POST-DIGITAL - European Training Network on Post-Digital Computing [GA860360]

Document Details

Title	D2.1 Report on the design of photonic- and electronic- based analogue computing
Deliverable number	D2.1
Deliverable Type	Report (public)
Deliverable title	Report on the design of photonic- and electronic- based analogue computing
Work Package	WP2 – Implementation and characterization
Description	Document describing the different approaches to implement analogue computing in the POST-DIGITAL European Training Network and the level of progress of the ESRs
Deliverable due date	30/09/2021
Actual date of submission	29/09/2021
Lead beneficiary	CSIC
Version number	V1.0
Status	Final

Dissemination level

PU	Public	X
CO	Confidential, only for members of the consortium (including Commission Services	

Project Details

Grant Agreement	860360
Project Acronym	POST-DIGITAL
Project Title	POST-DIGITAL - European Training Network on Post-Digital Computing
Call Identifier	H2020-MSCA-ITN-2019
Project Website	https://postdigital.astonphotonics.uk/
Start of the Project	1 April 2020
Project Duration	48 months

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860360.

Executive Summary

This deliverable is the first of three deliverables in the POST-DIGITAL project describing the progress in different approaches to build photonic- and electronic-based analogue computing devices. The present deliverable reports the main concepts as well as the way to implement these concepts in hardware designs. Using photonic substrates, designs based in spatial, temporal or spatial multiplexing are shown. An alternative approach to use an electronic re-configurable neuromorphic chip that uses analogue processing with spiking neurons is also considered. Along this deliverable, the level of progress of the ESRs is documented, and further steps towards reaching the final designs are discussed.

TABLE OF CONTENTS

List of Figures

List of Acronyms

ANN Artificial Neural Network
DMD Digital Micro-mirror Device
ESN Echo State Network
ESR Early Stage Researcher
ETN European Training Network
FPGA Field Programmable Gate Array
MINDS Modeling Intelligent Dynamical Systems
OFDM Orthogonal Frequency Division Multiplexing
OPU Optical Processing Unit
PAM Pulse-Amplitude Modulation
POST-DIGITAL Project 'POST-DIGITAL - European Training Network on Post-Digital Computing' EC GA 860360
QAM Quadrature Amplitude Modulation
RC Reservoir Computing
RUG Rijksuniversiteit Groningen (University of Groningen)
VCSEL Vertical-Cavity Surface-Emitting Laser

1 INTRODUCTION

This report covers the progress in the design of hardware implementations of analogue computing devices based on photonic and electronic platforms. The ESRs of the POST-DIGITAL ETN are exploring several hardware platforms and information processing strategies in the context of analogue computing. In the photonic domain, this report covers the advances in the exploitation of the physical substrate by multiplexing techniques, either in space (ESR3, ESR6, ESR7, and ESR8), time (ESR4, ESR5, ESR13, and ESR14) or frequency (ESR12). In the electronic domain, ESR1 and ESR2 are evaluating the properties of electronic mixed-signal neuromorphic hardware.

Regarding potential applications, ESR13 and ESR14 are, e.g., working on the design and development of photonic reservoir computing for ultra-fast information processing using semiconductor lasers.

ESRs are therefore obtaining a wide skill set, including expertise on analogue computational concepts, system design, and hardware implementations. This report shows that the ESRs are in the position to make a measurable impact in the field of post-digital computing. After the initial designs, in-depth experimental characterization will follow with the ultimate goal of high-performance applications.

List of contributors (alphabetical order): Steven Abreu (ESR1, Groningen), Diego Argüello (ESR7, Aston), Ivan Boikov (ESR6, Thales), Mirko Goldmann (ESR5, CSIC), Daniel Hesslow (ESR8, LightOn), Alessandro Lupo (ESR12, ULB), Long Nguyen (ESR13, Aston), Enrico Picco (ESR4, ULB), Guillaume Pourcel (ESR2, Groningen), Anas Skalli (ESR3, Femto-ST), Lucas Talandier (ESR14, CSIC). Document compiled by Miguel C. Soriano, IFISC (CSIC-UIB).

