

**International Conference:
Differential Equations for Data Science 2024
(DEDS2024)**

Monday, February 19 — Wednesday, February 21, 2024

Online (Zoom)

<https://scheme.hn/deds2024/>

Organizers:

Hayato Chiba (Tohoku University, JP)

Thomas de Jong (Kanazawa University, JP)

Yoshikazu Giga (The University of Tokyo, JP)

Lyudmila Grigoryeva (University of St. Gallen, CH)

Boumediene Hamzi (California Institute of Technology, US)

Masato Kimura (Kanazawa University, JP)

Hiroshi Kokubu (Kyoto University, JP)

Kohei Nakajima (The University of Tokyo, JP)

Hirofumi Notsu (Kanazawa University, JP, Chair)

Juan-Pablo Ortega (Nanyang Technological University, SG)

Julius Fergy Rabago (Kanazawa University, JP)

Registration:

Registration link for Zoom meeting (shortened): <http://tinyurl.com/dj23s9fe>

After registering, you will receive a confirmation email containing information about joining the meeting. (Once you register using the link above, you will be 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.

Supports:

MIRS, Kanazawa University and JST CREST, JPMJCR2014

Program

* All lectures are given by invited speakers.

Monday, February 19

16:55 – 17:00 Opening

Session 1 (5 talks) Chair: H. Notsu: JST 17:00 – 20:05

(= UTC 08:00 – 11:05 = CET 09:00 – 12:05 = PST 00:00 – 03:05)

17:00 – 17:50 Enrique Zuazua (University of Erlangen-Nuremberg, DE)

Control and Machine Learning

17:50 – 18:15 Romit Maulik (Pennsylvania State University, US)

Turbulence modeling for large-eddy simulations using neural differential equations

18:15 – 18:40 Marius Zeinhofer (Simula Research Laboratory, NO)

Error analysis for physics-informed neural networks

18:40 – 18:50 Break

18:50 – 19:15 Cristopher Salvi (Imperial College London, UK)

Scaling limits of random recurrent-residual neural networks

19:15 – 20:05 Pantelis R. Vlachas (ETH Zurich, CH)

Adaptive online learning of effective dynamics for complex systems across scales

Tuesday, February 20

Session 2 (5 talks) Chair: TBD: JST 17:00 – 20:05

(= UTC 08:00 – 11:05 = CET 09:00 – 12:05 = PST 00:00 – 03:05)

17:00 – 17:50 Boumediene Hamzi (California Institute of Technology, US)

Bridging Machine Learning, Dynamical Systems, and Algorithmic Information Theory: Insights from Sparse Kernel Flows and PDE Simplification

17:50 – 18:15 Guglielmo Gattiglio (University of Warwick, UK)

Nearest Neighbor GParareal: Improving Scalability of Gaussian Processes for Parallel-in-Time Solvers

18:15 – 18:40 Jianyu Hu (Nanyang Technological University, SG)

A structure-preserving kernel method for the learning of Hamiltonian systems

18:40 – 18:50 Break

18:50 – 19:15 James Louw (Nanyang Technological University, SG)

Error bounds for forecasting causal dynamics with universal reservoirs

19:15 – 20:05 Massimiliano Tamborrino (University of Warwick, UK)

Network inference in a stochastic multi-population neural mass model via approximate Bayesian computation

Wednesday, February 21

Session 3 (6 talks) Chair: TBD: Session 3: JST 17:00 – 20:05

(= UTC 08:00 – 11:05 = CET 09:00 – 12:05 = PST 00:00 – 03:05)

17:00 – 17:25 Yasamin Jalalian (California Institute of Technology, US)

Data-efficient kernel methods for PDE Identification

17:25 – 17:50 Theo Bourdais (California Institute of Technology, US)

Computational Hypergraph Discovery for the data-driven recovery of differential equations

17:50 – 18:15 Matthieu Darcy (California Institute of Technology, US)

One-shot learning of stochastic differential equations with Gaussian processes

18:15 – 18:40 Samuel Mercer (Delft University of Technology, NL)

Discrete to continuum: total variation flow

18:40 – 18:50 Break

18:50 – 19:15 Masato Hara (Kyoto University, JP)

A reservoir computing method for dynamical systems on general differentiable manifolds

19:15 – 20:05 Alessandro Corbetta (Eindhoven University of Technology, NL)

Machine learning turbulent cascades: inference and closure

20:05 – 20:10 Closing

Abstract

T. = Title, A. = Abstract.

1. Enrique Zuazua (University of Erlangen-Nuremberg, DE)

T. Control and Machine Learning

A. In this lecture we shall present some recent results on the interplay between control and Machine Learning, and more precisely, Supervised Learning and Universal Approximation.

