International Conference:
Differential Equations for Data Science 2022 (DEDS2022)

Monday, March 21 — Wednesday, March 23, 2022
Online (Zoom)
`https://scheme.hn/deds2022/`

Organizers:
Hayato Chiba (Tohoku University, Japan)
Ingo Fischer (IFISC (UIB-CSIC), Spain)
Lyudmila Grigoryeva (University of Warwick, UK)
Masato Kimura (Kanazawa University, Japan)
Hiroshi Kokubu (Kyoto University, Japan)
Kohei Nakajima (The University of Tokyo, Japan)
Hirofumi Notsu (Kanazawa University, Japan, Chair)
Juan-Pablo Ortega (Nanyang Technological University, Singapore)

Registration:
`https://zoom.us/meeting/register/tJIsd-mtrjItGtaYF3-FK_8gpryQdRjxFMsf`
After registering, you will receive a confirmation email containing information about joining the meeting. (Once you register from the link above, you are able to join all the sessions.) Note that the Zoom meeting will be available from 30 minutes before each session start time and that the maximum number of registrants is 500.

Program

∗ All lectures are given by invited speakers.

Monday, March 21

16:55 – 17:00   Opening

Session 1 (4 talks) Chair: H. Notsu: JST 17:00 – 20:00
(= UTC 08:00 – 11:00 = CET 09:00 – 12:00 = PST 00:00 – 03:00)

17:00 – 17:50   Peter Ashwin (University of Exeter, UK)

Excitable network attractors and computational properties of recurrent neural networks

17:50 – 18:00   Break

18:00 – 18:30   Masato Hara and Hiroshi Kokubu (Kyoto University, Japan)

Degenerate reservoir for studying learning mechanism of reservoir computing

18:30 – 19:00   Koichi Taniguchi (Tohoku University, Japan)

Spectral pruning for recurrent neural networks

19:00 – 19:10   Break

19:10 – 20:00   Allen Hart (University of Bath, UK)

Solving PDEs with random neural networks

Tuesday, March 22

Session 2 (4 talks) Chair: M. Kimura and H. Notsu: JST 17:00 – 20:20
(= UTC 08:00 – 11:20 = CET 09:00 – 12:20 = PST 00:00 – 03:20)

17:00 – 17:50   Serhiy Yanchuk (Potsdam Institute for Climate Impact Research
& Humboldt-Universität zu Berlin, Germany)

Deep neural networks using a single neuron: folded-in-time architecture using feedback-modulated delay loops

17:50 – 18:00   Break

18:00 – 18:50   André Röhm (The University of Tokyo, Japan)

Model-free inference of unseen attractors using reservoir computing

18:50 – 19:20   Hayato Chiba (Tohoku University, Japan)

Bifurcation of the Kuramoto model on networks based on the generalized spectral theory

19:20 – 19:30   Break

19:30 – 20:20   Manjunath Gandhi (University of Pretoria, South Africa)

Models from dynamical data: Part I

Wednesday, March 23

Session 3 (4 talks) Chair: H. Notsu: Session 3: JST 17:00 – 20:00
(= UTC 08:00 – 11:00 = CET 09:00 – 12:00 = PST 00:00 – 03:00)

17:00 – 17:30   Tomoyuki Kubota (The University of Tokyo, Japan)

Information processing capacity for reservoir computing

17:30 – 18:00   Kengo Nakai (Tokyo University of Marine Science and Technology, Japan)

Constructing differential equations using only a chaotic time-series

18:00 – 18:10   Break

18:10 – 19:00   Boumediene Hamzi (Imperial College London, UK)

Machine learning and dynamical systems meet in reproducing kernel Hilbert spaces

19:00 – 19:10   Break

19:10 – 20:00   Manjunath Gandhi (University of Pretoria, South Africa)

Models from dynamical data: Part II

20:00 – 20:05   Closing

Abstract

T. = Title, A. = Abstract.

1. Peter Ashwin (University of Exeter, UK)

T. Excitable network attractors and computational properties of recurrent neural networks

A. Recurrent neural networks (RNNs) are input-driven nonlinear dynamical systems that can be trained to undertake a variety of computational tasks depending on input. This ranges from simple classification tasks (where it is desirable to have an echo state property (ESP) such that memory of internal states are lost) to finite-state computations where internal states need to be stored at intermediate steps. In this talk I will outline recent work using excitable network attractors to describe (a) input-dependence of the ESP and a generalized "echo index" that quantifies the degree of multistability depending on input (b) design of RNNs with arbitrary excitable network attractors and (c) use of excitable network attractors to understand coupled dynamical systems that realise arbitrary Turing machines. (Joint work with A. Ceni, L. Livi and C. Postlethwaite.)

2. Masato Hara and Hiroshi Kokubu (Kyoto University, Japan)

T. Degenerate reservoir for studying learning mechanism of reservoir computing

A. We study learning by reservoir computing of chaotic dynamical systems such as logistic maps. Our previous numerical studies show that, if the learning of chaotic dynamics is successful, a dynamical system that is (semi-)conjugate to the target chaotic dynamics is formed in the phase space of the reservoir. The question then arises as to how there can be appropriate output weights that provide such a (semi-)conjugacy.

