Xavier Porte, Anas Skalli, Nasibeh Haghighi, Stephan Reitzenstein, James A. Lott, Daniel Brunner

Neural network computing using a large-area VCSEL

2021 27th International Semiconductor Laser Conference (ISLC) | 978-1-6654-4133-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/ISLC51662.2021.9615664

Neural network computing using a large-area VCSEL

Xavier Porte[1,*], Anas Skalli[1], Nasibeh Haghighi[2],
Stephan Reitzenstein[2], James A. Lott[2], and Daniel Brunner[1]

[1]*Institut FEMTO-ST, Université Bourgogne Franche-Comté CNRS UMR 6174, Besançon, France*
2 Institut für Festkörperphysik, Technische Universität Berlin, Hardenbergstraße 36, 10623 Berlin, Germany

*Email: javier.porte@femto-st.fr

Abstract – We implement a fully parallel photonic neural network based on the spatially distributed modes of a large-area vertical-cavity surface-emitting laser. All photonic connections are realized in hardware and the system is capable of autonomous operation.

I. Introduction

Over the past decade, artificial Neural Networks (NNs) have revolutionized computing. High-performance computing hardware is crucial for modern NN schemes. Photonics promises strong advantages in terms of parallelism, yet until now scalable and integrable concepts are scarce and partially rely on exotic substrates. The majority of large scale and parallel photonic NN demonstrations are neither standalone nor autonomous [1], usually lacking fundamental NN constituents or requiring substantial interaction with a classical electronic computer. In this contribution, we implement a fully parallel photonic reservoir computer based on the spatially distributed modes of an efficient and fast semiconductor laser [2]. Crucially, all neural network connections are realized in hardware, and our laser-based and fully parallel NN comprising ~ 100 neurons produces results without pre- or post-processing.

II. Results

As photonic neuron substrate we use the complex multimode field of an injection locked large-area vertical-cavity surface-emitting laser (LA-VCSEL) of ~20 µm diameter emitting around 920 nm. Figure 1(a) depicts the device we use and its free-running emission profile at a bias current of 1.3 times its lasing threshold. This LA-VCSEL follows a minimalistic design principle that optimizes operation efficiency and bandwidth [3] and were fabricated via standard commercial technology. A detailed description of the device characteristics can be found in [2].

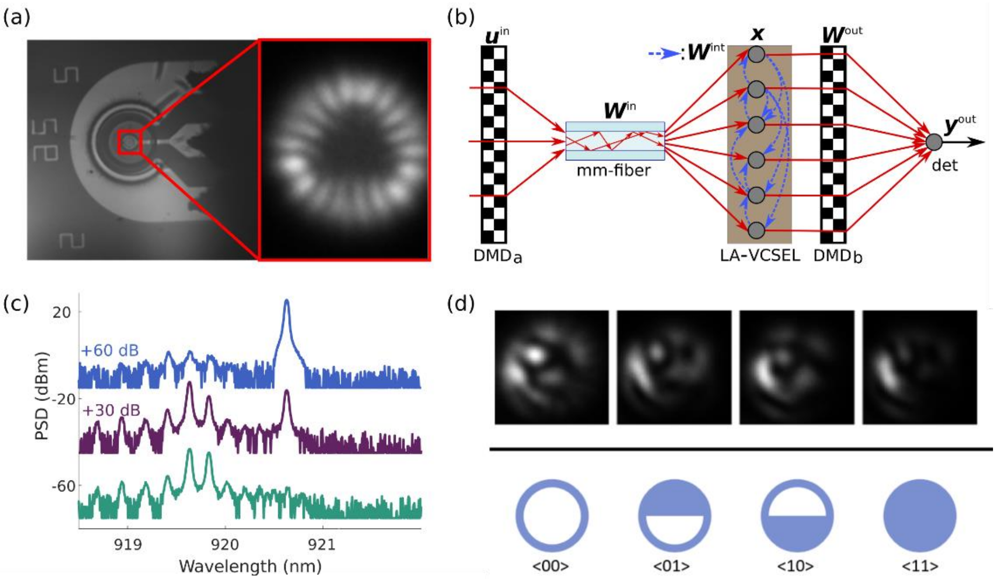

Fig. 1. (a) Left: White light image of the LA-VCSEL (~20 µm in diameter); Right: Magnified multimode free lasing emission for a bias current of 1.3 the LA-VCSEL threshold. (b) Schematic illustration of the photonic NN's sections linked to their corresponding physical devices. A digital micromirror device (DMD_a) encodes input information $\boldsymbol{u}^{in}$, which is mixed through the complex transfer matrix of a multimode fiber (mm-fiber). The LA-VCSEL acts as recurrent reservoir with state **x**, providing device-inherent internal coupling $\boldsymbol{W}^{int}$. DMD_b implements programmable Boolean readout weights $\boldsymbol{W}^{out}$, and a detector records computational result y^{out}. (c) Optical spectra under DC optical injection. For clarity, offsets of 30dB and 60dB have been respectively added to the middle and top spectra. (d) Perturbed mode profiles (upper panels) under different injection patterns (lower panels) showing the highly nonlinear nature of the LA-VCSEL response.

Our reservoir computing scheme is illustrated in Fig. 1(b). All the photonic NN connections are implemented in hardware: the complex transfer matrix of a multimode (mm) fiber ($\boldsymbol{W}^{in}$) couples the LA-VCSEL to the injected information ($\boldsymbol{u}^{in}$), which is Boolean encoded on a digital micro-mirror device (DMD_a). The VCSEL then transforms the injected information non-linearly yielding mode profiles such as in Figure 2. This transformation is comparable to the action of a neural network: similar inputs result in vastly different VCSEL responses, and we can use the laser for pattern recognition. Intra-cavity fields and carrier diffusion intrinsic to LA-VCSELs recurrently couple the photonic neurons ($\boldsymbol{W}^{\mathbf{int}}$), and trainable readout weights ($\boldsymbol{W}^{\mathbf{out}}$) are encoded on the DMD_b and photo-detected to directly provide the computational result (y^{out}). We operate our recurrent photonic NN in its steady state, and its bandwidth is limited by the input DMD's frame rate to around 100 inferences per second.

