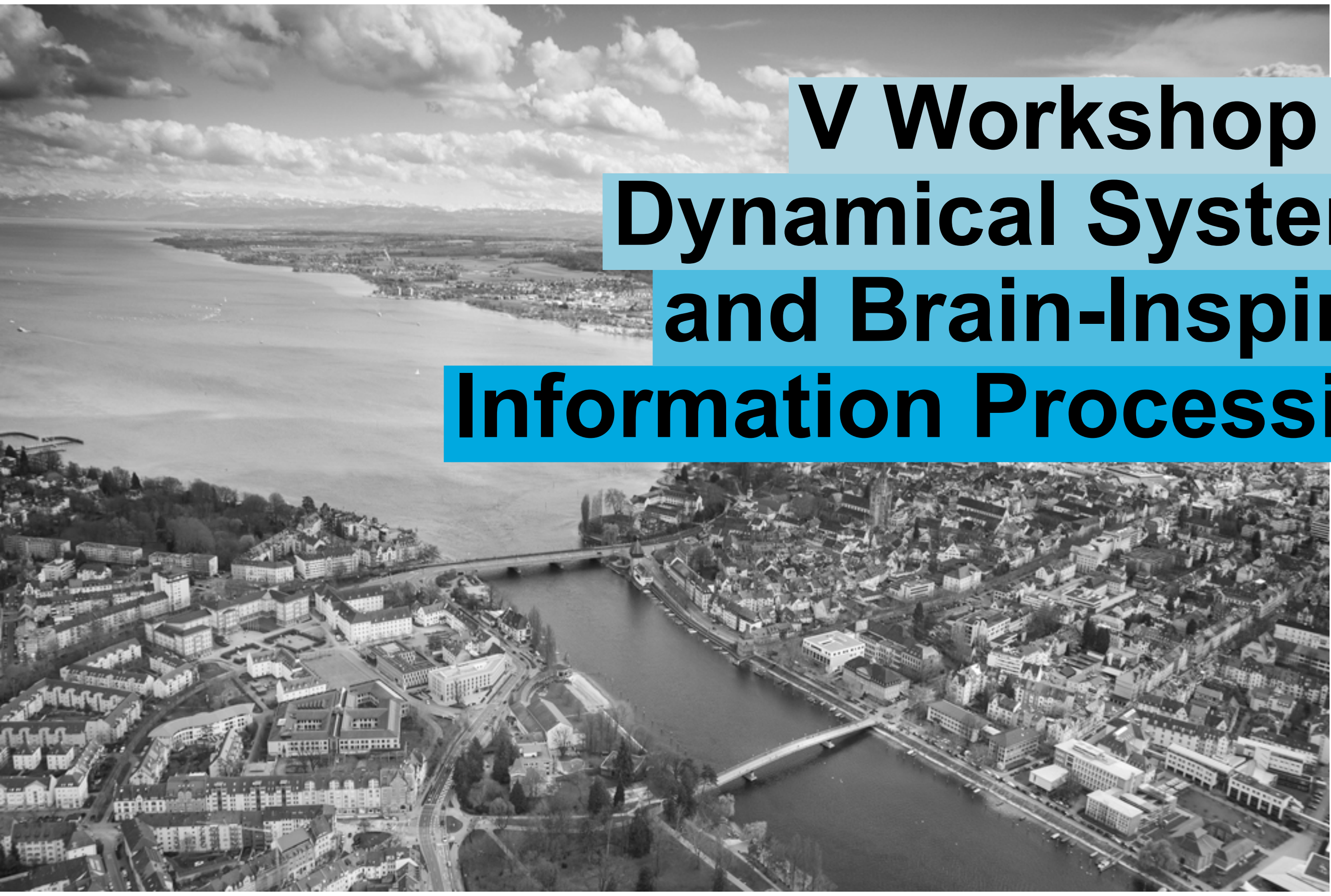
Universität
Konstanz



V Workshop on Dynamical Systems and Brain-Inspired Information Processing

July 29 - 31, 2019

@ Universität Konstanz, A704

Support

CarlZeissStiftung

Zukunftskolleg
@ Universität Konstanz

Graduate School of Decision Sciences
@ Universität Konstanz

Organizers

Daniel Brunner
CNRS (France)

Lukas Gonon
Universität St. Gallen (Switzerland)

Lyudmila Grigoryeva
Universität Konstanz (Germany)

Oleksandra Kukharenko
Universität Konstanz (Germany)

