in various machine learning applications. Similar responses are seen in Fig. 19(b) where the frequency of the burst is now

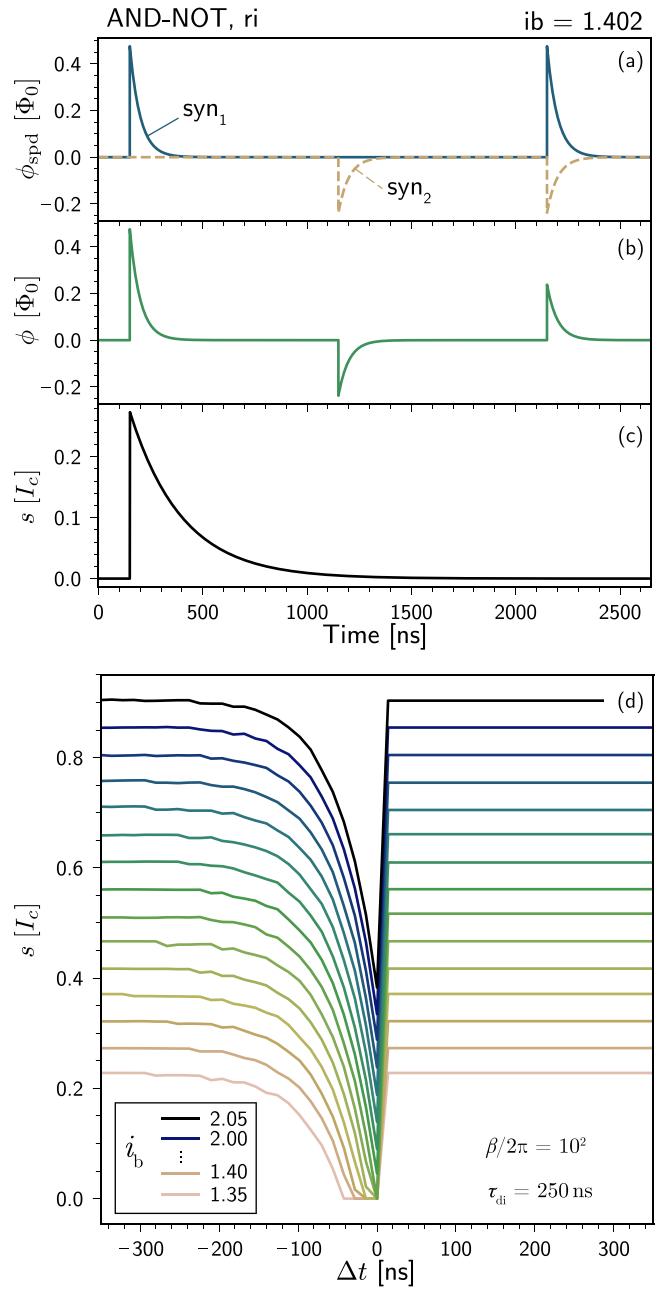


FIG. 17. Operation of the AND-NOT gate. (a) Input flux from the two synapses separately. (b) Combined input flux. (c) Dendrite output signal. (d) Analysis of the gate response as a function of time delay between the two synapse events for several values of i_b .

the quasicontinuous variable on the x axis, and various curves have different numbers of input pulses. Similar thresholding and saturation behavior is evident, and the same qualitative behavior has been demonstrated in experiments of related circuits [24]. If generated with conventional circuit models, the plots of Fig. 19 would have required so much computational time as to be impracticable with the computer used for this study. With the phenomenological model they are produced in minutes so that many scenarios may be investigated and insight across parameter space can be quickly gained.

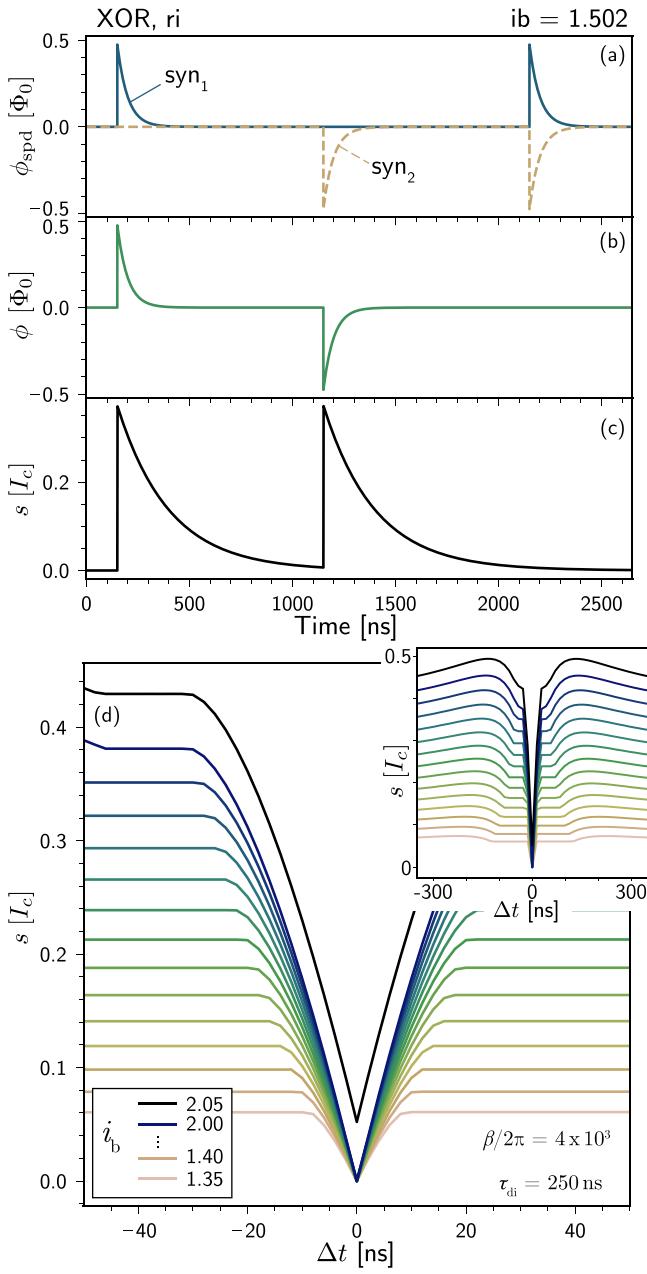


FIG. 18. Operation of the XOR gate. (a) Input flux from the two synapses separately. (b) Combined input flux. (c) Dendrite output signal. (d) Analysis of the gate response as a function of time delay between the two synapse events for several values of i_b . Here the time axis is narrower than Figs. 15–17. The inset shows the full time axis of the other logic gate figures.

V. NEURONS

Whereas a synapse was shown to be formed where a single-photon detector couples to a dendrite, a loop neuron can be as simple as a dendrite with a transmitter circuit providing output. Such a soma circuit is shown in Fig. 2(d). While all dendrites have a bias-dependent threshold for activation in the receiving loop below which applied flux elicits no response, the soma is distinguished in that it also has a threshold in the integration loop at its output. Embedded in the integration loop is a superconducting thresholding element referred to

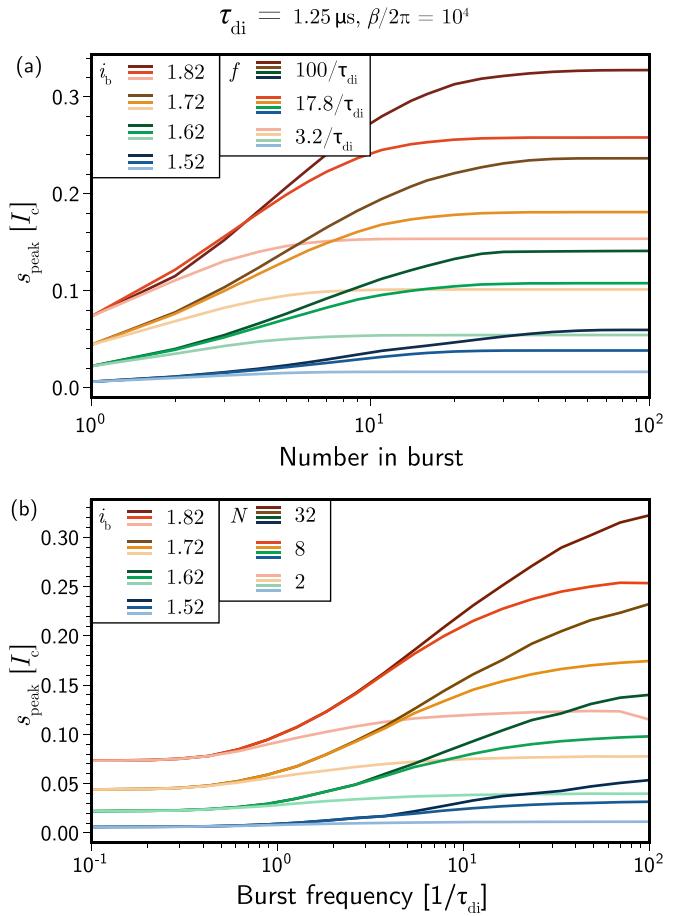


FIG. 19. Transfer functions for the RI dendrite. (a) Peak of integrated signal as a function of the number of synapse events in a burst for several values of i_b and frequency of burst inputs. (b) Peak of integrated signal as a function of the frequency of input pulse train for several values of i_b and number of pulses in the train.

as a tron that switches from a zero-resistance state below threshold to a high-resistance state when the current in the integration loop reaches a critical value. Such comparators have been demonstrated as an interface between superconductor and semiconductor electronics [23,36]. In the present case, accumulation of sufficient signal in the soma's integration loop will drive the tron to the voltage state, at which point a semiconductor-based transmitter circuit will generate photons from a light-emitting diode (LED). This is a spike event or action potential of a superconducting optoelectronic loop neuron. When this threshold is reached and the tron switches, the current in the somatic integration loop is purged, and integration must begin again. In most dendrites in the system, any accumulated signal in the integration loop is immediately coupled as flux into other dendrites. However, the soma has no output other than the transmitter circuit, so its state is only communicated to other elements of the network when threshold is reached and light is produced.

Several approaches to transmitter circuits for loop neurons have been explored [20], and here we base the phenomenological model on a circuit concept that leverages MOSFETs in an essentially digital configuration. Upon reaching threshold, activity is transferred to these semiconductor circuits,

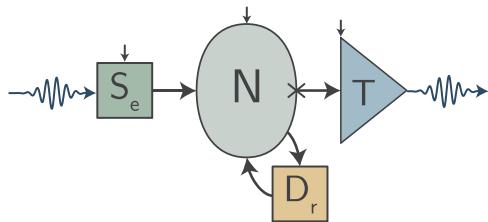


FIG. 20. Schematic of a monosynaptic point neuron with excitatory synapse (S_e), neuron cell body (N), transmitter (T), and refractory dendrite (D_r).

and to understand their operation we must work backward from the light source itself. To maintain “brevity” in the main thread of this article, we relegate the details to Appendices. Appendix E describes a rate-equation model to capture the primary behavior of the light emitters, while Appendix F treats the transmitter circuit from the thresholding element through the LED. Because the subject of this work is to establish a phenomenological model, here we simply extract the key message from those Appendices: when a soma reaches threshold, photons are produced with a few nanosecond delay followed by an exponential probability distribution. The distribution is obtained numerically through simulation of the transmitter circuit combined with the source rate equation model. In the reduced phenomenological model that is the subject of this work, each time a somatic integration current reaches threshold, a specified number of photon-production times are drawn randomly from this distribution of a delay followed by an exponential decay. These photon times are then assigned randomly to the neuron’s downstream synapses and added to a list of times of input synapse events. With this approach, the number of ODEs which must be treated remains unchanged, and the added computational burden of treating the transmitter circuit and sources is reduced to dealing with random number generation only on the time steps at which the soma’s threshold is reached. The indefatigable reader is directed to Appendices E and F for more details.

In addition to this output thresholding behavior, a loop neuron requires a means of establishing a refractory period so that spike events are indeed discrete and their rate does not exceed the roughly 30 MHz at which the synaptic SPDs can respond. A refractory period is accomplished by adding a second dendrite to the soma. A schematic of this minimal point neuron is shown in Fig. 20, where the refractory dendrite is labeled D_r . This refractory dendrite is driven by a flux pulse identical in form to the SPD response [Eq. (6)] when the soma reaches threshold, and the signal accumulated in the refractory integration loop is coupled back to the soma’s receiving loop as inhibition. The time constant of the refractory dendrite’s integration loop therefore sets the refractory period and maximum neuronal firing rate. To achieve robust refraction, the coupling from the refractory dendrite back to the soma is designed so that the inhibitory flux is as strong as possible without driving activity on the negative-flux side of the r -shell. This can be accomplished with knowledge of the values of s_{\max} for the refractory dendrite at its bias point as well as the values of threshold flux ϕ_{th}^+ and ϕ_{th}^- . The refractory dendrite is chosen to have a relatively small value of β so that it saturates quickly at s_{\max} each time the soma fires. The

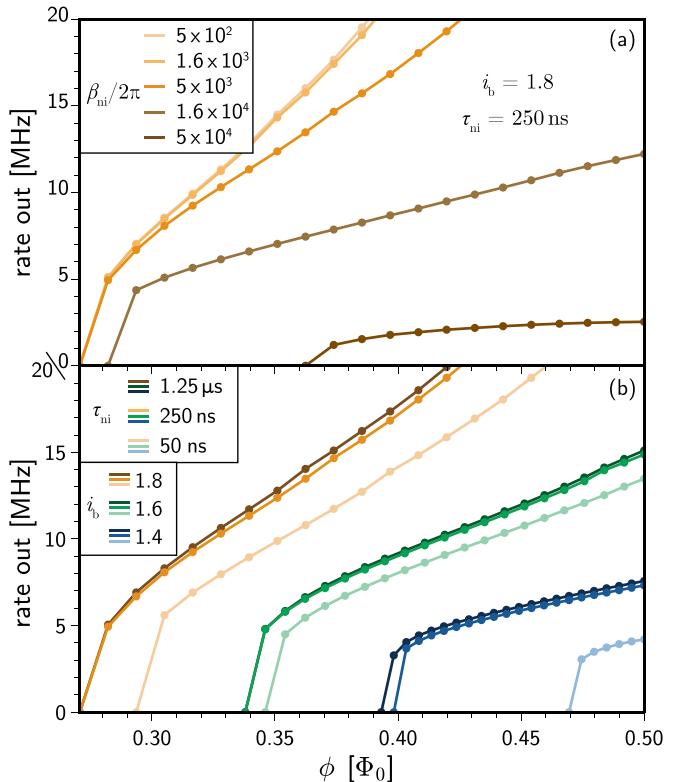


FIG. 21. Neuron firing rate versus applied flux. (a) Comparing several values of the neuronal integration loop inductance parameter, $\beta_{\text{ni}}/2\pi$. (b) Comparing several values of neuronal integration loop time constant and bias current. For $i_b = 1.6$ and 1.8 , $\beta_{\text{ni}}/2\pi = 5 \times 10^3$. For $i_b = 1.4$, $\beta_{\text{ni}}/2\pi = 1.6 \times 10^4$.

coupling from the refractory to the somatic dendrite is then given by $J = (\phi_{\text{th}}^+ - \phi_{\text{th}}^-)/s_{\max}$.

