Universal Self-Correcting Computing with Disordered Exciton-Polariton Neural Networks

Huawen Xu,[1,*] Sanjib Ghosh,[1] Michal Matuszewski,[2] and Timothy C.H. Liew[1,†]

[1] *School of Physical and Mathematical Sciences, Nanyang Technological University 637371, 637371, Singapore*
[2] *Institute of Physics, Polish Academy of Sciences, Al. Lótnikow 32/46, Warsaw PL-02-668, Poland*

(Received 1 November 2019; revised manuscript received 11 February 2020; accepted 13 May 2020; published 30 June 2020)

We show theoretically that neural networks based on disordered exciton-polariton systems allow the realization of Toffoli gates. Noise in input signals is self-corrected by the networks, such that the obtained Toffoli gates are in principle cascadable, where their universality would allow for arbitrary circuits without the need of additional error-correcting codes. We further find that the exciton-polariton reservoir computers can directly simulate composite circuits, such that they are a highly efficient platform allowing circuits to operate in a single step, minimizing the delay of signal transport between elements and error-correction overhead.

DOI: 10.1103/PhysRevApplied.13.064074

I. INTRODUCTION

Electronic systems have been highly successful in following Moore's law for decades; however, several challenges affect the future continuation of the trend. More densely packed transistor chips have led to heating becoming a key limitation to clock speed, while efficiency is limited by a significant error-correction overhead and the delay time needed to pass signals between memory and processing elements (the von Neumann bottleneck). Optical systems are promising for making fast interconnections to alleviate these problems [1], however, the use of optics to actually process information requires improvements of energy consumption to subfemtojoule scale or less to become competitive [2].

Polaritonic devices [3] hybridize photons and artificial atoms in the form of excitons in semiconductors. This captures both the weak dephasing and ultrafast response of an optical system together with the nonlinearity of a condensed-matter system. Research has focused on using polaritonic systems for the development of information-processing elements and methods to link them. Polaritonic logic gates were experimentally demonstrated over a decade ago [4]. Switches [5], transistors [6], memories [7], and routers [8,9] have been developed since. Polaritonic time and energy-consumption scales are in the picosecond (ps) [10] and subfemtojoule range [11], while a transistor operating at room temperature was demonstrated recently [12] (previous demonstrations requiring cryogenic cooling). Solitons [13,14], topological states [15], waveguided modes [16], and special wave patterns [17] advanced as robust mechanisms of signal propagation. In theory, universal logic is available [18].

Most recently, polaritons were proposed as a platform for neuromorphic computing [19]. Rather than aiming to perform as universal (classical) information processors, the aim was to specialize in pattern-recognition tasks, including character recognition, speech recognition, and the recognition and forecasting of nonlinear time series. These are the typical tasks at which neural networks excel. In principle, neural networks are efficient architectures as they process information in parallel and avoid the von Neumann bottleneck, with no physical separation of memory and processing elements.

Here, we consider the use of polariton-based neural networks to realize Toffoli logic gates. These gates operate on binary information and are typically sought for reversible computing, which reduces heating by Landauer's principle; regular irreversible logic gates (e.g., AND, XOR, etc.) must generate a small amount of entropy when they operate as they erase bits of information. The heating saved by using reversible gates, given by the von Neumann-Landauer limit [20], is relatively small compared to the overall heating coming from other processes (considering modern standards). Meanwhile, quantum Toffoli gates are realised with trapped ions [21] and photonic architectures [22]. The Toffoli gate is universal, and a valid gate on which to build a complete architecture for classical binary information processing.

One of the most remarkable features of neural networks is their ability to operate with noisy data. We find that a polariton-based neural-network approach to Toffoli gates

*huawen001@e.ntu.edu.sg
†timothyliew@ntu.edu.sg

allows operation with noisy analog signals, provided the noise is smaller than the digital logic level (used to define a binary from an analog signal). Furthermore, the size of analog noise is even reduced during operation. Thus, the obtained Toffoli gates are *self-correcting*, meaning that this architecture is both scalable (assuming the availability of external connections) and efficient in avoiding any need for separate error-correcting overhead.

Finally, although it is possible to construct arbitrary operations from cascaded Toffoli gates, we show that our network can directly represent complete circuits. For example, we consider the full adder circuit, designed for adding two bits of information. In principle, the circuit could be constructed with seven Toffoli gates, however, direct realization is possible in a *single processing step*, avoiding time delays that would otherwise occur when transferring information between cascaded elements. As in the case of the single Toffoli gate, we find that such a design is not only tolerant to a significant level of analog noise, but it also self-corrects for this noise.

II. NEURAL NETWORKS AND RESERVOIR COMPUTING

Neural networks are unconventional computational architectures formed by neurons, which represent the processing elements, and their interconnections, which define the network. Individual neurons receive an input, which can be a weighted combination of signals from many other neurons. The individual neuron shows a nonlinear response and then generates a signal, which is passed to subsequent neurons in the network. Such an architecture is massively parallel and avoids the von Neumann bottleneck, which has motivated the development of artificial hardware neural networks, where platforms include the following: inorganic synapses [23], memristors [24], spintronic systems [25,26], spiking silicon circuits [27,28], and microring resonators [29]. A key challenge in developing hardware neural networks is the need to control the neuron connections, which determine the functionality.

Reservoir computing is an alternative neuromorphic architecture [30], which works with a random network of nodes with arbitrarily fixed bidirectional coupling. The need for control of the internode connections is completely lifted and the network function is rather determined by varying a set of weight connections in a single output layer, which connects the neurons in the network to a final result. Hardware implementations of reservoir computers [31] have been based on memristor arrays [32,33], photonic arrays on silicon chips [34], microwaves [35], electronic circuits [36], and nonlinear optical elements coupled to delay lines [37–40]. Reservoir computing was recently considered theoretically for polariton systems, where relatively high success rates for standard benchmark tasks such as character recognition were predicted (95% success rate for recognition of the MNIST set of hand-written digits) [19]. Generally, systems described by discrete nonlinear Schrödinger equation are good candidates for reservoir computing [41].

III. POLARITON RESERVOIR NETWORK IN RECIPROCAL SPACE

In Ref. [19] a polariton reservoir network was constructed with a lattice of nodes separated in real space, with each node subjected to an effective gain, corresponding to the relaxation of particles from a higher energy excitation. In this work, our first objective is to use a reservoir network to create a universal logic gate that can couple to many subsequent gates, implying that the input and output need the same frequency. To ensure this, we introduce and develop an alternative geometry where the network nodes are defined via coherent resonant excitation of polaritons at specific in-plane wavevectors in a semiconductor microcavity [see Fig. 1(a)]. The energy of the system in reciprocal space (their dispersion) is illustrated in Fig. 1(b). There is a one-to-one coupling of polaritons in reciprocal space to the far field outside of the system, which allows excitation of specific points in reciprocal space. For example, by passing light through a spatial light modulator, or mask, and applying a Fourier lens, one can engineer arbitrary optical field patterns to excite the microcavity. In reciprocal space, we consider the network nodes around a circle centered on the origin, such that all nodes have the same energy.

