Multi-Fidelity Neural Network Surrogate Modeling for Large-Scale Bayesian Inverse Problems with Applications to Inverse Acoustic Scattering by Random Domains

Bioli Ivan[1], Dölz Jürgen[2] and Henríquez Fernando[3]

[1]École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland
[2]Institute for Numerical Simulation, University of Bonn, Friedrich-Hirzebruch-Allee 7, 53115 Bonn, Germany
[3]MCSS, Chair of Computational Mathematics and Simulation Science, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland

The computational approximation of large-scale Bayesian Inverse Problems (BIPs), i.e. characterized by a large number of available observations, involves utilizing techniques such as Markov Chain Monte Carlo (MCMC) to generate samples from the posterior distribution of the BIP. The primary obstacle in this approach arises from the necessity of repeatedly evaluating the high-fidelity forward model. This becomes cost-prohibitive for complex computational models, such as parametric PDEs with high-dimensional inputs, prompting the need for effective computational surrogates to facilitate MCMC execution.

In this work, we introduce a Multi-fidelity Neural Network (MF-NN) surrogate construction for Bayesian Inverse Problems (BIPs), inspired by the approach presented in [3]. Initially, an offline step involves constructing a low-fidelity prior-based surrogate solver using a neural network (NN) following the Galerkin-POD-NN technique [2]. Subsequently, a local online refinement process utilizes a NN to improve the low-fidelity ap- proximation by incorporating newly computed high-fidelity data. Unlike existing methods, the MF-NN approach enables an accurate approximation of the posterior distribution of the BIP by adaptively refining the model only where needed, specifically in regions where the posterior concentration phenomenon occurs. By focusing the computational effort only where necessary, the number of high-fidelity model evaluations is significantly reduced.

As a concrete application, we consider the time-harmonic acoustic scattering by random domains, as in [1]. Numerical experiments validate the effectiveness, efficiency and accuracy of the AMF-NN approach, especially in challenging settings for BIPs, including both the low noise and large data limits.

REFERENCES

[1] Dölz, J., Harbrecht, H., Jerez-Hanckes, C., and Multerer, M. (2022). Isogeometric multilevel quadrature for forward and inverse random acoustic scattering. Computer Methods in Applied Mechanics and Engineering 388, p. 114242.

[2] Hesthaven, J. S., and Ubbiali, S. (2018). Non-intrusive reduced order modeling of non-linear problems using neural networks. Journal of Computational Physics 363, pp. 55-78.

[3] Li, Y., Wang, Y., and Yan, L. (2023). Surrogate modeling for Bayesian inverse problems based on physics-informed neural networks. Journal of Computational Physics 475, p. 111841.

Solving the Bateman Equation using Physics Informed Neural Networks

Boiger Romana

Paul Scherrer Institute: Laboratory for Waste Management

In the context of nuclear energy production, waste management, and nuclear safeguards, the precise calculation and quantification of spent nuclear fuel characteristics are important for safety, efficiency, and sustainability. The abundances and activities of nuclear composites characterizing spent nuclear fuel can be mathematically modeled with the Bateman equation. This is a linear first-order ordinary differential equation that describes the time evolution of the nuclear assemblies based on decay rates. The complexity of solving this equation arises from the stiffness of the matrix and the potentially large number of considered nuclides. Moreover, inherent uncertainties of the input quantities from theory and measurements require reliable uncertainty quantification, especially for safety assessment.

This talk presents a novel approach to solving the Bateman equation using Physics Informed Neural Networks (PINN). Several different versions of PINN with various loss functions and network architectures, including classical PINN, Hard Boundary Method, Extreme Learning Machine Method, Exponential Method, and Domain Decomposition Method were implemented and validated. Their performance was tested on a simplified real use case considering the decay of Plutonium241. The performances of the PINNs were validated against the current state-of-the-art solution method, the Chebyshev Rational Approximation, and showed comparable accuracy.

Furthermore, the best PINN model was used to perform uncertainty quantification, demonstrating how the use of transfer learning can result in competitiveness with state-of-the-art solvers. The talk will present the advantages as well as the limits of the considered PINN methods for solving the Bateman Equation.