The role of the refractory dendrite is to inhibit the neuron for a brief duration after the neuron spikes. Yet similar inhibition in response to activity is often desirable on longer timescales. Homeostatic plasticity [37–39] is one means by which the activity of neurons can be maintained in a useful dynamic range. This self-regulatory behavior can be accomplished with feedback analogous to refraction if further dendrites receive input when the soma fires and couple inhibitively to the soma’s receiving loop. Such a homeostatic dendrite is likely to have a longer time constant and also a larger integration loop inductance than the refractory dendrite. In this way, each time the neuron fires, a small inhibitory signal is fed back to the soma, effectively increasing the threshold for the next firing. This signal will decay over a longer time period, on the order of many interspike intervals. Such a plasticity mechanism can be treated with the same phenomenological model presented here.

To begin analysis of neurons with this framework, we consider a neuron with no dendritic tree in which all inputs are received directly at the soma. Such a neuron is referred to as a point neuron. As a first means of gaining intuition about the design of a point neuron, we consider the spiking behavior of a simple loop neuron with a steady-state flux coupled directly into the soma and a single refractory dendrite to induce pulsatile behavior at the transmitter. In Fig. 21, we

show the rate of neuron spiking as a function of the steady-state applied flux ϕ . Figure 21(a) shows the response for five values of the neuronal integration loop inductance parameter, β_{ni} . In all cases, the time constant of the refractory integration loop was 50 ns and the value of the somatic integration loop threshold was set to 0.3 s_{max} . Smaller values of β_{ni} lead to rapidly refilling of the integrated signal after it is purged upon reaching threshold, while larger values require more time to accumulate signal back to threshold, leading to slower firing activity for the same flux drive.

By contrast, in Fig. 21(b), the somatic integration loop inductance was fixed at an intermediate value, while the operating bias point and somatic integration loop time constant were adjusted. Lower biasing requires higher input flux to achieve threshold and also leads to slower firing rates, while higher bias reduces threshold and increases the output rate. Such a control parameter is the primary means by which the soma can be reconfigured dynamically during network operation, either in an unsupervised manner by activity within the system or in a supervised manner by a user or control system. The time constant cannot be changed dynamically, as it is set by the resistance and inductance of the loop, which are fixed in fabrication. However, the time constant can have an appreciable effect on the transfer characteristics, as is most pronounced for the low-bias case where reducing the time constant to 50 ns dramatically increases the flux threshold and reduces the output rate at a given value of applied flux. All of these parameters must be considered when designing neurons to play various roles in a network.

As in the transfer functions of Fig. 19, the responses in Fig. 21 would have been quite cumbersome to obtain with the conventional circuit equation approach. Even at this basic level, the phenomenological model provides utility in guiding the design principles of loop neurons across a broad parameter space.

We now consider an example of employing the phenomenological model to treat a loop neuron with a more elaborate dendritic arbor to solve the standard benchmark problem of the nine-pixel image classifier [40]. Figure 22 shows the nine-pixel input and the manner in which the pixels can be argued to depict three letters: “z”, “v”, and “n”. This benchmark problem was chosen to illustrate the use of some of the machinery developed in this paper, including the two-input dendrites. This nine-pixel classification task can be readily handled with digital systems employing Boolean operations, so it is not meant to illustrate the supremacy of neural systems for this type of computation. Consideration of more complex problems exploiting the analog circuits, temporal dynamics, and highly interconnected networks of superconducting optoelectronic neurons will be the subject of future work.

This nine-pixel classification problem is typically formulated for small neural networks with the goal of training the network to identify the image as the correct letter even when the state of any one pixel is allowed to switch. The problem is often solved with small neural networks trained with conventional techniques such as backpropagation. Here we solve the problem using a single neuron with a small dendritic tree, showing that such a tree can serve a similar role to a feed-forward neural network [41] and to illustrate how even

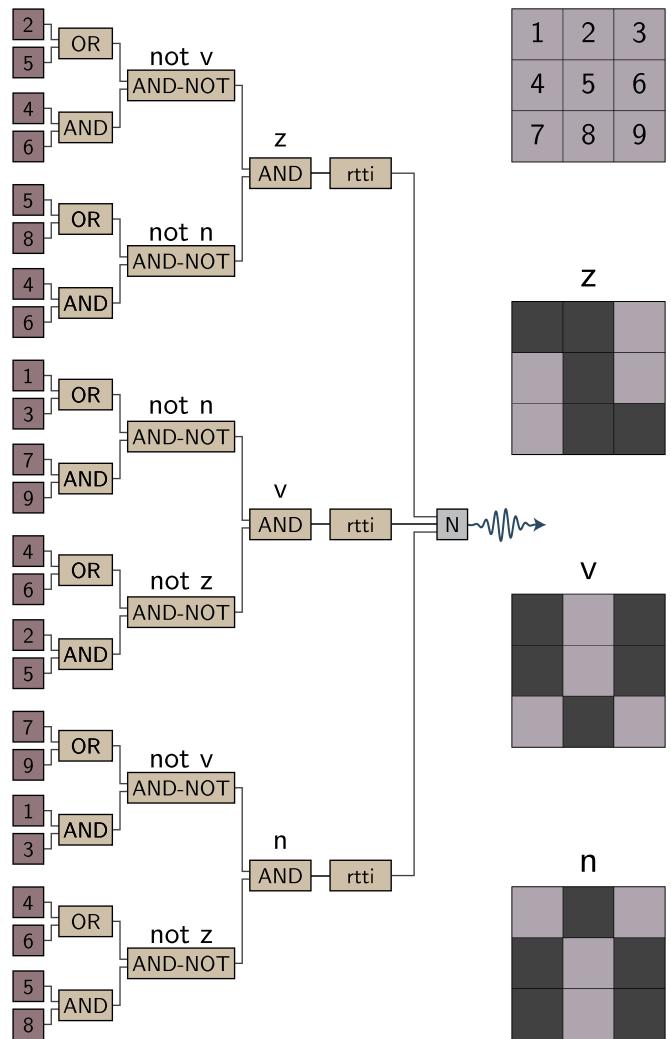


FIG. 22. Schematic of the nine pixel classifier. The dendritic tree is illustrated as are the three classes of inputs, z , v , and n .

relatively simple loop neurons become sophisticated computational processors, as inspired by their biological counterparts [42]. Also in contrast to published approaches [40,43], we achieve the desired neural operation through inspection without a training algorithm, meaning we analyze the active pixels in each of the three letters, including the cases with one-pixel variation, and write down the logic that allows identification of the letter based on two-pixel gates. This approach to completion of the task is in the spirit of biological vision systems that “hard-code” basic computational primitives, such as Gabor filters matched to common spatial frequencies in natural scenes [44]. While this approach of solution by inspection is sufficient for this small problem, more powerful learning techniques are required to train spiking neural networks in general. We discuss some of these techniques in Sec. VI.

The arbor that solves the problem is shown in Fig. 22. Each pixel is input to a synapse as a single spike event: a spike event occurs at synapse i if pixel i is active in the image under consideration, and no spike event is input at that synapse if the corresponding pixel is not active. The various dendrites are labeled by the logic operations they perform (Sec. IV). The

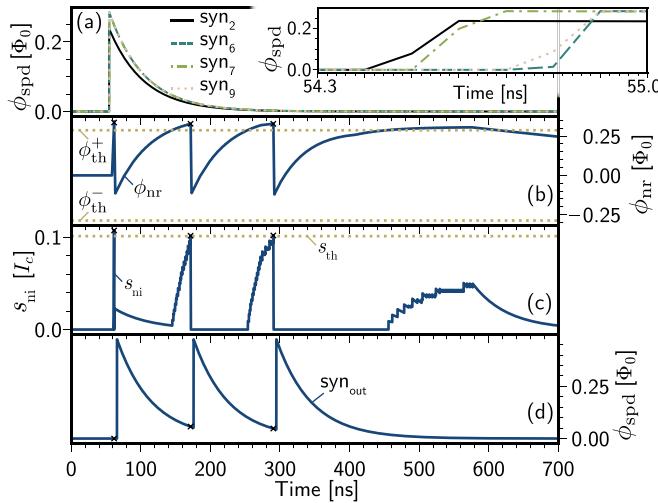


FIG. 23. Time traces for the nine-pixel classifier neuron for the case of n_4 , the image of n with pixel four switched. (a) The input synaptic flux from all four active synapses. The inset shows higher temporal resolution to display the timing jitter due to the model of the quantum-dot light source. (b) The total applied flux to the neuronal receiving loop, including the refractory contribution. (c) The integrated current in the neuronal integration loop. (d) The flux generated by the downstream SPD receiving photons upon neuronal firing.

knowledge accessible at various stages of the tree is written above the dendrites whose signal represents that information. The basic reasoning that solves the problem can be followed by stepping through the tree. For example, starting at the top left, if pixel two or five is active but pixels four and six are not both active, the letter cannot be v . Similar reasoning can be applied to rule out n . If it is not v and it is not n , it must be z . At the last stage of the tree before the soma, three different dendrites know whether their letter is present. By assigning these three dendrites different time constants [α parameters in Eq. (2)], they each evoke different numbers of pulses from the neuron. An output of one spike event from the neuron informs us a z has been presented, two correspond to v , and three to n . The neuron under consideration is shown to accomplish the task even in the presence of 1 ns timing jitter due to the transmitter circuit and quantum dot light source (Appendices E and F). Time traces from a single classification instance (n_4) are shown in Fig. 23. The four input synapse events are shown in Fig. 23(a), and the timing jitter due to the transmitter circuit and light source is evident in the inset. The total flux input to the neuronal receiving loop is shown in Fig. 23(b), as are the thresholds for activity due to positive and negative flux. The accumulated signal in the neuronal integration loop is shown in Fig. 23(c), along with the value of the neuronal integration loop threshold. The times of neuronal spike events are shown as black crosses. In this model, the neuron is given one output synapse, and the flux present at the output of that synapse due to photon detection events from neuronal firing are shown in Fig. 23(d). The full set of 30 inputs is shown in Fig. 31 in Appendix G. The results of all instances of the classification task are shown in Fig. 24. The maximum level of integrated signal in all three final dendrites is shown in Fig. 24(a) for all thirty presented images. The number of spikes produced

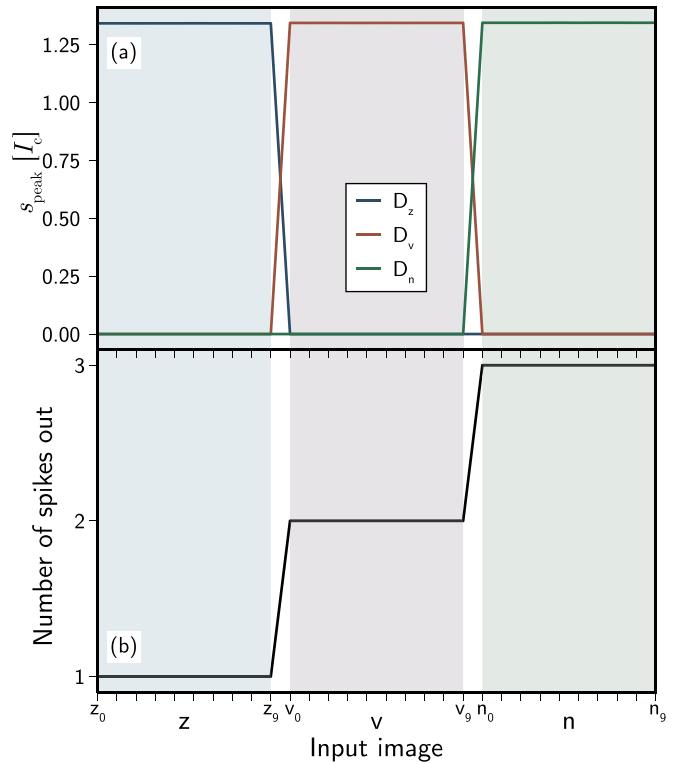


FIG. 24. Output of the nine-pixel classifier neuron versus input image. (a) The peak of the integrated signal in the final dendrites that represent the three letters. (b) The number of spikes out as a function of the input image.

by the neuron is shown in Fig. 24(b). It can be seen that no signal is generated in any of the dendrites except the one corresponding to the presented letter, even in the presence of one-pixel noise, and the neuron correctly produces the number of spike outputs to communicate the result of its calculation to other neurons that may be present in the network.

VI. DISCUSSION

We have seen that the computational circuitry of loop neurons consists of a network of interacting dendrites. Within this framework, a dendrite is a SQUID with output coupled to a current-storage loop. The bias point of the SQUID determines the threshold input flux required to initiate activity, and the storage capacity of the output loop provides a saturating nonlinearity. The various circuit parameters provide means to adjust the response characteristics across broad operating ranges. Learning and homeostatic adaptation can be accomplished with similar circuits providing feedback, dynamically adjusting bias points through coupled flux that can be stored perpetually and adjusted in small or large increments. The final stage of dendritic processing culminates in the soma, the neuron cell body. When this final dendrite reaches threshold, something different happens with regard to physical hardware: an amplifier drives a semiconductor light source. The threshold signal is the final stage of the computational process; the production of light is a binary action potential, and the physical transduction to photons is chosen to enable communication to many destinations across length scales that

result in wiring parasitics that render electronic signals impracticable. A neuron is a network of coupled dendrites with local electronic communication. A spiking neural network extends interconnectivity to longer distances with optical communication.

Simulating loop neurons with the phenomenological model brings significant speed increases relative to the conventional method of simulating the first-principles circuit model. The speed of simulation can increase by a factor of ten thousand while retaining a χ^2 of 10^{-4} when simulating a single dendrite for $12\ \mu\text{s}$. When simulating a large network, the circuit-model approach is impracticable due to the time required to run calculations over even a relatively short duration of network activity. The phenomenological model thus enables design of more components, exploration of a wider range of parameter space, and treatment of larger networks than would otherwise be possible. Still, simulating $12\ \mu\text{s}$ of activity from a single dendrite requires around $1\ \text{s}$ of CPU time. A physically instantiated network on a small chip (roughly $1\ \text{cm} \times 1\ \text{cm}$) would comprise roughly 10 000 neurons and half a million synapses and dendrites. Linearly extrapolating the time requirements, it would take over a year to simulate $1\ \text{ms}$ of activity of this network on the same workstation used in this work. Two conclusions can be drawn: (1) the phenomenological model must be used as a tool to facilitate further increases in speed of design and (2) constructing the physical hardware of loop neurons could bring a computational platform of immense power and utility. Regarding the first point, by using the phenomenological model to design a variety of dendrites, coupled-dendrite computational blocks, and a variety of neurons, libraries of optimized components can be assembled. The model can be used to efficiently design these components and reduce their behavior to stored transfer functions to achieve a further level of abstraction that can be used to model behavior of larger networks without explicitly stepping through the ODEs, as was done for the logic gates, burst, and rate transfer functions in Sec. III. Regarding the second point, it would take a year on a workstation with a $3.7\ \text{GHz}$ processor to simulate one millisecond of a network that would fit on a chip of similar size and power consumption to the processor, including cryogenic cooling. The increase in computational speed that would result from realizing SOEN hardware would be 32 billion for even a modest network. SOENs therefore appear to offer a path to neuromorphic supercomputing.