We have recently introduced the idea of "degenerate reservoir" and considering an ordinary reservoir as its singular perturbation. In this talk, we will present numerical examples of learning by degenerate reservoir computing of the logistic map, and discuss the existence of output functions of the degenerate reservoir and its singularly perturbed reservoir that give (semi-)conjugacies to the target chaotic dynamics.

3. **Koichi Taniguchi (Tohoku University, Japan)**

T. Spectral pruning for recurrent neural networks

A. Recurrent neural networks (RNNs) are a class of neural networks used in sequential tasks. However, in general, RNNs have a large number of parameters and involve enormous computational costs by repeating the recurrent structures in many time steps. As a method to overcome this difficulty, RNN pruning has attracted increasing attention in recent years. However, most existing methods of RNN pruning are heuristic. In this talk, we propose a pruning algorithm for RNNs from the viewpoint of the generalization error bounds, in which we use "information loss" to quantify less effective nodes of hidden state of RNNs. Our pruning is inspired by spectral pruning for DNNs (T. Suzuki, et al., 2020). We also provide numerical experiments to demonstrate our theoretical results and show the effectiveness of our pruning method compared with the existing methods. This talk is based on the joint work with T. Furuya, K. Suetake, H. Kusumoto, R. Saiin, and T. Daimon (arXiv:2105.10832).

4. **Allen Hart (University of Bath, UK)**

T. Solving PDEs with random neural networks

A. When using the finite element method, we approximate the solution of a PDE with a linear combination of basis functions, which are for example: little hats. In this talk, the basis functions will instead be random functions, like neural networks with random weights and biases, or Gaussians with random means and variances. In either case, given a PDE and a random neural network, we can determine a so-called readout layer that is combined with the network to approximate the solution. We obtain the readout layer using linear regression, so that we avoid the non-convex optimisation involved in training a vanilla neural network. We demonstrate the idea on 3 problems of escalating difficulty: Poisson's equation, Schrödinger's equation, and then the Navier–Stokes equations.

5. **Serhiy Yanchuk (Potsdam Institute for Climate Impact Research & Humboldt-Universität zu Berlin, Germany)**

T. Deep neural networks using a single neuron: folded-in-time architecture using feedback-modulated delay loops

A. We present a method for folding a deep neural network of arbitrary size into a single neuron with multiple time-delayed feedback loops [1]. This

single-neuron deep neural network contains only a single nonlinearity and appropriately adjusted modulations of the feedback signals. The network states emerge in time as a temporal unfolding of the neuron's dynamics. By adjusting the modulations within the feedback loops, we adapt the network's connection weights. These connection weights are determined via a back-propagation algorithm, where both the delay-induced and local network connections must be taken into account. Our approach can fully represent standard Deep Neural Networks (DNN), encompasses sparse DNNs, and extends the DNN concept toward dynamical systems implementations. The new method, which we call Folded-in-time DNN (Fit-DNN), exhibits promising performance in a set of benchmark tasks.

[1] Stelzer, F., Röhm, A., Vicente, R., Fischer, I., & Yanchuk, S. (2021). Deep neural networks using a single neuron: folded-in-time architecture using feedback-modulated delay loops. *Nature Communications*, 12(1), 5164.

6. André Röhm (The University of Tokyo, Japan)

T. Model-free inference of unseen attractors using reservoir computing

A. A reservoir computer can be trained to act as a surrogate for a target dynamical system using a sample time series. The state-of-the-art for such autonomous-mode reservoirs can reproduce a large variety of statistical properties. However, real-world dynamical systems can exhibit more than one long-term stable dynamical behavior, called an attractor. In this talk, we will show that a properly trained reservoir computer can not only infer the dynamics of the target dynamical system in the region from which the training data was taken, but that valid reconstructions can be obtained even for unseen attractors - i.e. those for which no time series data was available in training. These attractors do not even need to share a type, i.e. a limit cycle could be used to infer the existence of a torus solution.

7. Hayato Chiba (Tohoku University, Japan)

T. Bifurcation of the Kuramoto model on networks based on the generalized spectral theory

A. For the mean-field limit of a system of globally coupled phase oscillators defined on networks, a bifurcation from the incoherent state to the partially locked state at the critical coupling strength is investigated based on the

generalized spectral theory. This reveals that a network topology affects the dynamics through the eigenvalue problem of a certain Fredholm integral operator which defines the structure of a network.