We adopt the perspective of the simultaneous or ensemble control of systems of Residual Neural Networks (ResNets). Roughly, each item to be classified corresponds to a different initial datum for the Cauchy problem of the ResNets, leading to an ensemble of solutions to be driven to the corresponding targets, associated to the labels, by means of the same control.

We present a genuinely nonlinear and constructive method, allowing to show that such an ambitious goal can be achieved, estimating the complexity of the control strategies.

This property is rarely fulfilled by the classical dynamical systems in Mechanics and the very nonlinear nature of the activation function governing the ResNet dynamics plays a determinant role. It allows deforming half of the phase space while the other half remains invariant, a property that classical models in mechanics do not fulfill.

This viewpoint opens up interesting perspectives to develop new hybrid mechanics-data driven modelling methodologies.

This lecture is inspired in joint work, among others, with Borjan Geshkovski (MIT), Carlos Esteve (Cambridge), Domenec Ruiz-Balet (IC, London) and Dario Pighin (Sherpa.ai).

2. Romit Maulik (Pennsylvania State University, US)

T. Turbulence modeling for large-eddy simulations using neural differential equations

A. Differentiable fluid simulators are increasingly demonstrating value as useful tools for developing data-driven models in computational fluid dynamics (CFD). Differentiable turbulence, or the end-to-end training of machine learning (ML) models embedded in CFD solution algorithms, captures both the generalization power and limited upfront cost of physics-based simulations, and the flexibility and automated training of deep learning meth-

ods. We develop a framework for integrating deep learning models into a generic finite element numerical scheme for solving the Navier-Stokes equations, applying the technique to learn a sub-grid scale closure using a multi-scale graph neural network. We demonstrate the method on several realizations of flow over a backwards-facing step, testing on both unseen Reynolds numbers and new geometry. We show that the learned closure can achieve accuracy comparable to traditional large eddy simulation on a finer grid that amounts to an equivalent speedup of 10x. As the desire and need for cheaper CFD simulations grows, we see hybrid physics-ML methods as a path forward to be exploited in the near future.

3. Marius Zeinhofer (Simula Research Laboratory, NO)

T. Error analysis for physics-informed neural networks

A. In this talk, we discuss error estimates for physics-informed neural networks (PINNs) for a wide range of linear PDEs, including elliptic, parabolic and hyperbolic equations. For the analysis, we propose an abstract framework in the language of bilinear forms, and we show the required continuity and coercivity estimates for the mentioned equations. Our results illustrate that the L2 penalty approach that is commonly employed for boundary and initial conditions provably leads to a pronounced deterioration in convergence mode.

4. Cristopher Salvi (Imperial College London, UK)

T. Scaling limits of random recurrent-residual neural networks

A. I will present some scaling limit results for random recurrent and residual neural networks when width and depth tend to infinity. When the activation function is the identity, I will show that the limiting object is a Gaussian measure on some space of paths and its covariance agrees with the so-called signature kernel.

5. Pantelis R. Vlachas (ETH Zurich, CH)

T. Adaptive online learning of effective dynamics for complex systems across scales

A. Predictive simulations are crucial in various applications, including weather forecasting, material design, and understanding complex dynamic systems. The effectiveness of these simulations largely depends on their ability to

accurately model and predict the dynamics of the systems they represent. Traditional approaches to simulation face challenges: high-fidelity, massively parallel simulations, while detailed and accurate, are computationally expensive and limit the scope for experimentation. On the other hand, reduced-order models, which are computationally less demanding, often oversimplify the dynamics through linearization and heuristic closures, compromising accuracy.

This presentation explores how advancements in machine learning (ML) technologies, particularly Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Mixture Density Networks (MDNs), can overcome these limitations. These ML models, enhanced with novel training strategies, can forecast the high-dimensional dynamics of chaotic systems, identify and propagate reduced-order latent dynamics over time with minimal loss in accuracy, and capture complex stochastic behaviors in molecular dynamics. The efficacy of these data-driven approaches is demonstrated through standard benchmarks, including the Kuramoto-Sivashinsky equation for chaotic systems, the Lorenz-96 system for atmospheric dynamics, and Alanine Dipeptide for molecular dynamics simulations, showcasing their potential as predictive tools.