If the aspiration of loop neurons is to constitute systems of exceptional scale and complexity, might the simplicity of the dendritic building block limit the dynamical repertoire? We hope this model helps answer this question. It is known that even simple systems with simple rules for propagation, such as cellular automata, can give rise to behavior of great sophistication [45,46]. It has been argued that similar concepts can be applied as a starting point for physics, with the rich, natural world emerging based on the interactions of fundamental nodes [47]. Similarly, elements as simple as two-state spins interacting with nearest neighbors (Ising models) can give rise to phase transitions and critical phenomena, including crucial long-range correlations [48,49] as well as attractor dynamics for associative memory storage and retrieval [50–53]. By comparison, the dendrites studied here are

multidimensional and nuanced. The nonlinear spatiotemporal convolutions occurring in each neuron's synaptodendritic tree provide a deep repository that can inform the neuron's behavior. Transmission of action potentials to many destinations at light speed enables complex network topologies far beyond nearest-neighbor interactions. The use of superconducting circuits allows signal retention across a broad range of timescales through the choice of the τ_{di} parameters (the τ distribution) that specify the leak rates in Eq. (2), as has been demonstrated experimentally [24]. The range of integration loop inductances (the β distribution) in conjunction with the bias currents will establish the dynamic range of the network. With near and distant connectivity as well as a wide spectrum of dissipation times and an enormous dynamic range, SOENs are likely to achieve the long-range spatiotemporal correlations present in the critical states that optimize neural information processing [54–57]. The hierarchy of spatial connectivity in conjunction with the hierarchy of information retention times and response magnitudes appears excellent for enabling the fractal use of space and time that supports information integration and cognition [58–63]. Yet the simplicity of constructing the majority of computational grey matter from similar building-block components brings an advantage in modeling and technological implementation.

The significance of dendrites in neural computation is well documented [41,64–69], yet the incorporation in neuromorphic hardware has received proportionally less attention [70–74]. The dendrites studied here are capable of achieving many of the computational functions known from neuroscience, and they may also be capable of implementing numerous learning functions. As mentioned above in the context of the nine-pixel classification problem, general techniques are required to train spiking neural networks to solve diverse computational problems. Superconducting optoelectronic neurons and networks appear equipped to implement many of the training algorithms introduced for spiking neural networks in the literature [75–78], yet we have a keen focus on algorithms employing dendrites, as these allow simple circuit implementation, adaptation with only local information, and are commensurate with the concepts of reinforcement learning known to be employed by the brain [79]. In particular, dendrites may provide valuable means for implementing credit assignment in training algorithms that utilize spike-timing signals in conjunction with population activity [80–84], potentially leveraging much of what is known from neuroscience about the key role of dendrites in learning [85–92] and overcoming a major obstacle to widespread adoption of spiking neural networks for artificial intelligence. Further, the presence of dendritic signals that are continuous in time without erasure following a spike may provide new methods for training spiking neural networks that are not available based on spiking activity alone. Perhaps these continuous signals will be useful for constructing cost functions and training networks with variants of the backpropagation technique, a feat which has been difficult in the case of spiking neurons outside of the rate-coding domain, and a subject which is significant in bridging machine learning and neuroscience [93].

While the model presented here is interesting to explore, it is intended as a tool in a larger project. The objective of such research at this stage is to determine if SOENs are

indeed as promising for future study and worthy of appreciable investment as our hypothesis contends. We hope this model will help answer several questions: do the neuromorphic circuit principles demonstrated in these superconducting optoelectronic embodiments bestow systems with the dynamic interplay of structure and function that scaffolds successful neural systems? Do device and circuit features such as the achievable breadth of the τ and β distributions, the particular dendritic nonlinearities, and the available forms of plasticity bring benefits in support of cognition? Does refraction of the neuron without erasure of information in the dendritic tree offer advantages in information processing, lead to new forms of neural coding, or offer new means of training? How should the structure of the dendritic trees and the structure of the network be co-designed to achieve desired operations? The objective of the model presented here is to enable computational studies that answer these questions.

With the phenomenological model as a guide, we can compare and contrast loop neurons with other complex systems. Spin systems have received considerable attention over decades due to both their simplicity as well as the emergence of interesting phenomena such as phase transitions and spontaneous symmetry breaking. Common spin models include the Ising model, where each spin s takes a binary, scalar value of -1 or 1 . Coupling is typically between nearest neighbors and is also a scalar value, J_{ij} , typically with the symmetry $J_{ij} = J_{ji}$. The $x - y$ model and the Heisenberg model extend the spin to have two or three vector components, maintaining $|s| = 1$. In the present context, we can think of the current stored in the integration loop of dendrite i , s_i , as a spin. It is more like an Ising model than an x - y or Heisenberg in that it is a scalar value, but in general it is analog instead of binary and can take a large number of values between zero and one. Coupling is through the scalar quantities J_{ij} [Eq. (3)] and is manifest through transformers. In this case, there is no requirement that $J_{ij} = J_{ji}$, and the spatial limitations regarding which spins can couple are relaxed. Coupling is mediated by a transformer circuit rather than an exchange interaction, so flexibility in the adjacency matrix is available. The values of J_{ij} are fixed in time, but the sign can be positive or negative so that excitation or inhibition are possible. While the J_{ij} are fixed in circuit fabrication, the interaction between any two spins is dynamically reconfigurable either by an external influence (experimenter) or through network activity (plasticity) by changing applied bias currents to change the r -shell on which the dendrite resides. Furthermore, the magnitude of the spins in loop neurons is not fixed in time, but is a dynamical quantity that grows through interactions and decays passively, following the leaky integrator formalism. In general, the threshold for one spin to induce another is nonzero, an important nonlinearity familiar from neuroscience.

Perhaps the most interesting connection to spin systems is to glasses in which disorder and frustration play central roles. Frustration in spin glasses occurs when any spin cannot achieve a configuration in which the interaction energy to all of its neighbors is minimized. This phenomenon leads to energy landscapes with many peaks and valleys and a large number of metastable configurations [94]. Competing interactions are a hallmark of complex systems and result in interesting and often computationally useful dynamics as the

system traverses the landscape of metastable states [95]. In the present case, the dendritic tree of each loop neuron can be compared to a spin glass in which each dendrite has multiple interactions with competing signs of coupling. Frustration can be quantified for dendrite i as

$$f_i = \frac{(\sum_{j=1}^{n_+} J_{ij} s_j) (\sum_{j=1}^{n_-} |J_{ij}| s_j)}{(\sum_{j=1}^{n_+} J_{ij} s_j^{\max}) (\sum_{j=1}^{n_-} |J_{ij}| s_j^{\max})}, \quad (7)$$

where the sum to n_+ runs over the excitatory inputs, the sum to n_- runs over inhibitory inputs, and s_j^{\max} is the saturation level of the signal in dendrite j . This value for frustration is zero when the sum of excitatory or inhibitory inputs is zero and reaches a maximum value of one when all excitatory and inhibitory inputs are at their maximum values. The frustration of neuron p can be quantified by summing over its arbor, $\mathcal{F}_p = \sum_i f_i$. Each neuron's dendritic tree has qualitative features in common with a dynamical spin glass, with randomness and transitory inhibitory and excitatory interactions leading to ever-shifting states of competition and frustration, as is also evident in biological neural systems [96]. The frustration of the network can be obtained by summing once more over neurons, $\mathcal{F}_N = \sum_p \mathcal{F}_p$. Future work will explore the relationship between frustration, critical phenomena, and useful neural computation in the context of SOENs.

Above the level of the dendritic tree, at the level of coupled neurons, analogies to spin systems are less relevant, and comparison to pulse-coupled oscillators and their more complex cousins, spiking neurons, are straightforward. Superconducting optoelectronic loop neurons are spiking neurons that produce binary, pulsatile communication signals upon being driven to an internal threshold. The optical communication signals that facilitate inter-neuron interaction are not strictly based on spatial location (nearest neighbors, next-nearest neighbors, etc.), but can be engineered with great flexibility through an adjacency matrix physically realized by photonic interconnects. As compared to typical point neurons, a complex, active dendritic tree appears central to the function and construction of loop neurons [26]. Yet at the base of that tree resides a soma that sums inputs and registers a threshold. While the signal in that final integration loop of the soma is purged upon reaching threshold, the signals stored in the loops of the rest of the arbor need not change upon neuronal firing. As compared to archetypal leaky integrate-and-fire neurons, the lack of erasure following firing is distinct. This trait allows a single synapse event to push a neuron into a dynamical orbit and allows information retention regardless of firing state.

In computational neuroscience, a list of discrete spike times is sometimes considered a complete description of the behavior of a neuron or population in a given context. Such a list is available for networks of loop neurons, yet continuous variables specifying the states of all dendritic loops can also be used to gain more granular information about the system, much like retaining the time-continuous values of the membrane potentials on all dendrites of all neurons throughout the network. While loop neuron spike times, bursts, and rates are crucial to network activity, the state of any given neuron or population may also be specified by the signal in all loops, providing richer depiction of state space, high temporal resolution even between spike events, and continuous time-series representation suitable for defining distances

and overlaps between network states as well as correlation functions between components at different times. The state of dendrite i attached to neuron p is specified by s_i . This state description can be extended to neurons as $S_p = [s_1, \dots, s_n]_p$: The state of the neuron is the state of all of its dendrites arranged as a vector. The state of the network can then be defined in full detail as the full vector including the state of all dendrites, $\mathcal{S} = [S_1, \dots, S_N]$, or the state of the each neuron can be reduced to an average for computational efficiency when necessary. The distance between two states can then be defined as the Euclidean distance between two vectors, applicable to single neurons or networks. Such a metric is valuable in many contexts, including in quantifying sensitivity to variation in initial conditions to determine when a network is in a chaotic regime.

We can then discuss correlations between dendrites with themselves at different times, with other dendrites in the same neuronal arbor, and with other dendrites attached to different neurons. For example, a self-correlation function of the form

$$G_i(t, t') = s_i(t) \cdot s_i(t') \quad (8)$$

may be useful for identifying temporal coherence related to recurrent activity. Similarly, a cross-correlation function between two dendrites,

$$G_{ij}(t, t') = s_i(t) \cdot s_j(t'), \quad (9)$$

may be useful in identifying functional coalitions within a neuron with dendrites i and j within the same arbor, or it may be useful for identifying synchronized activity across the network when they are attached to different neurons. Extending this concept to neurons, we can write down a temporal correlation function of a neuron with itself,

$$G_p(t, t') = S_p(t) \cdot S_p(t'), \quad (10)$$

or with other neurons,

$$G_{pq}(t, t') = S_p(t) \cdot S_q(t'). \quad (11)$$

In Eqs. (10) and (11), the vector dot product can be taken to reduce the information to a single, scalar metric for temporal correlations. The cross-correlation function of Eq. (11) contains information about correlations across space and time, as does the dendritic cross-correlation function of Eq. (9) on a shorter length scale. Such functions inform us as to how inputs are represented and stored in correlations, and from these functions we can obtain quantities analogous to the susceptibility which provide information about phase transitions and critical phenomena. These concepts may be further extended to a continuum field theory, potentially conducive to formal analysis that is gaining traction in neural-network applications [97,98].

In spin [49], neural [99], and other complex systems, free energy is a powerful concept capturing a system's competing drives to minimize energy and maximize entropy. In an elementary construction, the free energy is given by $\mathcal{F} = \mathcal{U} - T\mathcal{S}$, with \mathcal{U} the total energy of the system, T the temperature, and \mathcal{S} the entropy. Within the present framework, the energy of a single dendrite is given by $L I^2/2 \propto \beta s^2$, and the total energy can be obtained by summing over dendrites. The entropy can be obtained from the logarithm of the number of microstates at a given value of total energy, which is a well-defined if tedious problem in combinatorics. Alternatively, an approach in terms of a partition function may

be tractable. The temperature of the system may be conceived as the actual, physical temperature in units of kelvins, with the $T\mathcal{S}$ term growing due to stochastic switching of JJs close to I_c , or temperature may be engineered as a more abstract quantity introduced via random photonic inputs to synapses. The conceptual foundation of the loop-neuron system in terms of simple component state variables as formulated with this phenomenological model offers multiple routes to connect to a deeper free-energy formulation.

From this list of comparisons and extensions, it is evident that loop neurons and superconducting optoelectronic networks have potential to lead to physical systems of immense complexity and useful computational functionality. The phenomenological model developed here may serve in the near term to facilitate design of useful circuits and networks for computation. In the longer term, the framework may enable more complete theoretical development of the hardware as a complex physical system, as a tool for artificial intelligence, as a platform for hypothesis testing in neuroscience, and as a means to devise systems with intelligence exceeding our own.

ACKNOWLEDGMENT

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

APPENDIX A: CIRCUIT EQUATIONS FOR THE RI DENDRITE

A circuit diagram of the RI dendrite is shown in Fig. 25. Kirchoff's current law gives $I_b = I_1 + I_2 + I_{di}$. Summing voltages around the R loop gives

$$\frac{\Phi_0}{2\pi} \frac{d\delta_1}{dt} + L_1 \frac{dI_1}{dt} - L_2 \frac{dI_2}{dt} - \frac{\Phi_0}{2\pi} \frac{d\delta_2}{dt} - \frac{d\Phi_r}{dt} = 0, \quad (A1)$$

where δ_i is the phase of the superconducting wave function across the i th junction. We introduce the following dimensionless variables [29]:

$$i_x \equiv I_x/I_c, \quad \beta_x \equiv \frac{2\pi L_x I_c}{\Phi_0}, \quad \phi \equiv \Phi_r/\Phi_0, \\ \alpha \equiv R_{di}/R_j, \quad \tau \equiv \frac{2\pi R_j I_c}{\Phi_0} t \equiv \omega_c t. \quad (A2)$$

ω_c is the characteristic Josephson frequency, and R_j is the Josephson junction shunt resistance in the resistively and capacitively shunted junction (RCSJ) model [29–32]. A di-

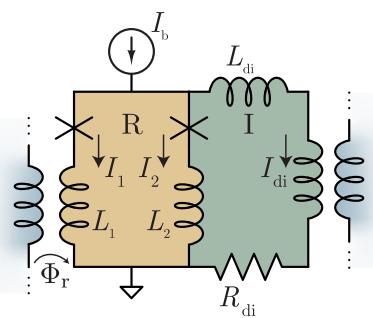


FIG. 25. Circuit diagram of the RI dendrite with relevant circuit parameters labeled.

dimensionless current i_x is defined for each branch of the circuit (i_b , i_1 , i_2 , and i_{di}), while a dimensionless screening parameter, β_x , is defined similarly for each inductor. We also use the notation $\beta_r = \beta_1 + \beta_2$ and $\bar{\beta} = \beta_1\beta_2 + \beta_1\beta_{di} + \beta_2\beta_{di}$. We assume all JJs have identical critical current, I_c . Algebraic manipulation of Eq. (A1) leads to

$$\frac{d i_1}{d\tau} = \frac{1}{\beta_r} \left(\frac{d \delta_2}{d\tau} - \frac{d \delta_1}{d\tau} + 2\pi \frac{d \phi}{d\tau} \right) + \frac{\beta_2}{\beta_r} \left(\frac{d i_b}{d\tau} - \frac{d i_{di}}{d\tau} \right), \quad (\text{A3})$$

where the relation $d/dt = \omega_c d/d\tau$ has been used. Summing voltages around the I loop gives

$$\frac{\Phi_0}{2\pi} \frac{d \delta_2}{dt} + L_2 \frac{d I_2}{dt} - L_{di} \frac{d I_{di}}{dt} - R_{di} I_{di} = 0. \quad (\text{A4})$$

Moving to dimensionless units and making use of Eq. (A3) we can state the primary equation of motion governing the circuit of Fig. 25:

$$\bar{\beta} \frac{d i_{di}}{d\tau} = \beta_1 \frac{d \delta_2}{d\tau} + \beta_2 \frac{d \delta_1}{d\tau} - 2\pi \beta_2 \frac{d \phi}{d\tau} + \beta_1 \beta_2 \frac{d i_b}{d\tau} - \alpha \beta_r i_{di}. \quad (\text{A5})$$

Equation (A5) must be solved in a system with an ODE for each of the JJs of the form

$$\beta_c \frac{d^2 \delta_x}{d\tau^2} = i_x - \sin(\delta_x) - \frac{d \delta_x}{d\tau}. \quad (\text{A6})$$

Here $\beta_c = 2\pi I_c R_j^2 C_j / \Phi_0$ is the Stewart-McCumber parameter, with C_j the capacitance of a JJ in the circuit. Such an expression can be derived in the framework of the RCSJ model [29–32]. This second-order ODE must be converted to two first order ODEs for each of the JJs. The system of coupled ODEs combining Eqs. (A5) and (A6) thus results in a system of five ODEs that can be solved numerically for arbitrary parameter values, bias conditions, and flux drives. After the system of ODEs has been integrated in time, i_1 can be obtained as

$$i_1 = \frac{1}{\beta_r} (\delta_2 - \delta_1 + 2\pi\phi) + \frac{\beta_2}{\beta_r} (i_b - i_{di}), \quad (\text{A7})$$

and $i_2 = i_b - i_1 - i_{di}$. This numerical model has been implemented in PYTHON and the system of ODEs solved with the `solve_ivp` function from SCIPY, employing a Runge-Kutta integration method with fifth-order accuracy and an adaptive time grid.