Physics-Aware Deep Nonnegative Matrix Factorization

Chinellato Erik

Department of Mathematics, University of Padova

In this talk we present the Physics-Aware Deep Nonnegative Matrix Factorization (PAD-NMF), a neural network stemming from recent "deep-unfolding" techniques with the goal of creating more interpretable networks by unfolding the iterations of traditional algorithmic schemes, here the well-known Nonnegative Matrix Factorization (NMF) algorithm. In general, NMF can be applied to a variety of dictionary learning problems as well as signal detection tasks. It can be unfolded to exploit a neural training and made "Physics-Aware" by embedding in the network's architecture intrinsic properties of the expected output. Indeed, the two- factor iterative paradigm typical of NMF, once unfolded, allows great control on the way the quantities of interest are propagated through the network, e.g. sparsity, smoothness, etc. Its interpretability is then guaranteed by a nonnegativity- preserving back-propagation algorithm. Furthermore, the physics-awareness is introduced by three main aspects: discriminative initialization of the dictionary based on our prior knowledge of the physical system, enforcement of the time- correlation by projecting the forward- and back-propagated factors into a block- Hankel space, and a modified loss-function that shifts the focus away from Wiener filtering in favor of a more accurate separation of the mixture contents.

We will show some applications of PAD-NMF in the context of audio source separation and detection, where the signals to be recognized within general mixtures are transients from impulses applied to a mechanical system.

PINN-PAD: Physics Informed Neural Networks in PADova
22–23 February 2024, Padova, Italy

Computational Paradigms in Scientific Machine Learning

D'Inverno Giuseppe Alessio[1], Brien D.[2], Varbella A.[2] and Gjorgiev B.[2]

[1] SISSA, Trieste
[2]ETH Zurich

The Alternating Current Optimal Power Flow (AC-OPF) is a highly non-linear non-convex optimization problem. The AC-OPF is often used to monitor real-time grid operations and thus the operators demand efficient and accurate algorithms. However, solving this problem is often computationally expensive for large power grids. In this work, we integrate the methods of Physics Informed Neural Networks (PINNs), Graph Neural Networks (GNNs) and Augmented Lagrangian methods in one algorithm to obtain highly accurate solutions. This approach takes advantage of the natural graph description of the power grid and ensures that the underlying physical rules and limitations are respected. We show via benchmark test cases this unsupervised approach provides a near-optimal solution that can be applied to different grids without having to adapt its architecture or hyperparameters.

PINN-PAD: Physics Informed Neural Networks in PADova
22–23 February 2024, Padova, Italy

Physics Informed Neural Networks for an Inverse Problem in Peridynamic Models

Difonzo Fabio V.[1], **Lopez Luciano**[2] **and Pellegrino Sabrina F.**[3]

[1]Istituto per le Applicazioni del Calcolo "Mauro Picone" - Consiglio Nazionale delle Ricerche
[2]Università degli Studi di Bari Aldo Moro
[3]Politecnico di Bari

Deep learning is a powerful tool for solving data driven differential problems and has come out to have successful applications in solving direct and inverse problems described by PDEs, even in presence of integral terms. In this paper, we propose to apply radial basis functions (RBFs) as activation functions in suitably designed Physics Informed Neural Networks (PINNs) to solve the inverse problem of computing the peridynamic kernel in the nonlocal formulation of classical wave equation, resulting in what we call RBF-iPINN. We show that the selection of an RBF is necessary to achieve meaningful solutions, that agree with the physical expectations carried by the data. We support our results with numerical examples and experiments, comparing the solution obtained with the proposed RBF-iPINN to the exact solutions.

Application of Physics-Informed Neural Networks in Nonlinear Systems Identification and Parameter Estimation

Forootani Ali

Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany

Physics-Informed Neural Networks (PINNs) represent a powerful approach for solving inverse problems, including identification and parameter estimation in nonlinear systems. The integration of physics-based knowledge into neural networks allows for more efficient and accurate learning, particularly in scenarios where limited and noisy data are available.

