PAPER • OPEN ACCESS

A complete, parallel and autonomous photonic neural network in a semiconductor multimode laser

To cite this article: Xavier Porte *et al* 2021 *J. Phys. Photonics* **3** 024017

View the article online for updates and enhancements.

Journal of Physics: Photonics

PAPER

A complete, parallel and autonomous photonic neural network in a semiconductor multimode laser

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 UMR6174, Besançon, France
2 Institut für Festkörperphysik, Technische Universität Berlin, Hardenbergstraße 36, 10623 Berlin, Germany
* Author to whom any correspondence should be addressed.

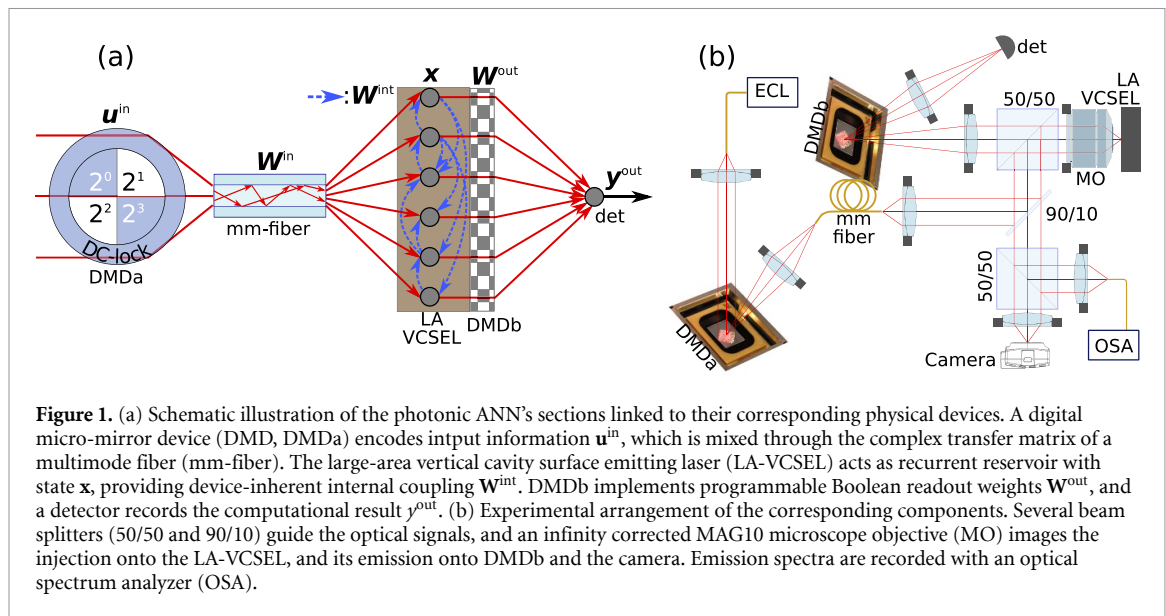E-mail: daniel.brunner@femto-st.fr

Abstract

Neural networks are one of the disruptive computing concepts of our time. However, they fundamentally differ from classical, algorithmic computing. These differences result in equally fundamental, severe and relevant challenges for neural network computing using current computing substrates. Neural networks urge for parallelism across the entire processor and for a co-location of memory and arithmetic, i.e. beyond von Neumann architectures. Parallelism in particular made photonics a highly promising platform, yet until now scalable and integratable concepts are scarce. Here, we demonstrate for the first time how a *fully parallel* and *fully implemented* photonic neural network can be realized by spatially multiplexing neurons across the complex optical near-field of a semiconductor multimode laser. Discrete spatial sampling defines $\sim$90 nodes on the surface of a large-area vertical cavity surface emitting laser that is optically injected with the artificial neural networks input information. Importantly, all neural network connections are realized in hardware, and our processor produces results without pre- or post-processing. Input and output weights are realized via the complex transmission matrix of a multimode fiber and a digital micro-mirror array, respectively. We train the readout weights to perform 2-bit header recognition, a 2-bit XOR logical function and 2-bit digital to analog conversion, and obtain $< 0.9 \times 10^{-3}$ and 2.9×10^{-2} error rates for digit recognition and XOR, respectively. Finally, the digital to analog conversion can be realized with a standard deviation of only 5.4×10^{-2}. Crucially, our proof-of-concept system is scalable to much larger sizes and to bandwidths in excess of 20 GHz.

1. Introduction

Artificial neural networks (ANNs) were conceptualized already in 1943, with providing a 'logical calculus of the ideas immanent in nervous activity' as the original motivation [1]. In line with this spirit, nonlinear nodes standing in as neurons form the ANN with the help of connections, much like synapses, dendrites and axons connect biological neurons. Such a *parallel* architecture distributes computing across a usually large number of simple nonlinear nodes, and programming becomes optimizing the network's topology. Besides several important breakthroughs such as error-back propagation to optimize, i.e. train an ANN [2], it was the increasing availability of powerful computing hardware [3] which turned ANNs into a technology of primary societal [4], scientific [5] and economic relevance.

