Brain-Inspired Photonic Hardware Platforms

Bhavin J. Shastri, Alexander N. Tait, Mitchell A. Nahmias, Thomas Ferreira de Lima,
and Paul R. Prucnal

Department of Electrical Engineering, Princeton University, Princeton NJ, 08544 USA

shastri@ieee.org

Abstract: Photonic integrated circuit technology has the potential to revolutionize optical information processing, beyond conventional binary-logic approaches, granting the capacity of complex, ultrafast categorization and decision making. We will discuss the progress and requirements of scalable and reconfigurable emerging brain-inspired photonic hardware platforms.

Photonics has revolutionized information communication, while electronics, in parallel, has dominated information processing. Recently, there has been a determined exploration of the unifying boundaries between photonics and electronics on the same substrate, driven in part as Moore's Law approaches its long-anticipated end. For example, the computational efficiency (multiply-accumulate (MAC) operations per joule) for digital processing has leveled-off around 100 pJ per MAC [1]. As a result, there has been a widening gap between computational efficiency and the next generation needs, such as Big Data applications which require advanced pattern matching and analysis on voluminous data in real-time. This in turn, has led to expeditious advances in: (1) emerging devices that are called beyond-CMOS or More-than-Moore, (2) novel processing or unconventional computing architectures called beyond-von Neumann, that are bio-inspired i.e. *neuromorphic*, and (3) CMOS-compatible photonic interconnect technologies. Collectively, these research endeavors have opened opportunities for emerging photonic hardware platforms that are reconfigurable i.e. programmable optical integrated circuits (POIC) such as *photonic spike processors* [2–7]. Such chips in the near future could combine ultrafast operation, moderate complexity, and full programmability, extending the bounds of computing for applications such as navigation control on hypersonic aircrafts, and real-time sensing and analysis of the radio frequency (RF) spectrum. We will discuss the current progress and requirements of such a platform.

In a photonic spike processor, information is encoded as *events* in the temporal and spatial domain of spikes (or optical pulses). This hybrid coding scheme is digital in amplitude but analog in time and benefits from the bandwidth efficiency of analog processing and the robustness to noise of digital communication. Optical pulses are received, processed, and generated by certain class of semiconductor devices that exhibit *excitability*—a nonlinear dynamical mechanism underlying all-or-none responses to small perturbations [8]. Optoelectronic devices operating in the excitable regime are dynamically analogous with the spiking dynamics observed in *neuron biophysics* but roughly eight orders of magnitude faster.

Using a gain and saturable absorber (SA) excitable laser, we recently demonstrated [2] the first unified, experimental demonstration of low-level spike processing functions in an optical platform. We showed that this platform can simultaneously exhibit logic-level restoration, cascadability and input-output isolation—fundamental challenges in optical information processing [9, 10]. We also implemented the classic spike processing tasks of temporal pattern detection and stable recurrent memory—while simple, these fundamental behaviors underly higher level processing. We will review the recent surge of interest in the information processing abilities of such excitable optoelectronic devices.