In constructing the phenomenological model, we have essentially replaced Eq. (A5) with the leaky integrator Eq. (2) and abstracted away the driving terms into the phenomenological rate function. By comparing Eqs. (A5) and (2) and noting that in the main text, we have made the replacement $i_{di} \rightarrow s$, we see that, roughly speaking,

$$\frac{1}{2} \left(\frac{d \delta_2}{d\tau} + \frac{d \delta_1}{d\tau} \right) - \pi \frac{d \phi}{d\tau} \rightarrow r(s, \phi; i_b), \quad (\text{A8})$$

if we let $\beta_1 = \beta_2$, take the limit $\beta_{di} \gg \beta_1$, and ignore the temporal derivative of the bias current. While the two differential equations tracking the current in the dendritic integration loop are qualitatively similar, by moving to the phenomenological model we no longer require tracking the time derivatives of

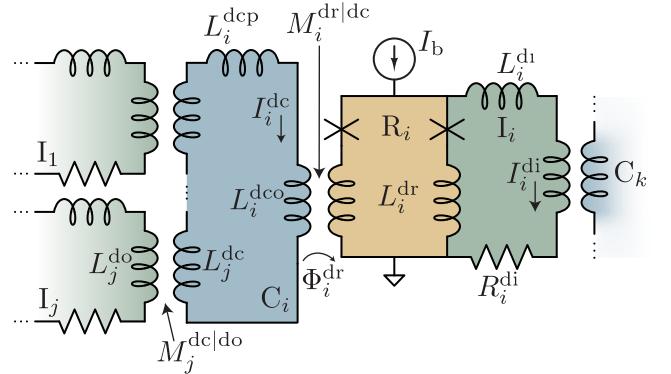


FIG. 26. Full circuit diagram for the RTTI dendrite labeling additional parameters relevant to the derivation of the model in dimensionless units.

the phases of the JJs and the applied flux, which evolve on the picosecond timescale, and instead require only the use of previously calculated rate arrays as look-up tables. The demonstrated speed benefits arise from this replacement.

APPENDIX B: CIRCUIT EQUATIONS FOR THE RTTI DENDRITE

The primary equation of motion for the RTTI dendrite is

$$\beta_{di} \frac{d i_{di}}{d\tau} = \frac{d \delta_4}{d\tau} - \alpha i_{di}. \quad (\text{B1})$$

Equation (B1) must be solved in a system with an ODE of the form given by Eq. (A6) for each of the four JJs in the circuit. Upon reducing the second-order odes to first order, the system comprises nine coupled ODEs. The remaining currents in the circuit can be obtained from

$$\begin{aligned} i_4 &= \beta_4^{-1} (\delta_3 - \delta_4) + i_{b3} - i_{di}, \\ i_3 &= \bar{\beta}^{-1} [\beta_1 (\delta_2 - \delta_3) + \beta_2 (\delta_1 - \delta_3 - 2\pi\phi) + \beta_1 \beta_2 i_{b1}] \\ &\quad + i_{b2} + i_{b3} - i_4 - i_{di}, \\ i_2 &= \frac{1}{\beta_1 + \beta_2} (\delta_1 - \delta_2 - 2\pi\phi) \\ &\quad + \frac{\beta_1}{\beta_1 + \beta_2} (i_{b1} + i_{b2} + i_{b3} - i_3 - i_4 - i_{di}), \\ i_1 &= i_{b1} + i_{b2} + i_{b3} - i_2 - i_3 - i_4 - i_{di}. \end{aligned} \quad (\text{B2})$$

APPENDIX C: OBTAINING THE PHENOMENOLOGICAL MODEL IN DIMENSIONLESS UNITS

In this Appendix, we begin with slightly different and more explicit variable names than used in the main text. We proceed with cluttered notation through the derivation to reduce ambiguity, and all notation will be simplified at the end. The key dynamical variable of each dendrite is the integrated current in the DI loop, I_{di} (see Fig. 26). The foundational postulate of this work is that time evolution of this quantity can be captured in a phenomenological model of the form

$$\frac{d I_{di}}{dt} = I_{fq} R_{fq}(\Phi_r, I_{di}; I_b) - \frac{I_{di}}{\tau_{di}}, \quad (\text{C1})$$

where $I_{\text{fq}} = \Phi_0/L_{\text{di}}$ is the current associated with each fluxon, R_{fq} is the rate of fluxon production as a function of Φ_r , the flux applied to the receiving loop, and I_{di} , the current accumulated in the integration loop. $\tau_{\text{di}} = L_{\text{di}}/R_{\text{di}}$ is the time constant governing the leak of the integrated signal. As in Appendix A, we convert this equation to dimensionless units:

$$\beta_{\text{di}} \frac{d i_{\text{di}}}{d \tau} = r_{\text{fq}}(\phi, i_{\text{di}}; i_b) - \alpha_{\text{di}} i_{\text{di}}. \quad (\text{C2})$$

The dimensionless parameters i_{di} , β_{di} , α_{di} , and τ have been defined in Eq. (A2). The dimensionless rate of fluxon production is given by

$$r_{\text{fq}} = \frac{R_{\text{fq}}}{\omega_c/2\pi}, \quad (\text{C3})$$

with $\omega_c = 2\pi R_j I_c / \Phi_0$.

Each dendrite in the system will obey an equation of the form of Eq. (C2), and knowledge of the system is complete when the values of i_{di} for all dendrites are obtained. However, to solve the complete system of coupled ODEs, we need to calculate how the values of i_{di} from input dendrites convert to coupling flux in the receiving dendrite. Thus the next step is to determine the total applied flux to the DR loop of dendrite i in terms of the values of i_{di} present in the input dendrites, which we index with j . To make space for indices, we move the part labels to superscripts.

The total applied flux to the DR loop of dendrite i is

$$\Phi_i^{\text{dr}} = M_i^{\text{dr}|dc} I_i^{\text{dc}}. \quad (\text{C4})$$

The notation $M_i^{\text{dr}|dc}$ refers to the mutual inductance coupling from the DC loop to the DR loop on dendrite i . Explicitly, $M_i^{\text{dr}|dc} = k_i^{\text{dr}|dc} (L_i^{\text{dr}} L_i^{\text{dco}})^{1/2}$, with $k_i^{\text{dr}|dc}$ the transformer efficiency. In Eq. (C4), I_i^{dc} is the net induced current in the collection coil input to dendrite i . This current obeys the differential equation

$$\frac{d \Phi_i^{\text{dc}}}{dt} - L_i^{\text{dc}} \frac{d I_i^{\text{dc}}}{dt} = 0, \quad (\text{C5})$$

which can be integrated to obtain

$$I_i^{\text{dc}} = \frac{1}{L_i^{\text{dc}}} \Phi_i^{\text{dc}}, \quad (\text{C6})$$

where $L_i^{\text{dc}} = L_i^{\text{dcp}} + L_i^{\text{dco}} + \sum_{j=1}^n L_j^{\text{dc}}$, with L_j^{dc} the inductance of the transformer input to the DC coil connecting dendrite j to dendrite i , L_i^{dco} the output inductance of the DC loop, which is the input to the transformer to the DR loop, and L_i^{dcp} any additional parasitic inductance on the DC coil. The sum is over the n inputs to the DC loop of dendrite i . The total flux applied to the DC loop is

$$\Phi_i^{\text{dc}} = \sum_{j=1}^n M_j^{\text{dc}|do} I_j^{\text{di}}, \quad (\text{C7})$$

with $M_j^{\text{dc}|do} = k_j^{\text{dc}|do} (L_j^{\text{dc}} L_j^{\text{do}})^{1/2}$. Inserting Eq. (C7) into Eq. (C6), we have

$$I_i^{\text{dc}} = \frac{1}{L_i^{\text{dc}}} \sum_{j=1}^n k_j^{\text{dc}|do} (L_j^{\text{dc}} L_j^{\text{do}})^{1/2} I_j^{\text{di}}. \quad (\text{C8})$$

Now inserting Eq. (C8) into Eq. (C4), we obtain

$$\Phi_i^{\text{dr}} = k_i^{\text{dr}|dc} \frac{(L_i^{\text{dr}} L_i^{\text{dco}})^{1/2}}{L_i^{\text{dc}}} \sum_{j=1}^n k_j^{\text{dc}|do} (L_j^{\text{dc}} L_j^{\text{do}})^{1/2} I_j^{\text{di}}. \quad (\text{C9})$$

We now define the dimensionless quantities

$$s_i \equiv i_i^{\text{di}} = \frac{I_i^{\text{di}}}{I_c},$$

$$\phi_i \equiv \phi_i^{\text{dr}} = \frac{\Phi_i^{\text{dr}}}{\Phi_0},$$

$$J_{ij} \equiv \frac{k_i^{\text{dr}|dc} k_j^{\text{dc}|do}}{2\pi} \frac{(\beta_i^{\text{dr}} \beta_i^{\text{dco}} \beta_j^{\text{dc}} \beta_j^{\text{do}})^{1/2}}{\beta_i^{\text{dc}}}, \quad (\text{C10})$$

where the various β parameters are related to the inductances as in Eq. (A2). For simplicity of notation we also specify $\beta_i^{\text{di}} \rightarrow \beta_i$ and $r_{\text{fq}} \rightarrow r$, and we drop the subscript i . We can now recast the phenomenological model of Eq. (C2) in its final form:

$$\beta \frac{ds}{d\tau} = r(\phi, s; i_b) - \alpha s, \quad (\text{C11})$$

where the coupling flux from the n dendrites indexed by j into dendrite i is given by

$$\phi_i = \sum_{j=1}^n J_{ij} s_j. \quad (\text{C12})$$

Equations (C11) and (C12) are the complete phenomenological model of a system of coupled superconducting loop dendrites. The J_{ij} are coupling parameters that can be chosen across a broad range in the circuit layout. To extend the model to account for neurons, certain dendrites are assigned to be somas, wherein a transmitter circuit is activated and light is produced when the value of s reaches a specified threshold. The times of these threshold events are spike times and lead to synapse events on that neuron's downstream synapses. To extend the model to account for synapses, certain dendrites are assigned to receive input flux from an SPD as in Eq. (6) at a time corresponding to the input neuron's spike times plus a delay to account for the activity of the transmitter circuits and light sources, as discussed in Appendices E and F.

APPENDIX D: ADDITIONAL DATA COMPARING PHENOMENOLOGICAL AND CIRCUIT MODELS

Table I relates to Fig. 6, Table II relates to Fig. 7, and Table III relates to Fig. 8. Figure 27 shows further data from square pulse driving functions. The data points come from cases of 10, 20, 40, 80, and 160 square pulses. The x axis is the duration of the simulated time.

APPENDIX E: RATE EQUATIONS FOR THE TRANSMITTER LIGHT SOURCE

To accurately model a superconducting optoelectronic loop neuron, the formalism of the computational circuitry—the dendrites—must be accompanied by a treatment of the transmitter circuit and light source that produce photons when a soma reaches threshold. We begin by treating the light

TABLE I. Values obtained from the simulation of Fig. 6, the case of the linear ramp input with 100 ps time step.

$\beta_{\text{di}}/2\pi$	τ_{di}	χ^2	$t_{\text{ode}}/t_{\text{soen}}$
10 ²	10 ns	6.76×10^{-5}	1.40×10^3
10 ²	50 ns	8.01×10^{-5}	1.29×10^3
10 ²	250 ns	9.95×10^{-5}	1.17×10^3
10 ²	1.25 μ s	8.14×10^{-5}	1.16×10^3
10 ³	10 ns	3.83×10^{-6}	1.78×10^3
10 ³	50 ns	5.07×10^{-6}	1.59×10^3
10 ³	250 ns	1.35×10^{-6}	1.71×10^3
10 ³	1.25 μ s	9.78×10^{-7}	1.37×10^3
10 ⁴	10 ns	2.44×10^{-6}	3.29×10^3
10 ⁴	50 ns	1.37×10^{-6}	2.90×10^3
10 ⁴	250 ns	7.26×10^{-7}	2.48×10^3
10 ⁴	1.25 μ s	8.75×10^{-7}	2.20×10^3

emitters themselves in a rate-equation framework and subsequently consider the transmitter circuit that will provide current to the emission medium. We consider two types of light emitters: quantum dots and silicon emissive centers. The quantum dots we have in mind are group III-V self-assembled dots grown by molecular beam epitaxy or metallorganic chemical vapor deposition. Such emitters have the advantages of high efficiency and fast radiative lifetime. However, for the application at hand scalable manufacturing and integration with semiconductor and superconductor electronics are crucial considerations, and III-V light emitters have a long history of difficulty integrating with silicon. The quantum dot model is also applicable to group IV, Ge-based quantum dots [100], which may prove adequate for the present application. We additionally consider a rate-equation model of silicon emissive centers, as they are easy to fabricate and integrate with electronics, although to date their efficiency is low.

1. Quantum dots

The areal density of quantum dots is n_D ($600/\mu\text{m}^2$). We define the concentration (number per area which can be converted to volume) in the uncharged, unpopulated, ground state

TABLE II. Values obtained from the simulation of Fig. 7, the case of the square pulse inputs with 100 ps time step.

$\beta_{\text{di}}/2\pi$	τ_{di}	χ^2	$t_{\text{ode}}/t_{\text{soen}}$
10 ²	10 ns	3.32×10^{-4}	1.66×10^3
10 ²	50 ns	1.86×10^{-4}	1.48×10^3
10 ²	250 ns	1.17×10^{-4}	1.69×10^3
10 ²	1.25 μ s	7.87×10^{-5}	1.49×10^3
10 ³	10 ns	6.53×10^{-5}	2.31×10^3
10 ³	50 ns	2.20×10^{-5}	2.00×10^3
10 ³	250 ns	1.74×10^{-5}	1.53×10^3
10 ³	1.25 μ s	1.82×10^{-5}	1.62×10^3
10 ⁴	10 ns	4.06×10^{-5}	2.77×10^3
10 ⁴	50 ns	1.12×10^{-5}	2.47×10^3
10 ⁴	250 ns	1.04×10^{-5}	2.11×10^3
10 ⁴	1.25 μ s	9.18×10^{-7}	1.87×10^3

TABLE III. Values obtained from the simulation of Fig. 8, the case of square pulse inputs with 1 ns time step.

