Poster session
Workshop on Computational and Mathematical Methods in Data Science 2024

- **Junaid Akhter**
  Pitfalls to avoid while using multiobjective optimization for machine learning
- **Amaury Bélières Frendo**
  Geometric Shape Optimization for Dirichlet Energy With Physics Informed and Symplectic Neural Networks
- **Roberto Bentivoglio**
  Multi-scale hydraulic-based graph neural networks: generalizing rapid flood mapping to irregular meshes and time-varying boundary condition
- **Jan Blechschmidt**
  A physics-informed DeepONet model for the solution of quantum graphs
- **Piero Deidda**
  Nonlinear joint spectral radius of cone order preserving functions
- **Antoine Lechevallier**
  Hybrid Newton method for the acceleration of well events handling in the numerical simulation of CO_2 Storage
- **Erion Morina**
  On the growth of parameters of approximating neural networks
- **Moreno Pintore**
  An introduction to the lowest-order Neural Approximated Virtual Element Method
- **Yong Shang**
  Randomized Neural Networks with Petrov–Galerkin Methods for Solving Linear Elasticity problem
- **Corné Verburg**
  High-resolution image segmentation with U-Net-based segmentation CNN on multiple GPUs

Pitfalls to avoid while using multiobjective optimization for machine learning

Junaid Akhter[1]

[1] University of Paderborn
junaid.akhter@uni-paderborn.de

Recently, there has been an increasing interest in exploring the application of multiobjective optimization (MOO) in machine learning (ML), the reason being the numerous situations in real-life applications where multiple objectives need to be optimized simultaneously. A key aspect of MOO is the existence of a Pareto set rather than a single optimal solution, illustrating the inherent trade-offs between objectives. Despite its potential, there is a noticeable lack of satisfactory literature that could serve as an entry-level guide for ML practitioners who want to use MOO. Hence, our goal in this paper is to produce such a resource. We critically review previous studies, particularly those involving MOO in deep learning (using Physics-Informed Neural Networks (PINNs) as a guiding example), and identify misconceptions that highlight the need for a better grasp of MOO principles in ML. Using MOO of PINNs as a case study, we demonstrate the interplay between the *data loss* and the *physics loss* terms. We highlight the most common pitfalls one should avoid while using MOO techniques in ML. We begin by establishing the groundwork for MOO, focusing on well-known approaches such as the weighted sum (WS) method, alongside more complex techniques like the multiobjective gradient descent algorithm (MGDA). We emphasize the importance of understanding the specific problem, the objective space, and the selected MOO method, while also noting that neglecting factors such as convergence can result in inaccurate outcomes and, consequently, a non-optimal solution. Our goal is to offer a clear and practical guide for ML practitioners to effectively apply MOO, particularly in the context of DL.

Geometric Shape Optimization for Dirichlet Energy With Physics Informed and Symplectic Neural Networks

Amaury Bélières Frendo[1]

[1] Université de Strasbourg - IRMA
belieres@unistra.fr

In this work, we are interested in optimizing the boundary of a given domain under a volume constraint, with respect to the solution of a partial differential equation. We focus on the numerical aspect of this question, and we propose to apply recent approaches based on neural networks. For simplicity, the problem under consideration is the minimization of the Dirichlet energy for the Poisson equation in R2. We first quickly recall results on the mathematical analysis of this problem. We select physics-informed neural networks to approximate the solution of the Poisson equation in a given shape. To represent the shape with a neural network, we parametrize a volume-preserving transformation from an initial shape to an optimal one. Both processes are combined in a single optimization algorithm, which only relies on minimizing one physical loss function, the Dirichlet energy. We conclude with the presentation of the open source code and its numerical validation.