High-performance computing hardware is therefore a must for ANNs. The reason lies in a core aspect of networks: the number of possible connections scales quadratically with the number of network nodes and every new computational result requires querying all network connections. Integration of parallel ANNs in electronic circuits is highly challenging due to: (a) the strong energy penalty for switching the large number of signaling wires; and (b) a fundamental incompatibility with two-dimensional (2D) integration [6]. Thus,

Figure 1. (a) Schematic illustration of the photonic ANN's sections linked to their corresponding physical devices. A digital micro-mirror device (DMD, DMDa) encodes intput information **u**^{in}, which is mixed through the complex transfer matrix of a multimode fiber (mm-fiber). The large-area vertical cavity surface emitting laser (LA-VCSEL) acts as recurrent reservoir with state **x**, providing device-inherent internal coupling **W**^{int}. DMDb implements programmable Boolean readout weights **W**^{out}, and a detector records the computational result y^{out}. (b) Experimental arrangement of the corresponding components. Several beam splitters (50/50 and 90/10) guide the optical signals, and an infinity corrected MAG10 microscope objective (MO) images the injection onto the LA-VCSEL, and its emission onto DMDb and the camera. Emission spectra are recorded with an optical spectrum analyzer (OSA).

it has so far been impossible to integrate parallel and large ANNs, and application specific circuits still heavily rely on serial routing.

Photonics is a promising alternative as it mitigates much of the energy cost associated to information transduction. This potential advantage over electronics was realized decades ago [7, 8] and is detailed in [9]. Reducing energy deposition into the substrate could enable novel three-dimensional integration [10] to ultimately realize scalable integration of parallel ANNs [11]. Optical ANNs did outperform their electronic counterparts already at a previous stage [12], and their superior scaling has recently again been demonstrated [13]. Finally, ANNs implemented in silicon photonics [14–16] promise ultra-high speed, ultra low latency and the benefits of integration leveraging the tools of silicon and 2D lithography.

However, the vast majority of large scale and parallel photonic ANN demonstrations are not standalone and autonomous. Most proof-of-concepts lack one or several of an ANN's constituents, require substantial involvement of a traditional computer in the creation of the ANN's state, are not parallel or run on substrates too exotic for realistic integration in mid-term technological platforms. Here, we address all these points and experimentally demonstrate a fully parallel photonic reservoir computer [17, 18] with 90 neurons that we train and operate in realtime with minimal interference from a control computer. Neurons are defined at periodic and adjacent spatial sampling positions across the complex multimode field of an injection locked large area vertical cavity surface emitting laser (LA-VCSEL). The LA-VCSEL has a diameter of $\sim$20 μm, lases around 920 nm and follows design principles developed for high bandwidth and high efficiency VCSEL arrays [19]. It was fabricated via standard ultraviolet contact mask lithography, dry (inductively coupled plasma-reactive ion etching) mesa etching, and thin-film metal evaporation and lift-off methods in a university cleanroom [20, 21]. All the ANN's connections are implemented in photonic hardware. The complex transfer matrix of a multimode fiber (mm-fiber) [22] couples injected information, which is Boolean encoded on a DMD, to the LA-VCSEL. Intra-cavity fields and carrier diffusion intrinsic to LA-VCSELs recurrently connect the photonic neurons [23]. Finally, trainable readout weights are encoded on a digital micro-mirror device (DMD) [24] which is imaged on a large-area detector to directly provide the ANN's computational result.

We operate our recurrent photonic ANN in its steady state, and its bandwidth is limited by the input DMD's frame rate to $\sim$130 inferences per second. However, such semiconductor lasers can be modulated extremely fast, often in excess of 20 GHz [20, 25], which therefore corresponds to the potential intrinsic speed of our system. We train the readout weights to perform 2-bit header recognition, a 2-bit XOR and 2-bit digital to analog conversion (DAC). For the 2-bit header recognition our systems achieves an excellent error rate smaller than 0.9×10^{-3}, for the more demanding XOR it is still a respectable 2.9×10^{-2}. Finally, the digital to analog conversion can be realized with a standard deviation of only 5.4×10^{-2}. In this context it is important to highlight the fully analog nature of our photonic ANN.

2. A photonic ANN in a complex and continuous nonlinear media

Parallelism in our photonic ANN is based on spatial mode multiplexing, and the correspondence between the ANN's conceptual sections and the relevant photonic hardware is schematically illustrated in figure 1(a).

The collimated beam of a tunable external cavity laser (ECL, Toptica CTL 950) illuminates an area of DMDa (Vialux XGA 0.7" V4100) with a Gaussian cross-section of 90 pixels in diameter. A DMD comprises mirrors which can be discretely switched between $\pm 12°$, but only mirrors in $-12°$ orientation inject information into the ANN. Input information $\mathbf{u}^{\text{in}}$ is therefore a Boolean matrix. Within the illuminated area we define a ring that is constantly active. This DC-locking signal $\mathbf{u}^{\text{DC}}$ stabilises the LA-VCSEL's potentially autonomous free-running dynamics and hence allows consistency of response [24]. Inside the remaining illuminated area we define n sub-sections, much like wedges of a pie, with which we encode input information $\mathbf{u}^{\text{d}}$ for digit d in an n-bit binary representation.

In figure 1(a), input information $\mathbf{u}^{\text{in}} = \mathbf{u}^{\text{DC}} + \mathbf{u}^{\text{d}}$ corresponds to injecting digit $d = 9$ in a $n = 4$ bit representation. The spatially-modulated beam is collected by a mm-fiber that implements complex input weights $\mathbf{W}^{\text{in}}$ through optical mode mixing induced by birefringence and scattering [22]. Based on the fiber's V number, we approximate that our input fiber supports between 30 and 40 modes, i.e. an input matrix of equal size.