$\beta_{\text{di}}/2\pi$	τ_{di}	χ^2	$t_{\text{ode}}/t_{\text{soen}}$
10 ²	10 ns	3.60×10^{-2}	5.09×10^3
10 ²	50 ns	2.34×10^{-2}	5.43×10^4
10 ²	250 ns	2.70×10^{-2}	1.41×10^4
10 ²	1.25 μ s	2.45×10^{-2}	1.48×10^4
10 ³	10 ns	2.84×10^{-3}	1.75×10^4
10 ³	50 ns	9.25×10^{-4}	1.41×10^4
10 ³	250 ns	2.50×10^{-4}	$>4 \times 10^4$
10 ³	1.25 μ s	1.20×10^{-4}	1.37×10^4
10 ⁴	10 ns	2.32×10^{-3}	2.54×10^4
10 ⁴	50 ns	6.62×10^{-4}	2.82×10^4
10 ⁴	250 ns	3.07×10^{-4}	2.14×10^4
10 ⁴	1.25 μ s	1.72×10^{-4}	1.92×10^4

as n_0 . The concentration that have trapped a hole is n_1^p , while the concentration that have a trapped electron is n_1^n . The concentration that have trapped a hole and an electron is n_2 . Assuming quantum dots are not created or destroyed during operation, we have

$$n_D = n_0 + n_1^p + n_1^n + n_2. \quad (\text{E1})$$

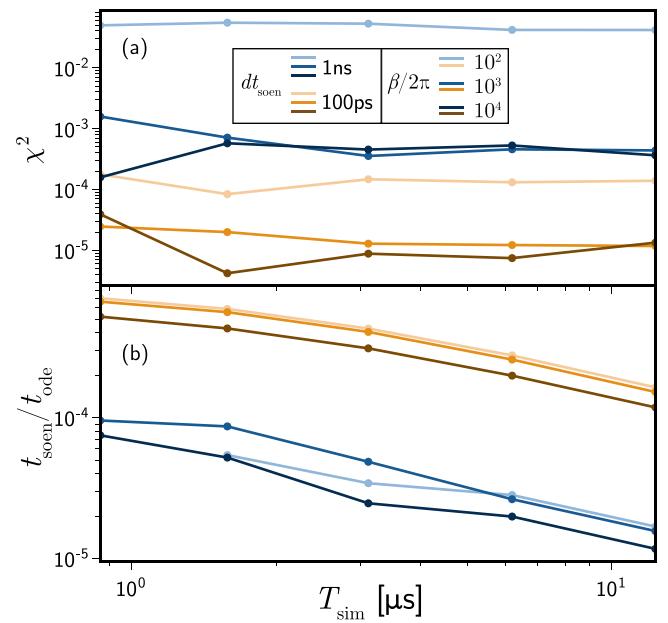


FIG. 27. Summary of errors in comparing the phenomenological model to the circuit equations as a function of the simulated time, T_{sim} . In these simulations, $\tau_{\text{di}} = 250 \mu\text{s}$, and the drive signals were random square-pulse inputs with 10, 20, 40, 80, and 160 pulses. (a) The values of χ^2 for several values of time step and integration loop inductance. (b) The ratio of required time for the computation. t_{soen} is the time required to step through the phenomenological model, while t_{ode} is the time required to solve the system of circuit ODEs. The longest simulation treated $T_{\text{sim}} = 12.32 \mu\text{s}$ of simulated time.

The population of quantum dots in these states is modeled by the following rate equations:

$$\begin{aligned}\frac{d n_0}{dt} &= -n_0 p c_{01}^p - n_0 n c_{01}^n + n_2 e_{20}, \\ \frac{d n_1^n}{dt} &= n_0 n c_{01}^n - n_1^n p c_{12}^p, \\ \frac{d n_2}{dt} &= n_1^n p c_{12}^p + n_1^p n c_{12}^n - n_2 e_{20}. \end{aligned}\quad (\text{E2})$$

In Eqs. (E2), c_{01}^p is the capture coefficient for an unpopulated dot to obtain a hole, and c_{01}^n is the capture coefficient for an unpopulated dot to obtain an electron. The quantities c_{12}^p and c_{12}^n are defined analogously, and e_{20} is the emission probability, related to the radiative lifetime by $\tau_{\text{rad}} = 1/e_{20}$. The populations of free electrons and holes in the semiconductor medium are given by

$$\begin{aligned}\frac{dn}{dt} &= \frac{\eta_{\text{inj}} I}{q V_{\text{LED}}} - n_0 n c_{01}^n - n_1^n p c_{12}^n \\ &\quad - \frac{v_s A_{\text{LED}}}{V_{\text{LED}}} n - (C_n n^2 p + C_p p^2 n), \\ \frac{dp}{dt} &= \frac{\eta_{\text{inj}} I}{q V_{\text{LED}}} - n_0 p c_{01}^p - n_1^n p c_{12}^p \\ &\quad - \frac{v_s A_{\text{LED}}}{V_{\text{LED}}} p - (C_n n^2 p + C_p p^2 n). \end{aligned}\quad (\text{E3})$$

In Eqs. (E3), η_{inj} is the injection efficiency, which could be determined by a spatial carrier transport model, v_s is the surface recombination velocity, A_{LED} is the surface area of the LED, V_{LED} is its volume, and C_n and C_p are the Auger recombination coefficients. Throughout Eqs. (E2), n_1^p can be eliminated with $n_1^p = n_{\text{qd}} - n_0 - n_1^n - n_2$.

The rate of photon production from electroluminescence under current injection is given by

$$\frac{dN_{\text{ph}}}{dt} = V_{\text{LED}} e_{20} n_2(t). \quad (\text{E4})$$

The total number of photons emitted is obtained by integrating:

$$N_{\text{ph}} = V_{\text{LED}} e_{20} \int_{t_0}^{t_f} n_2(t) dt. \quad (\text{E5})$$

The primary quantity of interest for our application is the efficiency of light production. In the context of the present model, we can define the efficiency as the number of photons generated divided by the number of electron-hole pairs injected into the structure:

$$\eta_i = \frac{N_{\text{ph}}}{N_{\text{eh}}}, \quad (\text{E6})$$

where N_{eh} is the total number of electron-hole pairs injected into the intrinsic region of the diode.

To simulate the QDs we use a radiative lifetime of 1 ns, giving $e_{20} = 1$ GHz. The capture coefficients are expected to be quite fast for quantum dots, and we take $c_{01} = c_{12} = 10^{-10}$ m³/s. The precise values of c_{01} and c_{12} have only minor impact on the numerical results as the capture rates are always much smaller than the recombination times. For the parameters governing nonradiative recombination, we $v_s = 2.5 \times 10^3$ cm/s [101] and $C_n = C_p = 7 \times 10^{-30}$ cm⁶/s [102].

The system of ODEs given by Eqs. (E2) and (E3) was again solved using `solve_ivp`. Complete analysis of these light sources is beyond the scope of the present work. The important takeaway is that the efficiency of the source can be quite high, provided the number of carriers injected is commensurate with the number of dots in the diode. For this spiking neuromorphic application, we are interested in cases wherein an initial pulse of electrons and holes is injected into the diode and left to decay. If the pulse of injected carriers contains appreciably more electrons and holes than the number of quantum dots, the excess carriers recombine nonradiatively through surfaces or Auger before the populated dots have time to emit photons and trap additional carriers. Thus, for a neuron to efficiently produce light upon reaching threshold, the light source must have a number of quantum dots that is chosen in accordance with the number of synapses that are intended to receive photons, and the transmitter circuit that injects the light source must be designed to inject the appropriate number of electrons and holes in a brief burst of charge. Such a transmitter circuit is summarized in Appendix F. Time-domain simulations of quantum dots being driven by such a driver circuit will be presented in that Appendix. First, we briefly introduce the rate equation model of silicon emissive centers.

2. Silicon emissive centers

We can expand the quantum dot model to treat silicon emissive centers. A similar model was presented in Ref. [103], but in that work, steady state emission was investigated to seek a continuous-wave laser. Here we require the full dynamical equations to investigate transient behavior. The total concentration of the particular emissive center of interest is n_D . We define the concentration (number per volume) in the uncharged, unpopulated, ground state as n_0 . The concentration that have trapped a hole is n_1 . The concentration that have trapped a hole and an electron is n_2 . Following Ref. [103], we focus on W centers [22] and assume a hole is always trapped before an electron. Assuming emissive centers are not created or destroyed during operation, we have

$$n_D = n_0 + n_1 + n_2. \quad (\text{E7})$$

The equations governing the population of emissive centers are

$$\begin{aligned}\frac{dn_0}{dt} &= -n_0 p c_{01} + (n_D - n_0 - n_2)c_{01}k_{01} + n_2 e_{20}, \\ \frac{dn_2}{dt} &= -n_2 (e_{20} + c_{12} k_{12}) + (n_D - n_0 - n_2) n c_{12}. \end{aligned}\quad (\text{E8})$$

In Eqs. (E8), c_{01} is the capture coefficient for an unpopulated center to obtain a hole. As was argued in Ref. [103], W centers appear to always trap a hole before an electron. The capture coefficient for trapping an electron after the center has been populated with a hole is c_{12} . e_{20} is the emission coefficient. The emission coefficients are related to the capture coefficients through the relations

$$\begin{aligned}e_{10} &= c_{01} k_{01} = c_{01} \frac{n_0^* n_c^*}{n_1^*}, \\ e_{21} &= c_{12} k_{12} = c_{12} \frac{n_1^* n_c^*}{n_2^*}. \end{aligned}\quad (\text{E9})$$

Asterisks refer to equilibrium values. The relations can be further specified by

$$\begin{aligned} k_{01} &= \frac{n_0^* n_c^*}{n_1^*} = \frac{N_v}{g_n} e^{-E_h/kT}, \\ k_{12} &= \frac{n_1^* n_c^*}{n_2^*} = \frac{N_c}{g_e} e^{(E_e - E_g)/kT}. \end{aligned} \quad (\text{E10})$$

Here N_v and N_c are the valence and conduction band densities of states, g_n and g_e are the degeneracies of the states of the emissive center, E_g is the band gap, and E_h and $E_g - E_e$ are the hole and electron binding energies. The energy of the emitted photon is $E_e - E_h$. For the case of the W center, this equals 1.018 eV. The top of the valence band is set to zero energy. See Ref. [104], Ch. 5 for further explanation.

In the second of Eqs. (E8), the term $e_{20} + c_{12} k_{12}$ plays a prominent role in determining the efficiency of emission. This term represents competition between emission (e_{20}) and release of the electron in the bound exciton back to the conduction band ($c_{12} k_{12}$). The ratio $e_{20}/(c_{12} k_{12})$ is roughly 10^5 at 4.2 K, so emission is much more likely than exciton dissociation at this temperature.

In addition to the rate equations for the emissive center populations, the model for Si light sources should include the effects of a population of nonradiative centers that will inevitably be produced when the implants for the emissive centers are performed. Currently, knowledge of the number of different types of nonradiative centers, their capture coefficients, and their lifetimes is limited. For simplicity, we assume there is one type of dominant nonradiative recombination center, we assume it can be populated first by an electron or hole, we assume the capture cross section is c_{01} , and we assume the secondary capture of a hole or electron is governed by c_{12} . We also assume after capture of an electron, hole, or both, no carriers are ever re-emitted to the bands, which is reasonable at the low temperatures necessary for the superconducting electronics present in the system. The total concentration of nonradiative defects is $n_D^{\text{nr}} = n_0^{\text{nr}} + n_1^{\text{nr}} + n_2^{\text{nr}}$. The population rate equations for nonradiative centers then can be written as

$$\begin{aligned} \frac{dn_0^{\text{nr}}}{dt} &= -n_0^{\text{nr}} c_{01}(p + n) + n_2^{\text{nr}} e_{20}^{\text{nr}}, \\ \frac{dn_2^{\text{nr}}}{dt} &= -n_2^{\text{nr}} e_{20}^{\text{nr}} + c_{12} (n_D^{\text{nr}} - n_0^{\text{nr}} - n_2^{\text{nr}}) (p + n). \end{aligned} \quad (\text{E11})$$

The coefficients c_{ij} are different here than in the emissive center equations, but we expect them to be fast, so this is of little consequence. The lifetime of the nonradiative transition of the center (which may be a trap-assisted Auger process or, more likely, phonon recombination) is given by $\tau_{\text{nr}} = 1/e_{20}^{\text{nr}}$.

Equations (E8) and (E11) are coupled to the charge-carrier rate equations. For the electron concentration, we have

$$\begin{aligned} \frac{dn}{dt} &= \frac{\eta_{\text{inj}} I}{q V_{\text{LED}}} - (n_D - n_0 - n_2) n c_{12} + n_2 c_{12} k_{12} \\ &\quad - n_0^{\text{nr}} n c_{01} - (n_D^{\text{nr}} - n_0^{\text{nr}} - n_2^{\text{nr}}) n c_{12} \\ &\quad - \frac{v_s A_{\text{LED}}}{V_{\text{LED}}} n - (C_n n^2 p + C_p p^2 n), \end{aligned} \quad (\text{E12})$$

and for the hole concentration, we have

$$\begin{aligned} \frac{dp}{dt} &= \frac{\eta_{\text{inj}} I}{q V_{\text{LED}}} - n_0 p c_{01} \\ &\quad - n_0^{\text{nr}} p c_{01} - (n_D^{\text{nr}} - n_0^{\text{nr}} - n_2^{\text{nr}}) p c_{12} \\ &\quad - \frac{v_s A_{\text{LED}}}{V_{\text{LED}}} p - (C_n n^2 p + C_p p^2 n). \end{aligned} \quad (\text{E13})$$

To numerically solve Eqs. (E8), (E11), (E12), and (E13), we need to specify c_{01} , c_{12} , and e_{20} . For the process of photon emission, related to e_{20} , we take $e_{20} = \tau_{ec}^{-1} = 1/40 \text{ ns} = 2.5 \times 10^7$ [22,105]. To our knowledge, c_{01} , c_{12} have not been measured. We take $c_{01} = 5 \times 10^{-14} \text{ m}^3/\text{s}$ and $c_{12} = 1.8 \times 10^{-11} \text{ m}^3/\text{s}$. These numbers are motivated by a simple geometrical model as well as a hydrogenic model of an impurity screened by the dielectric [106]. Numerical studies show that the results are not sensitive to these values across several orders of magnitude.

To specify the nonradiative terms in the rate equations, we take the Auger recombination coefficients as $C_n = 2.8 \times 10^{-31} \text{ cm}^6/\text{s}$ and $C_p = 1 \times 10^{-31} \text{ cm}^6/\text{s}$. These are the values at 300 K [102], and they will reduce slightly at low temperature (scaling with temperature as $T^{0.6}$ [107]), but this is not appreciable given the low level of accuracy in the present model, so we ignore the temperature dependence of Auger and use 300 K values. The surface recombination can be made quite low in properly treated Si and Si-SiO₂ interfaces, with values as low as 0.25 cm/s in the literature [108]. Throughout these simulations we use $v_s = 2.5 \text{ cm/s}$.

