PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Silicon photonic neural network applications and prospects

Shastri, Bhavin, Huang, Chaoran, Tait, Alexander, Ferreira de Lima, Thomas, Prucnal, Paul

Bhavin J. Shastri, Chaoran Huang, Alexander N. Tait, Thomas Ferreira de Lima, Paul R. Prucnal, "Silicon photonic neural network applications and prospects," Proc. SPIE 12019, AI and Optical Data Sciences III, 120190K (2 March 2022); doi: 10.1117/12.2614865

SPIE.

Event: SPIE OPTO, 2022, San Francisco, California, United States

Silicon Photonic Neural Applications and Prospects

Bhavin J. Shastri*[a,b,c], Chaoran Huang[d], Alexander N. Tait[e], Thomas Ferreira de Lima[b], Paul R. Prucnal[b]

[a]Department of Physics, Engineering Physics & Astronomy, Queen's University, Kingston, ON K7L 3N6, Canada; [b]Vector Institute for Artificial Intelligence, Toronto, ON M5G 1M1, Canada; [c]Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA; [d]Department of Electronic Engineering, the Chinese University of Hong Kong, Shatin, Hong Kong; [e]Department of Electrical and Computer Engineering, Queen's University, Kingston, ON K7L 3N6, Canada

ABSTRACT

Neural networks have enabled many applications in artificial intelligence and neuromorphic computing ranging from scientific computing, intelligent communications, security etc. Neural networks implemented in on digital platforms are limited in speed and energy efficiency. Neuromorphic (i.e., neuron-isomorphic) photonics aims to build processors in which optical hardware mimic neural networks in the brain. These processors promise orders of magnitude improvements in both speed and energy efficiency over purely digital electronic approaches. However, integrated optical neural networks are much smaller (hundreds of neurons) than electronic implementations (tens of millions of neurons). This raises a question: what are the applications where sub-nanosecond latencies and energy efficiency trump the sheer size of processor? We provide an overview of neuromorphic photonic systems and their real-world applications to machine learning and neuromorphic computing.

Keywords: Silicon photonics, neuromorphic computing, machine learning, photonic integrated circuits

1. INTRODUCTION

Neuromorphic computing promises to create radical new hardware platforms that can emulate the underlying neural structure of the brain. The general idea is to build circuits composed of physical devices that mimic the neuron biophysics interconnected by massive physical interconnects with co-integrated non-volatile memories [1], [2]. In doing so, neuromorphic hardware could break performance limitations inherent in von Neumann architectures and gain advantages in speed and efficiency in solving intellectual tasks [3], [4]. Achieving this goal requires significant advances in a wide range of technologies, including materials, devices, device fabrication, system integration, platform co-integration, packaging etc.

By combining the high bandwidth and parallelism of photonic devices with the adaptability and complexity attained by methods like those seen in the brain, photonic neural networks (PNNs) have the potential to be orders of magnitude faster than state-of-the-art electronic processors while consuming less energy per computation [5]. The goal of neuromorphic photonic processors is not to replace conventional computers, but to enable applications that are unreachable at present by conventional computing technology—those requiring low latency, high bandwidth and low energies [6], [7]. As shown in Figure 1, examples of applications for ultrafast neural networks include: 1) Enabling fundamental physics breakthroughs: qubit read-out classification, high-energy-particle collision classification, fusion reactor plasma control; 2) Nonlinear programming: solving nonlinear optimization problems (robotics, autonomous vehicles, predictive control) and partial differential equations; 3) Machine learning acceleration: vector–matrix multiplications, deep learning inference, ultrafast or online learning; 4) Intelligent signal processing: wideband radio-frequency signal processing, fibre-optic communication.

2. NEUROMORPHIC PHOTONICS APPROACHES

Over the 10 years, several neuromorphic photonic [6], [8] approaches have been proposed as shown in Figure 2. This can be divided into feedforward and recurrent (including random recurrent i.e. reservoir computing [9]–[11]), or coherent

*shastri@ieee.org

AI and Optical Data Sciences III, edited by Bahram Jalali,
Ken-ichi Kitayama, Proc. of SPIE Vol. 12019, 120190K
© 2022 SPIE · 0277-786X · doi: 10.1117/12.2614865

(single wavelength) [12], [13] and incoherent (multiwavelength) [14], [15] approaches, or continuous time networks and spiking networks, or integrated approaches and free-space. We briefly highlight some of these.

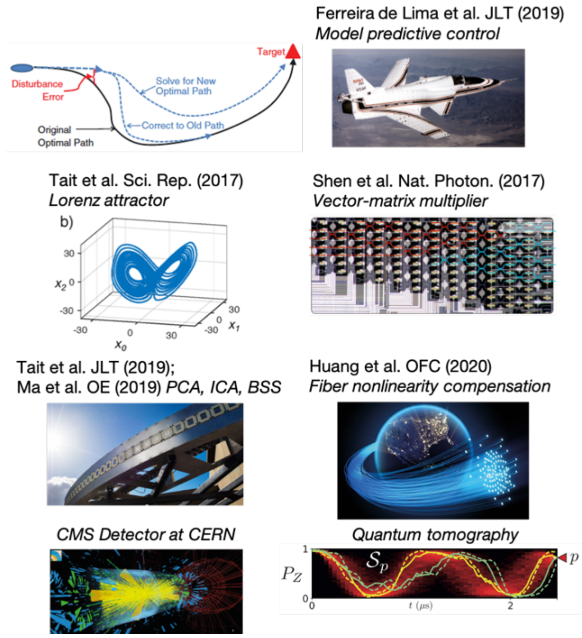


Figure 1. Applications for photonic neural networks that require low latency, high bandwidth, and real-time processing.