8. **Manjunath Gandhi (University of Pretoria, South Africa)**

    T. Models from dynamical data (Part I & Part II)

    A. A discrete-time dynamical system could arise naturally while modeling a real-world phenomenon or as a time-t map from the flow of an ordinary differential equation. The left-infinite orbits of a discrete-time dynamical system form what is called the inverse-limit space of the discrete-time dynamical system. Reservoir computing methods employ driven dynamical systems like recurrent neural networks to map temporal data onto a different space. The talk would concern using driven dynamical systems with certain properties to topologically embed the inverse-limit space of a discrete-time dynamical into a space derived from its state space. Such an embedding helps obtain equations from data and renders long-term topologically and statistically consistent models from data.

9. **Tomoyuki Kubota (The University of Tokyo, Japan)**

    T. Information processing capacity for reservoir computing

    A. Reservoir computing is a framework that utilizes a dynamical system as a computational resource and has been widely applied to simulated or physical systems. To date, their computational capabilities have been evaluated by benchmark tasks; however, these tasks have a problem that its performance reflects a part of the capabilities and overlooks the rest of them. In this presentation, we introduce a computational measure called information processing capacity (IPC), which can investigate all the capabilities of the system without omission—i.e., the IPC is equivalent to a coefficient of a state expanded with orthonormal polynomials and provides comprehensive computational capabilities by expanding with a sufficient number of polynomials [1]. Finally, to demonstrate the usefulness of the IPC, we decompose the capabilities of echo state networks that solve the representative benchmark task called the NARMA10 task. (Joint work with H. Takahashi and K. Nakajima.)

    [1] Tomoyuki Kubota, Hirokazu Takahashi, and Kohei Nakajima. "Unifying framework for information processing in stochastically driven dy-

namical systems." *Physical Review Research* 3.4 (2021): 043135.

10. Kengo Nakai (Tokyo University of Marine Science and Technology, Japan)

T. Constructing differential equations using only a chaotic time-series

A. We propose a simple method of constructing a system of differential equations of chaotic behavior based on the regression only from an observable time-series. The novelty is the introduction of a set of Gaussian radial basis functions in addition to polynomials to capture local structures. With the proposed method, the estimated system enables us to reconstruct detailed statistical behavior as well as to infer short time-series. We apply it to a time-series of a variable of the well-known chaotic Lorenz system and a macroscopic fluid variable. This is the joint work with Yoshitaka Saiki and Natsuki Tsutsumi (Hitotsubashi U.).

11. Boumediene Hamzi (Imperial College London, UK)

T. Machine learning and dynamical systems meet in reproducing kernel Hilbert spaces

A. Since its inception in the 19th century through the efforts of Poincaré and Lyapunov, the theory of dynamical systems addresses the qualitative behaviour of dynamical systems as understood from models. From this perspective, the modeling of dynamical processes in applications requires a detailed understanding of the processes to be analyzed. This deep understanding leads to a model, which is an approximation of the observed reality and is often expressed by a system of Ordinary/Partial, Underdetermined (Control), Deterministic/Stochastic differential or difference equations. While models are very precise for many processes, for some of the most challenging applications of dynamical systems (such as climate dynamics, brain dynamics, biological systems or the financial markets), the development of such models is notably difficult. On the other hand, the field of machine learning is concerned with algorithms designed to accomplish a certain task, whose performance improves with the input of more data. Applications for machine learning methods include computer vision, stock market analysis, speech recognition, recommender systems and sentiment analysis in social media. The machine learning approach is invaluable in settings where no explicit model is formulated, but measurement data is available. This is frequently the case in many

systems of interest, and the development of data-driven technologies is becoming increasingly important in many applications.

The intersection of the fields of dynamical systems and machine learning is largely unexplored and the objective of this talk is to show that working in reproducing kernel Hilbert spaces offers tools for a data-based theory of nonlinear dynamical systems.

In this talk, we use the method of parametric and nonparametric kernel flows to predict some chaotic dynamical systems. When trained on geophysical observational data, for example, the weekly averaged global sea-surface temperature, considerable gains are also observed by the proposed technique in comparison to classical partial differential equation-based models in terms of forecast computational cost and accuracy. When trained on publicly available re-analysis data for the daily temperature of the North-American continent, we see significant improvements over classical baselines such as climatology and persistence-based forecast techniques. Although our experiments concern specific examples, the proposed approach is general, and our results support the viability of kernel methods (with learned kernels) for interpretable and computationally efficient geophysical forecasting for a large diversity of processes.

We then show how kernel methods can be used to approximate center manifolds, propose a data-based version of the centre manifold theorem and construct Lyapunov functions for nonlinear ODEs. We also 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 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.

This is joint work with Jake Bouvrie (MIT, USA), Matthieu Darcy

(Caltech), Edward DeBrouwer (KU Leuven), Peter Giesl (University of Sussex, UK), Christian Kuehn (TUM, Munich/Germany), Jonghyeon Lee (Caltech), Romit Malik (ANNL), Sameh Mohamed (SUTD, Singapore), Houman Owhadi (Caltech), Martin Rasmussen (Imperial College London), Kevin Webster (Imperial College London), Bernard Hasasdonk and Dominik Wittwar (University of Stuttgart), Gabriele Santin (Fondazione Bruno Kessler).