The near-field at the mm-fiber's output facet is injected into the LA-VCSEL. Semiconductor lasers react highly nonlinearly to optical injection, and previously we showed that this nonlinear response can be harvested to implement the nonlinear nodes of an ANN [26, 27]. The feasibility of this concept was shown numerically for 33 GHz [28]. However, here each neuron is physically implemented in parallel via the LA-VCSEL's complex and high-dimensional multimode field. Optical diffraction of the LA-VCSEL's intra-cavity field as well as carrier diffusion induce interactions between different locations, i.e. coupling between our photonic nodes [23]. Dynamics induced by optical injection are governed by an equation of the type

$$\tau \dot{\mathbf{x}}(t) = -\mathbf{x}(t) + \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \tag{1}$$

where τ is the system's response time and $\mathbf{x}(t)$ is the LA-VCSEL's state at time t distributed across its surface. State $\mathbf{x}(t)$ comprises the local optical field amplitude, the optical phase and the electronic carriers divided between spin up and spin down populations [23]. While equation (1) in its full expression is highly non-trivial, structurally it is comparable to the concept of Neural Ordinary Differential Equations [29], and we leverage state $\mathbf{x}(t)$ to implement a recurrent neural network. As we will operate $\mathbf{x}(t)$ orders of magnitude slower than any of a LA-VCSEL's inherent time scales (nanoseconds and below for relaxation oscillations and photon lifetimes), we only need to consider its steady state.

The spatially distributed emission power $\mathbf{P}(t)$ represents the neurons' readout states. This spatially continuous multimode optical field varies on scales of $\gtrsim 3\ \mu$m, which allows for ~ 60 features to fit inside the LA-VCSEL's aperture. For implementing programmable readout weights we image the laser's near-field onto DMDb (Vialux XGA 0.7" D4100), which applies Boolean output weights $\mathbf{W}^{\text{out}}$ to $\mathbf{P}(t)$. Mirrors of DMDb therefore define individual neurons through spatial sampling, and here we over-sample by ~ 1.5 which creates in 90 photonic neurons. The weighted state is converted in output signal y^{out} via an optical detector (det):

$$y^{\text{out}}(t) \propto \mathbf{W}^{\text{out}}\mathbf{P}(t). \tag{2}$$

We image DMDb on the detector, and hence optical fields of individual DMD-mirrors do not spatially overlap. This makes the detected signal proportional to the weighted sum of powers. Finally, $y^{\text{out}}(t)$ is normalized by its standard deviation and its mean is subtracted.

Evolutionary learning will be used to optimize only $\mathbf{W}^{\text{out}}$ [30]. $\mathbf{W}^{\text{in}}$ and $\mathbf{W}^{\text{int}}$ do not partake in task-specific optimization, and we therefore implement a reservoir computer [17]. Output $y^{\text{out}}(t)$ is recorded for t^{T} discrete samples, hence t^{T} is our learning batch-size and we posses labeled target values $y^{\text{T}}(t)$ with $t \in [1, \cdots, t^{\text{T}}]$. We compute mean square error $\epsilon_k = ||y_k^{\text{out}}(t) - y^{\text{T}}(t)||^2 / t^T$ at each learning epoch k, and at the transition from $k \rightarrow k+1$ the single weight $W_{l(k),k}^{\text{out}}$ at position $l(k)$ is inverted. If such weight inversion was beneficial ($\epsilon_{k+1} < \epsilon_k$) we keep this particular modification, if not we revert according to $W_{l(k+1),k+1}^{\text{out}} = W_{l(k+1),k}^{\text{out}}$. In previous experiments on a spatial light modulator based setup [30] we found that for random selection of $l(k)$ error $\varepsilon(k)$ converges exponentially on average [31], and we have analytically shown that this form of convergence is task independent [32].