Time-domain simulations of silicon emissive centers based on Eqs. (E8), (E11), (E12), and (E13) provides a similar message to that learned for quantum dots: the emitters will populate rapidly, but if the level of carrier injection is too large, excess carriers will go to waste. In the case of silicon emissive centers, loss is not due to surfaces or Auger, but rather to nonradiative centers. The relative concentration of nonradiative centers formed during the ion implantation used to realize W centers is likely to be a crucial quantity determining the ultimate efficiency of such a light source. Given the competing populations of emissive and nonradiative centers, it is not sufficient to inject the same number of carriers as emissive centers and desired photons; one must inject enough carriers to populate both the nonradiative centers and the emissive centers. However, if the concentration of nonradiative centers can be limited to something like 10 times the concentration of emitters, total light-production efficiency may stay within a useful range.

We next describe transmitter circuits capable of providing brief pulses of injected current to produce the desired number of photons necessary to serve a neuron's downstream synaptic connections.

APPENDIX F: CIRCUIT EQUATIONS FOR THE TRANSMITTER DRIVER CIRCUIT

The transmitter circuit under consideration is shown in Fig. 28. The circuit combines JJs, an hTron thresholding element [23], MOSFETs (PMOS and NMOS), and an LED. The hTron serves as the interface between the superconducting and semiconducting domains, enabling low-voltage JJs to

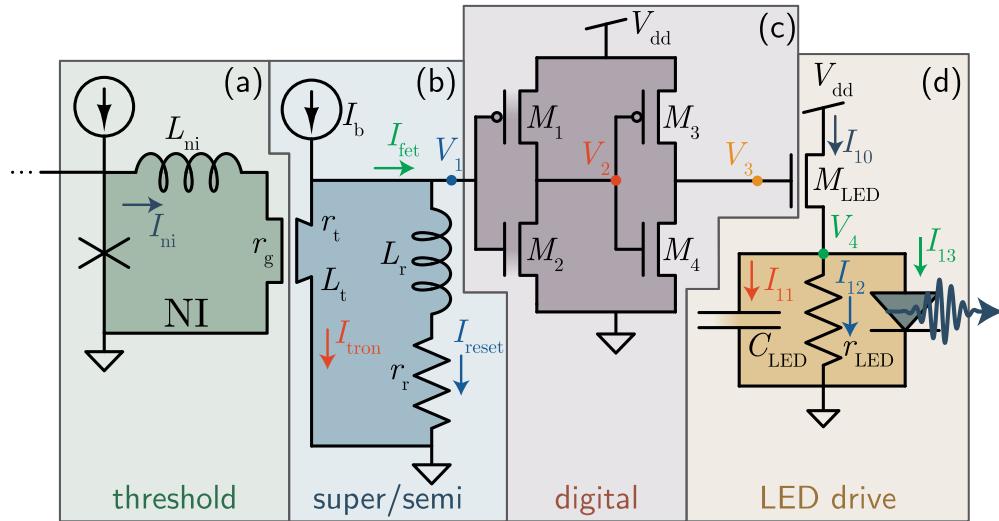


FIG. 28. Diagram of the transmitter circuit with relevant components labeled. (a) The neuronal integration loop (NI) and tron thresholding element. (b) The superconductor-semiconductor interface where the tron switches the first inverter. (c) The inverter pair that produces the digital drive signal. (d) The LED driver MOSFET and the LED itself.

switch a CMOS inverter. The basic operation of the circuit is as follows. Signal moves from left to right. When the JJ in the neuronal integration (NI) loop has added sufficient current to drive the hTron gate above threshold, the hTron channel switches from zero resistance to high resistance. The current bias to the hTron channel (I_b) is thus shunted from the hTron channel (I_{tron}) to the passive reset branch (L_r , r_r), temporarily inducing a voltage at node V_1 sufficient to switch the first CMOS inverter (M_1 , M_2). The two inverters serve to produce a digital signal with voltage V_{dd} applied to MOSFET M_{LED} when the neuronal threshold is reached and the hTron switches. This digital behavior ensures the superconducting components are well separated from the LED, providing consistent current biasing of the LED decoupled from the performance of the thresholding element. The LED driver MOSFET (M_{LED}) can then be designed with a width-to-length ratio and doping level commensurate with the number of photons desired from the LED, which is chosen based on the number of synaptic connections made by the neuron. M_{LED} will deliver a pulse of current to the LED while the hTron is keeping the voltage V_1 high, and when that voltage drops as the passive reset circuit performs its operation, the current to the LED will cease. As is shown below, the MOSFETs are active for a few nanoseconds. With this qualitative description of the circuit operation in mind, we now present the circuit model used in this work to provide a phenomenological treatment of the behavior.

To treat the thresholding component where the current added to the integration loop drives an hTron above threshold, we model the hTron as a device that switches from a zero resistance to a high-resistance state when a certain current is reached. Accurately modeling the electrothermal dynamics of the hTron is difficult and will do little to improve the model for the present purpose, so a simple resistive switch is employed.

To simulate the behavior of the MOSFETs, we use a charge-control model [109] with current-voltage

characteristics given by

$$I_{ds}(V_{ds}, V_{gs}) = \frac{W\mu_n c_i}{L} \times \begin{cases} (V_{gt} - V_{ds}/2)V_{ds}, & \text{for } V_{ds} \leq V_{sat}, \\ V_{gt}^2/2, & \text{for } V_{ds} > V_{sat}. \end{cases} \quad (F1)$$

V_{gt} is the voltage above threshold: $V_{gt} = V_{gs} - V_t$. We assume the drain-source current is zero for $V_{gs} < V_t$, i.e., we ignore subthreshold behavior. In this model, $V_{sat} = V_{gt}$. We treat the LED with a conventional diode model of the form

$$I_{LED} = eA[(D_p/L_p)p_n + (D_n/L_n)n_p] \times [\exp(eV/k_B T) - 1]. \quad (F2)$$

In the present work, we have considered two classes of LEDs. One is based on silicon emissive centers [22], as have been considered in the context of SOENs due to their unique process compatibility and feasibility for enabling low-cost manufacturing of large systems. The other is based on III-V quantum dots, specifically with material parameters for In-GaAs quantum dots on a GaAs platform. For integration with SOENs, wafer bonding of such substrates may be employed. The model presented here can take into account either of these light sources. Both will obey an equation of the form given by Eq. (F2), with the diode placed in parallel with a capacitor and a resistor as shown in Fig. 28. In reality the capacitance and parasitic shunt resistance will both be functions of applied voltage, but for simplicity, we fix these values. Based on the source rate equation model in Appendix E, we know that we would like to have one emitter (quantum dot or emissive center) per photon to be generated by the light source per neuronal firing pulse. We assume the emitters will have a specified density ρ_{ec} , so the value of the capacitance will depend on the number of emitters through the geometry of the junction. The number of emitters is in turn specified by the number of synaptic connections made by the neuron. Therefore we consider two values of capacitance per unit

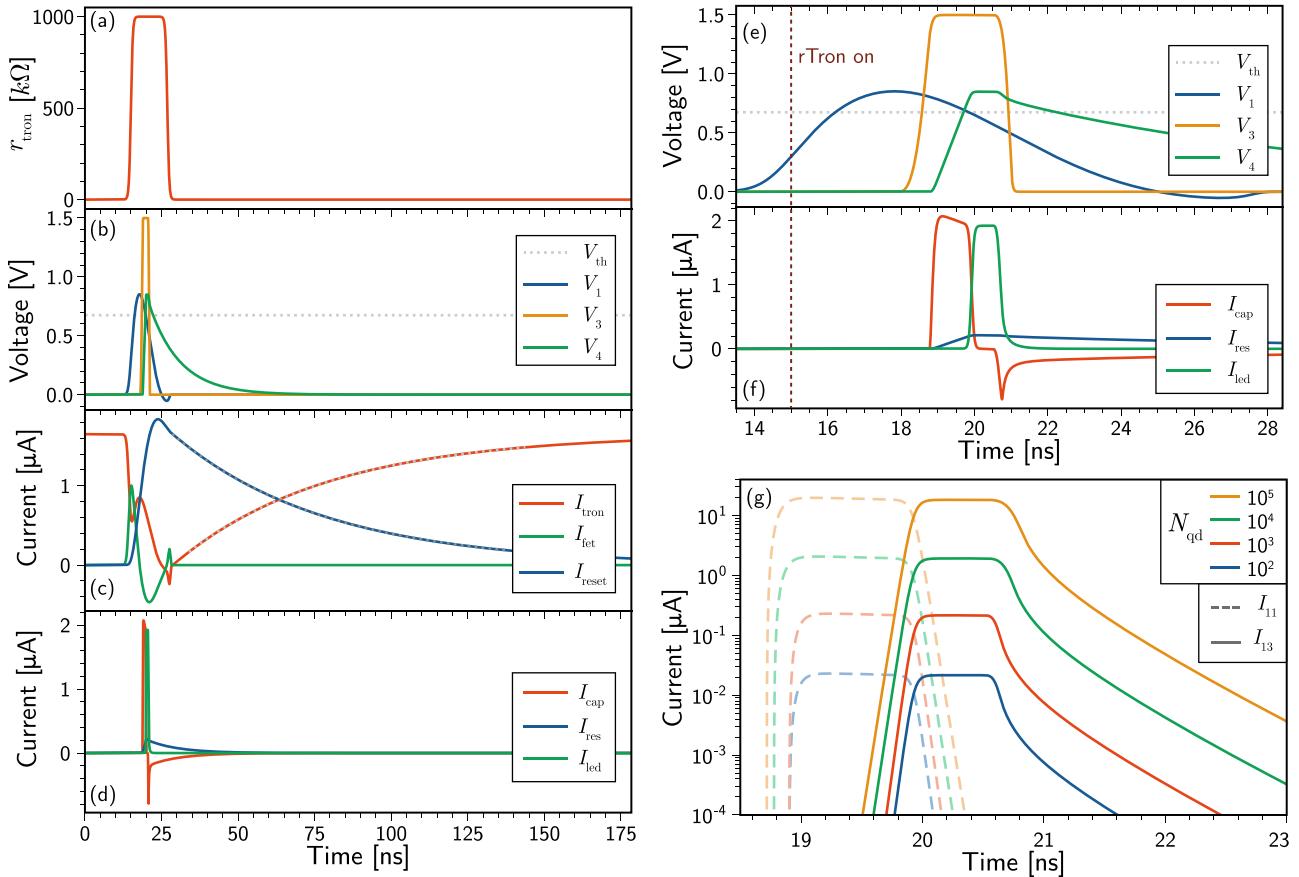


FIG. 29. Time traces from the transmitter circuit. (a) The resistance of the tron as a function of time. (b) The voltages of the MOSFETs. (c) The currents in the tron and related circuit elements. (d) Currents through the LED. (e) Temporal zoom of the MOSFET voltages. (f) Temporal zoom of the LED currents. (g) Currents into the LED and capacitor when different numbers of quantum dots are present in the light source.

area ($C_a = 1 \times 10^{-7}$ and 1×10^{-5} F/\$\mu\text{m}^2\$) to cover a range of values that may be found in various devices. The value of the capacitor, C_{LED} , is determined by $C_{\text{LED}} = C_a N_{\text{ph}} / \rho_{\text{ec}}$ with ρ_{ec} specified as a number of emitters per unit area. Given these models for the hTron, MOSFETs, and LED the following circuit equations can be derived:

$$\begin{aligned} \frac{d V_1}{dt} &= \frac{I_b - I_1 - I_3}{C_1}, \\ \frac{d V_2}{dt} &= \frac{-1}{C_1} [I_p(V_2 - V_{\text{dd}}, V_1 - V_{\text{dd}}) + I_n(V_2, V_1)], \\ \frac{d V_3}{dt} &= \frac{-1}{C_D} [I_p(V_3 - V_{\text{dd}}, V_2 - V_{\text{dd}}) + I_n(V_3, V_2)], \\ \frac{d V_4}{dt} &= \frac{1}{C_{\text{LED}}} \left[I_n(V_{\text{dd}} - V_4, V_3) - I_{\text{LED}}(V_4) - \frac{V_4}{r_{\text{LED}}} \right], \\ \frac{d^2 I_1}{dt^2} &= \frac{I_b - I_1 - I_3 - \dot{r}_t}{L_t C_1} I_1 - \frac{r_t I_1}{L_t} \frac{d I_1}{dt}, \\ \frac{d I_3}{dt} &= \frac{r_t}{L_t} I_1 + \frac{L_t}{L_r} \frac{d I_1}{dt} - \frac{r_r}{L_r} I_3. \end{aligned} \quad (\text{F3})$$

Here, $I_p(V_{\text{ds}}, V_{\text{gs}})$ and $I_n(V_{\text{ds}}, V_{\text{gs}})$ are the PMOS and NMOS forms of Eq. (F1) (i.e., current flows from source to drain in PMOS and drain to source in NMOS), C_1 is the capacitance of

an inverter, and C_D is the capacitance of the MOSFET driver to the LED. The time-dependent resistance of the hTron is r_t and its temporal derivative is \dot{r}_t . Time domain circuit simulations of the transmitter model given by Eqs. (F3) have been carried out using `solve_ivp`. A sample of results is given in Fig. 29. The resistance pulse of the hTron is shown in Fig. 29(a), and the relevant voltages are shown in Fig. 29(b). The currents are shown in Figs. 29(c) and 29(d). A temporal zoom of the voltage traces is shown in Fig. 29(e), where the MOSFET threshold is shown by the grey dotted line. The gates of the first inverter stage quickly rise above threshold, and a nearly ideal square pulse is delivered to the gate of M_{LED} , the driver MOSFET to the light source, as can be seen in the temporal zoom of the LED currents of Fig. 29(f). It is for this reason that the two-inverter digital circuit is employed. After the voltage to the first inverter drops below threshold, M_{LED} quickly shuts off, and the voltage across the LED block decays with the RC time constant of that circuit. The resistance, r_{LED} , has been chosen to establish a 100 ns time constant, and the actual value in a fabricated circuit will need to be investigated.