2 Design of Photonic- and Electronic- based Analogue Computing Devices

There is a rapid growth in the use of artificial neural networks (ANN) to process data and solve problems in areas where traditional algorithms fail to surpass the capabilities of the human brain. The drawbacks of using conventional digital hardware when running ANN methods has led to physically implement them in analogue systems [1]. In the following, the activities of the different ESRs involved in the development and design of photonic (section 2.1) and electronic (section 2.2) based analogue computing devices are detailed. A list of references is provided in Section 2.3.

2.1 Photonic Designs

The different photonic designs follow complementary strategies that exploit multiplexing in either space (section 2.1.1), time (section 2.1.2) or frequency (section 2.1.3).

In the case of spatial multiplexing, ESR7 and ESR3 are evaluating the use of multi-mode optical systems as artificial neural networks implementations. ESR8 is focusing on a design that exploits optical scattering for random projections and ESR6 is working towards the design of photonic crystal resonators for reservoir computing.

In the case of temporal multiplexing, ESR13 and ESR14 are closely working on the development of delay-based photonic reservoir computers for telecommunication fiber-optic signal recovery, while ESR5 and ESR4 are advancing the development of opto-electronic reservoir computers based on a delay loop.

Finally, ESR12 is exploring the possibility of using frequency multiplexing for reservoir computing.

2.1.1 Space multiplexing

In this section, the progress in the design of spatially-multiplexed optical systems for artificial neural networks and reservoir computing systems is presented. This covers the work of ESR3, ESR6, ESR7, and ESR8.

2.1.1.1 Multimode optical devices for artificial neural networks and reservoir computing

ESR7, who is located at Aston Institute of Photonics Technology, aims at developing novel implementations of ANNs using multi-core optical fibers (MCF) and other space division multiplexing technologies. In this approach, neurons and synapses can be realized as individual silica cores in the fiber. Moreover, the coupling effect of the outer cores to inner cores is going to be used as the hidden layers of an equivalent neural network. The nonlinear effects taking place in the fiber will work as activation functions. Although this approach has already been demonstrated [2], the number of hidden layers was constrained by the fiber design and the limited amount of input cores reduced the ability to perform real-world tasks. Therefore, at Aston novel designs of MCF are being investigated to make it possible to implement a wider range of ANNs architectures that can be used to solve relevant problems.

An additional objective of ESR7 is the development of novel optical implementations of reservoir computing. Like in the previous case of ANN implementations, the research that has already been carried out in the field of fiber optics technology will be applied to exploit fiber nonlinearities that allow to create systems with intrinsic optical memory. The final objective is using these new types of neural networks to create optical high-capacity communication systems.

ESR3 is working on the hardware implementation of a photonic neural network using spatially distributed modes of a large-area vertical cavity surface-emitting laser (LA-VCSEL). All neural network connections in this design are realized in hardware, and thus can produce results in real-time without pre- or post-processing.

The design of the experimental setup is shown in Figure 1. Information encoded in spatial patterns is injected into a VCSEL through the complex transmission matrix of a multi-mode fibre (input weights). The VCSEL then transforms the injected information non-linearly with a multi-mode profile. This transformation is comparable to the action of a neural network: similar inputs result in vastly different VCSEL responses that can be used for pattern recognition. The transformed mode profile is imaged onto a digital micro-mirror array (DMD), whose mirrors constitute programmable readout weights, connecting the VCSEL to the detector, the output node of this ANN implementation.

