Nonlinear dynamics in neuromorphic photonic networks: Physical simulation in Verilog-A

Hugh Morison[,1,*] Jagmeet Singh[,1] Nayem Al Kayed[,1] A. Aadhi[,1] Maryam Moridsadat[,1]
Marcus Tamura[,1] Alexander N. Tait[,2] and Bhavin J. Shastri[1]

[1]*Centre for Nanophotonics, Department of Physics, Engineering Physics, and Astronomy, Queen's University,
64 Bader Lane, Kingston, ON K7L 3N6, Canada*

[2]*Department of Electrical and Computer Engineering, Queen's University, 19 Union Street, Kingston,
ON K7L 3N6, Canada*

Advances in silicon photonics technology have enabled the field of neuromorphic photonics, where analog neuron-like processing elements are implemented in silicon photonics technology. Accurate and scalable simulation tools for photonic integrated circuits are critical for designing neuromorphic photonic circuits. This is especially important when designing networks with recurrent connections, where the dynamics of the system may give rise to unstable and oscillatory solutions which need to be accurately modeled. These tools must simultaneously simulate the analog electronics and the multichannel (wavelength-division multiplexed) photonics contained in a photonic neuron to accurately predict on-chip behavior. In this paper, we utilize a Verilog-A model of the photonic neural network to investigate the dynamics of recurrent integrated circuits. We begin by reviewing the theory of continuous-time recurrent neural networks as dynamical systems and the relation of these dynamics to important physical features of photonic neurons such as cascadability. We then present the neural dynamics of systems of one and two neurons in the simulated Verilog-A circuit, which are compared to the expected dynamics of the abstract continuous-time recurrent neural network model. Due to the presence of parasitic circuit elements in the Verilog-A simulation, it is seen that there is a topological equivalence, but not an exact isomorphism, between the theoretical model and the simulated model. The implications of these discrepancies for the design of neuromorphic photonic circuits are discussed. Our findings pave the way for the practical implementation of large-scale silicon photonic recurrent neural networks.

I. INTRODUCTION

As computers based on the von Neumann architecture—computers with separate processor and memory (central, graphics, and tensor processing units [1])—get faster, and as the datasets used to train machine learning models get larger, training time for state-of-the-art models is increasingly consumed by memory access operations rather than computations [2,3]. Neuromorphic engineering offers the potential to surpass this von Neumann bottleneck faced by digital processors, but also importantly may enable new application regimes by directly modeling neurons at the hardware level [4–8]. There has been much recent interest in the accurate modeling and control of complex systems of nonlinear differential equations. Notably, the Nobel Prize in Physics in 2021 was awarded to Parisi, Hasselmann, and Manabe for contributions to the understanding of nonlinear dynamical systems [9]. Moreover, the team at DeepMind has shown that neural networks and machine learning techniques can be used to tackle the real-time prediction and control of fast and highly nonlinear dynamical systems such as the magnetic confinement and shaping of plasma in nuclear fusion reactors [10]. Neuromorphic circuits have a behavioral repertoire including neural control algorithms, as well as emulation of differential equations. Analog silicon photonic circuits have emerged as a promising platform to implement neuromorphic architectures with high bandwidth and low latency operation [11–15]. While analog systems offer many advantages, they are sensitive to a wide variety of physical and environmental parameters so high-accuracy simulation relies on physically detailed models that are validated with experimental observations. It was first shown in [16] how a network of silicon photonic neurons can be modeled as a continuous-time recurrent neural network (CTRNN).

The CTRNN is a dynamical model that can be applied to problems including differential equation emulation and neural control. CTRNNs are able to take on a wide range of dynamics by programming their interconnection weights

*hugh.morison@queensu.ca

and the nonlinear behavior of each neuron. The most widespread example of a CTRNN is the Hopfield network [17,18]. Hopfield networks naturally minimize a programmable energy function, and it has been shown that many control and optimization problems can be mapped to such an energy function [19–21]. Hopfield networks are considered suitable candidates for a hardware implementation of optimization problems, but the length of the convergence time for such problems is directly related to the system's feedback latency.

Silicon photonics achieves massively parallel, low latency transmission through the use of passive optical waveguides and wavelength-division multiplexing [22–24]. Additionally, it supports high-bandwidth processing via an electro-optic modulation scheme using active modulators with demonstrated bandwidth greater than 50 GHz [25,26]. Figure 1 depicts a schematic of the silicon photonic neuron model investigated in this work. Silicon photonic neurons using this architecture have been fabricated and tested, implementing both feedforward and recurrent neural networks [16,27–29], although fully integrated experimental demonstrations of these systems, and fast iteration of system design, remain challenging, in part due to the lack of a simulation platform that can replicate the dynamics of the on-chip system to predict the behavior of networks of neurons including the effect of the electrical parasitics involved in the experimental setup.

Previous work has developed a set of Verilog-A-based models of photonic components composing the electro-optic transfer function of the photonic neuron, and fitted the parameters of these models using experimental measurements [30]. We believe that the incorporation of such models in established electronics simulators will enable large-scale cosimulation of the electronics and photonics that are faithful to the underlying physics and account for parasitic elements of on-chip implementations. In this paper, we will use the previously developed photonic Verilog-A models [30] to simulate the physical behavior of networks of neurons. This work will demonstrate important neural dynamics in the simulated system and demonstrate the deviations between these simulated dynamics and the expected behavior of an abstract CTRNN. We begin by reviewing the dynamics of analog neural networks and the argument for the isomorphism between the CTRNN model and a photonic implementation of such a system. We then demonstrate these neural dynamics in a simulated version of a photonic neural network that incorporates all the parasitics present in experimental systems.

II. DYNAMICS OF CONTINUOUS-TIME RECURRENT NEURAL NETWORKS

A. Linear feedforward dynamics

A CTRNN is a neural network with neurons modeled by differential equations, such that a network of these neurons can be treated as a dynamical system. Neurons are typically arranged in a layer consisting of N neurons, where all neurons have the same input vector **x** and share the same structure, but are independently parameterized.

The most common model of a continuous neuron can be described as a leaky integrator—or a low-pass filter (LPF)—that integrates the weighted sum of its presynaptic connections. The evolution of the internal state of the ith LPF neuron s_i with time constant τ_i, bias b_i, and weight vector $\mathbf{w}_i$, is described by

$$\tau_i \dot{s}_i(t) = -(s_i(t) - b_i) + \mathbf{w}_i \cdot \mathbf{x}(t)$$
$$y_i(t) = \sigma_i \left[s_i(t) \right], \tag{1}$$

where $\dot{s}_i$ indicates the time derivative of the state. The activation function σ_i defines the nonlinear relationship between the neuron state and its postsynaptic output y_i. For biological plausibility, this function is typically taken to be a monotonic function such as a sigmoid.

A layer of neurons can be treated as a system of equations by vectorizing Eq. (1). The scalar parameters become vectors and a fully connected matrix $\mathbf{W}_x$ is formed from the set of weight vectors $\{\mathbf{w}_i\}$. We use the subscript x to indicate the incoming signal that the matrix operates on. The system of equations representing a layer of neurons is then

$$\tau \dot{\mathbf{s}}(t) = -(\mathbf{s}(t) - \mathbf{b}) + \mathbf{W}_x \cdot \mathbf{x}(t)$$
$$\mathbf{y}(t) = \sigma \left[\mathbf{s}(t) \right], \tag{2}$$

where we have assumed that all neurons in the layer have identical activation functions σ and time constants τ. As linear first-order differential equations with negative linear coefficients, this system can be seen to evolve toward its stable fixed point

$$\mathbf{s}^* = \mathbf{W}_x \cdot \mathbf{x} + \mathbf{b},$$
$$\mathbf{y}^* = \sigma \left[\mathbf{W}_x \cdot \mathbf{x} + \mathbf{b} \right]. \tag{3}$$

This is easily seen to be a stable system for all parameter values with exponential convergence of the state toward its fixed point, governed by the response time of the neurons τ. Note that the fixed point remains valid even for a nonstationary input x that varies at a timescale much less than τ (this will be seen in Sec. III A). While the output of the system is a nonlinear transformation of the state, if it is connected to a subsequent layer of neurons, this next layer still only has linear dynamics with respect to the output variable **y**. This model can be seen as a continuous-time version of a dense feedforward neural network. For sufficiently large N, a feedforward network can act as a universal function approximator [31], but without recurrent connections it cannot represent systems of coupled equations with feedback effects.