3. Experimental setup

The experimental setup is schematically illustrated in figure 1(b). The output of our fiber-coupled ECL is collimated via a 7.5 mm focal distance lens (Thorlabs AC050-008-B-ML). The surface of DMDa is imaged via two lenses (Thorlabs AC254-150-B-M, C110TMD-B) onto a 25 μm core diameter, one meter long mm-fiber (Thorlabs M67L01). DMDa displays input information $\mathbf{u}^{\text{in}}$ with a frame rate of ~ 130 frames per second, which currently sets the speed of our photonic ANN. Using a 25 mm focal length lens
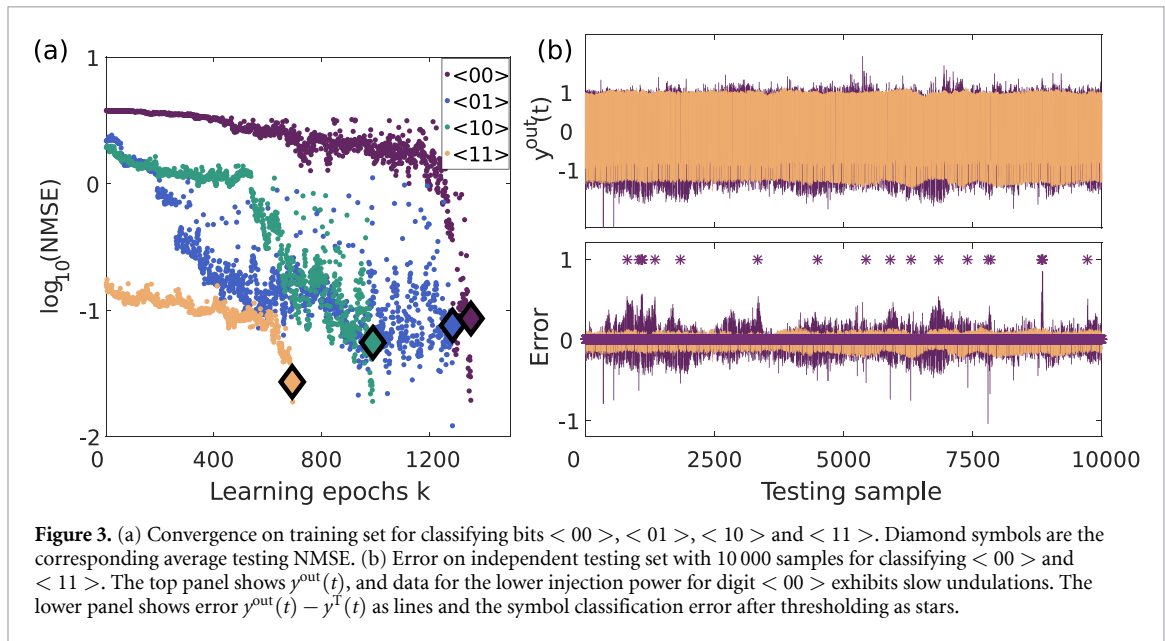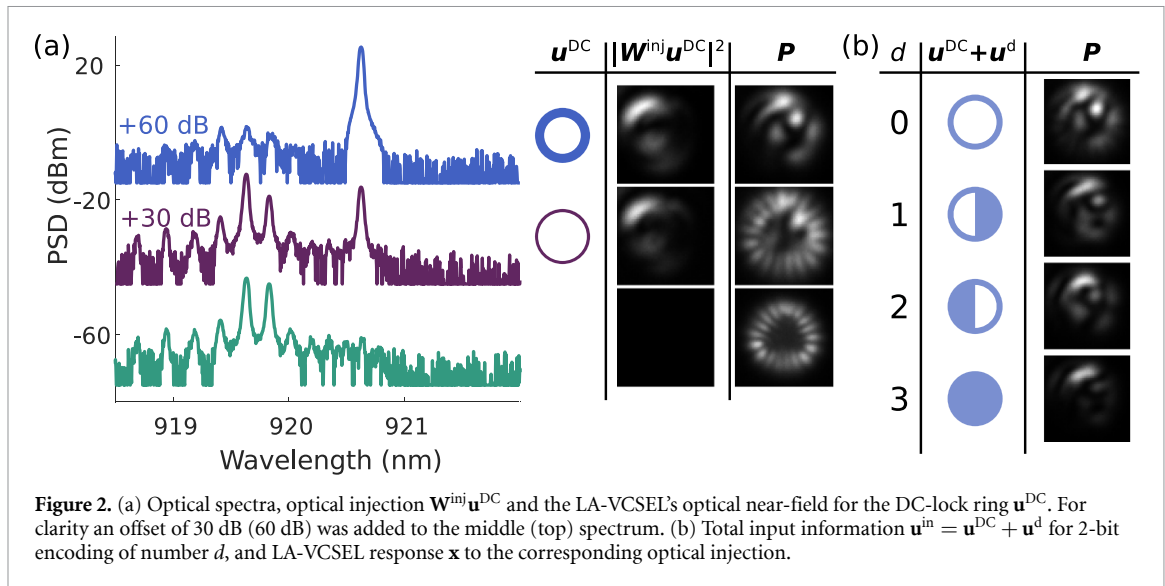
(AC127-025-B-ML) in combination with a MAG10 near infrared optimized microscope objective (MO, Olympus LMPLN10XIR), the mm-fiber's near-field at its output is imaged on top of our LA-VCSEL. The LA-VCSEL is imaged onto a CMOS camera (ids U3-3482LI-H) using a 100 mm focal distance lens (AC254-100-B-M). Within the same optical path we placed a non-polarizing 50/50 beam-splitter (Thorlabs CCM1-BS014/M), and the reflected signal is collected by a single mode fiber connected to an optical spectrum analyzer (OSA, Yokagawa AQ6370D). Additional non-polarizing 90/10 (Thorlabs BSX10R) and 50/50 (Thorlabs CCM1-BS014/M) beam-splitters direct optical injection and the LA-VCSEL's emission. Finally, we image the LA-VCSEL onto DMDb with a 100 mm focal length lens (Thorlabs AC254-100-B-M), resulting in a magnification of 5.6, and an additional lens (Thorlabs AC254-045-B-M) focuses the reflected signal on a detector (Thorlabs PM100A, S150C).