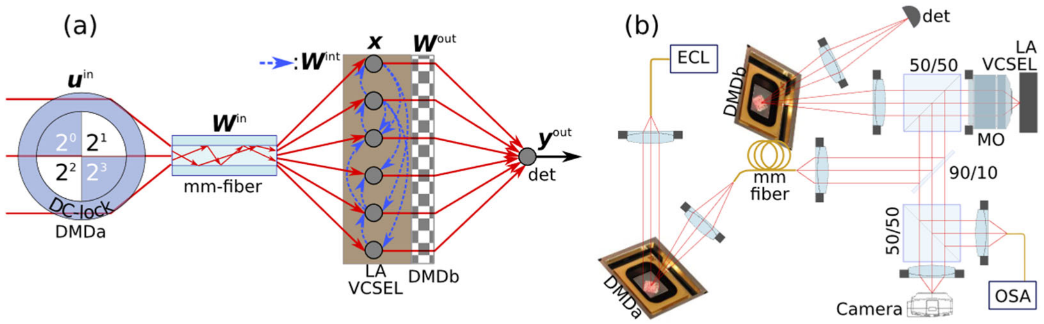

Figure 1 a) Schematic illustration of the different components of the ANN. b) Experimental setup of the VCSEL-based ANN.

The work of ESR3 demonstrates a fully parallel ANN consisting of 100+ neurons realised fully in hardware, based on an injection locked multimode, large area, VCSEL. This ANN can perform header recognition tasks as well as 2-bit XOR and digital to analog conversion tasks. In order to reach the full potential of this approach in terms of bandwidth, the current limit is the framerate of the injection DMD since the VCSEL itself can be modulated at bandwidths in the GHz range.

2.1.1.2 Optical Scattering for Linear Random Projections

ESR8 is working on optical methods for linear random projections with applications to randomized linear algebra at LigthOn. The optical processor unit (OPU) developed at LigthOn can perform random projections at high speed and is now commercially available [3]. This OPU is based on multiple optical scattering and measures the intensity of the output of the random projection but does not use any holographical techniques to recover the full complex field.

Since measuring the intensity is a nonlinear operation, applications of the simple OPU design could be limited. For instance, randomized linear algebra and direct feedback alignment (DFA) methods both require the projections to be linear. While holographical hardware can be added, and indeed has been added in the past, this complicates the construction of the OPU [4]. Instead, ESR8 has developed techniques to make use of the existing OPU for linear random projections without modifying the hardware. The results of this line of work have been presented in the HOT CHIPS 33 symposium that took place on the 23-24 of August 2021. Further, these algorithms are already available to customers of the LigthOn Cloud [5].

2.1.1.3 Photonic crystal resonators for reservoir computing

A photonic crystal is a low-loss semiconductor media with a periodic refractive index. By locally departing from a strict periodicity, it is possible to control confinement of photons inside it. In this manner, a circuit of photonic resonators and waveguides can be created. Unlike microring resonators, there is no degeneracy of resonator modes in photonic crystals, which allows for a better concentration of power in a resonator, which is important for a nonlinearity to make an impact. Moreover, lifetime of charge carriers (electrons and holes) could be considerably shorter in photonic crystals due to a high surface area, which ensures a quick reaction of nonlinearity to input, as it is dependent on concentration of charge carriers. Experiments show that a longer reaction time in the nonlinearity often results in systems being more prone to chaotic behavior, which harms computational performance. The Thales group has an extensive expertise on photonic crystals.

ESR6 is currently exploring the possibility of using photonic crystal resonators as reservoir computing media in a space-multiplexed topology. Preliminary simulations show visible performance benefits of using the intrinsic nonlinearity of photonic crystal resonators compared to the photodiode nonlinearity in [6] and currently a more detailed simulation of this setup is in progress. The questions that need to be answered are mostly related to scalability, the biggest concern being a power flow. Because a signal is injected only into a small fraction of resonators, some of them may not get enough power to perform nonlinear computation, or even receive no power at all. Depending on the simulation results, device fabrication will follow.