B. Nonlinear feedback dynamics

While the feedforward neuron is ensured to be stable, the dynamics of such a system becomes much richer through the introduction of feedback connections from the set of postsynaptic outputs of a layer, $\{y_i\}$, to all the presynaptic inputs of that same layer through a coupling matrix $\mathbf{W}_y$. In this case, there is a subsequent term added to Eq. (2) that introduces a nonlinear dependence on the current state of the layer into each neuron contained in the layer. The layer is now written as a system of coupled equations,

$$
\begin{aligned}
\tau \dot{\mathbf{s}}(t) &= -(\mathbf{s}(t) - \mathbf{b}) + \mathbf{W}_x \cdot \mathbf{x}(t) + \mathbf{W}_y \cdot \mathbf{y}(t) \\
&= -(\mathbf{s}(t) - \mathbf{b}) + \mathbf{W}_x \cdot \mathbf{x}(t) + \mathbf{W}_y \cdot \sigma \left[\mathbf{s}(t)\right].
\end{aligned} \quad (4)
$$

The capability of the system is now greatly increased; it has been shown that a CTRNN of the form of Eq. (4) can act as a universal approximator of not just any function, but of any smooth dynamics [32,33]. It is therefore unsurprising that it is no longer possible to seek closed-form solutions for the trajectory of the system in the general case, but it can be fruitful to investigate the fixed points and stability in specific well-defined cases. In particular, we will discuss the stability and bifurcations in the parameter space of single-neuron and two-neuron systems.

In this paper, unless otherwise specified, we will assume that all neurons have a sigmoidal activation function that can be parameterized with a scale factor α, a steepness β, and a baseline activation level γ:

$$
\sigma[s] = \frac{\alpha}{1 + \exp\left(-\beta(s - s_0)\right)} + \gamma \quad (5)
$$

The variable s_0 is used to center the activation function around a nonzero activation point. It is important to note that the forms of Eqs. (4) and (5) are different from other work on this photonic neuron architecture which has used Mach-Zehnder modulators and ac coupling between the balanced photodetector (BPD) and modulator [16]. The equations we have chosen more closely match the system under study but give rise to inelegant solutions, so we rely on numerical methods to solve the equations presented in the following sections. In many cases, we care about the dynamics of the output variable y rather than the state variable s (because it corresponds to a more readily observable variable—the optical power in a channel) but this can always be computed using the chain rule,

$$
\dot{y}_i \equiv \frac{\mathrm{d}y_i}{\mathrm{d}t} = \frac{\mathrm{d}y_i}{\mathrm{d}s_i}\frac{\mathrm{d}s_i}{\mathrm{d}t} = \sigma'_i[s_i]\,\dot{s}_i. \quad (6)
$$

1. Single-neuron networks

If we seek the fixed points to Eq. (4) in the case of a single neuron with a single scalar input x, using the activation function in Eq. (5), we have

$$
\begin{aligned}
0 &= -(s^* - b) + x + W_F \sigma\left[s^*\right] \\
&= -s^* + b + x + W_F \left(\frac{\alpha}{1 + \exp\left(-\beta(s^* - s_0)\right)} + \gamma\right),
\end{aligned} \quad (7)
$$

where we have used a feedback weight $\mathbf{W}_y = [W_F]$ and input weight $\mathbf{W}_x = [+1]$. This equation defines the nullcline for a single neuron. We cannot readily solve for s^* analytically but we can note certain properties in Eq. (7). Namely, this equation is the sum of a negative linear function of s^*, some constants, and the sigmoid. Since the sigmoid is an odd function, it can be approximated by a cubic, and solving for the fixed point s^* would be a matter of solving the resulting cubic equation. Under this cubic approximation, there are three roots; one of them is guaranteed to be real, while the other two may be imaginary. We can interpret the imaginary roots to be nonphysical solutions and the real roots to be the true fixed points. Whether the solutions are real or imaginary depends on the values of the parameters α, β, γ, and the weight W_F. In terms of the neural dynamics, this means that by varying the feedback weight, we can achieve between one and three different fixed points. In the case of three fixed points, only two will be stable while the third will be an unstable point. It has been shown in [16] how the observation of a transition from one stable fixed point (monostability) to two stable fixed points (bistability) is equivalent to a demonstration of the cascadability of the photonic neuron.

2. Two-neuron networks

Networks containing two neurons are the simplest testbed for demonstrating the coupled dynamics of the system. See [32] for a comprehensive overview of CTRNN dynamics; depending on the programmed weight matrix, two neuron networks can have up to nine qualitatively different phase portraits. Here, we will investigate two canonical examples of neural circuits that give rise to a stable system and an oscillatory system, respectively:

$$
\mathbf{W}_{\text{WTA}} = \begin{bmatrix} +1 & W_{\text{inh}} \\ W_{\text{inh}} & +1 \end{bmatrix}, \quad (8)
$$

$$
\mathbf{W}_{\text{Hopf}} = \begin{bmatrix} W_F & +1 \\ -1 & W_F \end{bmatrix}. \quad (9)
$$

Both matrices take the place of the recurrent matrix $\mathbf{W}_y$ in Eq. (4). In the symmetric matrix [Eq. (8)], the positive self-feedback and the inhibitory mutual feedback ($W_{\text{inh}} < 0$) give rise to the winner-take-all (WTA) system. The mutual inhibition guarantees that the neuron with the highest level of activation will suppress the activity of its counterpart

while the positive self-feedback ensures that it maintains its own level of activity. From a dynamics perspective, there are two basins of attraction in the state space around the points where one neuron is at maximum activity and the other is at minimum activity.

The asymmetric weight matrix [Eq. (9)] results in an unstable system. When the self-feedback parameterized by W_F is too low, the neurons will not have the activity required to reinforce themselves and they will settle at their stable fixed points, but as W_F increases, there is a bifurcation point where this stable state gives rise to an oscillatory solution (a stable limit cycle). This is known as the Hopf bifurcation.

C. Silicon photonic implementation

The model of the photonic modulator neuron seen in Fig. 1 is based on a microring modulator using electro-optic effects in p-n junction integrated silicon photonic waveguides. The nonlinearity σ_i is achieved through the electro-optic transfer function of the microring modulator [34]. Each neuron is tuned to modulate a different channel in the wavelength-division multiplexed waveguide bus, and the spectral position of the nonlinear Lorentzian transmission spectrum of the modulator is tuned by the applied voltage generated by the photocurrent.

In Eqs. (1)–(4), the state s_i is therefore related to the voltage across this p-n modulator V_i and the bias b_i is similarly related to a bias voltage $V_{b,i}$. The time constant of the system τ_i is governed by the bandwidth of the optical-electrical-optical conversion of the photonic neuron, which includes the p-n modulator and the photoreceiver impedance. When operating in reverse bias (carrier-depletion mode), p-n junction modulators act as a capacitive load with a cutoff frequency in the range of 40–70 GHz under typical bias conditions [25]. When operated in forward bias (carrier-injection mode), we can use the same dynamical equation but the time constant is instead limited to below 10 GHz by the slower rate of carrier recombination in the p-n junction [25,35]. In practice, a bandwidth of 10 GHz has been achievable with current technology [36].

All the weighted inputs to the LPF come from the microring weight bank [37] configuration seen in Fig. 1. The photocurrent generated by the BPD pair is proportional to the difference in the optical power applied to each photodetector, where the optical power is the sum over the power in each individual channel, P_j. The power is split between the two photodetectors using the transmissions at the THRU port (T_j) and DROP port (D_j) of a tunable filter bank implemented by the cascaded microring resonators (MRRs). In this way, the combination of the MRRs and BPD generate a photocurrent proportional to the weighted sum $\mathbf{w}_i \cdot \mathbf{x}$.