Joint work with: Emmanuel Franck, Victor Michel Dansac, Yannick Privat

Multi-scale hydraulic-based graph neural networks: generalizing rapid flood mapping to irregular meshes and time-varying boundary condition

Roberto Bentivoglio[1]

[1] TU Delft
r.bentivoglio@tudelft.nl

Deep learning models emerged as viable alternatives to rapid and accurate flood mapping, overcoming the computational burden of numerical methods. In particular, hydraulic-based graph neural networks present a promising avenue, offering enhanced transferability to domains not used for the model training. These models exploit the analogy between finite-volume methods and graph neural networks to describe how water moves in space and time across neighbouring cells. However, existing models face limitations, having been exclusively tested on regular meshes and necessitating initial conditions from numerical solvers. This study proposes an extension of hydraulic-based graph neural networks to accommodate time-varying boundary conditions, showcasing its efficacy on irregular meshes. For this, we employ multi-scale methods that jointly model the flood at different scales. To remove the necessity of initial conditions, we leverage ghost cells that enforce the solutions at the boundaries. Our approach is validated on a dataset featuring irregular meshes, diverse topographies, and varying input hydrograph discharges. Results highlight the model's capacity to replicate flood dynamics across unseen scenarios, without any input from the numerical model, emphasizing its potential for realistic case studies.

A physics-informed DeepONet model for the solution of quantum graphs

Jan Blechschmidt[1]

[1] TU Chemnitz
jan.blechschmidt@mathematik.tu-chemnitz.de

In this talk we focus on a machine learning approach for quantum graphs, i.e. metric graphs with an associated differential operator. In our case the differential equation is a drift-diffusion model. Computational methods for quantum graphs require a careful discretization of the differential operator that also incorporates the node conditions, in our case Kirchhoff-Neumann conditions. Traditional numerical schemes are rather mature but have to be tailored manually when the differential equation becomes the constraint in an optimization problem. Recently, physics informed deep operator networks (DeepONets) have emerged as a versatile tool for the solution of partial differential equations from a range of applications. We train physics-informed DeepONet models on a simple reference graph and show how to combine them for the solution of quantum graphs.

Joint work with: Jan-Frederik Pietschmann, Tom-Christian Riemer, Martin Stoll, Max Winkler

Nonlinear joint spectral radius of cone order preserving functions

Piero Deidda[1]

[1] Scuola Normale Superiore - Gran Sasso Science Institute piero.deidda@sns.it

A Neural Network can be seen as a discrete switched system that alternates maps from a class of nonlinear functions following a switching rule that is not known a-piori, but is determined by the training. In particular, stability properties of the NN can be studied in terms of the stability of the corresponding nonlinear switched system. Motivated by applications in machine learning, we study the stability of nonlinear switched systems that alternate homogeneous nonlinear functions that preserve the ordering induced by a cone. Such maps, admitting the notion of spectral radius, allow us to generalize, from the linear to the nonlinear case, the study of the joint spectral radius (JSR) of the system. In particular, we first prove that the value of the JSR yields information about the stability of the system. Hence, we investigate the properties of the nonlinear JSR, tracing analogies and differences from the linear case. Finally, we present an algorithm devoted to computing the nonlinear JSR.

Joint work with: Nicola Guglielmi, Francesco Tudisco.

Hybrid Newton method for the acceleration of well events handling in the numerical simulation of CO_2 Storage

Antoine Lechevallier[1]

[1] NORCE
anle@norceresearch.no

Numerical simulations are crucial for solving multi-phase flow equations in CO2 injection sites. However, simulating fluid flow in porous media is highly demanding computationally, often requiring several hours on an HPC cluster for a single injection scenario in a large CO2 reservoir. Well events, like opening and closure, pose significant challenges due to their sudden impact on the system, necessitating a drastic reduction in time step size to solve resulting nonlinear equations accurately. Yet, these events tend to exhibit spatial and temporal similarities, determined by factors such as injection conditions, reservoir state, boundary conditions, and porous media parameters.

This thesis work aims to employ recent advancements in physics-informed deep learning to mitigate the impact of well events in numerical simulations of multiphase flow in porous media. In practice, we suggest a hybrid method that complement the conventional nonlinear solver with a machine-learning model while maintaining numerical reliability. Our approach involves customizing the hybrid Newton methodology, which predicts a global initialization for Newton's method closer to the solution. We employ the Fourier Neural Operator machine-learning model for this prediction task.