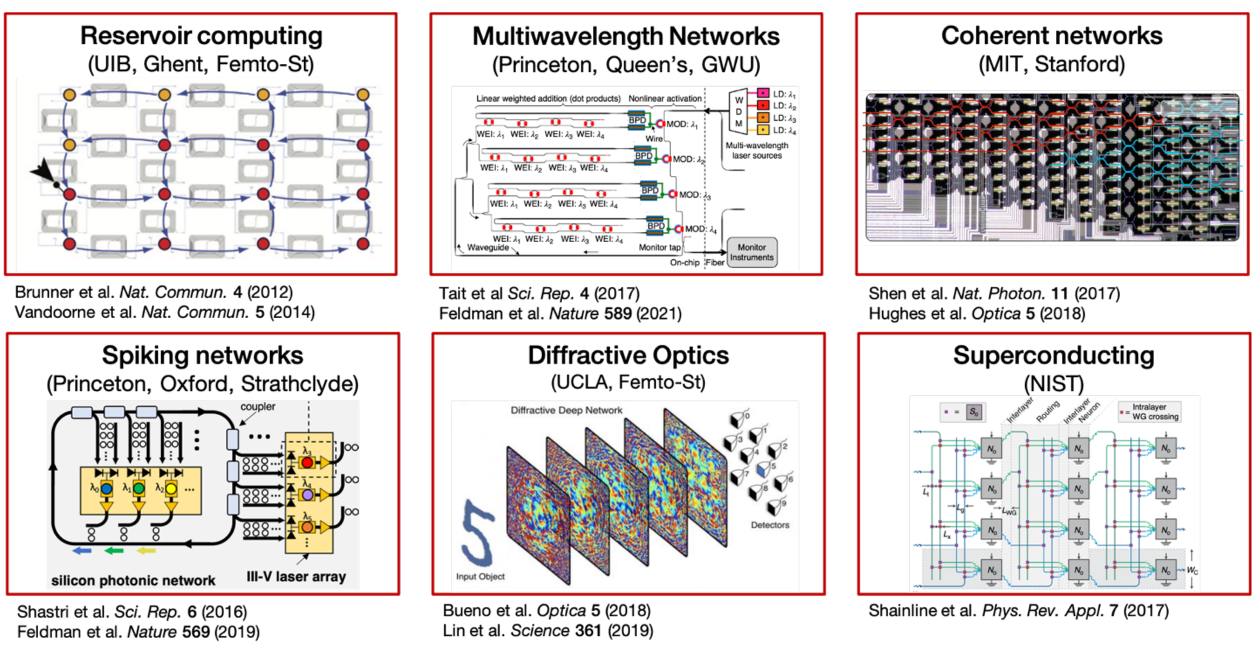


Figure 2. Examples of recently demonstrated photonic neural network architectures.

Broadcast-and-weight. In the authors' recent work [14] neurons are implemented using modulators that exploit the nonlinearity of the electro-optic transfer function. With the foundry-compatible silicon-on-insulator (SOI) technology for on-chip integrated photonics, OEO neurons are demonstrated by Tait et al. [15] using a silicon microring resonator (MRR) with embedded PN modulator. These neurons exploit the so-called broadcast-and-weight protocol [16] where signals are encoded onto multiple wavelengths and multiplexed to a single waveguide. WDM signals are weighted by groups of tunable filters such as MRRs These filters can be tuned on and off resonance which changes the transmission of each signal through that filter, effectively multiplying the signal with the desired weight. The resulting weighted signals travel into a photodetector, which can receive many wavelengths in parallel to perform a summing operation. Broadcast-and-weights have been investigated both theoretically and experimentally in the form of closely packed microring filters, prototyped in a silicon-photonic platform. The compact microring weight banks offer enormous information density. Broadcast-and-weight is also capable of operating across modes and polarizations.

Coherent. The coherent approach performs matrix multiplication within a single wavelength and mode for a given matrix [17]. To implement a matrix-vector operation on incoming signals, an array of beam splitters and phase shifters controls the destructive or constructive interference effects in a mesh Mach-Zehnder interferometers (MZIs). Weights are assigned by controlling the power of different waveguides through interference. In this architecture, there is no need to convert from the optical domain to the electrical domain; hence, interfacing a coherent system with both O/E/O neurons and all-optical neurons. The Mach-Zehnder interferometers have a much larger footprint than MRRs, putting a cap on the information density and power consumption in weight configuration.