The resultant dimensional version of Eq. (1) is given by

$$\tau_i \dot{V}_i(t) = -(V_i(t) - V_{b,i}) + \sum_j \eta(D_j - T_j)P_j(t), \quad (10)$$
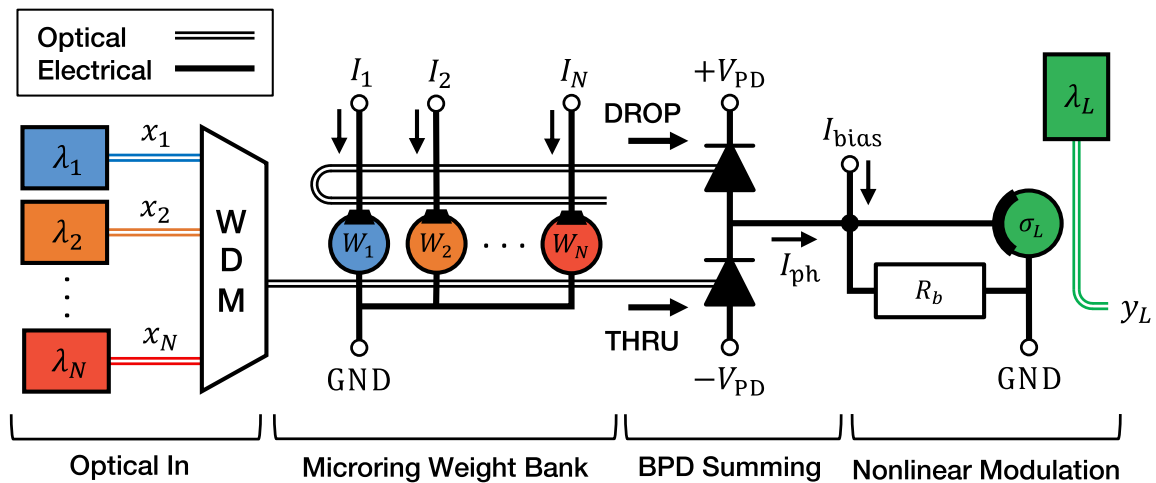

FIG. 1. Schematic of a silicon photonic neuron with N input channels. Each input signal x_i is amplitude modulated onto an optical carrier with wavelength λ_i (modulation not depicted) and then wavelength-division multiplexed (WDM) onto a single waveguide. Synapses are implemented with a bank of tunable add-drop microring filters. The labels I_i represent current sources applied to in-ring heaters used to tune each filter and give its corresponding channel weight W_i. The photocurrent I_{ph} generated by the balanced photodetector configuration is the result of the weighted sum across all the optical channels. Each neuron has electrical connections $\pm V_{\text{PD}}$ to bias the photodetectors. A p-n junction integrated microring modulator is used to apply a nonlinear activation function σ_L to the weighted sum, modulating a signal y_L onto a new optical carrier (here of wavelength λ_L, where $L \in [1, N]$). The modulator is biased using a current source with current I_{bias} to achieve a bias voltage $V_b = I_{\text{bias}}R_b$. The connection points show where bond pads introduce parasitics into the integrated circuit.

where η is the optical conversion efficiency (in units of volts per watt) which is related to the photodetector responsivity and the receiver impedance.

A layer of neurons is formed by tuning each neuron's modulator to an individual channel wavelength. An isomorphism between a photonic neural network using this architecture and the CTRNN system given in Eq. (4) has been shown in [16], although the accuracy of this result benefited from low-bandwidth operation and the filtering and modulation (and therefore the dynamics) being located off-chip. Fully integrated networks in silicon photonics technology will follow the same dominant dynamics, but experimental demonstrations deviate from the perfect CTRNN model due to a number of on-chip effects. Parasitic capacitance (from bond pads) and inductance (from wire bonds), as well as reactive components of the biasing circuitry, introduce higher-order terms with different time constants into Eq. (10). Feedback delays (assumed to be negligible in all the equations presented) can also produce spurious dynamics, and the electro-optic activation functions implemented by microring modulators are not arbitrarily programmable, so they deviate from idealized sigmoid and ReLU functions used in machine learning. For these reasons, using dynamical analysis of Eq. (4) is good as an approximation and to understand the capabilities of dynamical network models, but is insufficient for predicting on-chip behavior quantitatively.

III. VERILOG-A NEURAL NETWORKS

Our previous work has shown a methodology for modeling a photonic neuron in Verilog-A [30]. Verilog-A is a hardware description language for analog devices that is compatible with a broad range of circuit simulators (e.g., SPICE, Spectre [38,39]). Using Verilog-A enables accurate modeling of dynamic electro-optic effects and optimized cosimulation of the electronics present in photonic neuron circuits. We fit the parameter values of the electro-optic effect in p-n microring modulators and the thermo-optic effect in n-doped microring resonators to experimentally determined values so that the models reflect the behavior seen on-chip. The results of this simulation methodology proved to accurately model the on-chip behavior of a single photonic neuron.

In this section, we examine the signal-processing capability and dynamics of networks of simulated neurons in the presence of realistic parasitic components. All figures show a simplified circuit schematic. The full circuit includes bond pad capacitance, wire bond inductance, and a basic bias circuit; a detailed diagram is available in Appendix A. Section III A will show feedforward time-domain behavior of this neuron model with multiple input signals to demonstrate excitatory and inhibitory fan-in. Section III B will show the behavior of this single neuron in a self-feedback circuit to demonstrate the

indefinite cascadability of these neurons. Section III C will examine the dynamical properties of the neuron through the analysis of a Hopf bifurcation in a two-neuron system. Section III D shows that by reconfiguring the weights, the same two-neuron circuit can give rise to WTA dynamics.

A. Signal fan-in

The feedforward fixed point of a single neuron described by Eq. (3) implies that the system should be able to perform addition and subtraction of two input channels by configuring its weights. We can then apply a linear or nonlinear activation function to this computed result by configuring the bias to tune the transfer function of the modulator. Figure 2(a) shows the schematic for this simulation. We follow the procedure developed in [16] and prepare two input signals $x_1 = A(t)$ and $x_2 = B(t)$ on optical carriers with wavelengths λ_1 and λ_2, both amplitude modulated at 900 MHz. In this work, the weighting is implemented by the tunable thermo-optic effect of the microring weight bank. We configure W_1 with a weight of $+1$ and W_2 with a weight of ± 1 to implement addition and subtraction of the two inputs; the detailed weight mapping procedure is described in Appendix B. The modulator then implements either a linear or a quadratic activation function to converts the signal onto a carrier of wavelength λ_3.

Figure 2(b) shows the optical input signals $A(t)$ and $B(t)$ on the first two traces and the output signal $y(t)$ for a selection of different functions programmed into the neuron. The third trace $A + B$ shows the optical output of the neuron biased in the linear regime with $W_1 = W_2 = +1$. The fourth trace $A - B$ shows this same output after changing the weights to $W_1 = +1$ and $W_2 = -1$. The fifth $(A + B)^2$ and sixth $(A - B)^2$ traces show the same two cases for the weights, but with the neuron biased in the quadratic regime. Note the frequency doubling in the two quadratic cases.

Figure 2(c) shows the activation functions obtained by plotting the output optical power against the junction voltage and the linear and quadratic fits to this data. Qualitatively, we can see good agreement between the simulated results and the linear and quadratic functions. The results show slight nonlinearity in the linear activation function, as well as some hysteresis evident in the quadratic activation function, but we still achieve a root-mean-square error of 3.8% and 6.4% for the linear and quadratic fits, respectively.