Juan-Pablo Ortega
Universität St. Gallen, CNRS

Josef Teichmann
ETH Zürich (Switzerland)

Abstracts of Oral and Poster Presentations

Pau Aceituno

Max Planck Institute for Mathematics in the Sciences (MIS), Leipzig, Germany

Computational Effects of Synaptic Time-Dependent Plasticity: From Predictions to Frequency Adaptation

We study the effects of Synaptic Time-Dependent Plasticity (STDP), a biological process that adjusts the strength of synapses depending on the timing of pre- and postsynaptic spikes. We start by working on feed-forward networks, where we show that postsynaptic latency reduction, a well known effect of STDP, can generate predictions while increasing the signal-to-noise ratio and at the same time reduce the metabolic costs – in terms of spikes – of the neural code. We then extend this notion of efficiency to recurrent neural networks, where we show that STDP can decrease the costs of controlling a neural population by inducing cycles in the network structure. Finally, we turn to Reservoir Computing, a machine learning paradigm to show how the aforementioned cycles can adapt recurrent neural networks to a given task.

Xavier Porte, **Louis Andreoli**, Nadezhda Semenova, Vladimir Semenov, Maxime Jacquot, Laurent Larger, and Daniel Brunner

FEMTO-ST, Université Bourgogne Franche-Comté, France

Noise and Consistency of Analogue Spatio-Temporal Photonic Neural Networks

Photonic analogue implementations of neural networks represent a novel computational paradigm beyond classical Turing/Von Neumann architectures, which demonstrates extremely powerful performance for non-algorithmic problems. Similarly to deep learning schemes, efficient operation of analogue neural networks requires large scale complex network architectures. Nevertheless, in contrast to digital systems, analogue implementations will always be prone to noise of different origin. Therefore, understanding the role of noise and its propagation through the multiple network layers is essential for all connectionist computing schemes. However, no systematic studies in this direction have so far been reported yet. We analyze the impact of noise on the performance of analogue and large-scale neural networks implemented in photonic hardware.

We use a particular type of a photonic system as testbed to study the effects of noise in neural networks. Crucially, all readout and network-internal connections are photonically realized fully in parallel. We experimentally characterize noise in each component of the system and its impact on the final performance. Noteworthy, noise influences the systems consistency, which is a particularly important property for computing as it measures the reproducibility of an operation. We study short- and long-term consistency of our system under different noise scenarios. Moreover, we have developed a full-scale model of noise propagation, confirming

the aspects of noise found in the experiment. In this context, we derive analytical scaling laws for network size and number of networks layers, which allow us to define system-independent noise mitigation strategies.

Our work lays the foundation towards a comprehensive framework for consistency and performance-robustness of physical analogue networks. It therefore is relevant for all computing based on networks, among others the coherent annealers. Based on our large-scale photonic network we advance the understanding of vital principles relevant to beyond Turing/Von Neumann computing hardware substrates.

Igor Gershenzon, **Geva Arwas**, Chene Tradonsky, Nir Davidson, and Oren Raz

Weizmann Institute of Science, Israel

Towards Uniform Intensity Laser Network XY Machine

We investigate the properties of the degenerate cavity laser (DCL) system as an optimization computational machine. The DCL is utilized as a platform supporting a network of discrete coherently coupled laser oscillators. We find that different operation regimes give rise to different optimization target functions. Specifically, at the lasing threshold, we find that the laser network configuration finds the minimal loss cold-cavity mode. In contrast, we find that uniform intensity lasing states can be mapped to minima of the classical 2D XY model. The mapping from the laser network to the XY model consists of identifying the phase with the planar spin orientation angle and the laser network coherent coupling matrix with the spin-spin interaction matrix. Numerical tests are carried out to asses the accuracy as a function of the network and single laser parameters. We find that lasing intensity heterogeneity causes significant inaccuracies in finding the XY ground state. We also demonstrate parameter tuning to achieve uniform intensity.

Peter Bienstman, Joni Dambre, Andrew Katumba, Matthias Freiberger, Floris Laporte, Alessio Lugnan, Stijn Sackesyn, Chonghuai Ma, Nathan Coste, Rachele Catalano, Emmanuel Gooskens

Ghent University – imec, Belgium

Silicon Photonics Reservoir Computing at 32 Gbit/s

We present our latest results on passive reservoir computing using integrated silicon photonics chips. These new chips are designed to operate at 32 Gbit/s. We show these chips can be used to do nonlinear dispersion compensation of telecom signals, resulting in a lower error rate compared to standard linear tapped filter equalisers. Next, we present a novel training strategy for all-optical weights with limited resolution, and show how performance is much improved compared to a naïve discretisation of the weights. We also discuss several ways to combine smaller reservoirs in a larger hierarchy, and present a software framework for optimisation of large optical circuits using backpropagation.