The field is now reaching a critical juncture where there is shift from studying single excitable (spiking) devices to studying an interconnected network of such devices to build a photonic spike processor. We recently proposed [3] an on-chip networking architecture called *broadcast-and-weight* (Fig. 1) that could support massively parallel (all-to-all) interconnection between excitable lasers using wavelength division multiplexing (WDM). Within this network, an interface called *processing network node* (PNN) (Fig. 1), gives the ability for one laser to communicate with others. A PNN handles inputs from multiple sources by weighting and spatially summing (i.e. weighted addition), before sending the resulting signal to an excitable laser. PNN consist of spectral filter banks of continuously tunable microring resonator (MRR) weight banks that partially drop WDM channels that are present *without* demultiplexing. The use of MRRs as continuous-valued weights has recently been shown [11, 12]. In this way, each such PNN is connected with every other PNN in a *broadcast loop* using a broadcast-and-weight network. That is, the broadcast loop is fully multiplexed and capable of supporting N^2 interconnects in just one transparent link (waveguide) with N WDM channels. Recent channel density studies [13] quantified analysis for multiwavelength analog networks, and derived a limit of 108 PNNs per broadcast loop, of which many can co-exist on a single chip.
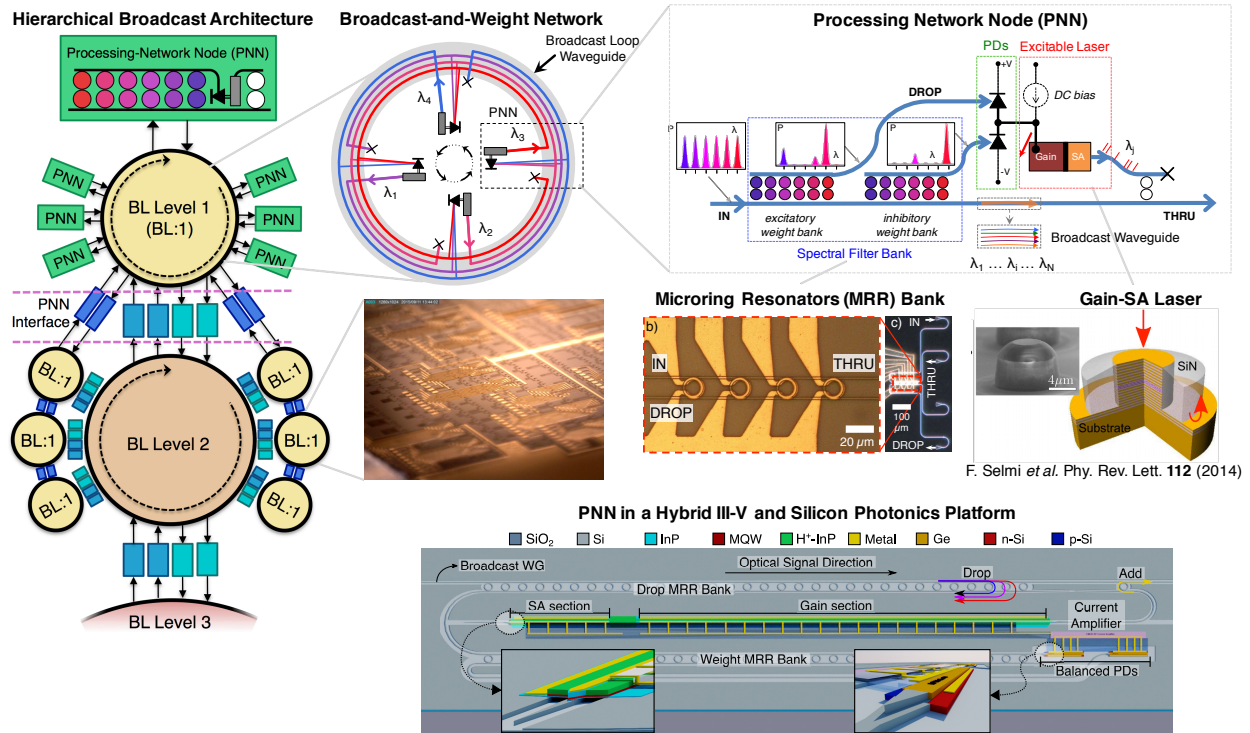
Fig. 1. Concepts, devices, and networks to build a photonic spike processor.

A hybrid III-V and silicon photonics platform (Fig. 1) is a candidate for an integarted hardware platform [14]. Scalable and fully reconfigurable networks of excitable lasers can be implemented in the silicon photonic layer of modern hybrid integration platforms, in which spiking lasers in a bonded InP layer are densely interconnected through a silicon layer. Such a photonic spike processor will potentially be able to support several thousand interconnected PNNs. It is predicted [15] that such a chip would have a computational efficiency of 260 fJ per MAC, which surpasses the energy efficiency wall by two orders of magnitude while operating at high speeds (i.e. signal bandwidths 10 GHz).

The emerging field of photonic spike processors has received tremendous interest and continues to receive further developments as PICs increase in performance and scale. As novel applications requiring real-time, ultrafast processing—such as the exploitation of the RF spectrum—become more critical, we expect that these systems will find use in a variety of high performance, time-critical environments.

References

1. B. Marr *et al*. IEEE Trans. Very Large Scale Integr. Syst. **21**, 147–151 (2013).
2. B. J. Shastri *et al*. Sci. Rep. **6**, 19,126 EP – (2016).
3. A. Tait *et al*. J. Lightw. Technol. **32**, 3427–3439 (2014).
4. M. A. Nahmias *et al*. IEEE J. Sel. Top. Quantum Electron. **19** (2013).
5. F. Selmi, *et al*. Phys. Rev. Lett. **112**, 183,902 (2014).
6. B. Romeira *et al*. Sci. Rep. **6**, 19,510 EP – (2016).
7. A. Hurtado and J. Javaloyes, Appl. Phys. Lett. **107**, 241103 (2015).
8. B. Krauskopf *et al*. Opt. Commun. **215**, 367–379 (2003).
9. R. S. Tucker, Nature Photon. **4**, 405–405 (2010).
10. D. A. B. Miller, Nature Photon. **4**, 3–5 (2010).
11. A. N. Tait *et al*. IEEE Photon. Technol. Lett. **28**, 887–890 (2016).
12. A. N. Tait *et al*. Opt. Express (2016).
13. A. N. Tait *et al*. IEEE J. Sel. Topics Quantum Electron. **(accepted)** (2016).
14. M. A. Nahmias *et al*. Opt. Express **23**, 26,800–26,813 (2015).
15. P. R. Prucnal *et al*. *Brain-Inspired Photonic Processing*. CRC Press, Taylor & Francis Group (2016).