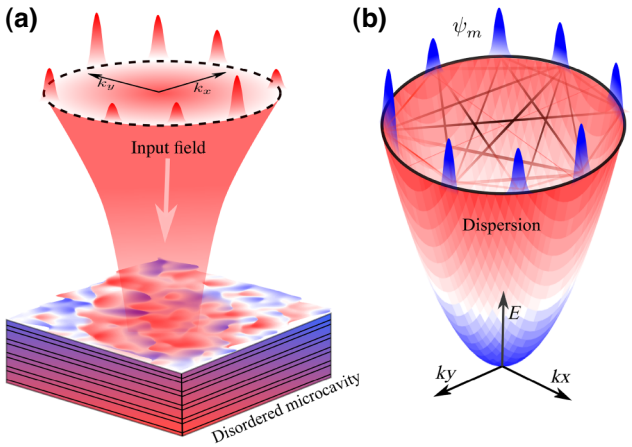

FIG. 1. (a) A monochromatic optical field excites a semiconductor microcavity polariton system with components at different wavevectors. (b) In-plane dispersion. Only modes along an equal energy circle in reciprocal space are excited, which form network nodes. The presence of disorder allows elastic collisions that effectively couple the different modes.

Theoretically, the behavior of coherently excited nodes is described by the Gross-Pitaevskii equation:

$$i\hbar\frac{\partial\psi}{\partial t}=\left(-\frac{\hbar^2\hat{\nabla}^2}{2m}+V(\mathbf{r})-\frac{i\Gamma}{2}+\alpha|\psi|^2\right)\psi+F(\mathbf{r})e^{i\omega t}, \tag{1}$$

where $\psi(\mathbf{r},t)$ is the two-dimensional polariton wave function; m is the effective mass; $V(\mathbf{r})$ is a spatially inhomogeneous potential representing disorder (a Gaussian random distribution with the strength V_0); Γ is the polariton dissipation rate; and α is the (Kerr) polariton-polariton interaction strength. $F(\mathbf{r})$ is a continuous in time spatially patterned driving optical field.

It is well established that polaritons in reciprocal space may change their wavevector by scattering with disorder [42], while the nonlinear term allows pair scattering [4,43]. These processes connect different network nodes, while scattering to states outside of the elastic circle is generally weak (for violating phase-matching conditions).

The last term in Eq. (1) represents the incident coherent optical field with frequency ω, taken to have a form:

$$\tilde{F}(\mathbf{k})=\sum_{nm}\mathbf{W}^{\text{in}}_{nm}\mathbf{u}_m\exp\left[-\frac{|\mathbf{k}-\mathbf{k}_n|^2}{\delta_k^2}\right], \tag{2}$$

in reciprocal space. $\tilde{F}(\mathbf{k})$ is the Fourier transform of $F(\mathbf{r})$ and $\mathbf{k}_n=(k_0\cos\theta_n,k_0\sin\theta_n)$. This form corresponds to a set of Gaussian spots in reciprocal space, centered at $\mathbf{k}=\mathbf{k}_n$ where n is an index labeling the nodes in reciprocal space. Following the scheme of reservoir computing, a set of inputs $\mathbf{u}_m$ are arranged in the input vector $\mathbf{u}$ and multiplied by a random mask matrix $\mathbf{W}^{\text{in}}$, before entering the network. This corresponds to modulating the amplitudes of the Gaussian spots (e.g., with spatial modulation or an optical vector matrix multiplier). The elements of $\mathbf{W}^{\text{in}}$ are uniformly distributed from 0 to $W^{\text{in}}_{\text{max}}$.

The network output is defined by the emission from each node, which is given by the internal fields in reciprocal space: $\psi_n=\int\tilde{\psi}(\mathbf{k})\exp[-|\mathbf{k}-\mathbf{k}_n|^2/\delta_k^2]d\mathbf{k}$. We define t_0 as the time at which this state is read. The final output $\mathbf{Y}$ is given by the vector formed by the emission fields $\boldsymbol{\psi}$ and the input vector $\mathbf{u}$ multiplied by the output weight matrix $\mathbf{W}^{\text{out}}$, i.e., $\mathbf{Y}=\mathbf{W}^{\text{out}}\mathbf{X}$, where $\mathbf{X}=[\mathbf{u},\boldsymbol{\psi}]$ and $\mathbf{W}^{\text{out}}$ is found by training. In principle, this matrix operation can be performed with a single layer of free-space optics applied to the microcavity transmission.

The functionality of the network depends on the choice of $\mathbf{W}^{\text{out}}$. We use a supervised learning technique where a set of known examples are used to construct the output weight matrix $\mathbf{W}^{\text{out}}$. For the ith known example, one knows the output layer $\mathbf{X}_{(i)}$ and the corresponding desired output $\mathbf{Y}_{(i)}$. With these known examples we form an output layer matrix $\boldsymbol{\mathcal{X}}=[\mathbf{X}_{(1)},\mathbf{X}_{(2)},\dots]$ and the corresponding final target output matrix $\boldsymbol{\mathcal{Y}}^{\text{target}}=[\mathbf{Y}_{(1)},\mathbf{Y}_{(2)},\dots]$, where each example satisfies $\mathbf{Y}_{(i)}=\mathbf{W}^{\text{out}}\mathbf{X}_{(i)}$. We use the ridge-regression method [44] to obtain a stable solution (optimum $\mathbf{W}^{\text{out}}$) for all i:

$$\mathbf{W}^{\text{out}}=\boldsymbol{\mathcal{Y}}^{\text{target}}\boldsymbol{\mathcal{X}}^T(\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{X}}^T+\beta\mathbf{I})^{-1}, \tag{3}$$

where β is the "regularization coefficient", which is used to control $\mathbf{W}^{\text{out}}$ in case of large output weights being very sensitive to tiny fluctuations of the input [30]; $\mathbf{I}$ is the identity matrix; and $(\dots)^T$ represents the transpose of a matrix. It is usually advisable to control the obtained output weight $\mathbf{W}^{\text{out}}$ so as to increase the precision of the solution produced by training. Too large $\mathbf{W}^{\text{out}}$ exploits and amplifies small differences among the dimensions of $\mathbf{X}(n)$, and can be extremely sensitive to tiny deviations of inputs. This phenomenon is also called overfitting, which is usually used for data with noise. For data without noise, the regularization part is not strictly necessary as there is no overfitting. However, we find similar results keeping the regularization coefficient.