Further, the presentation introduces a pioneering systematic framework termed Adaptive Learning of Effective Dynamics (AdaLED), which builds on the Equation-Free paradigm. AdaLED bridges the gap between detailed large-scale simulations and reduced-order models by adaptively extracting and forecasting the effective dynamics of multiscale systems. It employs autoencoders for dimensionality reduction and an ensemble of probabilistic RNNs for time-stepping, allowing for an efficient alternation between computational simulations and surrogate modeling. This process accelerates the simulation of known dynamics and facilitates the exploration of new dynamic regimes through continuous online adaptation. AdaLED's performance is validated on diverse systems, including the Van der Pol oscillator, 2D reaction-diffusion equations, and 2D Navier-Stokes flow, demonstrating its ability to dynamically learn and adjust to new conditions, thus offering a significant advantage for applications that require numerous complex simulations. This novel framework represents a significant leap forward in computational dynamics, providing a versatile and powerful tool for predictive modeling.

6. Boumediene Hamzi (California Institute of Technology, US)

- T. Bridging Machine Learning, Dynamical Systems, and Algorithmic Information Theory: Insights from Sparse Kernel Flows and PDE Simplification
- A. This presentation delves into the intersection of Machine Learning, Dynamical Systems, and Algorithmic Information Theory (AIT), exploring the connections between these areas. In the first part, we focus on Machine Learning and the problem of learning kernels from data using Sparse Kernel Flows. We draw parallels between Minimum Description Length (MDL) and Regularization in Machine Learning (RML), showcasing that the method of Sparse Kernel Flows offers a natural approach to kernel learning. By considering code lengths and complexities rooted in AIT, we demonstrate that data-adaptive kernel learning can be achieved through the MDL principle, bypassing the need for cross-validation as a statistical method.

Transitioning to the second part of the presentation, we shift our attention to the task of simplifying Partial Differential Equations (PDEs) using kernel methods. Here, we utilize kernel methods to learn the Cole-Hopf transformation, transforming the Burgers equation into the heat equation. We argue that PDE simplification can also be seen as an MDL and a compression problem, aiming to make complex PDEs more tractable for analysis and solution. While these two segments may initially seem distinct, they collectively exemplify the multifaceted nature of research at the intersection of Machine Learning, Dynamical Systems, and AIT, offering preliminary insights into the synergies that arise when these fields converge.

7. Guglielmo Gattiglio (University of Warwick, UK)

- T. Nearest Neighbor GParareal: Improving Scalability of Gaussian Processes for Parallel-in-Time Solvers
- A. With the advent of supercomputers, multi-processor environments and parallel-in-time (PiT) algorithms provide ways to integrate ordinary differential equations (ODEs) over long time intervals, a task often unfeasible with sequential time-stepping solvers within realistic timeframes. A recent approach, GParareal, combines machine learning (Gaussian Processes) with traditional PiT methodology (Parareal) to achieve faster parallel speed-ups. Unfortunately, the applicability of the model is limited to a small number of computer cores and ODE dimensions. We present Nearest Neighbor GParareal (NN-GParareal), a data-enriched parallel-in-

time integration algorithm that builds upon GParareal by improving its scalability properties for higher dimensional systems and increased processor count. Through data reduction, the model complexity is reduced from cubic in the sample size to loglinear, yielding a fast, automated procedure to integrate initial value problems over long intervals. The practical utility of NN-GParareal is demonstrated theoretically and empirically through its evaluation on nine different systems. Our analysis offers tangible evidence of NN-GParareal’s behavior, advantages, and validity.

8. Jianyu Hu (Nanyang Technological University, SG)

- T. A structure-preserving kernel method for the learning of Hamiltonian systems
- A. In this talk, we present a structure-preserving kernel method for the learning of Hamiltonian systems. In the presentation, we shall start by establishing reproducing properties of differentiable kernels on any subsets of R^{2d} , which enables us to embed the corresponding RKHS into the space of bounded differentiable functions with bounded derivatives. We then study the Hamiltonian learning problem using a kernel ridge regression, we provide an operator-theoretical framework to represent the structure-preserving kernel estimators, and we prove convergence results and error bounds for them. Finally, we present some numerical experiments.