Lennard Böselt, SEREINA RINIKER

Laboratory of Physical Chemistry, ETH Zurich, Switzerland

COMPUTING *AB INITIO* MOLECULAR-DYNAMICS TRAJECTORIES USING ARTIFICIAL NEURAL NETWORKS

Accurate calculation of the potential of mean force (PMF) of chemical reactions requires long ab initio molecular-dynamics (AIMD) trajectories using a computationally expensive description of the physical interactions like density-functional theory (DFT) or wave-function (WF) methods. Artificial neural networks (ANNs) belong to the supervised machine-learning approaches, which can be used to accelerate AIMD simulations and thus, can facilitate the computation of PMFs.

However, different ANN architectures and feature spaces can influence the training, the convergence, and the final performance of the ANN significantly. In this work, we compare, how the use of ANNs with different architectures and feature spaces reduces the complexity of AIMD simulations.

DANIEL BRUNNER

CNRS, FEMTO-ST, Université Bourgogne Franche-Comté, France

GENERAL CONSIDERATIONS FOR NEURAL NETWORKS IMPLEMENTED IN HARDWARE

The implementation of neural networks in analogue substrates is gaining substantial momentum in recent years. Besides encouraging progress, one needs to keep in mind and address the essential challenges: the scalability of such a computing architecture. I will highlight some of the relevant challenges and put them into relationships with recently demonstrated experimental systems.

Stéphane Chrétien

National Physics Laboratory, UK

An Overview of Sparse Polynomial Expansions and How They Might Explain the Expressivity of Reservoir Computers

Approximating high dimensional function is of primary importance at the age of machine learning. In this talk, we will relate approximation, sampling and compressed sensing in order to get a clearer picture of how recent activities in machine learning are connected. We then conclude on open questions related to the expressive power of reservoir computers.

Christa Cuchiero[1], Wahid Khosrawi[2], and Josef Teichmann[2]

[1]University of Vienna, Austria
[2]ETH Zürich, Switzerland

A Neural Network Approach to Calibration of Local Stochastic Volatility Models

Exactly calibrated local stochastic volatility models are practically important instances of McKean Vlasov SDEs whose existence and uniqueness is in general a challenging open question. For their calibration both PDE and Monte Carlo methods building on propagation of chaos have been proposed. We tackle the calibration problem by a novel approach, namely by parameterizing the leverage function with neural networks and learning it from the implied volatility surface without using Dupire's local volatility function.

Nir Davidson

Weizmann Institute of Science, Rehovot, Israel

Solving Hard Computational Problems with Coupled Lasers

Hard computational problems may be solved by realizing physics systems that can simulate them. Here we present a new a new system of coupled lasers in a modified degenerate cavity that is used to solve difficult computational tasks. The degenerate cavity possesses a huge number of degrees of freedom (300,000 modes in our system), that can be coupled and controlled with direct access to both the x-space and k-space components of the lasing mode. Placing constraints on these components can be mapped to different computational minimization problems. Due to mode competition, the lasers select the mode with minimal loss to find the solution. We demonstrate this ability for simulating XY spin systems and finding their ground state, for phase retrieval, for imaging through scattering medium and more.

Christine De Mol
Department of Mathematics and ECARES, Université Libre de Bruxelles, Belgium

On Learning and Sparsity

Most consistency results for the sparsity-enforcing lasso strategy have been derived in the case of fixed-design regression. We argue that the assumption made for identification are not quite realistic in the presence of correlated features. In the framework of learning theory, i.e. of random design, learning guarantees can be easily derived for the lasso in prediction. For identification, however, an additional quadratic penalty seems to be needed as a safeguard ensuring stability in the feature selection process.

Stephan Eckstein
Universität Konstanz, Germany

ANNs for (Distributionally) Robust Optimization, Representations of Probability Measures and Updating of Beliefs

Non-parametric and efficient representations of probability measures in high dimensions are a major challenge in many areas of optimization. Neural network based approaches have found use in this respect in areas like generative adversarial networks, optimal transport problems, and distributionally robust optimization. We present one of these approaches based on parametrization of densities, which comes up naturally when utilizing convex duality and penalization tools in many of the above mentioned optimization settings. Intrinsic to this approach is to start with an initial estimate for an optimizer, which can be seen as the counterpart to a prior in Bayesian settings. Similarly, one can utilize an updating procedure reminiscent of Bayesian updating of beliefs.