As a first application of the PINN idea, we introduce iNeuralSINDy as a novel approach in nonlinear system identification. SINDy is a widely adopted technique for identifying nonlinear dynamics. Despite numerous efforts, the presence of noisy and limited data continues to pose a significant hurdle to the effectiveness of the SINDy approach. We present a resilient approach for unveiling nonlinear governing equations from data characterized by noise and scarcity. Our strategy involves employing neural networks to assimilate an implicit representation based on measurement data. This representation not only generates outputs in proximity to the measurements but also encapsulates the time evolution of the output within a dynamic system. By harnessing the implicit representation through neural networks, we acquire derivative information, a prerequisite for SINDy, utilizing an automatic differentiation tool. To fortify the robustness of our methodology, we introduce an integral condition on the output of the implicit networks.

In the second application, we introduce the discrete empirical interpolation method (DEIM), specifically the QR-factorization-based variant known as Q-DEIM, as a strategic sampling technique for mitigating the computational complexities and time demands associated with parameter estimation for partial differential equations (PDEs) using PINN. Our methodology involves the judicious pre-selection of spatio-temporal data, thereby constructing a reduced dataset for training a neural network to estimate the coefficients of the underlying PDE governing the data. We establish that our proposed Q-DEIM-based sampling approach not only reduces the required training data for the neural network but also yields a commendable approximation of PDE coefficients in fewer training iterations.

Enhancing Biomechanical Impact Simulations through Physics-Informed Neural Networks

Gupta Karan

Sciences, Technologies and Measurements for Space (STMS)
Center of Studies and Activities for Space "G. Colombo" (CISAS)
University of Padova

Traditional biomechanical impact simulations, often relying on the Finite Element Method (FEM), face limitations in computational efficiency, parameter tuning complexity, and ability to extract data-driven insights. Physics-Informed Neural Networks (PINNs) emerge as a transformative approach to address these challenges and usher in a new era of advanced simulation techniques in computational biomechanics.

PINNs, seamlessly integrating physics principles into neural network architectures, provide a powerful tool for enhancing biomechanical impact simulations. By leveraging neural networks' ability to approximate complex relationships and learn from data, we can achieve significant improvements in several key areas:

- Computational Efficiency: PINNs significantly reduce computational time by approximating intricate biomechanical relationships, eliminating the need for extensive FEM simulations.
- Model Calibration and Optimization: PINNs expedite parameter tuning in biomechanical models, streamlining the calibration process and enhancing model accuracy.
- Data-Driven Insights: Neural networks' pattern recognition capabilities enable us to uncover hidden patterns and relationships within biomechanical datasets, leading to deeper understanding of biological structures and behaviors.
- Real-Time Decision Support: PINNs enable real-time predictions during simulations, enabling rapid decision-making in dynamic scenarios where timely responses are crucial.
- Integration with Finite Element Models: Hybrid approaches combining PINNs with FEM can synergistically enhance simulation accuracy and efficiency.

Wavenumber-robust deep ReLu neural network emulation in acoustic wave scattering

Henríquez Fernando[1] and Schwab Christoph[2]

[1]Chair of Computational Mathematics and Simulation Science, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland.
[2]Seminar for Applied Mathematics, ETH Zürich, 8092 Zürich, Switzerland

We present wavenumber-robust error bounds using deep neural networks to emulate the solution to the time-harmonic, sound-soft acoustic scattering problem in the exterior of a smooth, convex, two-dimensional obstacle.

The starting point of our analysis is the reduction of the scattering problem in the unbounded exterior region to its (bounded) boundary by means of the wavenumber-robust Combined Field Integral Equation (CFIE), yielding a second-kind boundary integral equation posed on the smooth surface Γ of the scatterer. This BIE is well-posed in $L^2(\Gamma)$, with explicit bounds on the continuity and stability constants that depend explicitly on the (non-dimensional) wavenumber κ.

Utilizing well-known wavenumber-explicit asymptotics of the solution to this problem, as introduced in the work of Melrose and Taylor [1], we explore the numerical approximation of the BIE using fully connected, deep feed-forward neural networks (DNNs) with the Rectified Linear Unit (ReLU) as the chosen activation function [2]. The results presented here can be straightforwardly extended to different activation functions such as the hyperbolic tangent or the Rectified Power Unit.