IV. TOFFOLI GATES

The Toffoli gate is defined by three inputs (A, B, and C) and three outputs ($A'=A$, $B'=B$, and $C'=AB\oplus C$). For convenience, a truth table is shown in Fig. 2(a). Without any noise in the input signal, eight input combinations and the corresponding outputs are enough for training $\mathbf{W}^{\text{out}}$. In testing, we then compare $\mathbf{Y}$ to the ideal output.

As polaritons are described by a continuous amplitude, their use to represent binary information requires defining a logic level, that is, a level of amplitude below which we consider the state as 0 and above which we consider the state as 1. We take the logic level as 0.5 in the input $\mathbf{u}_m$ and output $\mathbf{Y}$. Thus, $\mathbf{u}_m$ and $\mathbf{Y}$ are dimensionless variables, proportional to the amplitudes of input and output fields.

(a) Toffoli gate

A	B	C	A'	B'	C'
0	0	0	0	0	0
0	0	1	0	0	1
0	1	0	0	1	0
0	1	1	0	1	1
1	0	0	1	0	0
1	0	1	1	0	1
1	1	0	1	1	1
1	1	1	1	1	0

(b) Full adder

A	B	C_{in}	Sum	C_{out}
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

FIG. 2. Truth table of (a) a Toffoli gate and (b) a full adder. A Toffoli gate has three inputs (A, B, and C) and three outputs (A', B', and C'). A full adder has three inputs (A, B, and C_{in}) and two outputs (Sum and C_{out}).

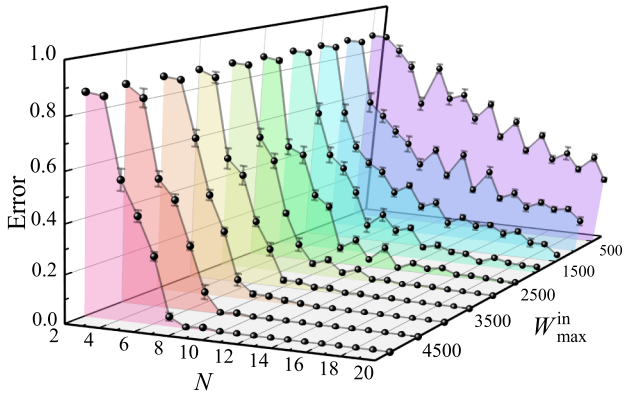

FIG. 3. Dependence of the maximum output error of the Toffoli gate on the system size N (number of nodes) and W^{in}_{max} (excitation strength). The error is defined as the largest absolute difference between the obtained output $\mathbf{Y}$ and ideal output $\mathbf{Y}_{ideal}$. Parameters: $t_0 = 3$ ps, $\Gamma = 0.06$ meV, $V_0 = 0.5$ meV, $\alpha = 1\ \mu\text{eV}^{-1}\ \mu\text{m}^{-1}$.

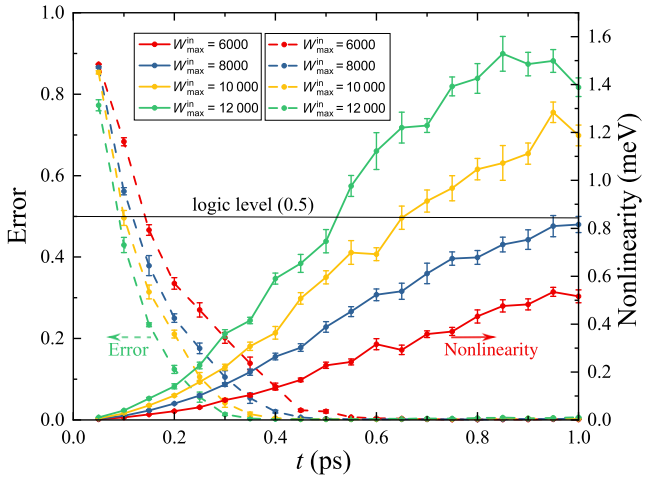

FIG. 5. Dependence of the maximum output error (dashed) and nonlinearity (solid) $[\max_{x,y}(\alpha|\psi|^2)]$ on operation time (t) for different input power (W^{in}_{max}) of the Toffoli gates. Parameters: $V(\mathbf{r}) = 0.5$ meV, $\alpha = 1\ \mu\text{eV}^{-1}\ \mu\text{m}^{-1}$, $\Gamma = 0.06$ meV, and ten nodes.

The physical units describing the real strengths of input and output fields are then assumed to be absorbed in the definitions of $\mathbf{W}^{in}$ and $\mathbf{W}^{out}$.

The results of the trained network from direct numerical simulation of Eq. (1) are illustrated in Fig. 3, where the output error is taken as $\max(|\mathbf{Y} - \mathbf{Y}_{ideal}|)$, where $\mathbf{Y}_{ideal}$ is the ideal output (max denotes the maximum element of the vector) and we average over all binary input combinations and different choices of $\mathbf{W}^{in}$ and $V(\mathbf{r})$. The error decreases with the system size and a network of just five nodes with $W^{in}_{max} = 1000$ is enough to keep the output error lower than the logic level (0.5).

An error-correction ability is also investigated in this system by adding complex noise to the input signals both in training and testing. We train $\mathbf{W}^{out}$ with each input (with 500 realizations of random noise) and test each input (with 200 realization of random noise). Figure 4 shows the reduction of the noise in the output for a reservoir network with 12 nodes. Increasing the number of nodes reveals that more noise is corrected [see Fig. 4(c)].

V. ANALYSIS OF PHYSICAL PARAMETERS FOR THE TOFFOLI GATES

To estimate the operation time and power requirement of our system, we first fix the nonlinear Kerr interaction strength $\alpha = 1\ \mu\text{eV}^{-1}\ \mu\text{m}^{-1}$ [7] and investigate the dependence of the maximum output error on the evolution time of the system and compare to the amount of nonlinearity $[\max_{x,y}(\alpha|\psi|^2)]$ in the system. As seen in Fig. 5, with a defined logic level and different driving-field amplitudes set by W^{in}_{max}, subpicosecond operation time with nonlinear interaction strength on the order of tenths of meV is predicted. Comparing to experimental works [7] achieving the same level of interaction energy, this suggests submilliwatt power requirement (we prefer to estimate power requirements in this way, comparing to experiments, as they automatically account for losses in the optical setup used to couple light into the polariton system).