9. James Louw (Nanyang Technological University, SG)

- T. Error bounds for forecasting causal dynamics with universal reservoirs
- A. For a few decades state-space systems have been used in the learning and prediction of input-output systems. In particular, the recent emergence of reservoir computing as a highly competitive learning strategy with numerous applications has motivated study in this area. While much work has been done in establishing universality properties of state-space systems in approximating input-output systems, and learnability of dynamical systems via embedding properties, not much research exists in establishing the accuracy of predictions implemented via this learning strategy. To this end we present our work, establishing bounds for the prediction error as a function of the forecasting horizon when learning input-output systems in the class of causal chains with infinite memory using reservoir computers. Causal chains include time series coming from the observations of a large class of dynamical systems and many other applications, such as

finite-dimensional observations from functional differential equations and the deterministic parts of stochastic processes. In our work we illustrate how the theory of nonuniform hyperbolicity and Lyapunov exponents plays a vital role in the rate at which estimation accuracy deteriorates. Most notably, the Multiplicative Ergodic Theorem of Oseledets is the cornerstone of the results, revealing the underlying structures and limitations of this prediction strategy.

10. Massimiliano Tamborrino (University of Warwick, UK)

- T. Network inference in a stochastic multi-population neural mass model via approximate Bayesian computation
- A. The aim of this work is to infer the connectivity structures of brain regions before and during epileptic seizure. Our contributions are fourfold. First, we propose a $6N$ -dimensional stochastic differential equation for modelling the activity of N coupled populations of neurons in the brain. This model further develops the (single population) stochastic Jansen and Rit neural mass model, which describes human electroencephalography (EEG) rhythms, in particular signals with epileptic activity. Second, we construct a reliable and efficient numerical scheme for the model simulation, extending a splitting procedure proposed for one neural population. Third, we propose an adapted Sequential Monte Carlo Approximate Bayesian Computation algorithm for simulation-based inference of both the relevant real-valued model parameters as well as the $0,1$ -valued network parameters, the latter describing the coupling directions among the N modelled neural populations. Fourth, after illustrating and validating the proposed statistical approach on different types of simulated data, we apply it to a set of multi-channel EEG data recorded before and during an epileptic seizure. The real data experiments suggest, for example, a larger activation in each neural population and a stronger connectivity on the left brain hemisphere during seizure.

11. Yasamin Jalalian (California Institute of Technology, US)

- T. Data-efficient kernel methods for PDE Identification
- A. For many problems in computational sciences and engineering, observational data exists for which the underlying physical models are not known. PDE identification methods provide systematic ways to infer these physical models directly from data. We introduce a framework for identifying and

solving PDEs using kernel methods. In particular, given observations of PDE solutions and source terms, we employ a kernel-based data-driven approach to learn the functional form of the underlying equation. We prove convergence guarantees and a priori error estimates for our methodology. Through numerical experiments, we demonstrate that our approach is particularly competitive in the data-poor regime where few observations are available.

12. Theo Bourdais (California Institute of Technology, US)

- T. Computational Hypergraph Discovery for the data-driven recovery of differential equations
- A. Most scientific challenges can be framed into one of the following three levels of complexity of function approximation. Type 1: Approximate an unknown function given input/output data. Type 2: Consider a collection of variables and functions, some of which are unknown, indexed by the nodes and hyperedges of a hypergraph (a generalized graph where edges can connect more than two vertices). Given partial observations of the variables of the hypergraph (satisfying the functional dependencies imposed by its structure), approximate all the unobserved variables and unknown functions. Type 3: Expanding on Type 2, if the hypergraph structure itself is unknown, use partial observations of the variables of the hypergraph to discover its structure and approximate its unknown functions. While most Computational Science and Engineering and Scientific Machine Learning challenges can be framed as Type 1 and Type 2 problems, many scientific problems can only be categorized as Type 3. Despite their prevalence, these Type 3 challenges have been largely overlooked due to their inherent complexity. Although Gaussian Process (GP) methods are sometimes perceived as well-founded but old technology limited to Type 1 curve fitting, their scope has recently been expanded to Type 2 problems.

In this talk, we introduce an interpretable GP framework for Type 3 problems, targeting the data-driven discovery and completion of computational hypergraphs. Our approach is based on a kernel generalization of (1) Row Echelon Form reduction from linear systems to nonlinear ones and (2) variance-based analysis. Here, variables are linked via GPs, and those contributing to the highest data variance unveil the hypergraph's structure. We illustrate the scope and efficiency of the proposed approach with applications to differential equations discovery.