These results demonstrate the substantial repertoire of behaviors of the feedforward photonic neuron. The MRR weights and BPD perform programmable excitatory and inhibitory fan-in of signals across multiple optical channels. The p-n junction MRR modulator demonstrates a configurable nonlinear conversion back into the optical domain operating at high bandwidth. Taken together, these
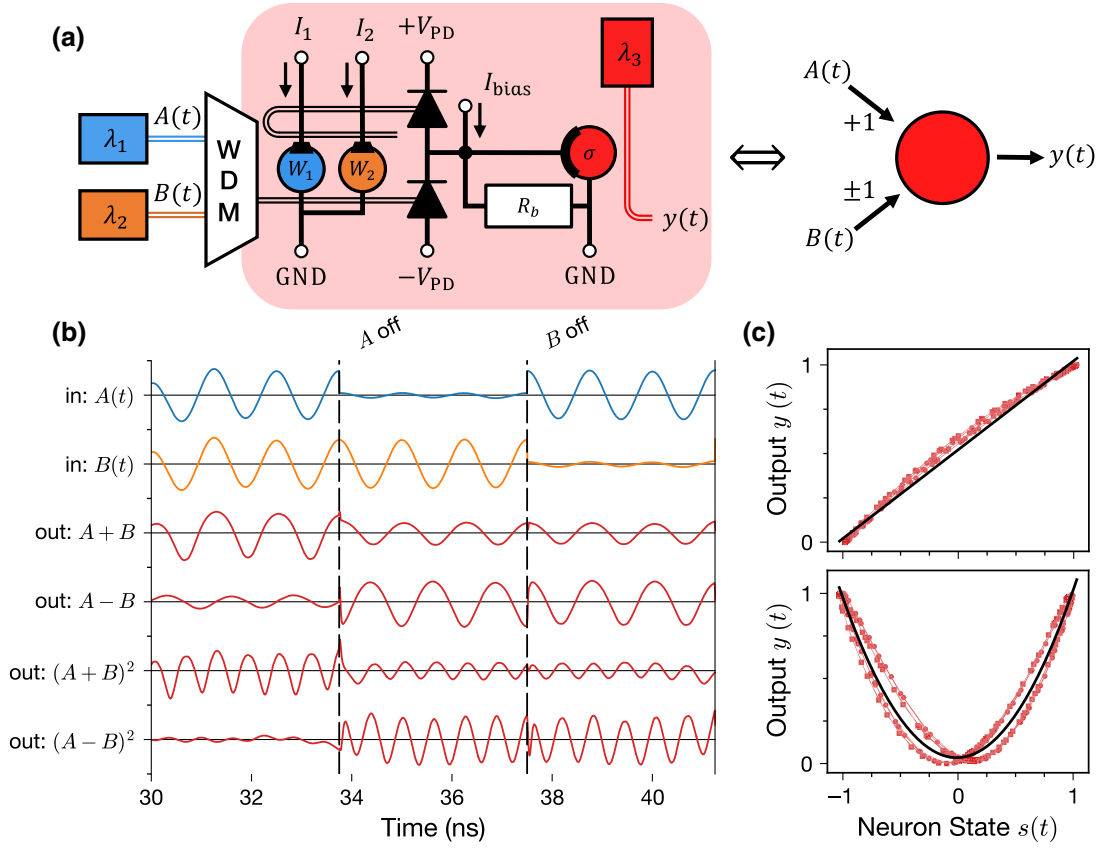
FIG. 2. Two-channel addition and subtraction, and nonlinear activation functions in a simulated photonic modulator neuron. (a) Schematic of Verilog-A simulation and equivalent conceptual schematic. (b) Input and output optical signals for various circuit configurations. A and B are input signals at λ_1 and λ_2. Output signals on λ_3 demonstrate four different operations applied to the inputs, achieved by programming the weights and activation function. All signals have a high-pass filter with a cutoff frequency of $400\,\mathrm{MHz}$ applied to remove the dc component and are scaled to emphasize the qualitative aspects of the results. (c) Linear and quadratic activation functions obtained by plotting the output power of the neuron against the voltage of the junction. Circular markers are used for $A + B$ and $(A + B)^2$ while square markers are used for $A - B$ and $(A - B)^2$. The black lines show fits to the data using a linear and quadratic functions.

results show that the feedforward photonic neuron can be used as building block of a CTRNN.

B. Cascadability

Using the same neuron circuit as in the feedforward case, connecting the output back to the input enables the self-feedback. For the following simulations, we configure the activation function to be a sigmoidal function, as described in Sec. II B. We configure the feedback delay to be $10\,\mathrm{ps}$, consistent with a $1-\mathrm{mm}$-long feedback waveguide. Figure 3(a) shows the simplified circuit diagram of this single-neuron CTRNN and its conceptual representation. We use a pump of wavelength λ_1 and tune the neuron modulator to this wavelength. To induce bistability in this circuit, we use a time-varying bias current and observe the optical output (postmodulation) on λ_1. Note that this is an atypical biasing configuration; static bias currents are typically used, but in this simulation it is simpler to use a

time-varying bias to highlight the dynamics that arises in the output optical power.

For this simulation, it is important that the signal be of sufficiently low frequency to ensure that the observed dynamics is a result of the self-feedback in the neuron and not a frequency-induced phase shift between the input and output. We consider three different cases for the signal on $I_{\mathrm{bias}}(t)$ to make this point. First, we use a $10-\mathrm{kHz}$ triangular wave; next, we use a $100-\mathrm{kHz}$ triangular wave to demonstrate that the effect is not frequency dependant; finally, we use a $10-\mathrm{kHz}$ asymmetric triangular wave to demonstrate that the effect has no dependence on the shape of the driving function. In all cases, we vary the self-feedback weight W_F from 0 to 1 to observe the effect of the feedback connection.