2.1.2 Time multiplexing

In this section, the advances on the design of delay-based photonic and opto-electronic reservoir computers are detailed. This entails the work of ESR4, ESR5, ESR13, and ESR14.

2.1.2.1 Delay-based reservoir computing and time multiplexing for fiber-optic channel equalisation

ESR13 is working on neuromorphic approaches for channel equalisation in advanced fibre-optic systems to obtain high spectral efficiency. His research will focus on the reservoir computing-based solutions for advanced modulation format (m-QAM), OFDM, multicarrier signals. To this end, ESR13 is designing and testing an experimental setup of an all-optical RC system. A time delay RC-based system is taken into consideration but more work needs to be performed to couple more reservoirs in the deep ESN concept. The system also requires dual signal processing abilities, and an approach of fibre echo state network analogue (FESNA) is considered taking into account that a reservoir temperature stabilization measure should be in place.

The RC framework and the photonic RC-based hardware systems have been demonstrated to efficiently compensate for deterministic signal impairments of PAM signals. However, the research on RC systems with similar equalization performance for m-QAM signals is still an unexplored area. Since m-QAM signals are susceptible to the fibre dispersion, the memory capacity required by the RC-based nonlinear equalizer also increases in this case and is proportional to the transmission distance.

The first task of ESR13 has been to evaluate the RC-based system in numerical simulations to set the theoretical performance limit compared to the least mean square (LMS) algorithm before implementing it in a hardware platform. The numerical modelling of the echo state network (ESN) on the m-QAM signal nonlinear equalization task shows a limited peak-to-peak value improvement of Q-factor. The proposed solution is to decrease the severe impact of the dispersion effect by using in-line dispersion compensating fibre (DCF) or come up with the concept of a deep ESN with higher memory capabilities.

The work of ESR14 is also focused on the photonic time delay reservoir computing for telecommunication signal recovery. Typically, the system is composed of three distinct parts. A response semiconductor laser is the main part of the reservoir, bringing nonlinearities and memory thanks to the feedback loop and the optical injection of external input. Secondly, the input part where input weights and sample and hold are applied to the signal, this pre-processing step is usually referred to as masking. It is done electronically such that electric to light conversion must be done at the output of the fiber transmission. In the last part, the output of the reservoir, a light to electric conversion is done enabling the linear classification in the electrical domain.

The activities of ESR14 in this first year have been focused on the design and numerical investigation of fiber-based time delay reservoir computing for post-processing signals coming from telecom transmission channels. He is currently exploring the possibilities to avoid the optoelectronic conversion at the input of the reservoir and feed the reservoir directly with the optical signal from the transmission channel, without any detection stage in between. This creates the necessity to deal with both the amplitude and phase properties of the transmission signal. This investigation was done thanks to simulations of the feedback laser with injection based on a modified version of the Lang-Kobayashi equations. The target is to process amplitude modulated signals (PAM and PAM4 encoding) that reach rates from 28 up to 46Gbauds and transmission lengths to at least 100km. Indeed, at the input of the fiber, the information is embedded only in the amplitude, the information carried by phase at the end of the signal is brought by optical effects, along with the propagation in the fiber. It has been found that a lower bit error rate (BER) can be achieved at the data recovery task by using the information carried by the phase.

Finally, ESR14 is working on building the transmission system in the laboratory, the first step being the implementation of an optical transmission line, from emitter to receptor to then connect it to the reservoir already mounted in the lab.

2.1.2.2 Opto-electronic reservoir computing based on a delay loop

The setup proposed by ESR4 is an implementation of an opto-electronic reservoir computer using an FPGA chip and an optical delay dynamical system, showed in Figure 2. The electronic part of the system is the FPGA itself, while the optical part is made by an incoherent light source, a delay loop (fiber spool) and a nonlinear node (Mach-Zehnder Intensity Modulator). The FPGA generates the input sequence in real time, collects the reservoir states and computes optimal readout weights using different training algorithms.