Gerhard Fechteler
Universität Konstanz, Germany

Marginal Effect based Inference in the Deep Learning Framework

In the recent years, deep learning has proven to be a powerful tool for nonlinear function approximation and modeling. If interested in the effect of the input parameters on the output, commonly either sensitivity analysis is applied or the marginal effects are computed. However, as opposed to linear frameworks, no tests for the significance of the effects are available, which would be crucial for inference. This paper contributes to filling this gap in the literature. Given a fixed network structure, it is shown that the marginal effects are asymptotically normally distributed. Algorithms to compute the asymptotic covariance matrix are presented, which makes common hypothesis testing available. The theoretical results are confirmed by simulation studies.

Claudio Gallicchio, Alessio Micheli

Computational Intelligence and Machine Learning Group,
Dipartimento di Informatica, Università di Pisa, Italy

Deep Reservoir Computing

In many application contexts, such as text and speech processing, time-series data is crucially featured by temporal hierarchies, i.e. multiple time-scales. It is then not surprising that in recent years Deep Learning models for temporal data, and especially deep Recurrent Neural Networks (RNNs), have become the focus of extensive research efforts within the Machine Learning community.

In this context, Reservoir Computing (RC) offers a refreshing perspective to the analysis and design of deep RNN architectures. We illustrate the features of a novel RC model, called Deep Echo State Network (DeepESN), whose dynamics are based on the nested composition of multiple discrete-time dynamical reservoir systems. The study of DeepESN allows shedding light on the intrinsic advantages of depth in the architectural construction of dynamical neural systems, especially in terms of improved richness of the developed dynamics. At the same time, the DeepESN model is put forward as an efficient approach to designing and training deep RNNs, where a grounded architectural design can have a high impact on real-world applications.

Chene Tradonsky, **Igor Gershenzon**, Ronen Chriki, Vishwa Pal, Asher A. Friesem, Oren Raz, and Nir Davidson

Weizmann Institute of Science, Israel

Teaching an Old Laser New Tricks: Solving the Phase Retrieval Problem Rapidly

Reconstructing an object solely from its scattered intensity distribution is a common problem that occurs in many applications. Currently, there are no efficient direct methods to reconstruct the object, though in many cases, with some prior knowledge, iterative algorithms result in reasonable reconstructions. Unfortunately, even with advanced computational resources, these algorithms are highly time consuming. Here we present a novel rapid all-optical method based on a digital degenerate cavity laser, whose most probable lasing mode approximates the object. We present experimental results showing the high speed ($<$100 ns) and efficiency of our method. The method is scalable, and can be applicable to any two dimensional object with known compact support, including complex-valued objects.

Lukas Gonon[1], Lyudmila Grigoryeva[2], and Juan-Pablo Ortega[1,3]

[1]Universität Sankt Gallen, Switzerland
[2]Universität Konstanz, Germany
[3]Centre National de la Recherche Scientifique (CNRS), France

Generalization Error Bounds for Reservoir Computing

In this talk we present our recent mathematical results on universality and error bounds for reservoir computing. We study approximation properties of reservoir computing systems, putting particular emphasis on echo state networks (ESNs). For ESNs with output feedback we obtain high-probability bounds on the approximation error in terms of the network parameters. For a general class of reservoir computing systems and weakly dependent (possibly non-i.i.d.) input data, we then also derive generalization error bounds based on a Rademacher-type complexity.

Jake Bouvrie[1], Peter Giesl[2], **Boumediene Hamzi**[3], Bernard Hasasdonk[4], Christian Kuehn[5], Sameh Mohamed[6], Martin Rasmussen [3], Gabriele Santin[4], Kevin Webster[3], and Dominik Wittwar[4]

[1]MIT, USA
[2]University of Sussex, UK
[3]Imperial College London, UK
[4]University of Stuttgart, Germany
[5]TUM, Munich, Germany
[6]SUTD, Singapore

Kernel Methods for Dynamical Systems

We introduce a data-based approach to estimating key quantities which arise in the study of nonlinear autonomous, control and random dynamical systems. Our approach hinges on the observation that much of the existing linear theory may be readily extended to nonlinear systems - with a reasonable expectation of success- once the nonlinear system has been mapped into a high or infinite dimensional Reproducing Kernel Hilbert Space. In particular, we develop computable, non-parametric estimators approximating controllability and observability energies and Lyapunov functions for nonlinear systems. We apply this approach to the problem of model reduction of nonlinear control systems. It is also shown that the controllability energy estimator provides a key means for approximating the invariant measure of an ergodic, stochastically forced nonlinear system. We also show how kernel methods can be used to detect critical transitions for some multi scale dynamical systems. Finally, we show how kernel methods can be used to approximate center manifolds for nonlinear ODEs.