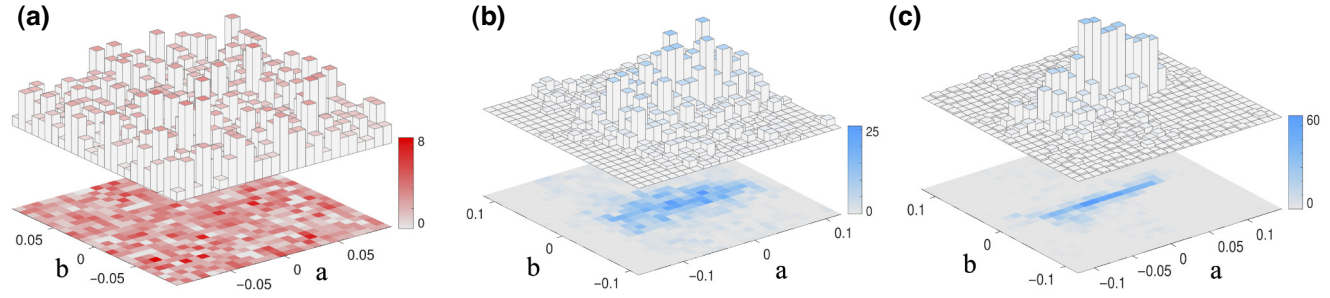

FIG. 4. Input error versus the output error for Toffoli gates. (a) Input complex noise (a + ib) added to the binary input of the Toffoli gate is randomly distributed from -0.1 to 0.1 for both a (real part) and b (imaginary part). (b),(c) Output complex noise distribution for the reservoir network consisting of 12 and 20 nodes, respectively.

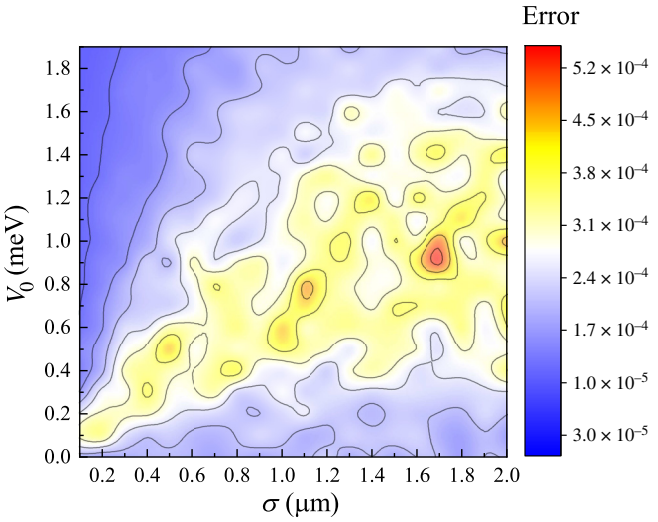

FIG. 6. Dependence of the maximum output error on the correlation length (σ) and strength of disorder (V) in the system corresponding to the Toffoli gate. Parameters: $\alpha = 1\ \mu\text{eV}^{-1}\ \mu\text{m}^{-1}$, $W^{\text{in}}_{\text{max}} = 6000$, $\Gamma = 0.06$ meV, $t_0 = 1$ ps, and ten nodes in total.

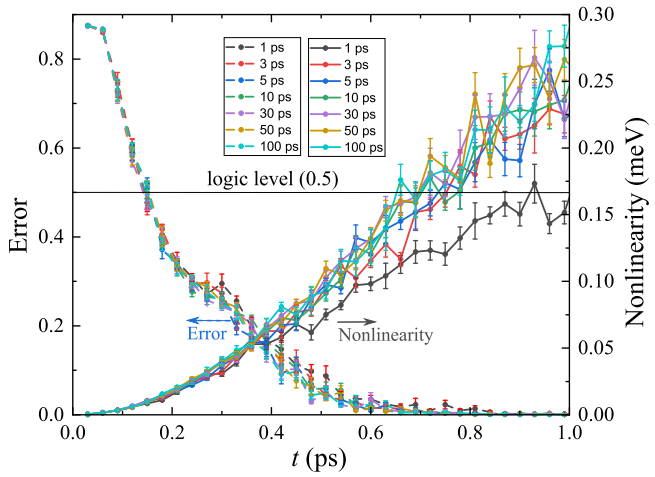

FIG. 7. Dependence of the maximum output error (dashed) and nonlinearity (solid) on the operation time (t) and lifetime of polaritons in the system corresponding to the Toffoli gate. Parameters: $V_0 = 0.5$ meV, $W^{\text{in}}_{\text{max}} = 4000$, $\alpha = 1\ \mu\text{eV}^{-1}\ \mu\text{m}^{-1}$, and ten nodes in total.

We also consider the influence of the strength and correlation length of the disorder in the polariton system on the Toffoli gate. We fix the operation time to 1 ps and set the upper limit of the input matrix 6000, and see how the maximum output error depends on the disorder (see Fig. 6). As shown in Fig. 6, disorder has only a small effect on the highest error (which is on the order of 10^{-4}) and, if analyzing, disorder improves the device performance. Note that even without disorder, our network nodes are coupled in reciprocal space due to the nonlinear scattering processes. The dependence of the maximum output error of our scheme on the lifetime of polaritons is also shown in Fig 7, which indicates that the lifetime only has a limited influence on the output, and longer lifetime increases the system performance slightly.

VI. COMPOSITE CIRCUITS

In principle, the previously considered Toffoli gate is a sufficient building block for composing arbitrary binary circuits. As an example, we consider the full adder circuit, which is designed to add three inputs and provide two outputs [see Fig. 2(b)]. Among the inputs, A and B are two binary digits and C_{in} is a carry signal (binary), which could come from another full adder in the case that the full adder is part of a larger circuit for adding multidigit numbers. At the output, Sum is the sum of the three inputs and C_{out} is a corresponding carry bit, which could be passed to a subsequent full adder circuit. As the Toffoli gate is universal, any logic circuit can be composed with it. Seven Toffoli gates are required to construct a full adder with seven constant inputs and 14 unused outputs. Each Toffoli gate would need at least five nodes in the reservoir network to ensure the output error is lower than the logic level. Thus, for the full adder construction with Toffoli gates, 35 network nodes are expected, in addition to the need of extra fixed inputs and unused outputs. The full adder can also be constructed with two XOR gates, two AND gates, and one OR gate. However, the same problems in cascading elements appear.

Remarkably, we find that the considered reservoir networks can be much more efficient as they can directly reproduce whole circuits. We show that the full adder can be directly realized with our considered reservoir network (see Fig. 8). The operation time and power requirements are investigated with the same method as considered for

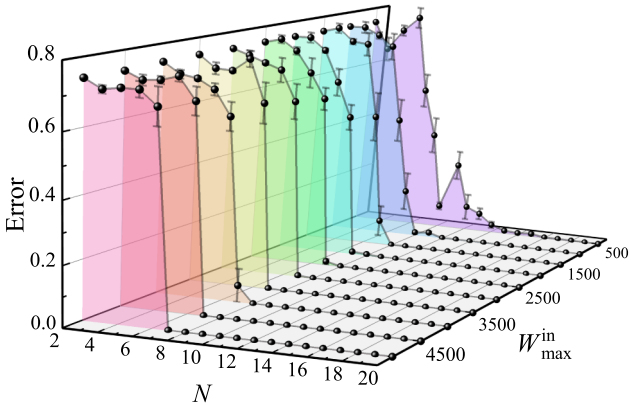

FIG. 8. Dependence of the maximum output error of the full adder on N and $W^{\text{in}}_{\text{max}}$. The error and parameters are defined in the same way as for the Toffoli gate.