Free-space diffractive network. This approach utilizes free-space diffraction (Lin et al., 2018). The information is encoded onto a spatial optical image, and the matrix operations are performed with spatially modulated diffractive optical elements. Spatial modulation renders a complex transmission map of the incoming optical image and sets the neural network weights through diffraction effects. Compared to the other two approaches that are implemented with photonic integrated circuits, diffractive neural networks have the advantages in performing large-scale matrix multiplication, but they suffer from the package cost and tight alignment tolerances. Diffractive neural networks have been demonstrated with conventional spatial light modulators for recurrent neural network and the state-of-the-art 3D-printing technologies for multi-layer feedforward configurations

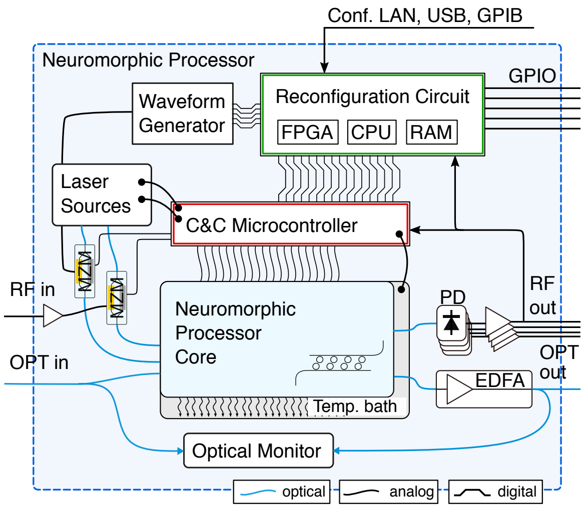
3. VISION OF A NEUROMORPHIC PROCESSOR

Recently, in our tutorial, Ref. [17], we proposed a vision for a neuromorphic processor. We discussed how such a neuromorphic chip could potentially be interfaced with a general-purpose computer (Figure 3), i.e. a CPU, as a coprocessor to target specific applications. In general, there are two levels of complexity associated with co-integrating a general-purpose electronic processor with an application-specific optical processor. Firstly, a CPU processes a series of computation instructions in an undecided amount of time and is not guaranteed to be completed. Neural networks, on the other hand, can process data in parallel and in a deterministic amount of time. CPUs have a concept of a ‘fixed’ instruction set on top of which computer software can be developed. However, a neuromorphic processor would require a hardware description language (HDL) because it describes the intended behavior of a hardware in real-time. Secondly, seamlessly interfacing a photonic integrated circuit with an electronic integrated circuit will take several advances in science and technology including on-chip lasers and amplifiers, co-integration of CMOS with silicon photonics, system packaging, high-bandwidth digital-to-analog converters (DAC) and analog-to-digital converters (ADCs).

4. MACHINE LEARNING APPLICATION: FIBER NONLINEARITY IMPAIRMENT COMPENSATION

The world is witnessing an explosion of internet traffic. The global internet traffic has reached 5.3 exabytes per day in 2020 and will continue doubling approximately every 18 months. Innovations in fiber communication technologies are required to sustain the long-term exponential growth of data traffic [18]. Increasing data rate and system crosstalk has imposed significant challenges on the digital signal processing (DSP) chips in terms of analog-to-digital converters (ADCs) performances, circuit complexity, and power consumption. A key to advancing the deployment of DSP relies on the consistent improvement in CMOS technology [19]. However, the exponential hardware scaling of ASIC based DSP chips, which is embodied in Moore’s law as other digital electronic hardware, is fundamentally unsustainable. In parallel, many efforts are focused on developing new DSP algorithms to minimize computational complexity, but usually at the expense of reducing transmission link performances [20].

Instead of embracing such a complexity-performance trade-off, an alternative approach is to explore new hardware platforms that intrinsically offer high bandwidth, speed, and low power consumption [6],[21],[22]. Machine learning algorithms, especially neural networks, have been found effective in performing many functions in optical networks, including dispersion and nonlinearity impairments compensation, channel equalization, optical performance monitoring, traffic prediction, etc [23].



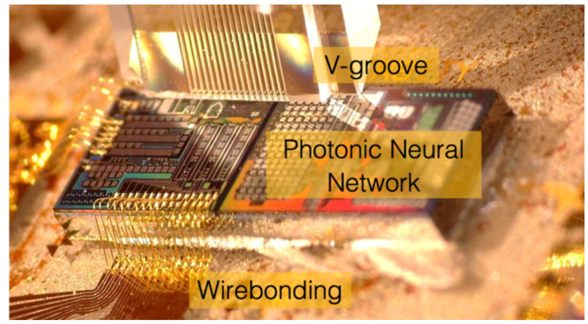


Figure 3. (Left) Simplified schematics of a neuromorphic processor. Thanks to integrated laser sources and photodetectors, it can input and output RF signals directly as an option to optically modulated signals. The waveform generator allows for programming arbitrary stimulus that can be used as part of a machine learning task. Reproduced from [17]. (Right) Micrograph image of a wirebonded photonic neural network chip from Princeton University's Lightwave Lab.