Although optical injection into multimode LA-VCSELs was extensively studied in the past [5], here we report injection-lock of such device to a complex optical input field for the first time. Figure 1(b) depicts the device's optical response under DC injection of ring-like optical pattern of different thicknesses generated with DMD_a. Lower spectrum (in green) corresponds to the free-running laser when biased at $1.28I_{th}$, meanwhile the middle and upper spectra respectively correspond to power injection ratios of P_{inj}/P_{VCSEL}=0.03 and P_{inj}/P_{VCSEL}=0.4 for identical bias condition. These latter spectra have been shifted upwards +30 dB and +60 dB for clarity reasons. Figure 1(d) shows the resulting photonic ANN state for the example of injecting the four possible configurations of 2-bit symbols. Individual responses significantly differ for each case, which is a prerequisite to differentiate individual digits.

The system learns the best configuration of output mirrors to differentiate between different input patterns. The DMD_b output mirrors each have two possible positions, implementing Boolean weights. We explore different Boolean learning strategies. The first, already described in [4], is a Markovian process where at each epoch a single mirror is flipped, the change is kept if it had a positive impact on the performance metric (NMSE error), otherwise the change is reversed, and another mirror is flipped. Following this learning strategy, we trained the readout weights to perform 2-bit header recognition, 2-bit XOR and 2-bit digital-analog conversion tasks. Figure 2(a) depicts exemplary convergence curves on a training set for classifying 2-bit patterns. The error rates achieved by our system for the different tasks are down to: $0.9x10^{-3}$ for the 2-bit header recognition; $2.9x10^{-2}$ for the 2-bit XOR task; and $5.4x10^{-2}$ for the 2-bit digital-to-analog conversion.

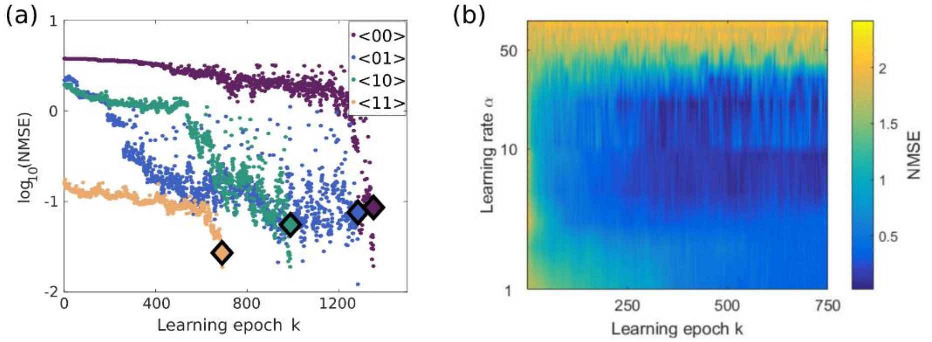

Fig. 1 (a) Convergence on training set for classifying bits < 00 >, < 01 >, < 10 > and < 11 >. Diamond symbols are the corresponding average testing NMSE. (b) 2D plot of the error (NMSE, color coded) as a function of the learning epochs and the learning rate.

Our second learning strategy adapts the previous Markovian process by taking inspiration for the widely used stochastic Gradient descent algorithm. In our adaption, at each epoch the number of flipped mirrors ($n_{mirrors}$) depends on the error via a constant learning rate α, which is a hyperparameter that we tuned as: $n_{mirrors}$=ceil(α*NMSE). The mirrors are still chosen randomly, but such improved strategy allows us to take big steps in the parameter space when the error is high, and progressively reduce the size of these steps as we start converging towards lower errors (note that if α=0, $n_{mirrors}$ is always 1). Figure 2(b) shows the clear dependence between learning rate (α), learning speed and NMSE performance. Thanks to this learning strategy, we successfully improve both the speed of convergence during learning as well as the final computational performance.

In conclusion, we demonstrate a fully analog spatially-extended photonic reservoir computer where each of the system's constituents implemented in hardware in readily available and cost-effective telecommunications components. Furthermore, our present system is scalable in size to much larger networks in excess of 1000 neurons per layer and to bandwidths in excess of 20 GHz, establishing a clear road map for future high-performance photonic hardware for NNs.

III. References

[1] G. Van Der Sande, D. Brunner and M. C. Soriano. "Advances in photonic reservoir computing," *Nanophotonics* **6**(3), 561-576 (2017).

[2] X. Porte, A. Skalli, N. Haghighi, S. Reitzenstein, J. A. Lott and D. Brunner, "A complete, parallel and autonomous photonic neural network in a semiconductor multimode laser," J. Phys. Photonics 3, 024017 (2021).

[3] N. Haghighi, P. Moser and J. A. Lott. "Power, bandwidth, and effciency of single VCSELs and small VCSEL Arrays". *IEEE JSTQE* **25**(6), 1-15 (2019).

[4] L. Andreoli, X. Porte, S. Chrétien *et al*. "Boolean learning under noise-perturbations in hardware neural networks". *Nanophotonics* **9**(13), 4139-4147 (2020).

[5] T. Ackemann, S. Barland, M. Giudici *et al*. "Patterns in broad-area microcavities". *Phys. Status Solidi b* **221**, 133–6 (2000).