The LA-VCSEL design and fabrication are optimized to improve operation efficiency and bandwidth. Our device wafers are produced by metal-organic vapor-phase epitaxy (MOVPE) on heavily n-doped (001) surface oriented GaAs substrates. The epitaxial VCSEL structure consists of an n-doped GaAs buffer layer, followed by a Si-doped (n-doped) 37-period hybrid (two section) bottom distributed Bragg reflector(DBR), a half-wavelength ($\lambda/2$ optically thick) cavity, a C-doped (p-doped) 14.5-period top DBR, and heavily p-doped $Al_{0.1}Ga_{0.9}As$ and GaAs current spreading layers each $\lambda/2$ optically thick. The bottom n-doped DBR is composed of 33 periods of $Al_xGa_{1-x}As$ with $x = 0.05$ high refractive index layers and $x = 1.0$ low refractive index layers, followed by four periods of $x = 0.05$ and $x = 0.92$ layers. The top p-doped DBR is similar but with alternating $x = 0.1$ and $x = 0.92$ $Al_xGa_{1-x}As$ layers. All DBRs include linearly graded interfaces between the alternating low and high refractive index layers to reduce series resistance. The optical cavity consists of five compressively strained ∼4.2 nm thick InGaAs quantum wells (QWs) surrounded by six strain-compensating ∼5.1 nm thick GaAsP barrier layers that are in tension. From double crystal x-ray diffraction measurements on QW calibration structures the net active region strain is close to zero (as designed). The ensemble of QWs and barriers is centered in the $\lambda/2$ optical cavity and the remaining layers of the optical cavity (on both sides of the QWs and barrier layers) are $Al_xGa_{1-x}As$ layers stepped from $x = 0.4$ to $x = 0.92$, surrounded by 20 nm thick (before oxidation) $Al_{0.98}Ga_{0.02}As$ layers and followed by linearly graded $Al_xGa_{1-x}As$ layers decreasing in mole fraction (x) from $x = 0.92$ toward $x = 0.1$ in the top p-doped DBR and from $x = 0.92$ toward $x = 0.05$ in the bottom n-doped DBR. The two $x = 0.98$ layers reside inside the optical cavity. The two $x = 0.98$ layers are selectively thermally oxidized during device processing to form wave-guiding and current confining oxide apertures with tapered ends (that minimize scattering losses). The estimated oxidation length (from the top mesa edge inward) is roughly 9 μm, based on a series of selective thermal oxidation experiments on cleaved slivers of the VCSEL material—yielding a plot of the oxidation length (in from a mesa edge) versus the oxidation time. On resonance (in our mode simulations) the optical field intensity nodes that define the $\lambda/2$ optical cavity fall outside the two $x = 0.98$ layers. We fabricate top-surface-emitting VCSELs onto quarter wafer pieces with high frequency co-planar ground-signal-ground (GSG) contact pads suitable for on-wafer testing via a standard GSG electrical probe head with pin to pin spacing of 150 μm using the fabrication methods reported in [20].

4. Optical injection and ANN states

Optical injection into multimode LA-VCSELs has been extensively studied [33, 34], yet here we injection-lock such a device to a complex optical input field for the first time. Figure 2(a) shows the device's optical response under DC injection using $\mathbf{u}^{DC}$ with different rings. The top, middle and bottom rows were recorded using a ring with 15, 5 and 0 DMD mirrors thickness, respectively. The bottom row therefore corresponds to the laser without optical injection. The table shows the configuration of $\mathbf{u}^{DC}$ (left column), the spatial distribution of the optical injection $|\mathbf{W}^{inj}\mathbf{u}^{DC}|^2$ (center column) and the injected LA-VCSEL's near-field $\mathbf{P}$ (right column).

We bias the laser with $I^{bias} = 9$ mA, which is 1.28 times above its lasing threshold, where it emits a total of $P = 1.68$ mW. The injection laser's wavelength was tuned to the VCSEL's highest susceptibility to optical injection locking at $\lambda^{inj} = 920.58$ nm. The free-running laser shows the expected multimode optical spectrum, and its near-field corresponds to high-order Laguerre–Gaussian modes [35]. A DC-lock ring width of 5 pixels results in injecting $P^{inj} = 0.05$ mW, for which the laser's emission spectrum as well as its spatial emission profile are already notably perturbed. Widening the DC-lock ring to 15 pixels increases the injection power to $P^{inj} = 0.7$ mW, for which the device is fully locked to the injected field and its near-field significantly differs from its free-running near-field and from the optically injected profile. This already demonstrates the non-trivial and non-uniform transformation of optically injected patterns. Figure 2(b) shows the resulting photonic ANN state for the example of injecting the four possible configurations $d \in [0, 1, 2, 3]$ of 2-bit symbols. Individual responses significantly differ for each case, which is a prerequisite to differentiate individual digits. Finally, we found the DC-lock ring of width 10 to be the smallest width

Figure 2. (a) Optical spectra, optical injection $\mathbf{W}^{\text{inj}}\mathbf{u}^{\text{DC}}$ and the LA-VCSEL's optical near-field for the DC-lock ring $\mathbf{u}^{\text{DC}}$. For clarity an offset of 30 dB (60 dB) was added to the middle (top) spectrum. (b) Total input information $\mathbf{u}^{\text{in}} = \mathbf{u}^{\text{DC}} + \mathbf{u}^{\text{d}}$ for 2-bit encoding of number d, and LA-VCSEL response $\mathbf{x}$ to the corresponding optical injection.