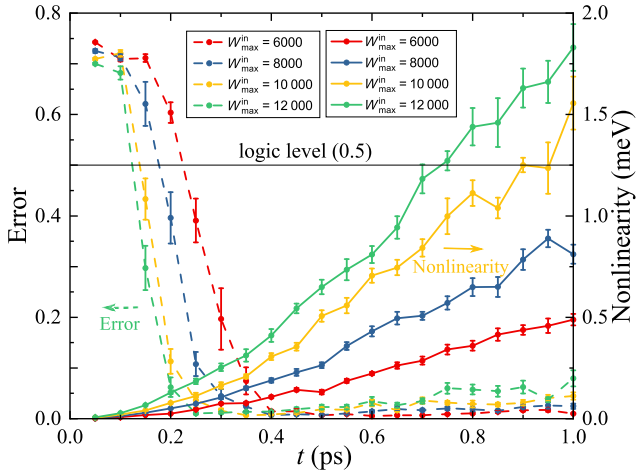

FIG. 9. Dependence of the maximum output error (dashed) and nonlinearity (solid) [$\max_{x,y}(\alpha|\psi|^2)$] on operation time ($t$) for different $W^{\text{in}}_{\text{max}}$ of the full adder. Parameters are the same as in Fig. 5.

the Toffoli gate, again setting $\alpha = 1\ \mu\text{eV}^{-1}\ \mu\text{m}^{-1}$. For ps-scale operation time, the full adder requires slightly more nonlinearity, implying slightly higher power than for the individual Toffoli gate. However, the required nonlinearity is still in the range of several tenths of meV (see Fig. 9). As pointed out earlier, experiments have reached this level of nonlinearity with submilliwatt power consumption (corresponding to subfemtojoule energy when integrated over picosecond operation time) [7]. Meanwhile, the same error-correction ability of the earlier considered Toffoli gates also holds for full adder circuit (see Fig. 10).

VII. DISCUSSION

The aforementioned results have been averaged over different random configurations of disorder and input weight matrices. If one allows selection of specific configurations, then we are aware that both the Toffoli gate and full adder circuit can be operated with just two network nodes. To understand the role played by the reservoir network in achieving this result, it is useful to consider Fig. 11. The horizontal axis shows the linear combination of inputs that contribute directly to the Toffoli gate output, while the vertical axis shows the linear combination of fields from the reservoir that contribute, for different input choices. In other words, if the matrix $\mathbf{W}^{\text{out}}$ is augmented in the form

$$\mathbf{W}^{\text{out}} = \left(\begin{array}{c|c} \mathbf{W}^{\text{out}}_{\text{u},A'} & \mathbf{W}^{\text{out}}_{\psi,A'} \\ \hline \mathbf{W}^{\text{out}}_{\text{u},B'} & \mathbf{W}^{\text{out}}_{\psi,B'} \\ \hline \mathbf{W}^{\text{out}}_{\text{u},C'} & \mathbf{W}^{\text{out}}_{\psi,C'} \end{array}\right), \tag{4}$$

then the output C' is determined by $\mathbf{W}^{\text{out}}_{\text{u},C'}\mathbf{u}$, plotted on the horizontal axis, and $\mathbf{W}^{\text{out}}_{\psi,C'}\boldsymbol{\psi}$, plotted on the vertical axis. Clearly, it is impossible to separate the inputs given solely their positions on the horizontal axis by defining a threshold. The reservoir performs a nonlinear transformation of the inputs, mapping them to different coordinates on the vertical axis. The introduction of a threshold is represented by a linear cut, which is able to distinguish the different input combinations. A similar plot holds for the full adder circuit (see Appendix B).

VIII. CONCLUSIONS

Disordered exciton-polariton neural networks based on a reservoir computing architecture can be used to process binary information. They can be used to realize Toffoli gates (which are universal and reversible) making use of small networks in reciprocal space. In addition to providing universality, our scheme automatically corrects for amplitude and phase errors in the input and output signals. Together with our use of an architecture where input and output have the same frequency, this implies that multiple Toffoli gates can be cascaded (using additional external connections). We further find that disordered exciton-polariton neural networks offer an efficient

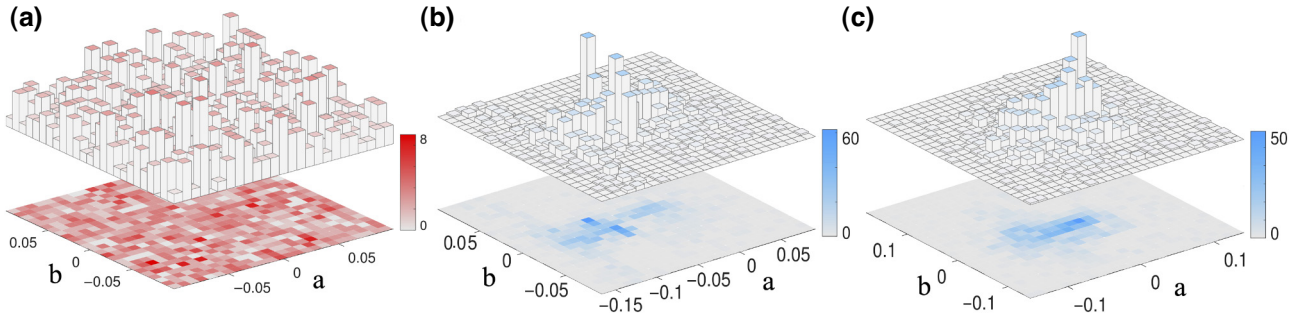

FIG. 10. Input error versus the output error for full adders realized with polariton neural networks. (a) Input complex noise (a + ib) added to the binary input of full adder, randomly distributed from −0.1 to 0.1 for both a (real part) and b (imaginary part). Output complex noise distribution for the reservoirs consisting of 16 nodes (b), 20 nodes (c).

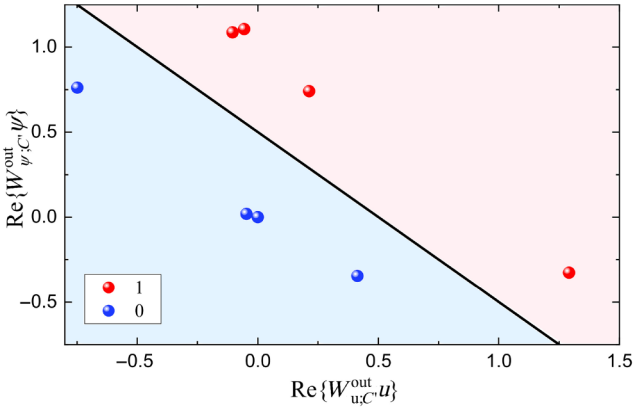

FIG. 11. Representation of how the reservoir network forms the C' output corresponding to a Toffoli gate. The diagonal line marks the logic level. We consider a five-node network.