Through a constructive argument, we prove the existence of DNNs affording an ϵ-error in the $L^\infty(\Gamma)$-norm with a fixed and small width and a depth that increases *spectrally* with the accuracy ϵ and polynomially with respect to $\log(\kappa)$. By *spectral accuracy*, we mean that there exists $\alpha > 0$ such that for each $n \in \mathbb{N}$, there exists a constant $C_n > 0$, such that for a prescribed accuracy $\epsilon > 0$, the complexity of the DNN is bounded by $C_n \epsilon^{\frac{\alpha}{n}}$.

REFERENCES

[1] Melrose, R. B., and Taylor, M. E. (1985). Near peak scattering and the corrected Kirchhoff approximation for a convex obstacle. Advances in Mathematics, 55(3), 242-315.

[2] Elbrächter D., Perekrestenko, D., Grohs, P., and Bölcskei, H. (2021). Deep neural network approximation theory. IEEE Transactions on Information Theory, 67(5), 2581-2623.

[3] Domínguez, V., Graham, I. G., and Smyshlyaev, V. P. (2007). A hybrid numerical-asymptotic boundary integral method for high-frequency acoustic scattering. Numerische Mathematik 106(3), 471-510.

[4] Ecevit, F., and Eruslu, H. H. (2019). A Galerkin BEM for high-frequency scattering problems based on frequency-dependent changes of variables. IMA Journal of Numerical Analysis 39(2), 893-923.

Parameter estimation in cardiac biomechanical models based on physics-informed neural networks

Höfler Matthias[1], Regazzoni Francesco[4], Pagani Stefano[4], Karabelas Elias[1,2,3], Quarteroni Alfio[4,5], Plank Gernot[2,3], Haase Gundolf[1,3] and Caforio Federica[1,2,3]

[1]Institute of Mathematics and Scientific Computing, NAWI Graz, University of Graz (Austria)
[2]Gottfried Schatz Research Center: Division of Biophysics, Medical University of Graz (Austria)
[3]BioTechMed-Graz, Austria
[4]MOX, Department of Mathematics, Politecnico di Milano (Italy)
[5]Institute of Mathematics, EPFL (Switzerland) (Professor Emeritus)

Biophysical models of the cardiac function are becoming increasingly popular due to their ability to predict patient outcomes and optimise treatment plans. However, the development and personalisation of these models is computationally expensive and requires extensive calibration, making them difficult to apply in clinical settings. In this presentation we study the application of a novel methodology [1] integrating physics-informed neural networks [2] with high-resolution three-dimensional nonlinear cardiac biomechanical models to reconstruct displacement fields and estimate patient-specific biophysical properties (s.a. passive stiffness and active contractility). The physics of the problem is represented by a mathematical model based on partial differential equations. Additionally, the learning algorithm incorporates displacement and strain data that can be routinely acquired in clinical settings. Various training methodologies are explored, e.g. different sampling strategies and adaptive weighting schemes for the individual loss terms as well as Fourier features. The presentation includes a series of benchmark tests that demonstrate the accuracy, robustness, and promising potential of this method for the precise and efficient determination of patient-specific physical properties in nonlinear biomechanical models.

REFERENCES

[1] Caforio, F., Regazzoni, F, Pagani, S., Karabelas, E., Augustin, C., Haase, G., Plank, G. and Quarteroni, A. (2023). Physics-informed neural network estimation of material properties in soft tissue nonlinear biomechanical models. arXiv preprint arXiv:2312.09787.

[2] Raissi, M., Perdikaris, P., and Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378:686-707.

PINN-PAD: Physics Informed Neural Networks in PADova
22–23 February 2024, Padova, Italy

Gauss-Newton Natural Gradient for Physics-Informed Computational Fluid Dynamics

Jnini Anas

Università degli Studi di Trento

We propose Gauss-Newton Natural Gradient (GNNG), a novel second-order optimization method based on Gauss-Newton's method in function space for the training of physics-informed neural networks (PINN). We give an infinite dimensional differential geometric interpretation of our method, explaining the algorithm's optimization dynamics in function space. More precisely, we have shown that the proposed method follows the function space updates up to an orthogonal projection on the model's tangent space. We demonstrate that – given appropriate integral discretization in the PINN formulation – the proposed method corresponds to the well-known Gauss-Newton method in parameter space. This leads to a matrix-free formulation that allows the applicability of the method to large network sizes.