PNNs are well suited for optical communications because the optical signals are processed directly in the optical domain. This innovation avoids prohibitive energy consumption overhead and speed reduction in ADCs, especially in data center applications. In parallel, many PNN approaches are inspired by optical communication systems, making PNNs naturally suitable for processing optical communication signals. For example, we proposed synaptic weights and neuron networking architecture based on the concept of WDM to enable fan-in and weighted addition [14]. This architecture can provide a seamless interface between PNNs and WDM systems, which can be applied as a front-end processor to address inter-wavelength or inter-mode crosstalks problems that DSP usually lacks the bandwidth or computing power to process (e.g., fiber nonlinearity compensation in WDM systems). Moreover, PNNs combine high-quality waveguides and photonic devices that have been initially developed for telecommunications. Therefore, PNNs, by default, can support fiber optic communication rates and enable real-time processing. For example, the scalable silicon PNN proposed by the authors is composed of microring resonator (MRR) banks for synaptic weighting and O/E/O neurons to produce standard machine learning activation functions. The MRR weight bank is inspired by WDM filters, and the O/E/O neurons use typical silicon photodetector and modulator. Therefore, the optimization of associated devices in PNNs can utilize the fruits of the entire silicon photonic ecosystem that is driven by telecommunications and data center applications.

To truly demonstrate photonics can excel over DSP, careful considerations are required to identify different application scenarios (i.e., long-haul, short reach) and system requirements (i.e., performances, energy). Continuous research is needed to improve photonic hardware and to develop hardware-compatible algorithms. Here, we discuss several approaches to train and apply PNNs for optical communications.

Long-haul communication systems prioritize high performances in terms of distance reach and spectral efficiency. This requirement allows the use of coherent technology, along with dense wavelength multiplexing and polarization multiplexing schemes, to maximize the fiber capacity. In long-haul fiber optic transmission systems, fiber nonlinearity remains a challenge to the achievable capacity and transmission distance. One reason is that the nonlinear interplay between signal, noises, and optical fibers negates the accuracy of conventional nonlinear compensation algorithms based on digital backpropagation. Another reason is the implementation of most nonlinear compensation algorithms in DSP chips demands excessive resources. In contrast, the neural network approach can learn and approximate the nonlinear perturbation from the abundant training data, rather than solely relying on the physical fiber model (known as stochastic nonlinear Schrodinger equation). Based on the perturbation methods, the derived neural network algorithm has enabled compensating the nonlinear distortion in a 10800 km fiber transmission link with 32 Gbaud signals [24]. In Ref. [15], we developed a PNN platform based on the so-called "neuromorphic" approach, aiming to map physical models of

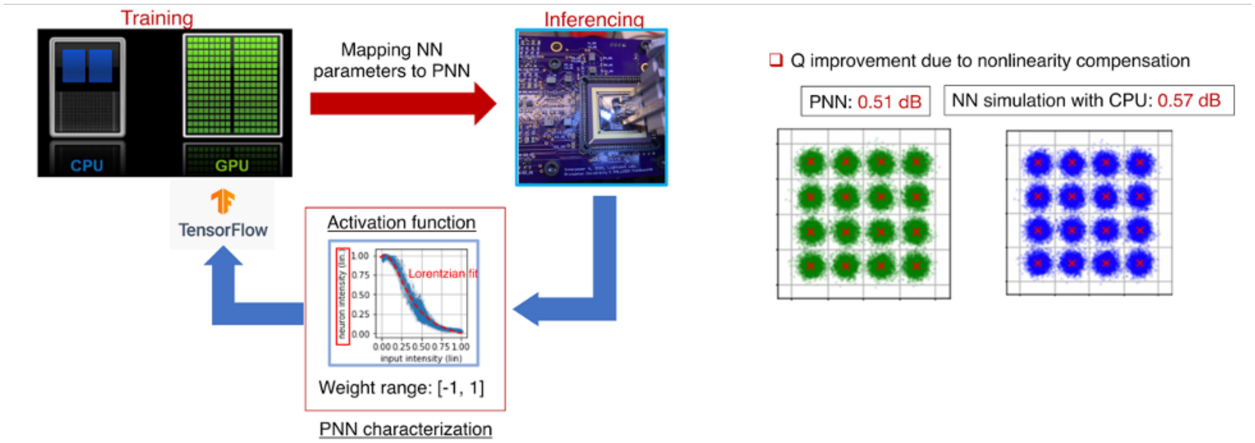


Figure 4. (left) Concept of training and implementing photonic neural networks. Inset shows a transfer function of the photonic neural network measured with real-time signals. (right) Constellations of X-polarization of a 32 Gbaud PM-16QAM, with the ANN-NLC gain of 0.57 dB in Q-factor and with the PNN-NLC gain of 0.51 dB in Q-factor [22].

optoelectronic systems to abstract models of neural networks (which differs from the reservoir approaches). By doing so, the PNN system can leverage existing machine learning algorithms (i.e., backpropagation) and map training results from simulations to heterogeneous photonic hardware. The concept is shown in Figure 4. A proof-of-concept experiment demonstrates the real-time implementation of a trained neural network model using an integrated silicon PNN chip [22]. In this work, the authors experimentally demonstrated that the silicon PNN can produce a similar Q factor improvement compared to the simulated neural network for nonlinear compensation as shown in Figure 4, but it promises to process the communication data in real-time and with high bandwidth and low latency.