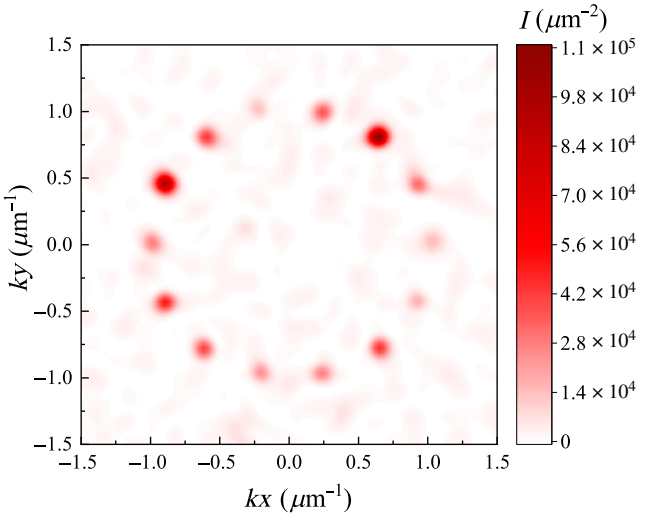

FIG. 12. Fourteen polariton modes in reciprocal space with radius 1 μm^{-1} in the disordered microcavity. Parameters: $V_0 = 0.5$ meV, $\alpha = 1$ μeV^{-1} μm^{-1}, $\Gamma = 0.06$ meV, $t_0 = 1$ ps, and 14 nodes in total.

alternative to cascading multiple elements, where they can directly reproduce composite circuits in a single operation step, essentially eliminating the delay time between individual gates. In principle, our results based on the Gross-Pitaevskii and nonlinear Schrödinger equation can also be relevant to other systems. Although here we consider a continuously driven system, we find similar results (in Appendix C) for pulsed systems with a prepared initial condition.

ACKNOWLEDGMENT

This work is supported by the Ministry of Education, Singapore (Grant No. MOE2018-T2-02-068). M.M. acknowledges support from the National Science Center, Poland via Grant No. 2017/25/Z/ST3/03032 under the QuantERA program.

APPENDIX A: EXCITON-POLARITON DISTRIBUTION IN RECIPROCAL SPACE

The resonant Rayleigh scattering for exciton polaritons in semiconductor microcavities has been well investigated [42]. In our scheme, exciton polaritons with a typical elastic circle (radius $= 1$ μm^{-1}) are arranged in momentum space (see in Fig. 12). As polariton modes could easily occupy less than 0.1 μm^{-1} in reciprocal space, we estimate that one could operate with 50 nodes around an elastic circle, which is considerably more than we find are needed for the considered tasks. The nonlinear term allows pair scattering (equivalently four-wave mixing) processes that can also connect different network nodes, while scattering to states outside of the elastic circle is generally weak (see Fig. 13).

APPENDIX B: NONLINEAR TRANSFORMATION

For the full adder, we also investigate how the reservoir network helps achieve the result, we here consider the Sum output. As can be seen in Fig. 14, the reservoir network performs a nonlinear transformation of the input, which allows separation of the different inputs by a linear cut. This is equivalent to defining a threshold given by the sums of the vertical and horizontal axes. Note that in this

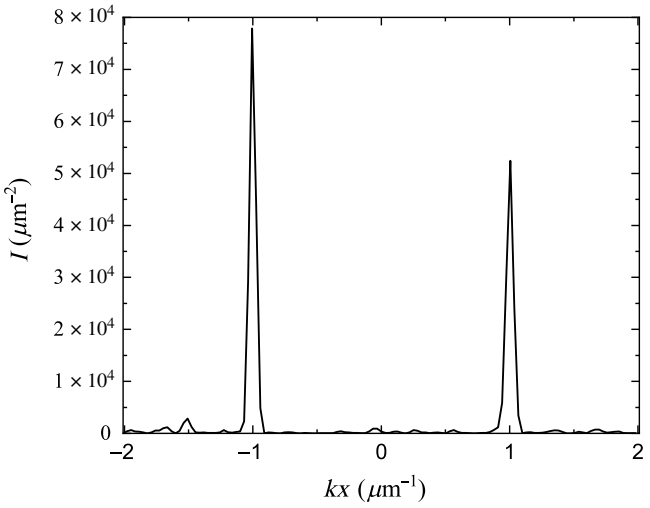

FIG. 13. Intensity $[I = |\psi(k_x)|^2]$ distribution in reciprocal space, along the cross section $k_y = 0$, which shows limited scattering outside the polariton ring in momentum space. The parameters are the same as in Fig. 12.

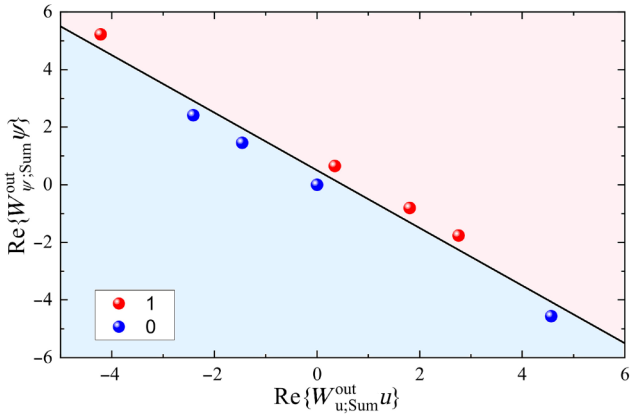

FIG. 14. Graphic representation of how the reservoir network forms the Sum output corresponding to a full adder. The horizontal and vertical axes correspond to the weighted sum of inputs in weighted sum of reservoir field amplitudes that contribute to the Sum output, respectively. The diagonal line shows the cut corresponding to the logic level. Here we consider a six-node reservoir network.

case the matrix $\mathbf{W}^{\text{out}}$ is augmented in the form:

$$\mathbf{W}^{\text{out}} = \left(\begin{array}{c|c} \mathbf{W}^{\text{out}}_{\text{u,Sum}} & \mathbf{W}^{\text{out}}_{\psi,\text{Sum}} \\ \hline \mathbf{W}^{\text{out}}_{\text{u},C_{\text{out}}} & \mathbf{W}^{\text{out}}_{\psi,C_{\text{out}}} \end{array} \right). \tag{B1}$$