Figure 3(b) shows the results of the simulation for the $10-\mathrm{kHz}$ input signal. In the absence of feedback ($W_F = 0$), the system is clearly monostable. When the feedback is introduced ($W_F = 1$), we can see bistability indicated
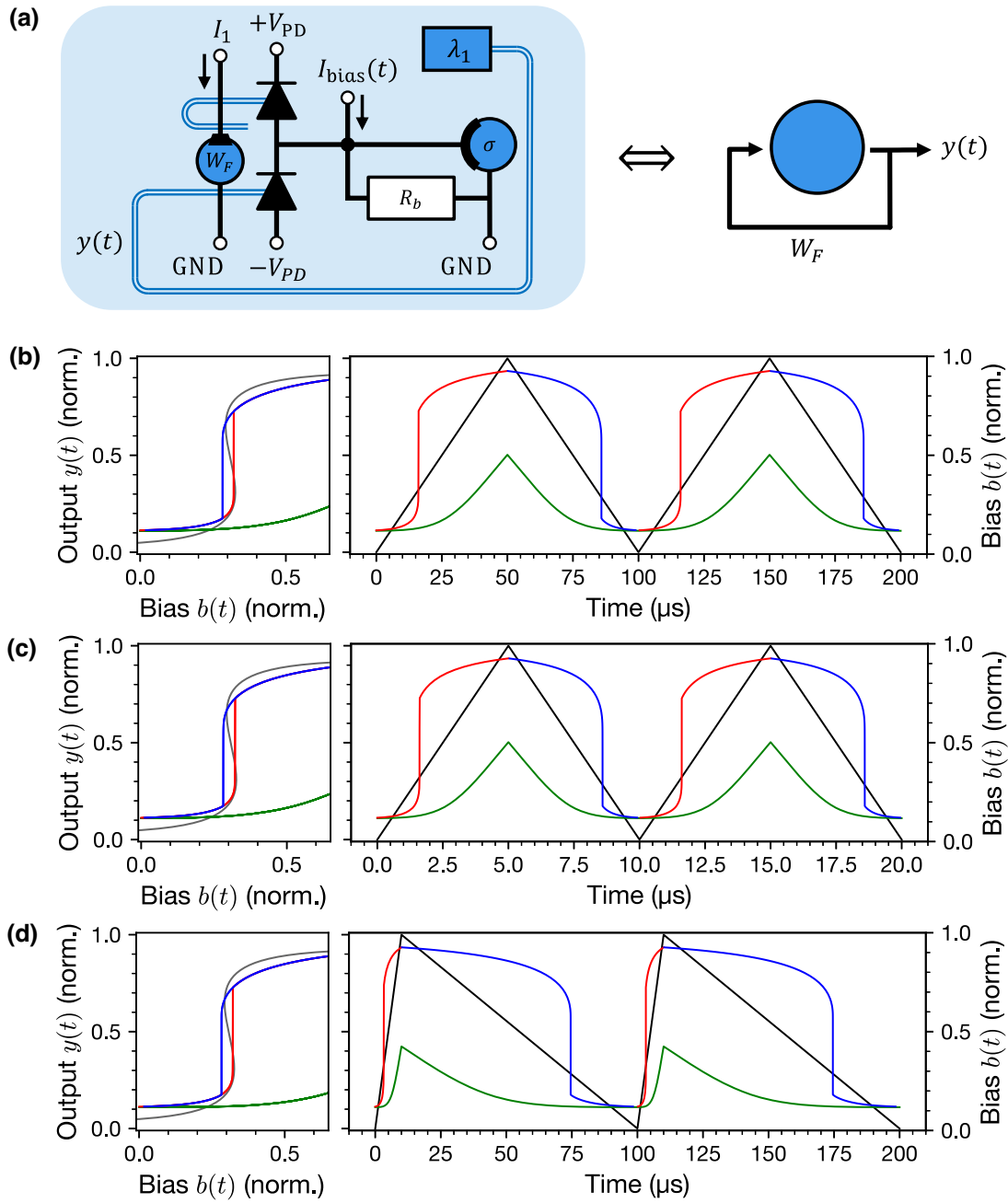
FIG. 3. Bistability in single self-afferent neuron. (a) Schematic of simulated recurrent circuit and equivalent conceptual schematic. (b)–(d) Relationship between neuron bias $b(t)$ and output $y(t)$ for three different bias signals. The output y on λ_1 when $W_F = 0$ is shown in green, while the red and blue traces show the rising and falling edges of y when $W_F = +1$. The y versus b plots on the left shows the signals along with the system's nullcline (gray trace). The transient plots on the right shows the same signals along with the bias waveform (black trace). Optical powers are normalized using $y(t) = P(t)/P_{\text{pump}}$, and the bias current is normalized using $b(t) = (I(t) - I_{\min})/(I_{\max} - I_{\min})$.

by the separation between the rising and falling edges of the output which follows the behavior predicted by Eq. (7). Figure 3(c) shows the results of the simulation for the $100-$kHz input signal. Both the monostable and bistable behavior are identical to the $10-$kHz case, as expected. Again, the bistability is evident from the hysteresis of

the bias-output relationship when $W_F = 1$. Figure 3(d) demonstrates how the time-domain behavior can vary considerably but the phase space representation again looks identical.

As expected, in all cases, the solution is monostable for $W_F \approx 0$. This is analogous to the feedforward results
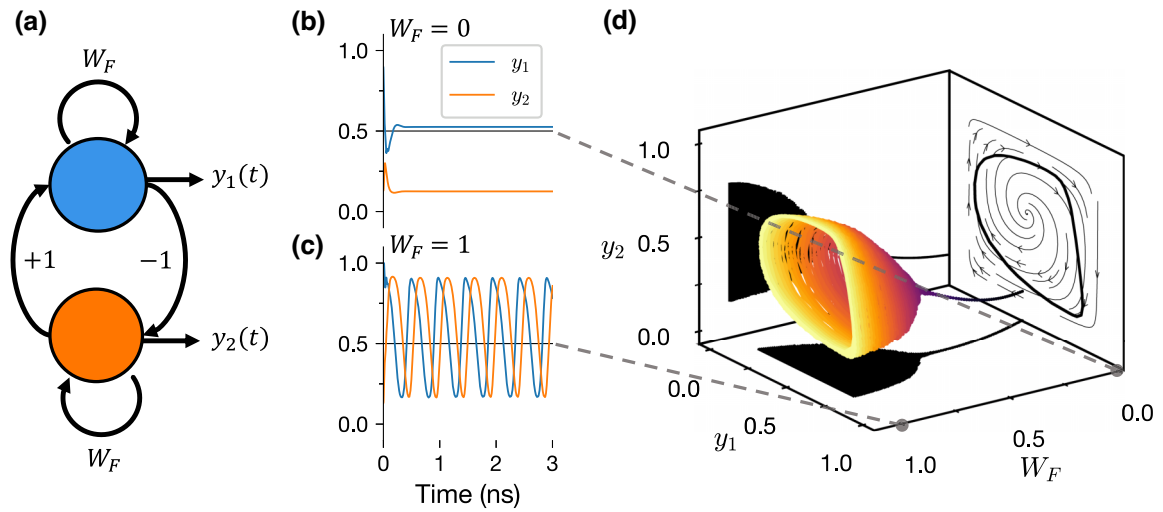
FIG. 4. (a) Conceptual representation of the two-neuron circuit showing the weight values. (b),(c) Transient response of neuron outputs for $W_F = 0$ and $W_F = 1$, respectively. The outputs y_i are given by the normalized output power from each modulator. (d) Three-dimensional bifurcation diagram of the Hopf system. The stability transitions from a stable fixed point to a stable limit cycle as W_F is varied. Black shadows show the projections of simulated data onto the y_i-W_F plane. The projection onto the y_1-y_2 plane shows both the expected dynamics as a vector field, and the slice of the simulated data where $W_F = 1$.

presented in Sec. III A; the output is just a nonlinear function of the neuron state defined by the neuron's activation function, and there is negligible frequency dependence. As W_F is increased, the results imply a bifurcation point where the monostable response gives way to this bistable response.

C. Hopf bifurcation

The two-dimensional system provides the simplest test circuit to demonstrate the interaction between neurons while still having easily interpreted phase portraits, as described in Sec. II B 2. In Fig. 4(a), we depict the simple conceptual diagram of the network, with weights configured as shown in Eq. (9). To observe the bifurcation in this circuit, we do not introduce any modulated input signal, as was done in the previous simulation. Instead, two lasers with wavelengths λ_1 and λ_2 are used as pumps and a time dependence of the modulation occurs due to the feedback dynamics of the circuit, governed by the weights and the activation functions. We observe the postmodulation signals over time while sweeping feedback weights W_F from 0 to 1.

The transient results for $W_F = 0$ and $W_F = 1$ are shown in Figs. 4(b) and 4(c), respectively. In Fig. 4(d), we see the bifurcation in the phase portrait of the system. When the feedback connection to each neuron is low, the system quickly decays to its single stable fixed point, regardless of its starting position.

Notice how the coordinate of the stable fixed point has a nonlinear relationship with W_F. Eventually, for large enough W_F, there is a bifurcation point where the fixed point evolves into a stable limit cycle. The amplitude of

this limit cycle increases by increasing the strength of each neuron's self-feedback W_F. The vector field projected on the y_1-y_2 plane in Fig. 4(d) is a plot of Eq. (4) using weight matrix Eq. (9) and an activation function fit to the results. The simulated trajectory follows the dynamics of the model but there are notable deviations and asymmetries which arise from the combination of all the physical effects present in the simulation. This deviation in the simple two-neuron system highlights the importance of physical simulation needed for accurately predicting on-chip experimental results.

D. Winner-take-all

Using the same simulated circuit, by reconfiguring the weights to the matrix in Eq. (8), we can change the system dynamics. The conceptual diagram in Fig. 5(a) depicts the system with the new weights. The mutual inhibitory weight W_{inh} is configured to be -1, so that each neuron strongly inhibits the other. For the WTA simulation, we add two more lasers at λ_3 and λ_4 for the inputs x_1 and x_2, which independently perturb neurons 1 and 2. Since the neurons are configured identically apart from their weights, the winning neuron is decided completely by which of them has the stronger input.