Giacomo Indiveri

University of Zurich, Switzerland

Analog/Digital Neuromorphic Circuits for Building Reservoirs of Spiking Neurons with Multiple Time Scales

In this talk I will present examples of neuromorphic VLSI chips that use ananlog neuron and synapse CMOS circuits and digital asynchronous circuits able to support the creation of ultra-low power spiking reservoirs. I will show case examples of applications of reservoir computing demonstrations using spiking architectures implemented with neuromorphic processor chips, discuss the challanges of achieving robust performance using noisy analog circuits and low precision synaptic weights and propose potential benefits of the approach presented for practical application scenarios.

Coralie Joucla[1], Damien Gabriel[1,2], Juan-Pablo Ortega[3,4], Emmanuel Haffen[1,2], and Lyudmila Grigoryeva[5]

[1]Université Bourgogne Franche-Comté, France
[2]CHRU Besançon, France
[3]Universität Sankt Gallen, Switzerland
[4]Centre National de la Recherche Scientifique (CNRS), France
[5]Universität Konstanz, Germany

Support Vector Machines Applied to EEG: A Reliable Pipeline for Tough Problems

Brain electrical activity can be recorded via placing ElectroEncaphaloGraphic (EEG) electrodes on the scalp surface. Identification of this activity allows to single out brain areas and temporal decay triggered by a specific stimulus. However, each and every activity in the brain of a subject independently of the experimental tasks elicits electrical activations within different parts of the brain. These represent a noise component often more pronounced than the signal we ought to detect. In conventional studies this problem is overcome by averaging a great number of trials across subjects which makes the activity of interest emerge. However, while these methods allow for fundamental studies, they are difficult to be applied in a clinical situation where a single subject examination is needed.

In recent years, machine learning (ML) methods gained popularity in medical applications (in particular, in the treatment of EEG signals) due to their success in discriminating patterns in data that are hardly or even not distinguishable with conventional techniques. Moreover, in the field of neuroscience ML models may potentially allow for moving from group-level statistical results to personalized diagnosis in neuro-psychiatric pathologies.

However, special care is needed when applying ML methods to EEG data suffering from intrinsic low signal-to-noise ratio and from a high inter-individual variability. We propose a pipeline which consists in a preliminary data preprocessing stage using the xDAWN method, subsequent exploration of the best way to extract relevant information from EEG recordings, and, finally, exploiting Support Vector Machines (SVM) in order to carry out individual subject-level classification task. It is customary in the literature to use the SVM in a out-of-the-box manner. Our study highlights the necessity for the SVM hyper-parameters to be tuned for each test subject of interest.

Results show a significant difference in classification accuracy and p-value depending on the extracted features and hyper-parameters from subject to subject. Selection of spatial regions and time window of interest will be later studied. Additionally, we use reservoir computing techniques (for example Echo State Networks) in the context of EEG signals classification.

Věra Kůrková

Czech Academy of Sciences, Czech Republic

Lower Bounds on Complexity of Shallow Networks

Although originally biologically inspired neural networks were introduced as multilayer computational models, shallow networks have been dominant in applications till the recent renewal of interest in deep architectures. Experimental evidence and successful applications of deep networks pose theoretical questions asking: When and why are deep networks better than shallow ones?

This lecture will present some probabilistic and constructive results showing limitations of shallow networks. It will show how geometrical properties of high-dimensional spaces imply probabilistic lower bounds on network complexity. Bounds depending on covering numbers of dictionaries of computational units will be derived and combined with estimates of sizes of some common dictionaries used in neurocomputing. Probabilistic results will be complemented by constructive ones built using Hadamard matrices and pseudo-noise sequences. The results will be iullustrated by an example of a class of functions which can be computed by two-hidden-layer perceptron networks of considerably smaller model complexities than by networks with one hidden layer. Connections with the No Free Lunch Theorem and the central paradox of coding theory will be discussed.