APPENDIX C: INJECTING INPUT SIGNAL VIA INITIAL CONDITION

In the main text, the input signal is fed into the reservoir network by a continuous excitation $F(\mathbf{r})$. In systems, such as cold atoms, continuous excitation $F(\mathbf{r})$ at specific points in reciprocal space can be challenging. It is however possible to prepare the initial state $\psi(\mathbf{r}, t=0)$ with specific distributions in reciprocal space [45]. We consider the initial state:

$$\tilde{\psi}(\mathbf{k}, t=0) = \frac{1}{\hbar} \sum_{nm} W^{\text{in}}_{nm} u_m \exp\left[-\frac{|\mathbf{k} - \mathbf{k}_n|^2}{\delta_k^2} \right]. \tag{C1}$$

where $\tilde{\psi}(\mathbf{k}, t=0)$ is the Fourier component of $\psi(\mathbf{r}, t=0)$. Given the initial state, the dynamics is given by the same nonlinear Schrödinger equation as considered in the main text:

$$i\hbar \frac{\partial \psi}{\partial t} = \left(-\frac{\hbar^2 \nabla^2}{2m} + V(r) - i\frac{\Gamma}{2} + \alpha |\psi|^2 \right) \psi. \tag{C2}$$

Note that, preparing an initial state $\psi(\mathbf{r}, t=0)$ is equivalent to exciting the system with a short pulse $F(\mathbf{r})\delta(t)$ where we choose $F(\mathbf{r}) = i\hbar\psi(\mathbf{r}, t=0)$. In Fig. 15, we show that the scheme of injecting an input signal via the initial state performs similarly to the continuous wave excitation case. As can be seen in Fig. 15, both methods (injecting input signal via initial state or via a continuous pump) show the same trend of decreasing output error with increasing system size.

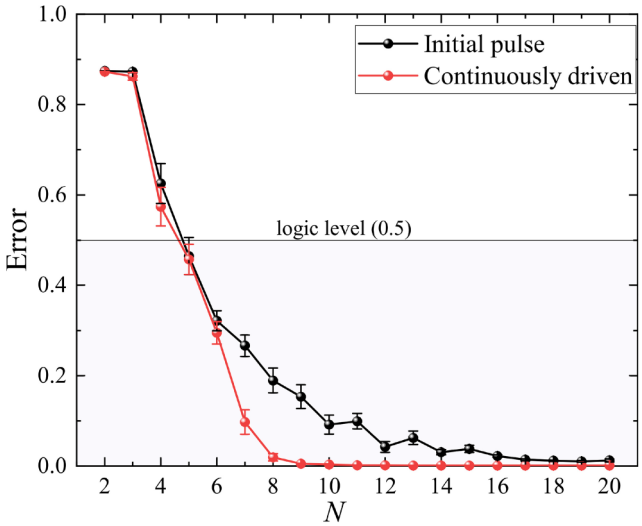

FIG. 15. Dependence of the maximum output error of the Toffoli gate on the network size N for systems with continuous excitation (red) and with initially prepared states (black). We find a similar behavior for both the curves. Here we consider $\mathbf{W}^{\text{in}}_{\text{max}} = 1000$.

[1] C. Sun, Single-chip microprocessor that communicates directly using light, Nature **528**, 534 (2015).

[2] D. A. B. Miller, Attojoule optoelectronics for low-energy information processing and communications, J. Lightwave. Technol. **35**, 346 (2017).

[3] D. Sanvitto and S. Kéna-Cohen, The road towards polaritonic devices, Nat. Mater. **15**, 1061 (2016).

[4] C. Leyder, T. C. H. Liew, A. V. Kavokin, I. A. Shelykh, M. Romanelli, J. P. Karr, E. Giacobino, and A. Bramati, Interference of Coherent Polariton Beams in Microcavities: Polarization-Controlled Optical Gates, Phys. Rev. Lett. **99**, 196402 (2007).

[5] T. Gao, P. S. Eldridge, T. C. H. Liew, S. I. Tsintzos, G. Stavrinidis, G. Deligeorgis, Z. Hatzopoulos, and P. G. Savvidis, Polariton condensate transistor switch, Phys. Rev. B **85**, 235102 (2012).

[6] D. Ballarini, M. De Giorgi, E. Cancellieri, R. Houdré, E. Giacobino, R. Cingolani, A. Bramati, G. Gigli, and D. Sanvitto, All-optical polariton transistor, Nat. Commun. **4**, 1778 (2013).

[7] R. Cerna, Y. Léger, T. K. Paraïso, M. Wouters, F. Morier-Genoud, M. T. Portella-Oberli, and B. Deveaud, Ultrafast tristable spin memory of a coherent polariton gas, Nat. Commun. **4**, 2008 (2013).

[8] F. Marsault, H. S. Nguyen, D. Tanese, A. Lemaître, E. Galopin, I. Sagnes, A. Amo, and J. Bloch, Realization of an all optical exciton-polariton router, Appl. Phys. Lett. **107**, 201115 (2015).