Figure 5(b) depicts this behavior. The inputs x_i are square waves that implement a truth table. Initially, by random chance, neuron 1 (blue) has a slightly higher output, so its activation function saturates (it wins) and it suppresses neuron 2 (orange). When x_1 goes low and x_2 remains high (around 5 ns), the outputs switch. When both inputs go low (around 15 ns), the system is already in the basin of attraction around neuron 1 and so it stays there;
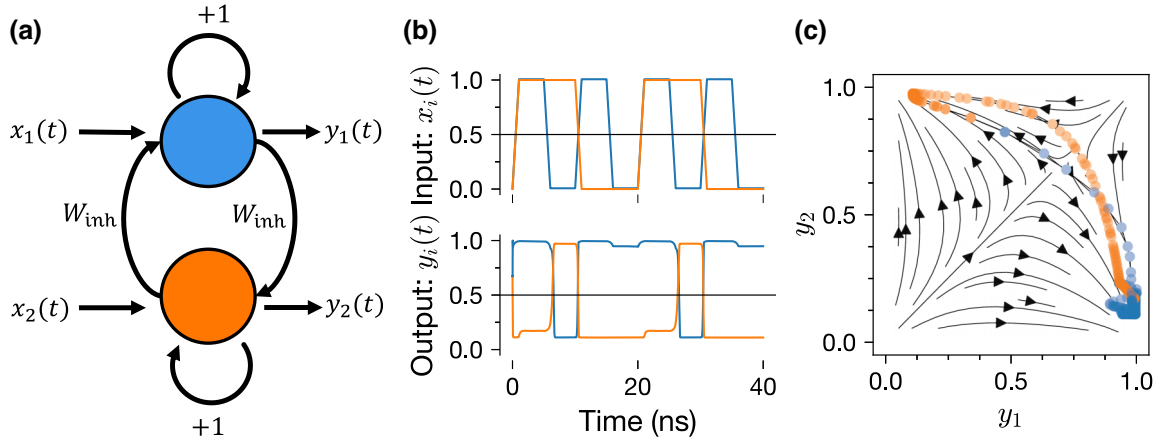
FIG. 5. (a) Conceptual circuit diagram of winner-take-all simulation showing programmed weights and output signals. (b) Transient input and output signals for each of the two neurons. (c) State-space plot of outputs, where the color of the marker corresponds to the neuron with the greater input (the expected winner). The vector field depicts the expected dynamics in the absence of any input signal.

the system has a memory of its last state. Figure 5(c) shows these results as a phase portrait. The vector field is a plot of Eq. (4) using weight matrix Eq. (8) in the absence of any input ($x_1 = x_2 = 0$) which shows the two basins of attraction on either side of the line $y_1 = y_2$. The data points show the trajectory of the system in response to the input perturbations; each point is colored based on the higher input at that time (the expected winner). The high density of orange and blue points in their respective basins of attraction demonstrate the WTA dynamics.

IV. CONCLUSION

In this work we have reviewed the dynamics of continuous-time recurrent neural networks implemented in silicon photonics. We have explored their fixed points and stability in both feedforward and feedback configurations, and presented a simplified model of how the dynamics of the electro-optic circuit can be mapped to the dynamics of a CTRNN. We then demonstrated these dynamics through a few example problems using previously developed behavioral models written in Verilog-A, including all relevant parasitic effects present in experimental demonstrations.

The results indicate that there can be significant deviations from the CTRNN model in the presence of these parasitics. There is qualitative agreement between the simulated results and the model, but we do not see quantitative agreement. Therefore, we can conclude that while the simulation model and the CTRNN are not isomorphic, they share the important dynamical features of their state space; we can say they have an approximate topological equivalence. These results motivate the suitability of silicon photonic neurons for the implementation of recurrent neural networks where the topological behavior is more important than the quantitative result. They also motivate the investigation of programming methods that consider the

parasitics to compensate for the differences and improve the quantitative reliability of the system. Finally, these deviations support using physical simulation as a tool for predicting and analyzing the results of future on-chip experiments.

Future work on the Verilog-A based models should focus on three main areas: first, the implementation of larger-scale systems and training of simulated neural networks to implement optimization and control problems; second, the incorporation circuit schematics for experimental packaging circuits and CMOS control systems; and third, the integration with chip layout tools. Automated generation of SPICE netlists using these Verilog-A models will allow for an interface between the simulator and established neural network packages, such as PyTorch or Nengo. Combining these simulations with schematics of packaging circuits will allow for a further level of physical verification between simulation and experiment. Finally, integration with layout tools will allow for automated generation of chip layouts based on simulated circuits, bringing the silicon photonics design flow closer to the established electronic design automation process in digital electronics. In sum, this work will pave the way for experimental realization of large-scale silicon photonic recurrent neural networks.

APPENDIX A: VERILOG-A PHOTONIC NEURON CIRCUIT

1. Photonic Verilog-A models

The simulation methodology used to simulate photonic circuits using Verilog-A is covered extensively in [30].

The approach taken is to treat the optical signals as three-element data structures containing the real and imaginary electric field amplitudes along with their wavelength,

$$E = \left(E_{\text{real}}, E_{\text{imag}}, \lambda\right).$$

This is consistent with the slowly varying envelope approximation. This approach discards information about the time-dependent oscillations of the electric field, but it is sufficient for the purposes of this work because it retains information about the optical power and the phase of the signal. Treating the optical signals of different wavelengths as distinct data structures allows for the simulation of wavelength-division multiplexed systems and is suitable so long as we are not dealing with nonlinear optical effects.

a. Waveguides

The waveguide model is assumed to be linear, with a complex index of refraction $\tilde{n}_0 = n_0 + i\alpha_0$, to account for phase shift and loss. The model converts the real and imaginary parts of the electric field into polar form, applies the loss and phase shift, and then converts back to rectangular form.

b. Phase shifters

The phase shifter model is an augmented waveguide where input voltages and currents can change the values of the real and imaginary parts of the index of refraction. Specifically, we use experimentally validated models for the change in refractive index due to either a heater current I_{heat} for the n-doped heater-integrated waveguides, or a modulator voltage V_{mod} for the p-n junction microring modulators. These are combined into a compound model,

$$\tilde{n}(I_{\text{heat}}, V_{\text{mod}}) = \tilde{n}_0 + \Delta\tilde{n}(I_{\text{heat}}) + \Delta\tilde{n}(V_{\text{mod}}),$$

where only one of the two effects is used at a time depending on the phase shifter being modeled.

c. Couplers

The coupler is a four-port model that uses a transfer matrix to calculate the output signals based on the input signals. The transfer matrix is assumed to be lossless and is parametrized by the coupling coefficient k,

$$\begin{bmatrix} E_{\text{out},1} \\ E_{\text{out},2} \end{bmatrix} = \begin{bmatrix} t & -ik \\ -ik & t \end{bmatrix} \begin{bmatrix} E_{\text{in},1} \\ E_{\text{in},2} \end{bmatrix},$$

where $t = \sqrt{1 - k^2}$ is the transmission coefficient, assuming no coupling loss.

d. Microring resonators

The MRR model is constructed using the previously described models of the coupler, waveguide, and phase
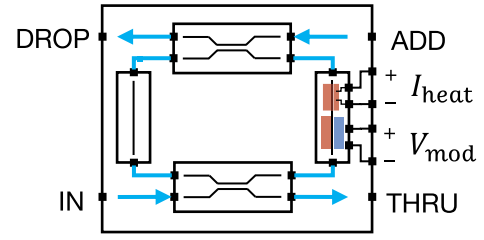

FIG. 6. A schematic representation of the Verilog-A model of a microring resonator. The model has four optical ports—two input ports (IN and ADD) and two output ports (DROP and THRU). The model consists of two directional couplers, a waveguide, and a phase shifter which has four electrical ports for the positive and negative terminals of I_{heat} and V_{mod}.

shifter. The schematic of the model is shown in Fig. 6. It is possible to use either the n-doped heater or the p-n junction modulator as the phase shifter, depending on which set of electrical ports on the phase shifter are used.

e. Photodetectors

The photodetector model performs the electro-optic conversion from optical power to electrical current. The model is constructed from a behavioral current source in parallel with a junction capacitance and a shunt resistor, all in series with a small resistance. The electro-optic conversion takes place in the current source, where the current is given by the sum across all incoming optical channels,

$$I_{\text{ph}}(t) = \eta \sum_i (|E_{\text{real},i}|^2 + |E_{\text{imag},i}|^2),$$

where η is the quantum efficiency of the photodetector.