We also proposed a photonic architecture enabling all-to-all continuous-time recurrent neural networks (RNN) [15]. Recurrent neural networks can resemble optical fiber transmission systems: the linear neuron-to-neuron connections with internal feedback is analog to linear multiple-input multiple-output (MIMO) fiber channel with dispersive memory. With neuron nonlinearity, RNNs can be ideally used to approximate all types of linear and nonlinear effects in a fiber transmission system and compensate for different transmission impairments. RNNs, consisting of many feedback connections, are computationally expensive for digital hardware and require at least milliseconds to conduct a single inference. Contrarily, in photonic RNN, the feedback operations are simply done by busing the signals on photonic waveguides, allowing photonic hardware capable of converging to the solution within microseconds. This architecture also adopts the neuromorphic approach and thus allows to train PNNs externally using standard machine learning algorithms.

5. NEUROMORPHIC COMPUTING APPLICATION: NONLINEAR PROGRAMMING—SOLVING OPTIMIZATION PROBLEMS AND ORDINARY DIFFERENTIAL EQUATIONS

Solving mathematical optimization problems lies at the heart of various applications present in modern technology such as machine learning, resource optimization in wireless networks, and drug discovery. Many optimization problems can be written as a quadratic program. For example, the least squares regression method can be mathematically mapped to a relatively easy quadratic program (with a positive definite quadratic matrix). Quadratic programming (QP) refers to algorithms related to solving the optimization problem of finding the extremes of a quadratic objective function subject to linear constraints. The general formulation of a quadratic program is usually solved iteratively, often requiring many time steps to reach the desired solution. The difficulty of QP grows exponentially with the dimension of the problem. Algorithms that can deal with large dimensions involve more computationally intensive techniques such as genetic algorithms or particle swarm optimization. As a result, conventional digital computers must either be limited to solving quadratic programs of very few variables, or to applications where the computation time is non-critical. Therefore, traditional computers are not appropriate to implement algorithms depending on QP for high-speed applications such as signal processing and control systems. In machine learning, many algorithms, such as SVM, require offline training because of the computational complexity of QP, but would be much more effective were they trained online.

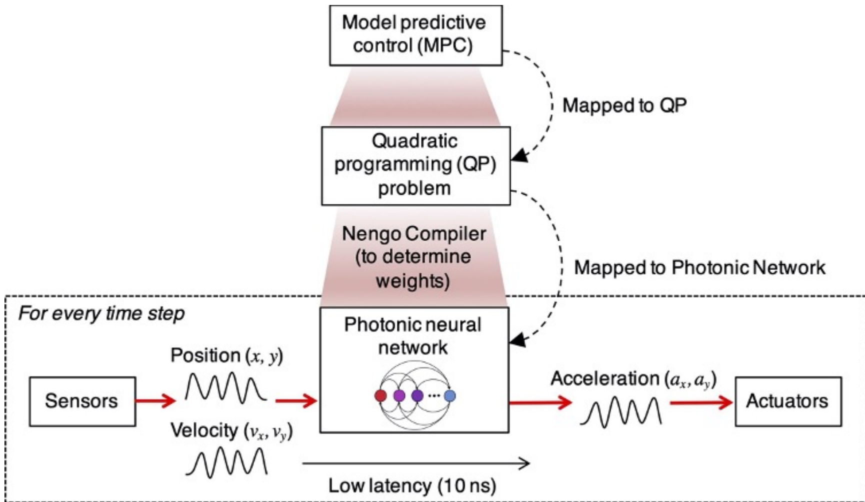


Figure 5. Schematic figure of the procedure to implement the MPC algorithm on a neuromorphic photonic processor. Firstly, map the MPC problem to QP. Then, construct a QP solver with continuous-time recurrent neural networks (CT-RNN) (Cichocki et al., 1993). Finally, build a neuromorphic photonic processor to implement the CT-RNN. The details of how to map MPC to QP, and how to construct a QP solver with CT-RNN are given in De Lima et al. (2019). Adapted from Ferreira De Lima et al. (2019) [7].

Several high-speed control problems, e.g., controlling plasma in aircraft actuations, fusion power plants, guiding of drones etc., are currently bottlenecked by the speed and latency of the control algorithms. Model predictive control (MPC) is an advanced technique to control complex systems, outperforming traditional PID control methods because it can predict control violations, rather than react to it. However, its control loop involves solving a quadratic problem at every control step, and therefore it is not computationally tractable for systems requiring speeds higher than kHz. Photonic neural networks help overcome this tradeoff by using techniques such as wavelength division multiplexing, which enables hundreds of high bandwidth signals (20 GHz) to be guided through a single optical waveguide.

Neural networks were demonstrated to solve general-purpose quadratic optimization problems by Hopfield and Tank [25],[26]. Until now, Hopfield networks have not been commonly implemented in hardware due to its all-to-all connectivity, which creates an undesirable tradeoff between neuron speed and neural network size — in an electronic circuit, as the number of connections increases, the bandwidth at which the system can operate decreases [14]. This means a photonic Hopfield network implementation can simultaneously tackle quadratic programs with large dimensions and converge in nanoseconds [7]. Implemented in photonic hardware, model predictive control (Figure 5) can be employed in systems operating in the MHz regime.