Figure 3. (a) Convergence on training set for classifying bits $< 00 >$, $< 01 >$, $< 10 >$ and $< 11 >$. Diamond symbols are the corresponding average testing NMSE. (b) Error on independent testing set with 10 000 samples for classifying $< 00 >$ and $< 11 >$. The top panel shows $y^{\text{out}}(t)$, and data for the lower injection power for digit $< 00 >$ exhibits slow undulations. The lower panel shows error $y^{\text{out}}(t) - y^{\text{T}}(t)$ as lines and the symbol classification error after thresholding as stars.

resulting in a stable LA-VCSEL near-field. We used this for all the following computing experiments, as the smallest width allows us to use the maximum input dimensionality for input information encoding.
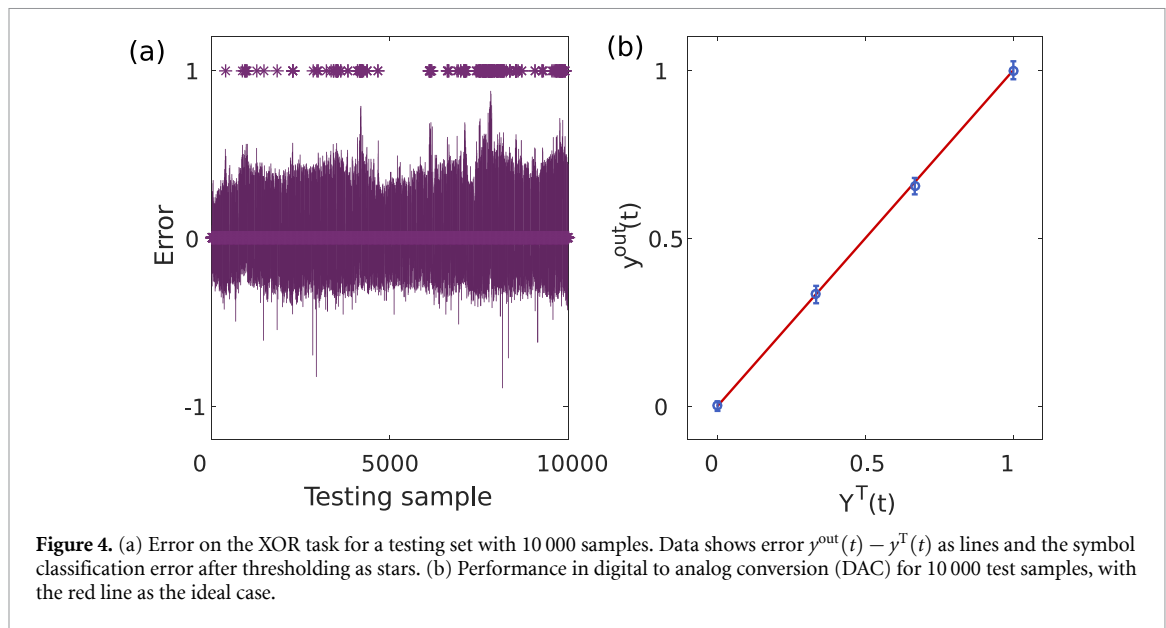
5. ANN performance of the LA-VCSEL

As starting point for our photonic ANN's performance evaluation we trained a single output classifier to identify one of the four digits in a 2-bit header. Training sequences of $t^{\text{T}} = 50$ samples were comprised to 50% of the digits to be classified, and to 50% of the remaining three digits. The system was trained until the training error ε_k dropped below $\epsilon < 2 \times 10^{-2}$, and in figure 3(a) we show the resulting convergence.

For all four tasks our system converged to below the targeted training error, however starting from strongly differing initial performances. As reason we assume the injection conditions consequence of the different $\mathbf{u}^{\text{in}}$ configurations, where P^{inj} is lowest for $d = 0$ and highest for $d = 3$. Additionally, for the case of classifying bits $< 00 >$ the LA-VCSEL ANN needs to invert the power-dependency: the digit with the lowest injection power needs to be transformed into the highest power in the ANN's output. This observation is consistent with the initial classification errors for bits $< 01 >$ and $< 10 >$ starting from comparable levels, and with digit $< 11 >$ starting from the lowest level.

After successful training, the system was evaluated with an independent test-sequence comprising 1000 samples repeated ten times. Average testing errors correspond to the filled diamonds in figure 3(a). Again, they reflect the stability of the different locking conditions. Strongest injection (bits $< 11 >$) produces a

Figure 4. (a) Error on the XOR task for a testing set with 10 000 samples. Data shows error $y^{out}(t) - y^T(t)$ as lines and the symbol classification error after thresholding as stars. (b) Performance in digital to analog conversion (DAC) for 10 000 test samples, with the red line as the ideal case.

testing error nearly identical to the training error, and lower injection powers result in stronger training vs. testing error divergence. Figure 3(b) shows the temporally resolved testing sequences for bits $< 00 >$ and $< 11 >$. Output $y^{out}(t)$ in the upper panel clearly shows stronger undulations for the case of classifying $< 00 >$. The lower panel shows error $y^{out}(t) - y^T(t)$ as lines, while the stars show a classification performance after threshold, i.e. $|(y^{out}(t) > 0) - y^T(t)|$. Therefore, stars in figure 3(b) located at amplitude 1 are samples for which our photonic ANN misclassified the corresponding digit. The number of misclassified samples normalized by the test sequence's length results in the symbol error rate (SER), and we obtained SER = $(2.5 \times 10^{-3}, < 1 \times 10^{-4}, 1.1 \times 10^{-3}, < 1 \times 10^{-4})$ for bits $< 00 >, < 01 >, < 10 >, < 11 >$, respectively. On average we therefore obtain SER$< 0.9 \times 10^{-3}$.