2. Neuron circuit model

The electro-optic circuits shown in the main text are simplified circuit diagrams to show the functionality, without the details of the biasing circuits and parasitic components. The complete circuit diagram of a single-neuron, including the custom Verilog-A models, and active and passive circuit components coming from the typical experimental setup, is shown in Fig. 7. The optical waveguides are not shown here.

The dominant parasitic elements come from the circuit nodes that are connected off-chip using bond pads. We model the bond pads as a capacitance C_p, and the off-chip wire bond connection as an inductance L_b. The bond pad capacitance is $C_p = 10\,\text{fF}$. The wire bond inductance is taken to be $L_b = 1\,\text{nH}$. These values are typical for the bond pads and wire bonds.

We also design a simple bias tee made from an inductor $L_T = 1\,\text{mH}$ and a capacitor $C_T = 2\,\mu\text{F}$. A bias tee is a three-port circuit that allows separation of the dc and ac components of a signal. This is necessary for using off-chip
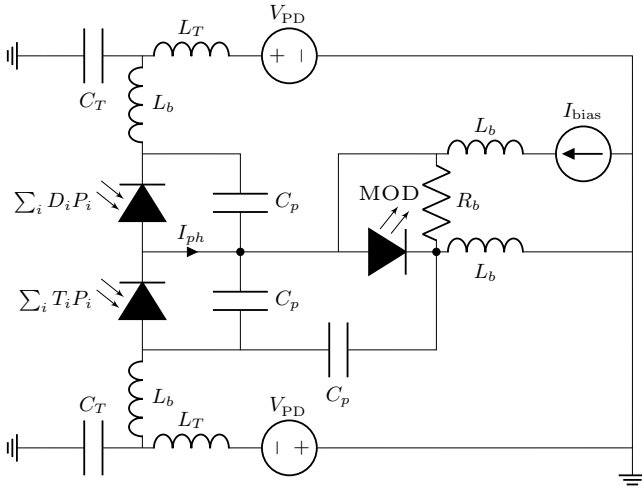
FIG. 7. Electrical circuit diagram of a single-neuron circuit. The optical interconnects of the system are omitted since these vary for each simulation. Diode symbols are used to represent the Verilog-A behavioral models of photonic components. The balanced photodetector configuration generates the current I_{ph} proportional to the optical power P_i on each channel. The microring modulator *MOD* is represented using an LED symbol for simplicity. C_p and L_b are the parasitic bond pad capacitance and wire bond inductance, respectively. R_b is the bias resistor. C_T and L_T are the bias tee capacitor and inductor, respectively. V_{PD} is the photodetector bias voltage. I_{bias} is the modulator bias current. I_{ph} is the photocurrent generated by the balanced photodetectors.

voltage and current sources to bias the photodetectors and the modulator, while still allowing the time-varying photocurrent to remain on-chip to modulate the output optical signal. This bias tee design is far from optimal and does not use realistic component selection, but is functional for the signals of interest in this work.

APPENDIX B: WEIGHT MAPPING PROCEDURE

In order to implement an arbitrary weight vector using a bank of microring resonators, it is necessary to perform a characterization of the microrings for a given channel (carrier wavelength) and ring radius of interest. In this Appendix, we demonstrate feedforward characterization of a five-channel MRR weight bank. We choose here the base radius of the first ring in the weight bank to be $R_0 = 8\,\mu m$ and increment the radius for each ring by $\delta R = 12\,nm$. Similarly, the carrier wavelengths of the first channel are chosen to be $\lambda_0 = 1548.7\,nm$ and increment the wavelength by $\delta\lambda = 2.35\,nm$.

After constructing a neuron as in Fig. 1 with these parameters, we set the heater current for each ring to 0, sweep the input wavelength and measure the voltage across the modulator. and obtain the spectrum in Fig. 8.

Next, for each channel independently, we can set the pump laser to λ_i, and sweep the input heater current from 0 to 1 mA and again measure the voltage at the neuron
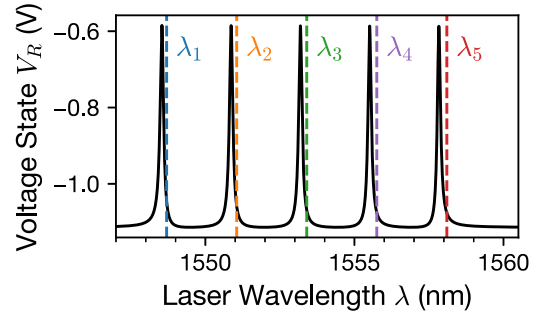

FIG. 8. Spectrum of voltage measured at the neuron junction for a weight bank with $R_i = R_0 + i\delta R$. The vertical lines correspond to the chosen channels at $\lambda_i = \lambda_0 + i\delta\lambda$.

junction to obtain the individual resonance curve of each ring. We can then determine the value of the applied current corresponding to bringing the ring onto resonance with the given channel, as seen in Fig. 9(a). This is the current for each channel that corresponds to the most light being coupled to the DROP port. Note the asymmetry in the response; there is no bias current supplied and yet most of the resonance curve lies below 0 V.

Finally, we discard all the data above the resonance current, truncate the curve so that it is symmetric about
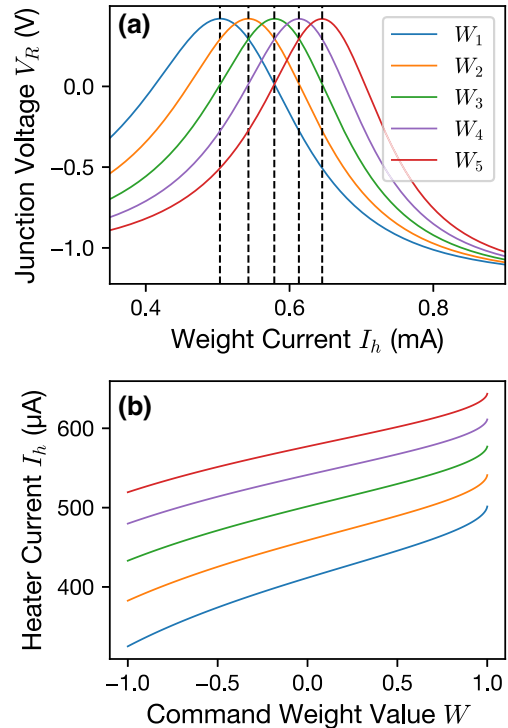

FIG. 9. (a) Junction voltage response to a sweep of weight tuning current for each channel. (b) Final weight mapping curve for each weight channel. For any desired weight $-1 <= W_i <= +1$, we have a unique heater current that we can apply to achieve this.

0 V, and invert the relationship between heater current and junction response. The final result can be seen in Fig. 9(b), where we have now normalized the response so that the maximum voltage corresponds to $W = -1$ and the minimum corresponds to $W = +1$.

It is important to note that this weight mapping procedure is the theoretically optimal way to map heater currents to MRR weights, and is fit for the purpose of simulation, but it does not account for many of the important physical effects that limit the precision in experiment. Fabrication variation, environmental temperature fluctuations, and photoconductivity of *n*-doped heaters are some of the effects that are not yet accounted for in the model, and these are left for future work.

[1] Y. E. Wang, G. Y. Wei, and D. Brooks, Benchmarking TPU, GPU, and CPU platforms for deep learning, ArXiv:1907.10701 (2019).