[9] H. Flayac and I. G. Savenko, An exciton-polariton mediated all-optical router, Appl. Phys. Lett. **103**, 201105 (2013).
[10] E. Cancellieri, A. Hayat, A. M. Steinberg, E. Giacobino, and A. Bramati, Ultrafast Stark-Induced Polaritonic Switches, Phys. Rev. Lett. **112**, 053601 (2014).
[11] A. Dreismann, H. Ohadi, Y. del Valle-Inclan Redondo, R. Balili, Y. G. Rubo, S. I. Tsintzos, G. Deligeorgis, Z. Hatzopoulos, P. G. Savvidis, and J. J. Baumberg, A sub-femtojoule electrical spin-switch based on optically trapped polariton condensates, Nat. Mater. **15**, 1074 (2016).
[12] A. V. Zasedatelev, A. V. Baranikov, D. Urbonas, F. Scafirimuto, U. Scherf, T. Stöferle, R. F. Mahrt, and P. G. Lagoudakis, A room-temperature organic polariton transistor, Nat. Photonics **13**, 378 (2019).
[13] M. Sich, F. Fras, J. K. Chana, M. S. Skolnick, D. N. Krizhanovskii, A. V. Gorbach, R. Hartley, D. V. Skryabin, S. S. Gavrilov, E. A. Cerda-Méndez, K. Biermann, R. Hey, and P. V. Santos, Effects of Spin-Dependent Interactions on Polarization of Bright Polariton Solitons, Phys. Rev. Lett. **112**, 046403 (2014).
[14] J. K. Chana, M. Sich, F. Fras, A. V. Gorbach, D. V. Skryabin, E. Cancellieri, E. A. Cerda-Méndez, K. Biermann, R. Hey, P. V. Santos, M. S. Skolnick, and D. N. Krizhanovskii, Spatial Patterns of Dissipative Polariton Solitons in Semiconductor Microcavities, Phys. Rev. Lett. **115**, 256401 (2015).
[15] S. Klembt, T. H. Harder, O. A. Egorov, K. Winkler, R. Ge, M. A. Bandres, M. Emmerling, L. Worschech, T. C. H. Liew, M. Segev, C. Schneider, and S. Höfling, Exciton-polariton topological insulator, Nature **562**, 552 (2018).
[16] G. Lerario, A. Cannavale, D. Ballarini, L. Dominici, M. De Giorgi, M. Liscidini, D. Gerace, D. Sanvitto, and G. Gigli, Room temperature Bloch surface wave polaritons, Opt. Lett. **39**, 2068 (2014).
[17] A. Gianfrate, L. Dominici, O. Voronych, M. Matuszewski, M. Stobińska, D. Ballarini, M. De Giorgi, G. Gigli, and D. Sanvitto, Superluminal x-waves in a polariton quantum fluid, Light: Sci. Appl. **7**, 17119 (2018).
[18] T. Espinosa-Ortega and T. C. H. Liew, Complete architecture of integrated photonic circuits based on and and not logic gates of exciton polaritons in semiconductor microcavities, Phys. Rev. B **87**, 195305 (2013).
[19] A. Opala, S. Ghosh, T. C. H. Liew, and M. Matuszewski, Neuromorphic Computing in Ginzburg-Landau Polariton-Lattice Systems, Phys. Rev. Appl. **11**, 064029 (2019).
[20] A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz, Experimental verification of Landauer's principle linking information and thermodynamics, Nature **483**, 187 (2012).
[21] T. Monz, K. Kim, W. Hänsel, M. Riebe, A. S. Villar, P. Schindler, M. Chwalla, M. Hennrich, and R. Blatt, Realization of the Quantum Toffoli Gate with Trapped Ions, Phys. Rev. Lett. **102**, 040501 (2009).
[22] B. P. Lanyon, M. Barbieri, M. P. Almeida, T. Jennewein, T. C. Ralph, K. J. Resch, G. J. Pryde, J. L. O'Brien, A. Gilchrist, and A. G. White, Simplifying quantum logic using higher-dimensional Hilbert spaces, Nat. Phys. **5**, 134 (2009).
[23] T. Ohno, T. Hasegawa, T. Tsuruoka, K. Terabe, J. K. Gimzewski, and M. Aono, Short-term plasticity and long-term potentiation mimicked in single inorganic synapses, Nat. Mater. **10**, 591 (2011).
[24] P. Krzysteczko, J. Münchenberger, M. Schäfers, G. Reiss, and A. Thomas, The memristive magnetic tunnel junction as a nanoscopic synapse-neuron system, Adv. Mater. **24**, 762 (2012).
[25] N. Locatelli, V. Cros, and J. Grollier, Spin-torque building blocks, Nat. Mater. **13**, 11 (2013).
[26] V. Quang Diep, B. Sutton, B. Behin-Aein, and S. Datta, Spin switches for compact implementation of neuron and synapse, Appl. Phys. Lett. **104**, 222405 (2014).
[27] G. Indiveri and T. Horiuchi, Frontiers in neuromorphic engineering, Front. Neurosci. **5**, 118 (2011).
[28] C.-S. Poon and K. Zhou, Neuromorphic silicon neurons and large-scale neural networks: Challenges and opportunities, Front. Neurosci. **5**, 108 (2011).
[29] J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. P. Pernice, All-optical spiking neurosynaptic networks with self-learning capabilities, Nature **569**, 208 (2019).
[30] M. Lukoševičius, *A Practical Guide to Applying Echo State Networks* (Springer, Berlin Heidelberg, 2012), p. 659.
[31] G. Tanaka, T. Yamane, J. B. Héroux, R. Nakane, N. Kanazawa, S. Takeda, H. Numata, D. Nakano, and A. Hirose, Recent advances in physical reservoir computing: A review, Neural. Netw. **115**, 100 (2019).
[32] N. Soures and D. Kudithipudi, Deep liquid state machines with neural plasticity for video activity recognition, Front. Neurosci. **13**, 686 (2019).
[33] L. Du, X. Li, W. Lou, G. Sullivan, K. Chang, J. Kono, and R.-R. Du, Evidence for a topological excitonic insulator in InAs/GaSb bilayers, Nat. Commun. **8**, 1971 (2017).
[34] K. Vandoorne, P. Mechet, T. Van Vaerenbergh, M. Fiers, G. Morthier, D. Verstraeten, B. Schrauwen, J. Dambre, and P. Bienstman, Experimental demonstration of reservoir computing on a silicon photonics chip, Nat. Commun. **5**, 3541 (2014).
[35] J. Torrejon, M. Riou, F. A. Araujo, S. Tsunegi, G. Khalsa, D. Querlioz, P. Bortolotti, V. Cros, K. Yakushiji, A. Fukushima, H. Kubota, S. Yuasa, M. D. Stiles, and J. Grollier, Neuromorphic computing with nanoscale spintronic oscillators, Nature **547**, 428 (2017).
[36] L. Appeltant, M. C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, Information processing using a single dynamical node as complex system, Nat. Commun. **2**, 468 (2011).
[37] F. Duport, B. Schneider, A. Smerieri, M. Haelterman, and S. Massar, All-optical reservoir computing, Opt. Express **20**, 22783 (2012).
[38] L. Larger, M. C. Soriano, D. Brunner, L. Appeltant, J. M. Gutierrez, L. Pesquera, C. R. Mirasso, and I. Fischer, Photonic information processing beyond turing: An optoelectronic implementation of reservoir computing, Opt. Express **20**, 3241 (2012).
[39] D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, Parallel photonic information processing at gigabyte per

second data rates using transient states, Nat. Commun. **4**, 1364 (2013).

[40] L. Larger, A. Baylón-Fuentes, R. Martinenghi, V. S. Udaltsov, Y. K. Chembo, and M. Jacquot, High-Speed Photonic Reservoir Computing Using a Time-Delay-Based Architecture: Million Words per Second Classification, Phys. Rev. X **7**, 011015 (2017).

[41] S. Borlenghi, M. Boman, and A. Delin, Modeling reservoir computing with the discrete nonlinear Schrödinger equation, Phys. Rev. E **98**, 052101 (2018).

[42] W. Langbein, E. Runge, V. Savona, and R. Zimmermann, Enhanced Resonant Backscattering of Excitons in Disordered Quantum Wells, Phys. Rev. Lett. **89**, 157401 (2002).

[43] C. Diederichs, J. Tignon, G. Dasbach, C. Ciuti, A. Lemaître, J. Bloch, P. Roussignol, and C. Delalande, Parametric oscillation in vertical triple microcavities, Nature **440**, 904 (2006).

[44] A. E. Hoerl and R. W. Kennard, Ridge regression: Biased estimation for nonorthogonal problems, Technometrics **12**, 55 (1970).

[45] F. Jendrzejewski, K. Müller, J. Richard, A. Date, T. Plisson, P. Bouyer, A. Aspect, and V. Josse, Coherent Backscattering of Ultracold Atoms, Phys. Rev. Lett. **109**, 195302 (2012).