As a final test on the 2-bit task we train the LA-VCSEL ANN to (a) implement an XOR logic operation; and (b) to implement a DAC. The XOR task is a classical benchmark to demonstrate that an ANN is capable to implement nonlinear transformations. As such it ultimately motivated the development of deep ANNs since single layer perceptrons [36] failed in this tasks [37]. DAC places a similar, yet slightly easier challenge to an ANN. Figure 4(a) shows the error of the XOR task, illustrating $y^{out}(t) - y^T(t)$ as a line and as previously the misclassifications after thresholding as stars. The obtained SER for the XOR task is 2.9×10^{-2}. Finally, the performance in DAC is shown in figure 4(b), with the red line as the ideal case. We find that the system excellently approximates the output target with a standard deviation of 5.4×10^{-2}.

6. Conclusions

We demonstrated a spatially-extended ANN based on an injection locked LA-VCSEL with $\sim$90 artificial photonic neurons. All ANN connections are implemented in photonic hardware, and the network's operation is fully parallel. Trainable readout connections are realized using a DMD, while static input connections are obtained via the complex transfer matrix of a mm-fiber. Finally, the information input was provided via another DMD.

We trained and tested our photonic ANN's performance against tasks involving sequences of 2-bits, i.e. pattern classification, the XOR task and DAC. Our LA-VCSEL based ANN shows excellent performance in all, achieving long-term stability of the trained states during the much longer testing sequences. Finally, our demonstration is only a proof-of-concept. The ANN's present speed is orders of magnitude below its potential bandwidth, and the speed of convergence can most likely be significantly improved with more advanced training concepts.

Our photonic ANN's future potential relies on the fact that it is capable of fully parallel computing in the multi-GHz range, a potential speed-up by around seven orders of magnitude that would enable direct applications in ultra-fast systems [38]. DMD's can be operated $> 10^3$ frames s^{-1} in realtime, making information injection the current speed bottleneck. In the future we will address this limitation by exploring alternative systems with same functionality, like e.g. ferroelectric spatial light modulators that readily enable $> 10^6$ frames s^{-1} or spatial light modulators based on arrays of vertical microcavities promising phase modulations at GHz bandwidths [39]. Crucially, such improvement is compatible with scaling up the

number output channels through creation of multiple, spatially non-overlapping copies of the LA-VCSEL's near-field. Noteworthy, the size of input, photonic ANN or output does not reduce the system's speed due to complete photonic parallelism. All weights implemented in our system are passive, making the system continuously operate at the DC power-levels of its components. $> 10^6$ inferences per second with 100 TBit/s input data rate at 1…10 W power consumption are therefore realistic, which is orders of magnitude beyond any electronic special purpose ANN computing hardware [40].

Data availability statement

The data generated and/or analysed during the current study are not publicly available for legal/ethical reasons but are available from the corresponding author on reasonable request.

Acknowledgments

ORCID iDs

Xavier Porte ⓘ https://orcid.org/0000-0002-9869-7170
Stephan Reitzenstein ⓘ https://orcid.org/0000-0002-1381-9838
James A Lott ⓘ https://orcid.org/0000-0003-4094-499X
Daniel Brunner ⓘ https://orcid.org/0000-0002-4003-3056

References

[1] McCulloch W S and Pitts W 1943 A logical calculus of the ideas immanent in nervous activity *Bull. Math. Biophys.* **5** 115–33
[2] Rumelhart D E, Hinton G E and Williams R J 1986 Learning representations by back-propagating errors *Nature* **323** 533–6
[3] Raina R, Madhavan A and Ng A Y 2009 Large-scale deep unsupervised learning using graphics processors *Proc. 26th Annual Int. Conf. on Machine Learning - ICML '09* (New York: ACM Press) pp 1–8
[4] LeCun Y, Bengio Y and Hinton G 2015 Deep learning *Nature* **521** 436–44
[5] Senior A W *et al* 2020 Improved protein structure prediction using potentials from deep learning *Nature* **577** 706–10
[6] Lin P *et al* 2020 Three-dimensional memristor circuits as complex neural networks *Nat. Electron.* **3** 225–32
[7] Farhat N H, Psaltis D, Prata A and Paek E 1985 Optical implementation of the Hopfield model *Appl. Opt.* **24** 1469
[8] Lohmann A W 1990 Principles of optical computing *Nonlinear Optics and Optical Computing* (Boston, MA: Springer US) pp 151–7
[9] Miller D A 2017 Attojoule optoelectronics for low-energy information processing and communications *J. Lightwave Technol.* **35** 346–96
[10] Moughames J, Porte X, Thiel M, Ulliac G, Jacquot M, Larger L, Kadic M and Brunner D 2020 Three dimensional waveguide-interconnects for scalable integration of photonic neural networks *Optica* **7** 640–6
[11] Dinc N U, Psaltis D and Brunner D 2020 Optical neural networks: the 3D connection *Photoniques* **114** 34–8
[12] Li H Y S, Qiao Y and Psaltis D 1993 Optical network for real-time face recognition *Appl. Opt.* **32** 5026
[13] Rafayelyan M, Dong J, Tan Y, Krzakala F and Gigan S 2020 Large-scale optical reservoir computing for spatiotemporal chaotic systems prediction *Phys. Rev.* X **10** 041037
[14] Tait A N, Member S, Nahmias M A, Shastri B J and Prucnal P R 2014 Broadcast and weight: an integrated network for scalable photonic spike processing *J. Lightwave Technol.* **32** 3427–39
[15] Shen Y *et al* 2017 Deep learning with coherent nanophotonic circuits *Nat. Photon.* **11** 441–6
[16] Feldmann J, Youngblood N, Wright C D, Bhaskaran H and Pernice W H P 2019 All-optical spiking neurosynaptic networks with self-learning capabilities *Nature* **569** 208–14
[17] Jaeger H and Haas H 2004 Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication *Science* **304** 78–80
[18] Van Der Sande G, Brunner D and Soriano M C 2017 Advances in photonic reservoir computing *Nanophotonics* **6** 561–76
[19] Haghighi N 2021 Vertical cavity surface emitting lasers (VCSELs) and VCSEL arrays for communication and sensing *Dissertation* Technical University Berlin, Germany accepted
[20] Haghighi N, Moser P and Lott J A 2020 40 Gbps with electrically parallel triple and septuple 980 nm VCSEL arrays *J. Lightwave Technol.* **38** 3387–94
[21] Haghighi N, Moser P and Lott J A 2019 Power, bandwidth and efficiency of single VCSELs and small VCSEL arrays *IEEE J. Sel. Top. Quantum Electron.* **25** 1–15
[22] Carpenter J, Eggleton B J and Schröder J 2016 Complete spatiotemporal characterization and optical transfer matrix inversion of a 420 mode fiber *Opt. Lett.* **41** 5580