[2] J. Backus, Can programming be liberated from the von Neumann style? A functional style and its algebra of programs, Commun. ACM **21,** 613 (1978).

[3] R. S. Williams, What's next? [The end of Moore's law], Comput. Sci. Eng. **19,** 7 (2017).

[4] J. Hasler and H. Marr, Finding a roadmap to achieve large neuromorphic hardware systems, Front. Neurosci. **7,** 1 (2013).

[5] K. Berggren, *et al.*, Roadmap on emerging hardware and technology for machine learning, Nanotechnology **32,** 012002 (2021).

[6] M. Davies, *et al.*, Loihi: A neuromorphic manycore processor with on-chip learning, IEEE Micro **38,** 82 (2018).

[7] P. A. Merolla, *et al.*, A million spiking-neuron integrated circuit with a scalable communication network and interface, Science **345,** 668 (2014).

[8] B. V. Benjamin, P. Gao, E. McQuinn, S. Choudhary, A. R. Chandrasekaran, Jean-Marie Bussat, R. Alvarez-Icaza, J. V. Arthur, P. A. Merolla, and K. Boahen, Neurogrid: A mixed-analog-digital multichip system for large-scale neural simulations, Proc. IEEE **102,** 699 (2014).

[9] A. R. Ravishankara, D. A. Randall, and J. W. Hurrell, Complex and yet predictable: The message of the 2021 Nobel Prize in Physics, Proc. Natl. Acad. Sci. **119,** e2120669119 (2022).

[10] J. Degrave, *et al.*, Magnetic control of tokamak plasmas through deep reinforcement learning, Nature **602,** 414 (2022).

[11] D. Thomson, A. Zilkie, J. E. Bowers, T. Komljenovic, G. T. Reed, L. Vivien, D. Marris-Morini, E. Cassan, L. Virot, J. M. Fédéli, and J. M. Hartmann, Roadmap on silicon photonics, J. Opt. **18,** 073003 (2016).

[12] P. R. Prucnal and B. J. Shastri, *Neuromorphic Photonics* (CRC Press, Boca Raton, FL, 2017).

[13] T. F. de Lima, B. J. Shastri, A. N. Tait, M. A. Nahmias, and P. R. Prucnal, Progress in neuromorphic photonics, Nanophotonics **6**, 577 (2017).

[14] T. F. de Lima, Hsuan-Tung Peng, A. N. Tait, M. A. Nahmias, H. B. Miller, B. J. Shastri, and P. R. Prucnal, Machine learning with neuromorphic photonics, J. Lightwave Technol. **37,** 1515 (2019).

[15] B. J. Shastri, A. N. Tait, T. Ferreira 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).

[16] A. N. Tait, T. F. de Lima, E. Zhou, A. X. Wu, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, Neuromorphic photonic networks using silicon photonic weight banks, Sci. Rep. **7,** 7430 (2017).

[17] J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proc. Natl. Acad. Sci. **79,** 2554 (1982).

[18] H. Ramsauer, B. Schäfl, J. Lehner, P. Seidl, M. Widrich, T. Adler, L. Gruber, M. Holzleitner, M. Pavlovic, G. K. Sandve, V. Greiff, D. Kreil, M. Kopp, G. Klambauer, J. Brandstetter, and S. Hochreiter, Hopfield networks is all you need, ArXiv:2008.02217 (2020).

[19] J. J. Hopfield and D. W. Tank, Computing with neural circuits: A model, Science **233,** 625 (1986).

[20] M. Kennedy and L. Chua, Neural networks for nonlinear programming, IEEE Trans. Circuits Syst. **35,** 554 (1988).

[21] U.-P. Wen, K.-M. Lan, and H.-S. Shih, A review of Hopfield neural networks for solving mathematical programming problems, Eur. J. Oper. Res. **198,** 675 (2009).

[22] J. W. Goodman, Fan-in and fan-out with optical interconnections, Opt. Acta: Int. J. Opt. **32,** 1489 (1985).

[23] A. N. Tait, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, Broadcast and weight: An integrated network for scalable photonic spike processing, J. Lightwave Technol. **32,** 4029 (2014).

[24] D. Pérez, I. Gasulla, L. Crudgington, D. J. Thomson, A. Z. Khokhar, K. Li, W. Cao, G. Z. Mashanovich, and J. Capmany, Multipurpose silicon photonics signal processor core, Nat. Commun. **8,** 636 (2017).

[25] L. Chrostowski and M. Hochberg, *Silicon Photonics Design: From Devices to Systems* (Cambridge University Press, Cambridge, UK, 2015).

[26] D. W. U. Chan, X. Wu, Z. Zhang, C. Lu, A. P. T. Lau, and H. K. Tsang, C-band 67 GHz silicon photonic microring modulator for dispersion-uncompensated 100 Gbaud PAM-4, Opt. Lett. **47,** 2935 (2022).

[27] W. Zhang, J. C. Lederman, T. Ferreira de Lima, J. Zhang, S. Bilodeau, L. Hudson, A. Tait, B. J. Shastri, and P. R. Prucnal, A system-on-chip microwave photonic processor solves dynamic RF interference in real time with picosecond latency, Light: Sci. Appl. **13,** 14 (2024).

[28] B. A. Marquez, J. Singh, H. Morison, Z. Guo, L. Chrostowski, S. Shekhar, P. Prucnal, and B. J. Shastri, Fully-integrated photonic tensor core for image convolutions, Nanotechnology **34,** 395201 (2023).

[29] C. Huang, S. Fujisawa, T. F. de Lima, A. N. Tait, E. C. Blow, Y. Tian, S. Bilodeau, A. Jha, F. Yaman, Hsuan-Tung Peng, H. G. Batshon, B. J. Shastri, Y. Inada, T. Wang, and P. R. Prucnal, A silicon photonic–electronic neural network for fibre nonlinearity compensation, Nat. Electron. **4,** 837 (2021).

[30] J. Singh, H. Morison, Z. Guo, B. A. Marquez, O. Esmaeeli, P. R. Prucnal, L. Chrostowski, S. Shekhar, and B. J. Shastri, Neuromorphic photonic circuit modeling in Verilog-A, APL Photonics **7,** 046103 (2022).

[31] K. Hornik, M. Stinchcombe, and H. White, Multilayer feedforward networks are universal approximators, Neural Netw. **2**, 359 (1989).

[32] R. D. Beer, On the dynamics of small continuous-time recurrent neural networks, Adapt. Behav. **3**, 469 (1995).

[33] X. D. Li, J. Ho, and T. Chow, Approximation of dynamical time-variant systems by continuous-time recurrent neural networks, IEEE Trans. Circuits Syst. II: Express Briefs **52**, 656 (2005).

[34] 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).

[35] K. Liu, C. R. Ye, S. Khan, and V. J. Sorger, Review and perspective on ultrafast wavelength-size electro-optic modulators, Laser Photon. Rev. **9**, 172 (2015).

[36] T. F. de Lima, A. N. Tait, H. Saeidi, M. A. Nahmias, Hsuan-Tung Peng, S. Abbaslou, B. J. Shastri, and P. R. Prucnal, Noise analysis of photonic modulator neurons, IEEE J. Sel. Top. Quantum Electron. **26**, 1 (2020).

[37] W. Zhang, C. Huang, Hsuan-Tung Peng, S. Bilodeau, A. Jha, E. Blow, T. F. de Lima, B. J. Shastri, and P. Prucnal, Silicon microring synapses enable photonic deep learning beyond 9-bit precision, Optica **9**, 579 (2022).

[38] R. Pratap, V. Agarwal, and R. K. Singh, in *2014 International Conference on Power, Control and Embedded Systems (ICPCES)* (IEEE, Allahabad, India, 2014), p. 1.

[39] C. C. McAndrew, in *2017 47th European Solid-State Device Research Conference (ESSDERC)* (IEEE, Leuven, Belgium, 2017), p. 22.