The system shown in Figure 2 can work with two different output configurations. The first is by means of a digital readout, i.e., averaging and collecting the values of the nodes of the reservoir that are then used to compute the weights; it has been already tested experimentally, for example on a Real-Time Channel Equalization task [7]. The second one is an analog readout layer, where the aim is to compute the sum of the weighted states analogically in order to obtain the output of the reservoir without further computation. This would allow to get rid of the slow offline data post-processing, and it would open the possibility of feeding the output signal back into the reservoir, thus significantly enriching its dynamics and making it capable of solving signal generation tasks. It has already been investigated numerically using a capacitor as readout layer [8], and it is now under experimental investigation using an op-amp based circuit which implements the integrating function.

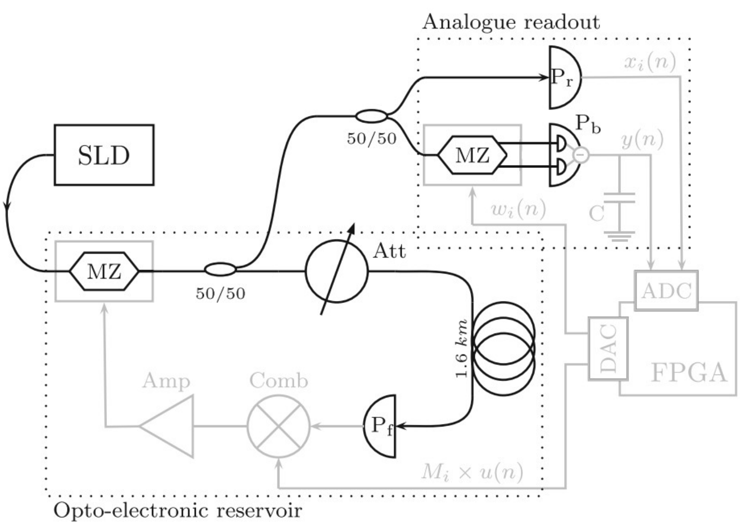

Figure 2 Scheme of the experimental setup for a delay-based opto-electronic reservoir computer, with both digital (Pr) and analog (Pb) readout configurations.

In turn, ESR5 is working on a concept for variable deep reservoir computing based on several interacting optoelectronic delay systems. This concept fuses the advantages of deep learning and reservoir computing to yield a variable, high-performing, and hardware implementable processor for machine learning. In practice, such a reservoir contains several unidirectional coupled delay systems that carry out transient information processing. Using a time-multiplexed input signal, each of the nonlinear nodes emulates a high-dimensional recurrent neural network. This concept already showed promising performances in chaotic time series prediction [9,10]. Nevertheless, there is no algorithm to optimize the hyperparameter of the single layers. By introducing a deep reservoir based on a modular design, the work of ESR5 can provide a simple optimization method. Here, several delay systems with different hyperparameters represent the modules. Related to their different hyperparameters setting, each layer will carry out different computations and provide different memory.

The recent work of ESR5 shows that, depending on the arrangement of the layers in the deep architecture, the interplay of the layers leads to variable computational abilities. The connections between the layers can be changed using simple switches. Accordingly, the flow of information and the interplay of the different layers can be optimized only by simple switches. Optimizing the order of the modules can improve the performance without doing hyperparameter scans and changing the layers themselves. The results for a two-layer system shows that even the arrangement of two similar nodes with different delays enables the reservoir optimization towards different tasks. In the next step, a larger ensemble of layers will be characterized to enhance the variability provided by the modules. So far, only the signal entering the first layer observes an external modulation during the time-multiplexing. Nevertheless, all the connections between the single layers could be modulated. In future investigations, the potential of such modulations and algorithms that optimize the amplitudes of the modulation will be analyzed.