Laurent Larger[1], Maxime Jacquot[1], Yuri Maistrenko[2], and Daniel Brunner[1,3]

[1] FEMTO-ST, Université Bourgogne Franche-Comté, France
[2]National Academy of Sciences of Ukraine, Ukraine
[3]CNRS, France

From Reservoir Computing to Chimera States... and Back to RC?

Reservoir Computing and chimera states appeared both at the same time, independently, in the early 2000. The first proposed a novel architecture for recurrent neural network with simplified learning and efficient processing of complex information, whereas the second dealt with the discovery of an unexpected occurrence of complex but organized symmetry breaking patterns, in fully homogeneous coupled network of oscillators. 10 years later both topics met in a tiny and fruitful way through the common implication of delay dynamics as a toy model model to explore them, and to contribute through this particular class of complex dynamics to a better understanding of their underlying dynamical mechanisms. We propose to report on both topics, how they involve some particular features revealed by a specific viewpoint on their modeling, and how they have also led to real world experimental demonstrations.

Andrea Ceni[1], Peter Ashwin[1], **Lorenzo Livi**[1,2], and Claire Postlethwaite[3]

[1] University of Exeter, UK
[2]University of Manitoba, Canada
[3]University of Auckland, New Zealand

Multi-Stability of Input-Driven Recurrent Neural Networks

If a recurrent neural network possesses the echo state property, this means that, for a given input, it will produce the same output regardless of initial conditions: it will "forget" any internal states and end up following a unique (possibly complex) trajectory. However, recurrent neural networks might take advantage of internal memory states that correspond to the simultaneous existence of multiple attractors in the absence of inputs (e.g. in classification and complex memory tasks).

How do these properties fit together? We introduce a generalisation of the echo state property that allows us to understand how the echo state property may be gained for a multi-stable system subjected to inputs.

Romain Alata, Jaël Pauwels, Marc Haelterman, and
Serge Massar

Université libre de Bruxelles, Belgium

Phase Noise Robustness of a Coherent Spatially Parallel Optical Reservoir

Reservoir computing is a machine learning algorithm particularly adapted to process time-dependent signals. It can be easily implemented experimentally with good performance. However experimental implementations are subject to noise which degrades performance. We develop strategies to mitigate the effects of noise. The specific system on which we illustrate our approaches – currently under development in our laboratory – is a coherent linear Fabry-Perot resonator in which neurons are encoded as a grid of spots on the input mirror plane. In this system changes in length of the resonator are the major source of noise, and can be modeled as phase noise. This can in principle be partially solved by active stablisation, but it is interesting to find other strategies to counter the effect of phase noise. We show that a completely unknown phase can be tolerated with only a small degradation of performance by using appropriate training and a readout layer architecture in which the output weights depend on the noisy parameter (the phase). Furthermore the phase can be estimated by the reservoir itself, leading to an architecture in which the reservoir has two outputs, one of which (the phase estimation) is used to control the other. Our approach should find applications in many experimental implementations of reservoir computing.

Alessio Micheli

Computational Intelligence and Machine Learning Group,
Dipartimento di Informatica, Università di Pisa, Italy

Reservoir Computing for Structured Domains

The processing of structured domains (sequences, trees and graphs data) is a mainstream in the current evolution of Neural Networks (NN) and Machine Learning (ML), which is rooted in foundational works developed in the last 20 years. The recent advancements within the current deep learning revolution often come at the price of a high computation cost, emphasizing the need for efficient approaches.

The Reservoir Computing naturally offers a paradigm for modeling untrained (hence extremely efficient) dynamical NN systems. We will show by a short excursus through different models, including deep architectures, how it is possible to extend such paradigm from sequential to hierarchical up to general graph data processing, combining potential benefits both in terms of computational cost and performance.

Lorenz K. Müller, Pascal Stark, Bert Offrein, and Stefan Abel

IBM Research Zurich, Switzerland

Reservoirs for Neuromorphics: Concepts for System Architectures

Reservoir Computing (RC) is an appealing computational framework for neuromorphic systems. This is because reservoirs use simple basic operations, they rely on fixed, random, sparse connections, and they are recurrent, i.e. perform computation incrementally, in-place. However, in the last years RC has faded in popularity in the Machine Learning community in favor of fully trained architectures, which are well suited for standard hardware. In this talk I present our recent efforts in finding a place for RC systems in the context of current deep learning architectures.