While neural networks are often used for their learning properties, they can also be programmed directly. Direct programming of analog systems has a major pitfall in that the components are unreliable and subject to parameter variation. One approach is to represent variables as population states that are robust to parameter variations. The neural engineering framework (NEF) [27] provides an algorithm to program ensembles of imperfect analog devices to perform operations on population coded variables.

Tait et al. [15] demonstrated a programmable network of two photonic neurons. The NEF algorithm was fed the responses of photonic devices, resulting in a weight matrix that allows the network to approximate variables, operations, and differential equations. It was shown in simulation how 24 neurons could emulate the Lorenz attractor. The approximation improves with more neurons. The example of the Lorenz attractor is an example of a task-based benchmark for a 24-neuron network (Figure 6). Task-based benchmarks play an important role in validating experimental systems as they begin to incorporate more neurons and parameters. The example also demonstrated the compatibility between PNNs and the NEF, including all the NEF's key principles and consequent functionality. Using this neural compiler provides a route to a variety of known applications and other benchmarks [28].

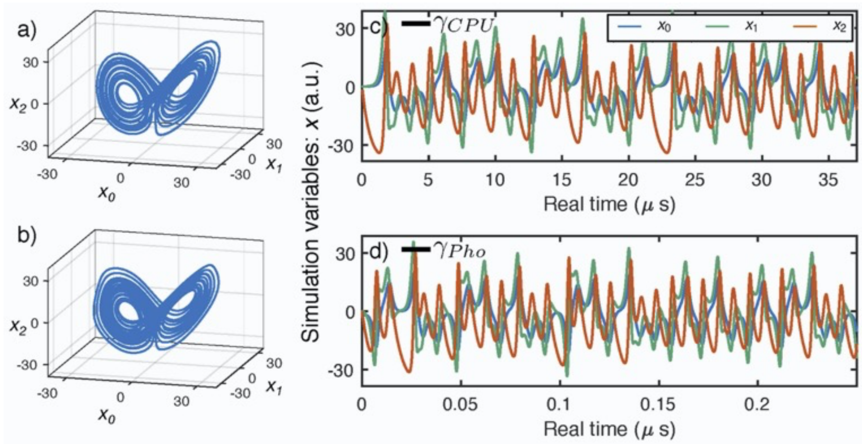


Figure 6. Photonic neural network benchmarking against a CPU. (a,b) Phase diagrams of the Lorenz at- tractor simulated by a conventional CPU (a) and a photonic neural network (b). (c,d) Time traces of simulation variables for a conventional CPU (c) and a photonic CTRNN (d). The horizontal axes are labeled in physical real time, and cover equal intervals of virtual simulation time, as benchmarked by γCPU and γPho. The ratio of real-time values of γ's indicates a 294-fold acceleration. From [15].

6. DISCUSSION AND OUTLOOK

The renaissance of field of neuromorphic photonics is enabled by the confluence of three area (Figure 7): technological advances in integrated photonics due to silicon photonics, the algorithmic advances in machine learning algorithms, and advances in analogue photonic signal processing. In the recent roadmap article [29], we outlined some scientific and technological advances necessary to meet the challenges to envision a practical neuromorphic processor outlined in the previous section.

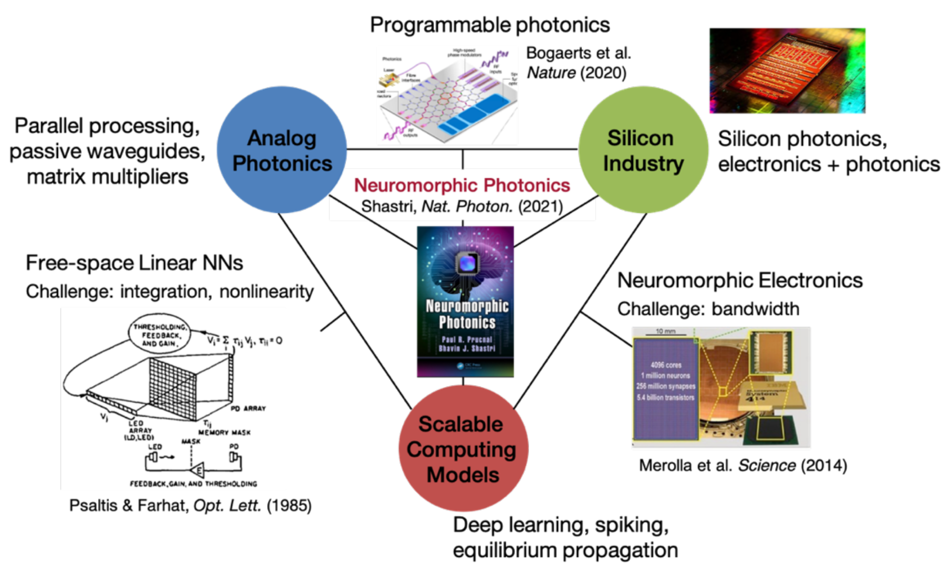


Figure 7. The advent of neuromorphic photonics is due to the convergence of recent advances in photonic integration technology, resurgence of scalable computing models (e.g., spiking, deep neural networks), and a large-scale silicon industrial ecosystem.