[23] Mulet J and Balle S 2002 Spatio-temporal modeling of the optical properties of VCSELs in the presence of polarization effects *IEEE J. Quantum Electron.* **38** 291–305

[24] Bueno J, Brunner D, Soriano M and Fischer I 2017 Conditions for reservoir computing performance using semiconductor lasers with delayed optical feedback *Opt. Express* **25** 2401–12

[25] Simpson T B, Liu J M and Gavrielides A 1995 Bandwidth enhancement and broadband noise reduction in injection-locked semiconductor lasers *IEEE Photonics Technol. Lett.* **7** 709–11

[26] Brunner D, Soriano M C, Mirasso C R and Fischer I 2013 Parallel photonic information processing at gigabyte per second data rates using transient states *Nat. Commun.* **4** 1364

[27] Vatin J, Rontani D and Sciamanna M 2019 Reservoir computing using VCSEL polarization dynamics *Opt. Express* **27** 18579–84

[28] Estébanez I, Schwind J, Fischer I and Argyris A 2020 Accelerating photonic computing by bandwidth enhancement of a time-delay reservoir *Nanophotonics* **9** 4163–71

[29] Chen R T Q, Rubanova Y, Bettencourt J and Duvenaud D 2018 Neural ordinary differential equations *32nd Conf. on Neural Information Processing Systems (NeurIPS 2018) (Montréal, Canada)* vol 1 pp 191–235

[30] Bueno J, Maktoobi S, Froehly L, Fischer I, Jacquot M, Larger L and Brunner D 2018 Reinforcement learning in a large scale photonic recurrent neural network *Optica* **5** 756–60

[31] Andreoli L, Porte X, Chrétien S, Jacquot M, Larger L and Brunner D 2020 Boolean learning under noise-perturbations in hardware neural networks *Nanophotonics* **9** 4139–47

[32] Chrétien S, Andreoli L, Porte X and Brunner D 2021 Convergence and scaling of Boolean-weight optimization for hardware reservoirs (in preparation) **1** 1–7

[33] Ackemann T, Barland S, Cara M, Giudici M and Balle S 1999 Spatial structures and their control in injection locked broad-area VCSELs *Nonlinear Guided Waves and Their Applications* (Washington, DC: OSA) p WC4

[34] Ackemann T, Barland S, Giudici M, Tredicce J, Balle S, Jaeger R, Grabherr M, Miller M and Ebeling K 2000 Patterns in broad-area microcavities *Phys. Status Solidi* b **221** 133–6

[35] Degen C, Elsaber W and Fischer I 1999 Transverse modes in oxide confined VCSELs: influence of pump profile, spatial hole burning and thermal effects *Opt. Express* **5** 38–47

[36] Rosenblatt F 1958 The perceptron: a probabilistic model for information storage and organization in the brain *Psychol. Rev.* **65** 386

[37] Minsky M and Papert S 1969 *Perceptrons: An Introduction to Computational Geometry*

[38] Genty G, Salmela L, Dudley J M, Brunner D, Kokhanovskiy A, Kobtsev S and Turitsyn S K 2020 Machine learning and applications in ultrafast photonics *Nat. Photon.* **15** 91–101

[39] Peng C, Hamerly R, Soltani M and Englund D R 2019 Design of high-speed phase-only spatial light modulators with two-dimensional tunable microcavity arrays *Opt. Express* **27** 30669

[40] Reuther A, Michaleas P, Jones M, Gadepally V, Samsi S and Kepner J 2020 Survey of machine learning accelerators (arXiv:2009.00993) pp 1–11