For example, when using Echo-State-Networks or other inherently sequential approaches for image classification, we need to transform two-dimensional image data into a one-dimensional data stream. This initially seems like a nuisance, but the freedom to choose this transformation can be a blessing in disguise: We propose to use a set of random scanning paths over input images to transform a given dataset of images to a set of data-sets of input streams. These datasets can be combined using multiple predictors in a form of bootstrapping.

Further we investigate the use of fixed random recurrent connections to complement sparse learned feedforward connections. We argue that given a computational substrate that implements the random recurrence cheaply, such a combination can give rise to resource efficient neural networks.

Lukas Gonon[1], Lyudmila Grigoryeva[2], and **Juan-Pablo Ortega**[1,3]

[1]Universität Sankt Gallen, Switzerland
[2]Universität Konstanz, Germany
[3]Centre National de la Recherche Scientifique (CNRS), France

Dynamic and Control Theoretical Aspects of Reservoir Computing

Reservoir computing (RC) carries out the learning of input/output systems by using families of state space systems in which the training of a small portion of their parameters suffices for excellent generalization properties. In this talk we shall analyze in detail the connection between various dynamic and control theoretical features of widely used families of RC families and their impact in basic learning theoretical goals, in particular in the development of bounds for the approximation and estimation errors and in the design of dimension reduction techniques.

Bruno Romeira[1], Jos Figueiredo[1], Julien Javaloyes[2], and **Ignacio Ortega Piwonka**[2]

[1]Universidade do Algarve, Portugal
[2]Universitat de les Illes Balears, Spain

Delay Dynamics of Optoelectronic Neuromorphic Nanoscale Resonators

Project ChipAI aims to develop an energy-efficient neuromorphic nanophotonic architecture technology using neuron-like nanoscale non-linear light sources and detectors to realize interconnected high-bandwidth spike-encoded synapses for optical neural networks, and hence capable of addressing the predicted future needs of AI systems and computing processors.

Edward Ott

University of Maryland at College Park, USA

Machine Learning for Discovery of Dynamical Processes of Chaotic Systems

In this talk we consider the situation where one is interested in gaining understanding of the dynamical properties of a chaotically time evolving system solely through access to a limited duration of time series measurements that depend on the evolving state of an, otherwise unknown, system. We will show that machine learning is an extremely effective tool for accomplishing this general type of task, and we discuss how the ability to do this can be of practical utility. Several examples illustrating this for different purposes will be given. We also turn the problem around and utilize chaos theory to explain the dynamical basis for how a machine learning system is able to do accomplish such tasks.

Xavier Porte[1], Louis Andreoli[1], Stéphane Chrétien[2], Maxime Jacquot[1], Laurent Larger[1], and Daniel Brunner[1,3]

[1] FEMTO-ST, Université Bourgogne Franche-Comté, France
[2]National Physics Laboratory, UK
[3]CNRS, France

Greedy Boolean Learning in Photonic Recurrent Neural Networks

We demonstrate the implementation of a large scale recurrent photonic neural network with up to 2025 photonic neurons. All network internal and readout connections are physically implemented with fully parallel photonic technology. Based on a digital micro-mirror array, we can train the Boolean readout weights using a greedy version of reinforcement learning paradigm. We find that the learning excellently converges. Furthermore, it appears to possess a conveniently convex-like cost-function and demonstrates exceptional scalability of the learning effort with system size.

Bruno Romeira[1], C. Maibohm[1], P. Moura[1,2], J. Lourenço[3], J. M. L. Figueiredo[3], J. Javaloyes[4], and Jana B. Nieder[1]

[1] INL – International Iberian Nanotechnology Laboratory, Portugal
[2]Universidade Nova de Lisboa, Portugal
[3]Universidade de Lisboa, Portugal
[4]Universitat de les Illes Balears, Spain

Micro- and Nanoscale Photonic Neuron Circuits

Artificial intelligence systems using computing algorithms of deep neural networks are emerging rapidly. However, reducing the energy consumption of the massively dense interconnects in existing CPUs needed to emulate complex brain functions is a major challenge. Therefore, there is a worldwide effort towards the development of an artificial brain using energy-efficient interconnected neuromorphic devices. The aim is to create the hardware elements capable of performing both memory and information processing analogous to the neural and synaptic functions of the brain.