The emerging photonic computing models reveal the need for low-cost, manufacturable, and versatile fabrication platforms. A complete optical processor needs the core optical processor performing computing functions. They also need light sources and complex on-chip electronic circuitry for calibration and control of the network parameters. To compete with electronics, one challenge is to increase the number of optical components in a single chip to extend the spectrum of information processing capabilities. Electronic control circuits, which typically scale with the number of optical components, are another limiter of scalability. The third challenge is to retain cascadablity, implying the need for low-loss and low-power optical devices, especially for optical nonlinearities.

Silicon photonic integration provides an unprecedented platform to produce large-scale and low-cost photonic systems. Silicon photonic chip containing thousands of components has been demonstrated [30]. Mainstream silicon photonic platforms offer a device library (modulators, waveguides, detectors) to implement computing functions that are vigorously pursued for different photonic computing models. However, silicon photonic platforms are still in a nascent period with many potential advantages yet to be realized. Some key components, such as light sources, are not yet widely available in mainstream silicon photonics foundries. The devices' efficiency, especially nonlinear devices, need to be optimized to compete with electronics. Modifications to standard and mainstream manufacturing processes promise to extend the capability of existing devices and platforms. Here we discuss the emerging technologies that integrate light sources, electronic circuits, and highly efficient devices necessary for an optical processor. These technologies carry the potential to reinvigorate investigation into advanced photonic information processing in the immediate future.

Electrical control circuits for reconfigurable optics: Photonic chips require DC analog signals (e.g. bias voltages/currents), control systems (e.g. feedback, algorithms etc.), interfaces with electronics (e.g. digital-to-analog converters (DACs) and analog-to-digital converter (ADCs)), and require stabilization (e.g. temperature). Dedicated analog electronic circuits are needed for these purposes. As a result, photonic chips generally require significantly more electrical ports than optical ones, leading to a high electronic interconnect density. This challenge can be overcome by co-integration of CMOS with photonic chips.

There are several technological routes with different tradeoffs, including wire-bonding, flip-chip bonding, and monolithic fabrication. Wire-bonding and flip-chip bonding suffer from high packaging costs [31]. Traditional wire-bonding on the edge of the chip hardly scale up to more than 1,000 wires. Approaches based on flip-chip bonding are promising solutions to the limited electrical I/O numbers [32]. Monolithic fabrication entails integrating electronics and photonics on the same substrate, which offers the closest integration and the most cost-effective packaging solution. So-called 'zero-change' platforms earn their name by offering a monolithic fabrication process that is minimally changed from an industry-standard fabrication process CMOS process, which trade-off photonic performance with electronic circuitry for analog or digital control (waveguide losses, limited choice in transistor nodes) [33]. Another direction is to adapt SOI photonic processes to enable some active integrated electronic logic [34].

Nonvolatile materials for optical memories and in-memory computing: In-memory computing is emerging as a promising paradigm to address the rapid increase in the processor to memory communication requirement as machine learning and neural network algorithms are intensively used. Nonvolatile materials allow implementing memory elements directly on the computing devices, thus promising to solve the processor to memory communication bottleneck. On-chip nonvolatile memories that can be written, erased, and accessed optically are rapidly bridging a gap toward on-chip photonic computing [35]. For example, current approaches with photonic neural networks are driven by electronic circuits or micro-controllers to load matrix. As discussed earlier, the integration and packaging of large-scale optical and electronic circuits can be challenging in terms of cost and power. In some machine learning applications such as deep learning inference) the weights, once trained, do not have to be updated often or at all. In these cases, the integration of novel photonic memory technologies can limit the need to read from and write to electronic memories with DACs and ADCs. Non-volatile photonic memories with phase change memories (PCMs) set and retain the weights without further holding power after being set [36], resulting in almost zero power consumption in performing matrix multiplication operations. The change in optical transmissions is also reversible, therefore enabling dynamic synaptic plasticity and online learning in photonic neural networks

REFERENCES

[1] Schuman, Catherine D., et al. "A survey of neuromorphic computing and neural networks in hardware." arXiv preprint arXiv:1705.06963 (2017).