2.1.3 Frequency multiplexing

ESR12 is working on frequency multiplexing schemes for Reservoir Computing. The current experimental setup encodes each network neuron in the complex amplitude of a line of a frequency comb. Frequency lines are made to interfere through phase modulation, which mixes neuron information. A fiber loop implements a recurrence path in the system, since each new state depends both on the current input and on the previous network state. The output variables of the reservoir are the optical powers of each comb line, which constitute a quadratic readout nonlinearity. This scheme has already been demonstrated to work [11]. Recent improvements in stabilization mechanism allowed the system to run in “optical weighting” mode, where multiplication by output weights is executed optically (setting the proper optical attenuation on each comb line) and the subtraction between positive and negative weighted neurons is still executed electronically (results to be published). Alternative schemes that would allow to perform the subtraction optically (through interference) are under study.

A variation of this setup, where the fiber loop is removed, has been studied. This results in a network with no recurrence (hence no memory), which constitutes an Extreme Learning Machine. This scheme has been demonstrated to work also in “optical weighting” mode, even if electronics is still necessary to perform subtractions between positive and negative quantities [12]. Improvements of this scheme, where multiple frequency combs propagate on the same setup and thus improving the dimension of the network, are currently under study.

2.2 Electronic Designs

The work of ESR1 and ESR2 revolves around the use of Electronic Mixed-Signal Neuromorphic Hardware, namely the DYNAPSE2 processor. The Dynamic Neuromorphic Asynchronous Processor (DYNAP) family of neuromorphic processors has been developed at the Institute of Neuroinformatics in Zurich and at the spin-off company SynSense as re-configurable neuromorphic chips that use analog processing with spiking neurons and digital event-based communication [13,14]. Previous members of the MINDS research group at RUG have already worked with the first generation of this chip [15], and ESR1 and ESR2 are currently starting their experimental work with the second generation of the chip, the DynapSE2 [14].

The DynapSE2 is implemented using existing (electronic) CMOS technology, operated in the sub-threshold regime as originally proposed by Carver Mead in the late 1980s as "neuromorphic engineering" [16,17]. This approach is particularly promising for edge computing, where low energy consumption is desirable, and sometimes even necessary. Furthermore, it allows to leverage existing VLSI technology, which has already been shown to scale efficiently.

A further difference that sets the DynapSE2 apart from other approaches to neuromorphic computing is that it uses spiking neuron models that are more biologically accurate than the rate-based neuron models that are familiar from current deep learning research. This allows not only for (massively) parallel computation but also for sparse, event-driven computation that helps to further reduce the energy consumption when compared against dense rate-based neural networks [18].

Another way of comparing the DynapSE2 approach to neuromorphic computing against other approaches is that the DynapSE2 is not primarily used as a hardware accelerator for existing algorithms (e.g., deep neural networks) but it also serves as a platform for experimentation with novel computing paradigms, i.e., event-driven spiking neural networks [19]. As the MINDS research group at RUG is primarily interested in the development of concepts and methods for unconventional computing, the DynapSE2 chip provides them with a suitable experimental platform to develop their theories.

2.3 References

[1] Marković, D., Mizrahi, A., Querlioz, D., & Grollier, J. (2020). Physics for neuromorphic computing. Nature Reviews Physics, 2(9), 499-510.

[2] Cohen, E., Malka, D., Shemer, A., Shahmoon, A., Zalevsky, Z., & London, M. (2016). Neural networks within multi-core optic fibers. Scientific reports, 6(1), 29080.

[3] Brossollet, C., Cappelli, A., Carron, I., Chaintoutis, C., Chatelain, A., Daudet, L., ... & Tommasone, G. L. (2021). LightOn Optical Processing Unit: Scaling-up AI and HPC with a Non von Neumann co-processor. *arXiv preprint arXiv:2107.11814*.

[4] Launay, J., Poli, I., Müller, K., Carron, I., Daudet, L., Krzakala, F., & Gigan, S. (2020). Light-in-the-loop: using a photonics co-processor for scalable training of neural networks. *arXiv preprint arXiv:2006.01475*.