Here, we present our work on the implementation of artificial photonic spiking neuron circuits with capabilities to process information in a way more analogous to the biological brain. We will discuss the fundamental aspects of solid-state photonic chips to implement artificial neural activity and regenerative memory operations. Firstly, we cover our recent results on neuromorphic excitable (spike-encoded) photonic chips using delayed feedback photonic resonators with applications in optical memories. Then, we discuss the implementation of energy-efficient and ultrafast on-chip neuromorphic photonic circuits and interconnects using nanoscale semiconductor nanostructures integrated with subwavelength cavities. These form a new class of optically interconnected nanophotonic spiking neuron circuits with potential for energy-efficient and high-speed brain-inspired neuromorphic computing.

Christa Cuchiero[1], Lukas Gonon[2], Lyudmila Grigoryeva[3], Martin Larsson[4], Juan-Pablo Ortega[2,5], and **Josef Teichmann**[4]

[1]University of Vienna, Austria
[2]Universität Sankt Gallen, Switzerlandl
[3]Universität Konstanz, Germany
[4]ETH Zürich, Switzerland
[5]CNRS, France

Randomness in Training Algorithms

We analyse from a control-theoretic viewpoint the role of randomness in training algorithms. An application of Chow's theorem on controllability together with a result on random vector (neural network type) vector fields then yields that mainly randomly chosen deep neural networks are expressive. Similar results adjusted to recurrent neural networks are shown by methods from compressed sensing. Parallels to reservoir computing are discussed.

Peter Tiňo

University of Birmingham, UK

Machine Learning in the Space of Dynamical Systems

In learning from "static" data (order of data presentation does not carry any useful information), one framework for dealing with such data is to transform the input items non-linearly into a feature space (usually high-dimensional), that is "rich" enough, so that linear techniques are sufficient. However, data such as EEG signals, or biological sequences naturally comes with a sequential structure. I will present a general dynamical state space model that effectively acts as a dynamical feature space for representing temporally ordered samples. I will then outline a framework for learning on sets of sequential data by building kernels based such dynamical filters. The methodology will be demonstrated in a series of sequence classification tasks and in an incremental temporal "regime" detection task.

Guy Van der Sande[1], Fabian Böhm[1], Krishan Harkhoe[1], Andrew Katumba[2], Peter Bienstma[2], and Guy Verschaffelt[1]

[1]Vrije Universiteit Brussel, Belgium
[2]Ghent University – imec, Belgium

Optical Systems with Delayed Feedback for Integrated Optical Reservoir Computing and Ising Machines

We present two optical systems based on delayed self-feedback: one to solve complex information processing tasks using reservoir computing, one to tackle NP-hard optimisation problems relying on the concept of ising machines.

First, delay-based reservoir computing offers a simple technological route to implement photonic neuromorphic computation. Its operation boils down to a time-multiplexing with the delay limiting the processing speed. As most optical setups end up to be bulky employing long fiber loops or free-space optics, the processing speeds are limited in the range of kSa/s to tens of MSa/s. In this work, we focus on external cavities which are far shorter than what has been realized before in experiment reaching computation speeds of up to 0.87 GSa/s. We present the results of an experimental validation of reservoir computing based on a semiconductor laser with a 10.8 cm delay line, both integrated on an active/passive InP photonic chip.

Second, to reduce the footprint and complexity of coherent ising machines (CIMs) and improve their stability, we propose a new scheme based on opto-electronic oscillators (OEOs) subjected to self-feedback. OEOs are easily build from few off-the shelf photonic components and are well known for their rich bifurcation structure and inherent stability. Like machines based on degenerate optical parametric oscillators (DOPOs), the OEO-based CIM uses bistable optical states to represent the spin up and spin down configuration of Ising spins. However, the artificial spins are encoded in the light intensity instead of the optical phase, making it more robust against external perturbations. Also contrary to DOPO-based systems, the spins are generated in a feedback-induced bifurcation, which removes the necessity for nonlinear optical processes and large external cavities and offers significant advantages regarding size, cost and stability.

Pietro Verzelli[1], Cesare Alippi[1,2], and Lorenzo Livi[3]

[1]Università della Svizzera italiana, Switzerland
[2]Politecnico di Milano, Italy
[3]University of Manitoba, Canada & University of Exeter, UK

Hyper-Spherical Reservoirs for Echo State Networks

We propose a model of ESNs that eliminates critical dependence on hyper-parameters, resulting in networks that provably cannot enter a chaotic regime and, at the same time, denotes nonlinear behaviour in phase space characterised by a large memory of past inputs, comparable to the one of linear networks. Our contribution is supported by experiments corroborating our theoretical findings, showing that the proposed model displays dynamics that are rich-enough to approximate many common nonlinear systems used for benchmarking.