On the growth of parameters of approximating neural networks

Erion Morina[1]

[1] University of Graz
morina.erion@hotmail.com

This talk focuses on the analysis of fully connected feed forward ReLU-neural networks as they approximate a given, smooth function. In contrast to conventionally studied universal approximation properties under increasing architectures, e.g. in terms of width or depth of the networks, we are concerned with the asymptotic growth of the parameters of approximating networks. Such results are of interest, e.g., for computing generalization errors or for proving consistency results for neural network training. The main result of our work is that, for a ReLU architecture which is known to achieve an optimal approximation error, the realizing parameters grow at most polynomially. The obtained rate with respect to normalized network size is compared to existing results and shown to be superior in most cases, in particular for high dimensional input.

Joint work with: Martin Holler (University of Graz).

An introduction to the lowest-order Neural Approximated Virtual Element Method

Moreno Pintore[1]

[1] Inria, Sorbonne University
moreno.pintore@inria.fr

Thanks to the rapidly growing interest in Scientific Machine Learning, numerous numerical methods relying on deep neural networks have been proposed by the scientific community in the last few years. One of such methods is the Neural Approximated Virtual Element Method (NAVEM), a polygonal method to solve partial differential equations (PDEs) recently introduced in [1]. The method combines the nonlinear properties and the efficiency of a neural network with the flexibility and the accuracy of the more classical Virtual Element Method (VEM) [2].

The VEM is a polygonal method relying on virtual functions, i.e. functions that are not known in a closed form. Therefore, suitable polynomial projectors and stabilization terms are necessary to evaluate the differential operators characterizing the PDE and to retrieve the coercivity of the discrete problem. The NAVEM replaces such unknown virtual basis functions by suitable neural network-based approximations with similar functional properties. This way, it is possible to define a new polygonal method without any projection or stabilization, which are problem-dependent and may limit the model accuracy in case of strongly anisotropic PDEs.

In order to approximate the local VEM basis functions, we employ linear combinations of suitable basis functions, where the coefficients are predicted by a deep neural network. This way, it is also possible to reduce spurious oscillations and discontinuities in the local and global basis functions respectively, and improve the NAVEM stability as a consequence. Two-dimensional numerical tests on distorted quadrilateral meshes and on Voronoi meshes are proposed to validate the method accuracy and flexibility.

Joint work with: Stefano Berrone, Oberto Davide, and Teora Gioana (Politecnico di Torino).

References

[1] S. Berrone, D. Oberto, M. Pintore, and G. Teora. The lowest-order Neural Approximated Virtual Element Method. ArXiv preprint arXiv:2311.18534 , (2023).

[2] L. Beirao Da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. D. Marini, and A. Russo. Basic principles of virtual element methods. Math. Models Methods Appl. Sci. 23, 199–214, (2013).

Randomized Neural Networks with Petrov–Galerkin Methods for Solving Linear Elasticity problem

Yong Shang[1]

[1] ETH Zurich
yong.shang@sam.math.ethz.ch

We develop the Randomized Neural Networks with Petrov-Galerkin Methods (RNN-PG methods), which use the Petrov-Galerkin variational framework, where the solution is approximated by randomized neural networks and the test functions can be employed in a flexible choice. Unlike conventional neural networks, the parameters of the hidden layers of the randomized neural networks are fixed randomly, while the parameters of the output layer are determined by the least squares method, which can effectively approximate the solution. We also develop mixed RNN-PG methods for linear elasticity problems, which ensure the symmetry of the stress tensor and avoid locking effects. We compare RNN-PG methods with different methods on several examples, and the numerical results demonstrate that RNN-PG methods achieve higher accuracy and efficiency.

High-resolution image segmentation with U-Net-based segmentation CNN on multiple GPUs

Corné Verburg[1]

[1] Delft University of Technology
c.verburg@student.tudelft.nl

We introduce a novel approach for segmenting high-resolution images across multiple GPUs, by combining the U-Net architecture with domain decomposition strategies. Our approach partitions high-resolution input images into non-overlapping patches, each stored and processed independently on different GPUs. A communication network facilitates global communication across subdomains, using deep, encoded feature maps, leading to minimal memory overhead.

Our method makes it possible to process high-resolution images across multiple GPUs without losing detailed contextual information or sacrificing global information from other patches. In contrast, the memory overhead due to communication is minimal. Extensive evaluation across diverse datasets, including synthetic data, Inria Aerial Images, and DeepGlobe Satellite Segmentation Dataset, demonstrates good performance compared to the baseline U-Net, particularly in consistently accurate class predictions along boundaries.