[2] Hasler, Jennifer, and Harry Bo Marr. “Finding a roadmap to achieve large neuromorphic hardware systems.” Front. Neurosci. 7 (2013): 118.
[3] Merolla, Paul A., et al. “A million spiking-neuron integrated circuit with a scalable communication network and interface." Science 345.6197 (2014): 668-673.
[4] Davies, Mike, et al. “Loihi: A neuromorphic manycore processor with on-chip learning.” IEEE Micro 38.1 (2018): 82-99.
[5] Ferreira De Lima, Thomas, et al. “Progress in neuromorphic photonics.” Nanophotonics 6.3 (2017): 577-599.
[6] Shastri, Bhavin J., et al. “Photonics for artificial intelligence and neuromorphic computing.” Nat. Photon. 15.2 (2021): 102-114.
[7] Ferreira De Lima, Thomas, et al. “Machine learning with neuromorphic photonics.” J. Lightwave Technol. 37.5 (2019): 1515-1534.
[8] Prucnal, Paul R., and Bhavin J. Shastri. Neuromorphic photonics. CRC Press, 2017.
[9] Brunner, Daniel, et al. “Parallel photonic information processing at gigabyte per second data rates using transient states.” Nat. Commun. 4.1 (2013): 1-7.
[10] Vandoorne, Kristof, et al. “Experimental demonstration of reservoir computing on a silicon photonics chip.” Nat. Commun. 5.1 (2014): 1-6.
[11] Larger, Laurent, et al. “Photonic information processing beyond Turing: an optoelectronic implementation of reservoir computing.” Opt. Express 20.3 (2012): 3241-3249.
[12] Shen, Yichen, et al. “Deep learning with coherent nanophotonic circuits.” Nat. Photon. 11.7 (2017): 441-446.
[13] Hughes, Tyler W., et al. “Training of photonic neural networks through in situ backpropagation and gradient measurement.” Optica 5.7 (2018): 864-871.
[14] Tait, Alexander N., et al. “Neuromorphic photonic networks using silicon photonic weight banks.” Sci. Rep. 7.1 (2017): 1-10.
[15] Tait, Alexander N., et al. “Silicon photonic modulator neuron.” Phys. Rev. Appl. 11.6 (2019): 064043.
[16] Tait, Alexander N., et al. “Broadcast and weight: an integrated network for scalable photonic spike processing.” J. Lightwave Technol. 32.21 (2014): 4029-4041.
[17] Ferreira De Lima, Thomas, et al. “Primer on silicon neuromorphic photonic processors: architecture and compiler.” Nanophotonics 9.13 (2020): 4055-4073.
[18] Cisco. (n.d.). Cisco annual internet report - cisco annual internet report (2018–2023) white paper. https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual- internet-report/white-paper-c11-741490.html.
[19] Pillai, Bipin Sankar Gopalakrishna, et al. “End-to-end energy modeling and analysis of long-haul coherent transmission systems.” J. Lightwave Technol. 32.18 (2014): 3093-3111.
[20] Agrell, Erik, et al. “Roadmap of optical communications.” J. Opt.18.6 (2016): 063002.
[21] Argyris, Apostolos, Julián Bueno, and Ingo Fischer. “Photonic machine learning implementation for signal recovery in optical communications.” Sci. Rep. 8.1 (2018): 1-13.
[22] Huang, Chaoran, et al. “A silicon photonic-electronic neural network for fibre nonlinearity compensation.” Nature Electronics 4.11 (2021): 837–844.
[23] Khan, Faisal Nadeem, et al. “An optical communication's perspective on machine learning and its applications.” J. Lightwave Technol. 37.2 (2019): 493-516.
[24] Zhang, Shaoliang, et al. “Field and lab experimental demonstration of nonlinear impairment compensation using neural networks.” Nat. Commun. 10.1 (2019): 1-8.
[25] Cochocki, A., and Rolf Unbehauen. Neural networks for optimization and signal processing. John Wiley & Sons, Inc., 1993.
[26] Hopfield, John J., and David W. Tank. “Computing with neural circuits: A model.” Science 233.4764 (1986): 625-633.
[27] Stewart, Terrence C., and Chris Eliasmith. “Large-scale synthesis of functional spiking neural circuits.” Proc. IEEE 102.5 (2014): 881-898.
[28] Stewart, Terrence C., et al. “Closed-loop neuromorphic benchmarks.” Front. Neurosci. 9 (2015): 464.
[29] Berggren, Karl, et al. “Roadmap on emerging hardware and technology for machine learning.” Nanotechnology 32.1 (2020): 012002.
[30] Sun, Jie, et al. "Large-scale nanophotonic phased array." Nature 493.7431 (2013): 195-199.

[31] Palesko, Chet A., and E. Jan Vardaman. "Cost comparison for flip chip, gold wire bond, and copper wire bond packaging." 2010 Proceedings 60th Electronic Components and Technology Conference (ECTC). IEEE, 2010.
[32] Patterson, Deborah, Isabel De Sousa, and Lousi-Marie Achard. "The future of packaging with silicon photonics." Chip Scale Rev 21.1 (2017): 1-10.
[33] Giewont, Ken, et al. "300-mm monolithic silicon photonics foundry technology." IEEE Journal of Selected Topics in Quantum Electronics 25.5 (2019): 1-11.
[34] Jha, Aashu, et al. "Lateral bipolar junction transistor on a silicon photonics platform." Optics Express 28.8 (2020): 11692-11704.
[35] Feldmann, Johannes, et al. "All-optical spiking neurosynaptic networks with self-learning capabilities." Nature 569.7755 (2019): 208-21
[36] Ríos, Carlos, et al. "In-memory computing on a photonic platform." Science Advances 5.2 (2019): eaau5759.