13. Matthieu Darcy (California Institute of Technology, US)

- T. One-shot learning of stochastic differential equations with Gaussian processes
- A. We consider the problem of learning the drift f and diffusion of stochastic differential equations (SDE) of the form $dX_t = f(X_t) dt + \sigma(X_t) dW_t$ from one sample trajectory. These types of equations are widely used in areas like finance or the geophysical and planetary sciences to model stochastic dynamics. This problem is more challenging than learning deterministic dynamical systems because one sample trajectory only provides indirect information on the unknown functions f and σ . We propose a method that places a Gaussian process prior on the unknown functions and computes their maximum a posteriori (MAP) estimator given the data. We also leverage efficient methods to learn the kernel (or covariance) functions of the Gaussian processes from the data with cross-validation or maximum likelihood estimation (MLE).

Our approach not only allows us to predict future dynamics but also provides an uncertainty quantification of such prediction. We illustrate the efficacy of our method through numerical experiments and an application to the prediction of laboratory earthquakes.

14. Samuel Mercer (Delft University of Technology, NL)

- T. Discrete to continuum: total variation flow
- A. In this talk we will present some results on discrete to continuum limits for Cauchy problems on a sequence of Banach spaces. In particular by investigating the structure of a discrete to continuum limit and using this to motivate a general framework we call Banach stacking. Our work is motivated by recent developments using the $TL^p(\Omega)$ metric for Γ -convergence results, inspired further from the theory of optimal transport.

We then apply these results to deduce uniform convergence of total variation flow along $TL^1(\Omega)$ from discrete to continuum.

15. Masato Hara (Kyoto University, JP)

- T. A reservoir computing method for dynamical systems on general differentiable manifolds
- A. Reservoir computing is a kind of machine learning method which can learn and reproduce various information on nonlinear dynamics. Researchers

have been trying to reveal its mechanism of learning dynamics. In the theoretical study of reservoir computing, it is natural and helpful to impose some “good” properties such as structural stability or ergodicity for target systems for learning. Taking into account that reservoir computing is usually defined on Euclidean spaces, however, those good properties are not satisfied by typical chaotic systems such as the Henon map and the Lorenz system, as these are known to be not structurally stable in the usual sense. On the other hand, typical dynamical systems that are structurally stable are often defined on a closed manifold, such as the torus or the sphere. We therefore would like to formulate the scheme of reservoir computing that allow the target dynamical systems to be defined on manifolds. In this talk, I will discuss the formulation of such a reservoir computing method and show some numerical examples including a hyperbolic toral automorphism.

16. Alessandro Corbetta (Eindhoven University of Technology, NL)

- T. Machine learning turbulent cascades: inference and closure
- A. Turbulence, the ubiquitous and chaotic state of fluid motions, is characterized by strong, multiscale, and statistically nontrivial fluctuations of the velocity field. This has opened longstanding fundamental challenges with vast technological relevance. For instance, turbulent fluctuations hinder convergence of statistical estimators, making even the bare quantification of the turbulence intensity or of the Reynolds number highly demanding in terms of data volumes. Also, high-statistical fidelity closure models, parametrizing the influence of small unresolved scales on the dynamics of large, resolved ones, remain outstanding. In this talk, I will discuss the capability of recent deep neural models at learning features of turbulent velocity signals. First, I will show how deep neural networks can accurately estimate the Reynolds number within 15% accuracy, from a statistical sample as small as two large-scale eddy turnover times. In contrast, physics-based statistical estimators are limited by the convergence rate of the central limit theorem and provide, for the same statistical sample, at least a hundredfold larger error. Second, I will present a closure, based on a deep recurrent network, that quantitatively reproduces, within statistical errors, Eulerian and Lagrangian structure functions and the intermittent statistics of the energy cascade, including those of subgrid fluxes. To achieve high-order statistical accuracy, and thus a stringent statistical test, I consider shell models of turbulence. These results encourage

the development of similar approaches for three-dimensional Navier-Stokes turbulence.

In collaboration with R. Benzi, V. Menkovski, G. Ortali, G. Rozza, F. Toschi.

Refs.

- G. Ortali, A. Corbetta, G. Rozza, F. Toschi. Numerical proof of shell model turbulence closure. *Phys. Rev. Fluids*. 7, L082401, 2022.
- A. Corbetta, V. Menkovski, R. Benzi, F. Toschi. Deep learning velocity signals allows to quantify turbulence intensity. *Sci. Adv.* 7: eaba7281, 2021.