In the temporal zooms of Figs. 29(e) and 29(f), the time at which the hTron switches to the resistive state is labeled (15 ns), and the total delay of the circuit from the time of the resistance to the time that current begins to enter the diode is 5 ns, with delays accrued due to charging up the various

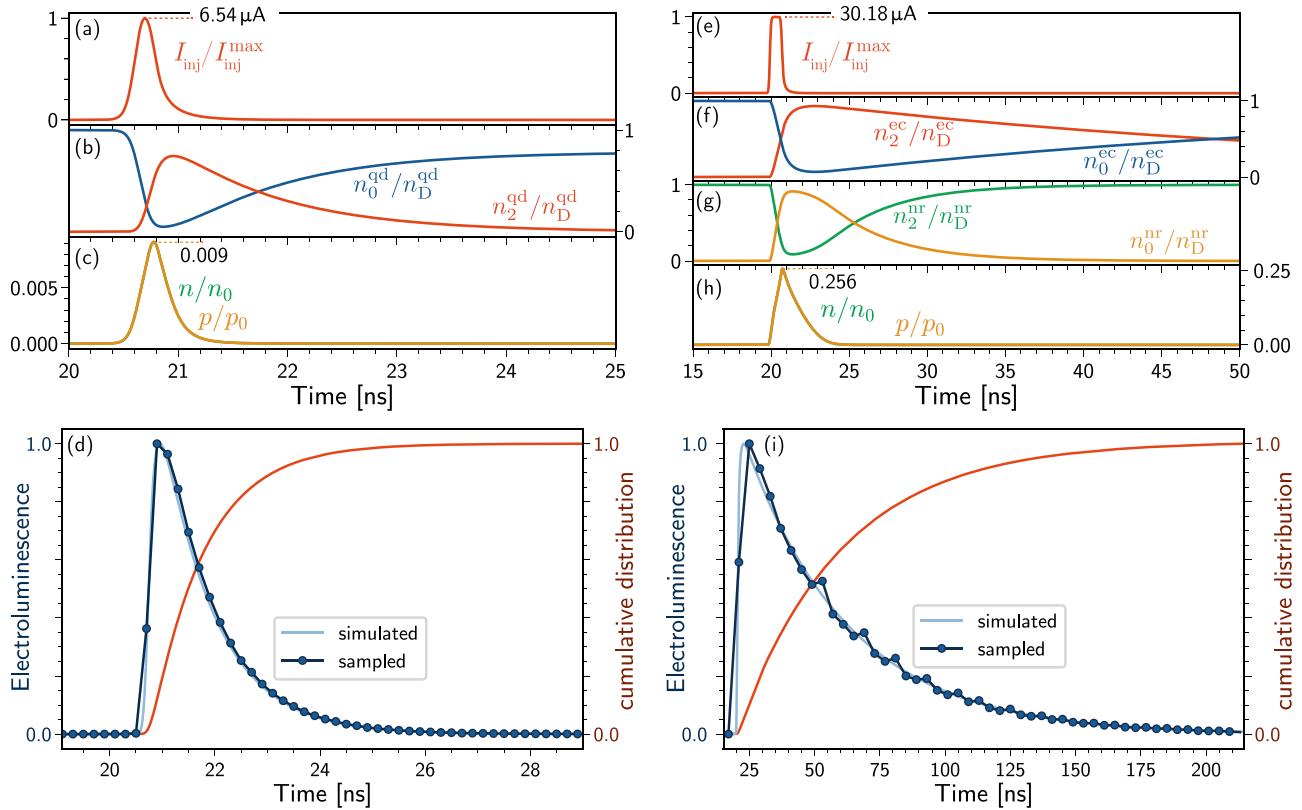


FIG. 30. Time traces for the light source model. [(a)–(d)] The quantum-dot model. [(e)–(i)] The emissive center model. (a) The current pulse injected from the transmitter circuit. (b) The populations of quantum dots in the ground and excited states normalized to the concentration of quantum dots in the LED volume. (c) The concentrations of free electrons and holes normalized to the total number injected into the LED volume. (d) The simulated electroluminescence signal, the cumulative distribution, and a histogram of sampled emission times. (e) The current pulse injected from the transmitter circuit. (f) The populations of emissive centers in the ground and excited states normalized to the concentration of emitters in the LED. (g) The populations of nonradiative recombination centers in the ground and excited states normalized to the concentration of nonradiative recombination centers in the LED. (h) The concentrations of free electrons and holes. (i) The simulated electroluminescence signal, the cumulative distribution, and a histogram of sampled emission times.

MOSFET and LED capacitances. The MOSFET model treated here assumed a 1 μm minimum feature size, compatible with the cleanroom at NIST, and shorter delays are achievable with the reduced capacitance of advanced CMOS. Full circuit simulations have not been run for MOSFET models corresponding to contemporary CMOS nodes, but estimates indicate the delay will reduce to 2–3 ns.

In running these simulations, the number of quantum dots or emissive centers was specified (N_{qd}), and the width-to-length ratio of M_{LED} as well as the channel doping of that MOSFET were iterated to obtain the appropriate current injection to populate that number of emitters. For each value of N_{qd} , the area of the diode was calculated based on the density of emitters, and from this area the capacitance was approximated using a parallel-plate model. Thus the capacitance increases linearly with the number of photons required from the source. Figure 29(g) shows the currents in the LED on a fine temporal scale with a logarithmic y axis for four values of the number of emitters ranging from 10^2 to 10^5 , covering a broad range of neuron types with varying degree of connectivity. The dashed curves at early times in the plot correspond to the current being driven into the LED capacitance (I_{11} in Fig. 28), while the solid traces at slightly later times correspond to the current into the active region of the LED (I_{13} in Fig. 28). It is evident

that for all cases current must be delivered to the capacitor to bring the voltage across the LED above threshold before current is driven into the diode itself, resulting in approximately 1 ns of additional delay beyond the MOSFET stages.

In the phenomenological model that is the subject of this work, the current through the diode is used as an input to the source rate equations. I_{13} from the transmitter model is input as the driving current, I , in Eqs. (E3), (E12), and (E13). The results of these calculations are shown in Fig. 30. Figures 30(a)–30(d) shows results from the quantum dot rate equations, while Figs. 30(e)–30(i) shows results from the emissive center rate equations. In Fig. 30(a), the current injection pulse is shown, while Fig. 30(b) shows the ground-and excited-state populations as a function of time. The sum of these two quantities does not return to unity at the end of the simulation because a finite fraction of the dots remains in an excited state with a trapped electron or hole. At low temperature, without a mechanism for decay of these states, the consequence would be that slightly less injected charge is required on subsequent pulses to populate the ensemble of quantum dots with excitons. Figure 30(c) shows the populations of electrons and holes normalized to the total number of carriers injected in the simulation. The simulated electroluminescence is shown in Fig. 30(d). Similar time traces for the

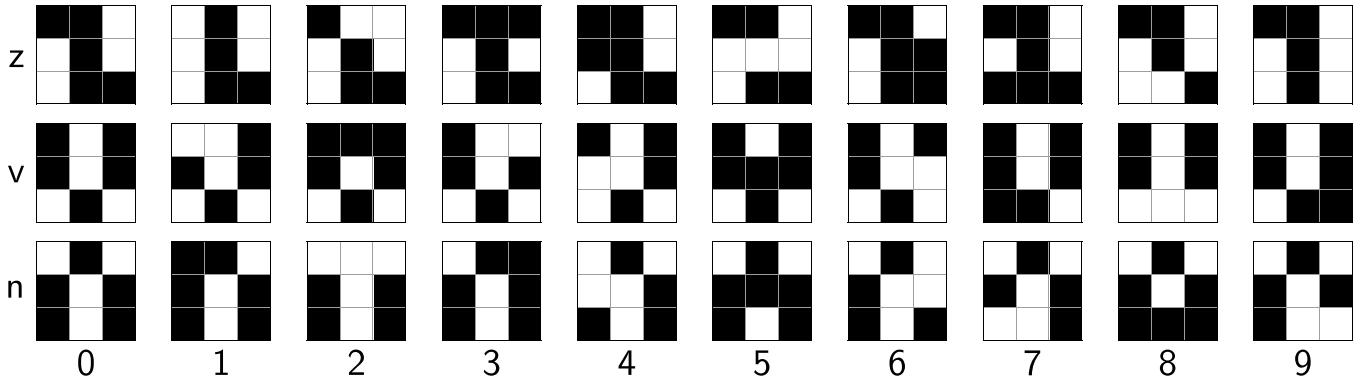


FIG. 31. All 30 inputs to the nine-pixel classifier.

silicon emissive center model are shown in Figs. 30(e)–30(i), with the addition of the populations of the nonradiative centers in Fig. 30(g). The result is very similar to the quantum dot case, except the longer lifetime of the emissive centers results in an electroluminescence signal that extends further in time. In both Figs. 30(d) and 30(i), a histogram of 10^6 samples drawn from the respective cumulative distributions are shown by the dark blue dots.

Only the electroluminescence output from the simulations enters the phenomenological model. From the electroluminescence signal, the cumulative distribution is formed, and from this, photon delay times can be sampled. To use these simulations in the phenomenological model under consideration, it is helpful that the four curves in Fig. 29(g) are essentially scaled versions of each other. When input to the source rate equations, the output electroluminescence spectrum is very similar. When implemented in the phenomenological model code, a neuron will have a given number of downstream synapses, and a multiplicative factor is specified to determine how many photons are generated when the neuron reaches threshold, which we refer to as N_{ph} . Each time a neuron spikes, N_{ph} samples are drawn from the electroluminescence distribution giving N_{ph} values of time delay. These N_{ph} values of time delay are randomly assigned across the receiving synapses. Typically, N_{ph} is larger than the number of synapses, so for each synapse the earliest time from its list is chosen as the spike time. This behavior is justified based on the binary

response of the single-photon detectors [110] that form the receivers at each synapse.

While Appendices E and F have been somewhat involved, the take-away message is simple: when a neuron reaches threshold, its synapses all receive spike events with times randomly sampled from the numerically determined probability distributions. The probability distributions depend on which light source is chosen, with the primary difference being the exponential decay time constant set by the lifetime of the radiative emission process. This transmitter circuit and source design completes the phenomenological treatment. The phenomenological model of the computational circuits—the dendrites, synapses, and somas—is given by a leaky-integrator equation, Eq. (2), with a nonlinear driving function, $r(\phi, s; i_b)$, and a simple form for coupling between dendrites, Eq. (3). The phenomenological model of communication between neurons is simply a probabilistic delay time drawn from a specified distribution obtained through circuit and source simulations.

APPENDIX G: NINE-PIXEL DRIVE MATRICES

The full set of 30 inputs to the nine-pixel image-classification task are shown in Fig. 31. The images are generated by starting with the ideal letter instances for z , v , and n . Subsequent variants are generated by letting one pixel at a time switch its state.

-
- [1] C. Mead, *Analog VLSI and Neural Systems*, 1st ed. (Addison-Wesley, New York, 1989).
 - [2] G. Indiveri, B. Linares-Barranco, T. J. Hamilton, A. van Schaik, R. Etienne-Cummings, T. Delbrück, S.-C. Liu, P. Dudek, P. Häfliger, S. Renaud *et al.*, Neuromorphic silicon neuron circuits, *Front. Neurosci.* **5**, 73 (2011).
 - [3] S.-C. Liu, T. Delbrück, G. Indiveri, A. Whatley, and R. Douglas, *Event-Based Neuromorphic Systems* (Wiley, West Sussex, UK, 2014).
 - [4] S. Furber, Large-scale neuromorphic computing systems, *J. Neural Eng.* **13**, 051001 (2016).
 - [5] C. S. Thakur, J. L. Molin, G. Cauwenberghs, G. Indiveri, K. Kumar, N. Qiao, J. Schemmel, R. Wang, E. Chicca, J. O. Hasler *et al.*, Large-scale neuromorphic spiking array processes: A quest to mimic the brain, *Front. Neurosci.* **12**, 891 (2018).
 - [6] B. J. Shastri, A. N. Tait, T. F. de Lima, W. H. P. Pernice, H. Bhaskaran, C. D. Wright, and P. R. Prucnal, Photonics for artificial intelligence and neuromorphic computing, *Nat. Photonics* **15**, 102 (2021).
 - [7] N. H. Farhat, D. Psaltis, A. Prata, and E. Paek, Optical implementation of the hopfield model, *Appl. Opt.* **24**, 1469 (1985).
 - [8] Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund *et al.*, Deep learning with coherent nanophotonic circuits, *Nat. Photon.* **11**, 441 (2017).
 - [9] A. N. Tait, T. F. De Lima, M. A. Nahmias, H. B. Miller, H.-T. Peng, B. J. Shastri, and P. R. Prucnal, Silicon Photonic Modulator Neuron, *Phys. Rev. Appl.* **11**, 064043 (2019).

- [10] J. Moughames, X. Porte, M. Thiel, G. Ulliac, L. Larger, M. Jacquot, M. Kadic, and D. Brunner, Three-dimensional waveguide interconnects for scalable integration of photonic neural networks, *Optica* **7**, 640 (2020).
- [11] J. Chiles, S. M. Buckley, S. W. Nam, R. P. Mirin, and J. M. Shainline, Design, fabrication, and metrology of 10×100 multi-planar integrated photonic routing manifolds for neural networks, *APL Photon.* **3**, 106101 (2018).
- [12] M. Hidaka and L. A. Akers, An artificial neural cell implemented with superconducting circuits, *Supercond. Sci. Technol.* **4**, 654 (1991).
- [13] Y. Harada and E. Goto, Artificial neural network circuits with josephson devices, *IEEE Trans. Magn.* **27**, 2863 (1991).
- [14] M. Schneider, E. Toomey, G. Rowlands, J. Shainline, P. Tschirhart, and K. Segall, Supermind: a survey of the potential of superconducting electronics for neuromorphic computing, *Supercond. Sci. Technol.* **35**, 053001 (2022).
- [15] F. Chiarello, P. Carelli, M. G. Castellano, and G. Torrioli, Artificial neural network based on squids: Demonstration of network training and operation, *Supercond. Sci. Technol.* **26**, 125009 (2013).
- [16] Y. Yamanashi, K. Umeda, and N. Yoshikawa, Pseudo sigmoid function generator for a superconductive neural network, *IEEE Trans. Appl. Supercond.* **23**, 1701004 (2012).
- [17] N. V. Klenov, A. E. Schegolev, I. I. Soloviev, S. V. Bakurskiy, and M. V. Tereshonok, Energy efficient superconducting neural networks for high-speed intellectual data processing systems, *IEEE Trans. Appl. Supercond.* **28**, 1 (2018).
- [18] T. Hirose, T. Asai, and Y. Amemiya, Spiking neuron devices consisting of single-flux-quantum circuits, *Phys. C: Supercond. Appl.* **445**, 1020 (2006).
- [19] K. Segall, M. LeGro, S. Kaplan, O. Svitelskiy, S. Khadka, P. Crotty, and D. Schult, Synchronization dynamics on the picosecond timescale in coupled josephson junction neurons, *Phys. Rev. E* **95**, 032220 (2017).
- [20] J. M. Shainline, S. M. Buckley, A. N. McCaughan, J. T. Chiles, A. J. Salim, M. Castellanos-Beltran, C. A. Donnelly, M. L. Schneider, R. P. Mirin, and S. W. Nam, Superconducting optoelectronic loop neurons, *J. Appl. Phys.* **126**, 044902 (2019).
- [21] J. M. Shainline, Optoelectronic intelligence, *Appl. Phys. Lett.* **118**, 160501 (2021).
- [22] S. Buckley, J. Chiles, A. N. McCaughan, G. Moody, K. L. Silverman, M. J. Stevens, R. P. Mirin, S. W. Nam, and J. M. Shainline, All-silicon light-emitting diodes waveguide-integrated with superconducting single-photon detectors, *Appl. Phys. Lett.* **111**, 141101 (2017).
- [23] A. N. McCaughan, V. B. Verma, S. M. Buckley, J. P. Allmaras, A. G. Kozorezov, A. N. Tait, S. W. Nam, and J. M. Shainline, A superconducting thermal switch with ultrahigh impedance for interfacing superconductors to semiconductors, *Nat. Electron.* **2**, 451 (2019).
- [24] S. Khan, B. A. Primavera, J. Chiles, A. N. McCaughan, S. M. Buckley, A. N. Tait, A. Lita, J. Biesecker, A. Fox, D. Olaya *et al.*, Superconducting optoelectronic single-photon synapses, *Nat. Electron.* **5**, 650 (2022).
- [25] J. M. Shainline, Fluxonic processing of photonic synapse events, *IEEE J. Sel. Top. Quantum Electron.* **26**, 1 (2019).
- [26] B. A. Primavera and J. M. Shainline, An active dendritic tree can mitigate fan-in limitations in superconducting neurons, *Appl. Phys. Lett.* **119**, 242601 (2021).
- [27] P. Dayan and L. F. Abbott, *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (MIT Press, 2005).
- [28] W. Gerstner and W. M. Kistler, *Spiking Neuron Models: Single Neurons, Populations, Plasticity* (Cambridge University Press, 2002).
- [29] *The SQUID Handbook*, edited by J. Clarke and A. I. Braginski (Wiley-VCH, 2006).
- [30] T. Van Duzer and C. W. Turner, *Principles of Superconductive Devices and Circuits*, 2nd ed. (Prentice Hall, USA, 1998).
- [31] A. M. Kadin, *Introduction to Superconducting Circuits* (Wiley-Interscience, 1999).
- [32] M. Tinkham, *Introduction to Superconductivity* (Dover, 2004).
- [33] J. K. W. Yang, A. J. Kerman, E. A. Dauler, V. Anant, K. M. Rosfjord, and K. K. Berggren, Modeling the electrical and thermal response of superconducting nanowire single-photon detectors, *IEEE Trans. Appl. Supercond.* **17**, 581 (2007).
- [34] C. M. Natarajan, M. G. Tanner, and R. H. Hadfield, Superconducting nanowire single-photon detectors: Physics and applications, *Supercond. Sci. Technol.* **25**, 063001 (2012).
- [35] A. Gidon, T. A. Zolnik, P. Fidzinski, F. Bolduan, A. Papoutsis, P. Poirazi, M. Holtkamp, I. Vida, and M. E. Larkum, Dendritic action potentials and computation in human layer 2/3 cortical neurons, *Science* **367**, 83 (2020).
- [36] Q.-Y. Zhao, A. N. McCaughan, A. E. Dane, K. K. Berggren, and T. Ortlepp, A nanocryotron comparator can connect single-flux-quantum circuits to conventional electronics, *Supercond. Sci. Technol.* **30**, 044002 (2017).
- [37] L. N. Cooper and M. F. Bear, The bcm theory of synapse modification at 30: Interaction of theory with experiment, *Nat. Rev. Neurosci.* **13**, 798 (2012).
- [38] K. Pozo and Y. Goda, Unraveling mechanisms of homeostatic synaptic plasticity, *Neuron* **66**, 337 (2010).
- [39] G. Turrigiano, Homeostatic synaptic plasticity: Local and global mechanisms for stabilizing neuronal function, *Cold Spring Harbor Perspect. Biol.* **4**, a005736 (2012).
- [40] M. Prezioso, F. Merrikh-Bayat, B. D. Hoskins, G. C. Adam, K. K. Likharev, and D. B. Strukov, Training and operation of an integrated neuromorphic network based on metal-oxide memristors, *Nature (London)* **521**, 61 (2015).
- [41] P. Poirazi, T. Brannon, and B. W. Mel, Pyramidal neuron as two-layer neural network, *Neuron* **37**, 989 (2003).
- [42] C. Koch and I. Segev, The role of single neurons in information processing, *Nat. Neurosci.* **3**, 1171 (2000).
- [43] M. L. Schneider, C. A. Donnelly, and S. E. Russek, Tutorial: High-speed low-power neuromorphic systems based on magnetic josephson junctions, *J. Appl. Phys.* **124**, 161102 (2018).
- [44] P. Sterling and S. Laughlin, *Principles of Neural Design* (MIT Press, 2015).
- [45] S. Wolfram, Statistical mechanics of cellular automata, *Rev. Mod. Phys.* **55**, 601 (1983).
- [46] S. Wolfram, *A New Kind of Science* (Wolfram Media, 2002).
- [47] S. Wolfram, A class of models with the potential to represent fundamental physics, *Complex Syst.* **29**, 107 (2020).
- [48] H. Nishimori and G. Ortiz, *Elements of Phase Transitions and Critical Phenomena* (Oxford University Press, 2010).
- [49] R. Shankar, *Quantum Field Theory and Condensed Matter: An Introduction* (Cambridge University Press, 2017).