[5] Hesslow, D., Cappelli, A., Carron, I., Daudet, L., Lafargue, R., Müller, K., ... & Poli, I. (2021). Photonic co-processors in HPC: using LightOn OPUs for Randomized Numerical Linear Algebra. *arXiv preprint arXiv:2104.14429*.

[6] Vandoorne, K., Mechet, P., Van Vaerenbergh, T., Fiers, M., Morthier, G., Verstraeten, D., ... & Bienstman, P. (2014). Experimental demonstration of reservoir computing on a silicon photonics chip. Nature communications, 5(1), 3541.

[7] Antonik Piotr, François Duport, Michiel Hermans, Anteo Smerieri, Marc Haelterman, and Serge Massar. ‘Online Training of an Opto-Electronic Reservoir Computer Applied to Real-Time Channel Equalisation’. IEEE Transactions on Neural Networks and Learning Systems 28, no. 11 (November 2017): 2686–98. https://doi.org/10.1109/TNNLS.2016.2598655.

[8] Antonik Piotr, Marc Haelterman, and Serge Massar. ‘Online Training for High-Performance Analogue Readout Layers in Photonic Reservoir Computers’. Cognitive Computation 9, no. 3 (June 2017): 297–306. https://doi.org/10.1007/s12559-017-9459-3.

[9] Penkovsky, B., Porte, X., Jacquot, M., Larger, L., & Brunner, D. (2019). Coupled nonlinear delay systems as deep convolutional neural networks. Physical review letters, 123(5), 054101.

[10] Goldmann, M., Köster, F., Lüdge, K., & Yanchuk, S. (2020). Deep time-delay reservoir computing: Dynamics and memory capacity. Chaos: An Interdisciplinary Journal of Nonlinear Science, 30(9), 093124.

[11] Butschek, L., Akrout, A., Dimitriadou, E., Haelterman, M., & Massar, S. (2020). Parallel photonic reservoir computing based on frequency multiplexing of neurons. *arXiv preprint arXiv:2008.11247*.

[12] Lupo, A., Butschek, L., & Massar, S. (2021). Photonic extreme learning machine based on frequency multiplexing. Optics Express, 29(18), 28257-28276.

[13] S. Moradi, N. Qiao, F. Stefanini, and G. Indiveri, “A Scalable Multicore Architecture With Heterogeneous Memory Structures for Dynamic Neuromorphic Asynchronous Processors (DYNAPs),” IEEE Transactions on Biomedical Circuits and Systems, vol. 12, no. 1, Art. no. 1, Feb. 2018, doi: 10.1109/tbcas.2017.2759700.

[14] SynSense “DynapSE2 a low power scalable SNN processor”, 2018, pdf: https://www.synsense-neuromorphic.com/wp-content/uploads/2018/10/wp-dynapse2-opt-20181031.pdf

[15] X. He, T. Liu, F. Hadaeghi, and H. Jaeger, “Reservoir Transfer on Analog Neuromor-phic Hardware,” Mar. 2019, doi: 10.1109/ner.2019.8716891.

[16] C. Mead, Analog VLSI and neural systems. Reading, Mass: Addison-Wesley, 1989.

[17] C. Mead, “Neuromorphic electronic systems,” Proceedings of the IEEE, vol. 78, no. 10, Art. no. 10, 1990, doi: 10.1109/5.58356.

[18] P. Blouw and C. Eliasmith, “Event-Driven Signal Processing with Neuromorphic Computing Systems,” May 2020, doi: 10.1109/icassp40776.2020.9053043.

[19] C. Frenkel, D. Bol, and G. Indiveri, “Bottom-Up and Top-Down Neural Processing Systems Design: Neuromorphic Intelligence as the Convergence of Natural and Artificial Intelligence,” Jun. 2021. arXiv:2106.01288.