- [50] W. A. Little, The existence of persistent states in the brain, *Math. Biosci.* **19**, 101 (1974).
- [51] J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, *Proc. Natl. Acad. Sci. USA* **79**, 2554 (1982).
- [52] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Spin-glass models of neural networks, *Phys. Rev. A* **32**, 1007 (1985).
- [53] D. J. Amit and D. J. Amit, *Modeling Brain Function: The World of Attractor Neural Networks* (Cambridge University Press, 1989).
- [54] O. Kinouchi and M. Copelli, Optimal dynamical range of excitable networks at criticality, *Nat. Phys.* **2**, 348 (2006).
- [55] W. L. Shew, H. Yang, T. Petermann, R. Roy, and D. Plenz, Neuronal avalanches imply maximum dynamic range in cortical networks at criticality, *J. Neurosci.* **29**, 15595 (2009).
- [56] W. L. Shew, H. Yang, S. Yu, R. Roy, and D. Plenz, Information capacity and transmission are maximized in balanced cortical networks with neuronal avalanches, *J. Neurosci.* **31**, 55 (2011).
- [57] J. Dambre, D. Verstraeten, B. Schrauwen, and S. Massar, Information processing capacity of dynamical systems, *Sci. Rep.* **2**, 514 (2012).
- [58] C. J. Honey, R. Kötter, M. Breakspear, and O. Sporns, Network structure of cerebral cortex shapes functional connectivity on multiple timescales, *Proc. Natl. Acad. Sci. USA* **104**, 10240 (2007).
- [59] M. G. Kitzbichler, M. L. Smith, S. R. Christensen, and E. Bullmore, Broadband criticality of human brain network synchronization, *PLoS Comput. Biol.* **5**, e1000314 (2009).
- [60] D. S. Bassett, D. L. Greenfield, A. Meyer-Lindenberg, D. R. Weinberger, S. W. Moore, and E. T. Bullmore, Efficient physical embedding of topologically complex information processing networks in brains and computer circuits, *PLoS Comput. Biol.* **6**, e1000748 (2010).
- [61] S. L. Bressler and V. Menon, Large-scale brain networks in cognition: Emerging methods and principles, *Trends Cognit. Sci.* **14**, 277 (2010).
- [62] B. J. He, J. M. Zempel, A. Z. Snyder, and M. E. Raichle, The temporal structures and functional significance of scale-free brain activity, *Neuron* **66**, 353 (2010).
- [63] M. Rubinov, O. Sporns, J.-P. Thivierge, and M. Breakspear, Neurobiologically realistic determinants of self-organized criticality in networks of spiking neurons, *PLoS Comput. Biol.* **7**, e1002038 (2011).
- [64] B. W. Mel, Information processing in dendritic trees, *Neural Comput.* **6**, 1031 (1994).
- [65] M. London and M. Häusser, Dendritic computation, *Annu. Rev. Neurosci.* **28**, 503 (2005).
- [66] J. Hawkins and S. Ahmad, Why neurons have thousands of synapses, a theory of sequence memory in neocortex, *Frontiers in Neural Circuits*, page 23, 2016.
- [67] 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, *Sci. Rep.* **7**, 18036 (2017).
- [68] J. H. Kirchner and J. Gjorgjieva, Emergence of synaptic organization and computation in dendrites, *Neuroforum* **28**, 21 (2022).
- [69] D. Beniaguev, I. Segev, and M. London, Single cortical neurons as deep artificial neural networks, *Neuron* **109**, 2727 (2021).
- [70] J. G. Elias, Artificial dendritic trees, *Neural Comput.* **5**, 648 (1993).
- [71] J. V. Arthur and K. Boahen, Recurrently connected silicon neurons with active dendrites for one-shot learning, in *Proceedings of the 2004 IEEE International Joint Conference on Neural Networks* (IEEE, 2004), Vol. 3, pp. 1699–1704.
- [72] J. Schemmel, L. Kriener, P. Müller, and K. Meier, An accelerated analog neuromorphic hardware system emulating nmda-and calcium-based non-linear dendrites, in *Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN)* (IEEE, 2017), pp. 2217–2226.
- [73] J. Kaiser, S. Billaudelle, E. Müller, C. Tetzlaff, J. Schemmel, and S. Schmitt, Emulating dendritic computing paradigms on analog neuromorphic hardware, *Neuroscience* **489**, 290 (2022).
- [74] N. Dahad, Exclusive: An interview with carver mead, 2022.
- [75] F. Zenke and S. Ganguli, Superspike: Supervised learning in multilayer spiking neural networks, *Neural Comput.* **30**, 1514 (2018).
- [76] A. Taherkhani, A. Belatreche, Y. Li, G. Cosma, L. P. Maguire, and T. M. McGinnity, A review of learning in biologically plausible spiking neural networks, *Neural Networks* **122**, 253 (2020).
- [77] G. Bellec, F. Scherr, A. Subramoney, E. Hajek, D. Salaj, R. Legenstein, and W. Maass, A solution to the learning dilemma for recurrent networks of spiking neurons, *Nat. Commun.* **11**, 3625 (2020).
- [78] T. Zhang, X. Cheng, S. Jia, M.-M. Poo, Y. Zeng, and B. Xu, Self-backpropagation of synaptic modifications elevates the efficiency of spiking and artificial neural networks, *Sci. Adv.* **7**, eabh0146 (2021).
- [79] H. Bergman, *The Hidden Life of the Basal Ganglia: At the Base of Brain and Mind* (MIT Press, 2021).
- [80] J. Guerguiev, T. P. Lillicrap, and B. A. Richards, Towards deep learning with segregated dendrites, *eLife* **6**, e22901 (2017).
- [81] B. A. Richards and T. P. Lillicrap, Dendritic solutions to the credit assignment problem, *Curr. Opin. Neurobiol.* **54**, 28 (2019).
- [82] F. A. Mikulasch, L. Rudelt, and V. Priesemann, Local dendritic balance enables learning of efficient representations in networks of spiking neurons, *Proc. Natl. Acad. Sci. USA* **118**, e2021925118 (2021).
- [83] S. Hodassman, R. Vardi, Y. Tugendhaft, A. Goldental, and I. Kanter, Efficient dendritic learning as an alternative to synaptic plasticity hypothesis, *Sci. Rep.* **12**, 1 (2022).
- [84] F. A. Mikulasch, L. Rudelt, M. Wibral, and V. Priesemann, Where is the error? hierarchical predictive coding through dendritic error computation, *Trends in Neurosciences*, 2022.
- [85] P. Poirazi and B. W. Mel, Impact of active dendrites and structural plasticity on the memory capacity of neural tissue, *Neuron* **29**, 779 (2001).
- [86] J. C. Magee and D. Johnston, Plasticity of dendritic function, *Curr. Opin. Neurobiol.* **15**, 334 (2005).
- [87] A. Losonczy, J. K. Makara, and J. C. Magee, Compartmentalized dendritic plasticity and input feature storage in neurons, *Nature (London)* **452**, 436 (2008).
- [88] P. J. Sjostrom, E. A. Rancz, A. Roth, and M. Häusser, Dendritic excitability and synaptic plasticity, *Physiol. Rev.* **88**, 769 (2008).

- [89] B. W. Mel, J. Schiller, and P. Poirazi, Synaptic plasticity in dendrites: Complications and coping strategies, *Curr. Opin. Neurobiol.* **43**, 177 (2017).
- [90] J. Bono and C. Clopath, Modeling somatic and dendritic spike mediated plasticity at the single neuron and network level, *Nat. Commun.* **8**, 706 (2017).
- [91] A. Payeur, J. Guerguiev, F. Zenke, B. A. Richards, and R. Naud, Burst-dependent synaptic plasticity can coordinate learning in hierarchical circuits, *Nat. Neurosci.* **24**, 1010 (2021).
- [92] S. dAquin, A. Szonyi, M. Mahn, S. Krabbe, J. Gründemann, and A. Lüthi, Compartmentalized dendritic plasticity during associative learning, *Science* **376**, eabf7052 (2022).
- [93] A. H. Marblestone, G. Wayne, and K. P. Kording, Toward an integration of deep learning and neuroscience, *Front. Comput. Neurosci.* **10**, 94 (2016).
- [94] K. Binder and A. P. Young, Spin glasses: Experimental facts, theoretical concepts, and open questions, *Rev. Mod. Phys.* **58**, 801 (1986).
- [95] M. Rabinovich, A. Volkovskii, P. Lecanda, R. Huerta, H. D. I. Abarbanel, and G. Laurent, Dynamical Encoding by Networks of Competing Neuron Groups: Winnerless Competition, *Phys. Rev. Lett.* **87**, 068102 (2001).
- [96] L. L. Gollo and M. Breakspear, The frustrated brain: From dynamics on motifs to communities and networks, *Philos. Trans. R. Soc. London B* **369**, 20130532 (2014).
- [97] M. Helias and D. Dahmen, *Statistical Field Theory for Neural Networks* (Springer, 2020), Vol. 413.
- [98] J. Halverson, A. Maiti, and K. Stoner, Neural networks and quantum field theory, *Mach. Learn. Sci. Technol.* **2**, 035002 (2021).
- [99] K. Friston, The free-energy principle: A unified brain theory?, *Nat. Rev. Neurosci.* **11**, 127 (2010).
- [100] V. A. Zinovyev, A. F. Zinovieva, Zh. V. Smagina, A. V. Dvurechenskii, V. I. Vdovin, A. K. Gutakovskii, L. I. Fedina, O. M. Borodavchenko, V. D. Zhivulko, and A. V. Mudryi, Si-based light emitters synthesized with Ge⁺ ion bombardment, *J. Appl. Phys.* **130**, 153101 (2021).
- [101] A. Higuera-Rodriguez, B. Romeira, S. Birindelli, L. E. Black, E. Smalbrugge, P. J. Van Veldhoven, W. M. M. Kessels, M. K. Smit, and A. Fiore, Ultralow surface recombination velocity in passivated ingaas/inp nanopillars, *Nano Lett.* **17**, 2627 (2017).
- [102] D. K. Schroder, *Semiconductor Material and Device Characterization* (Wiley, 2015).
- [103] D. Recht, F. Capasso, and M. J. Aziz, On the temperature dependence of point-defect-mediated luminescence in silicon, *Appl. Phys. Lett.* **94**, 251113 (2009).
- [104] R. F. Pierret and G. W. Neudeck, *Advanced Semiconductor Fundamentals* (Addison-Wesley Reading, MA, 1987), Vol. 6.
- [105] C. Chartrand, On the hunt for a near-infrared single photon emitter in silicon—Optical characterization of radiation damage defects in ²⁸Si, PhD. thesis, Science: Department of Physics, Simon Fraser University, 2019.
- [106] M. Lannoo, *Point Defects in Semiconductors I: Theoretical Aspects* (Springer Science & Business Media, 2012), Vol. 22.
- [107] K. G. Svantesson and N. G. Nilsson, The temperature dependence of the auger recombination coefficient of undoped silicon, *J. Phys. C: Solid State Phys.* **12**, 5111 (1979).
- [108] E. Yablonovitch, D. L. Allara, C. C. Chang, T. Gmitter, and T. B. Bright, Unusually Low Surface-Recombination Velocity on Silicon and Germanium Surfaces, *Phys. Rev. Lett.* **57**, 249 (1986).
- [109] T. Ytterdal, Y. Cheng, and T. A. Fjeldly, *Device Modeling for Analog and RF CMOS Circuit Design* (Wiley, 2003).
- [110] S. M. Buckley, A. N. Tait, J. Chiles, A. N. McCaughan, S. Khan, R. P. Mirin, S. W. Nam, and J. M. Shainline, Integrated-Photonic Characterization of Single-Photon Detectors for Use in Neuromorphic Synapses, *Phys. Rev. Appl.* **14**, 054008 (2020).