Numerically, the proposed method demonstrates unprecedented accuracy of PINN solutions for the Navier-Stokes equations, achieving relative L^2 errors up to two orders of magnitude lower than those obtained with standard optimizers like Adam or BFGS, and reaching as low as 10^{-8}.

PINN-PAD: Physics Informed Neural Networks in PADova
22–23 February 2024, Padova, Italy

Brain memory working. Optimal control behavior for improved Hopfield-like models

Cardin Franco, Lovison Alberto, Maritan Amos and Megighian Aram

University of Padova

The celebrated 1982 Hopfield Neural Network is a dynamic brain model with symmetric and constant synaptic weights. It has been recognized, recently even by Hopfield itself, the necessity of developing more realistic representations incorporating the asymmetry. Indeed, symmetry appears a too restrictive simplification. Dynamically, it is forcing –in an unrealistic way– the recognition of every incoming pattern, always going towards a minimum or a saddle of an underlying Lyapunov-like energy function, always existing in the symmetric case. Physiologically, symmetry is not reliable, since the role of axons and dendrites is not exchangeable. Some authors (e.g. Krotov, Parisi) also dispute the constancy of the synaptic matrix. First at all, we present a general framework producing Hopfield-like vector fields of gradient type, encompassing the existing symmetric models (2021 Hopfield, Krotov) as particular cases. Moreover, we introduce dynamic synaptic weights controlled during the dynamics according to a suitable multiobjective variational principle. As a result, the deriving dynamics incorporates relevant features observed in the real memory functioning: pattern recognition, pattern learning, pattern generating undetermined dynamics (wandering) and finally the possibility of forgetting and restoring memories.

PINN-PAD: Physics Informed Neural Networks in PADova
22–23 February 2024, Padova, Italy

Predicting coronal mass ejections' travel times by using physics-informed loss functions

Marchetti Francesco

Department of Mathematics, University of Padova

The problem of predicting the arrival of coronal mass ejections (CMEs) to Earth represents an important topic in the context of space weather, which has been addressed by many empirical, physics-based and machine learning strategies.

In this talk, we propose a physics-driven deep learning method, in which we include the widely-used drag-based physical model in the definition of the loss functions to minimize during the training process of a cascade of two neural networks. We show that including physical information in the architecture improves the predictive capabilities and robustness of the scheme with respect to the purely-data driven approach.

PINN-PAD: Physics Informed Neural Networks in PADova
22–23 February 2024, Padova, Italy

A deep neural network approach for parameterized PDEs and Bayesian inverse problems

Onwunta Akwum

Department of Industrial and systems Engineering
Lehigh University

We consider the simulation of Bayesian statistical inverse problems governed by large-scale linear and nonlinear partial differential equations (PDEs). Markov chain Monte Carlo (MCMC) algorithms are standard techniques to solve such problems. However, MCMC techniques are computationally challenging as they require a prohibitive number of forward PDE solves. The goal of this paper is to introduce a fractional deep neural network (fDNN) based approach for the forward solves within an MCMC routine. Moreover, we discuss some approximation error estimates. We illustrate the efficiency of fDNN on inverse problems governed by nonlinear elliptic partial differential equations and the unsteady Navier-Stokes equations. In the former case, two examples are discussed, respectively depending on two and 100 parameters, with significant observed savings. The unsteady Navier-Stokes example illustrates that fDNN can outperform existing DNNs, doing a better job of capturing essential features such as vortex shedding.

Spectral learning for solving molecular Schrödinger equations

Saleh Yahya[1,2], Iske Armin[2], Yachmenev Andrey[1,4] and Küpper Jochen[1,3,4]

[1]Department of Mathematics, Universität Hamburg, Hamburg, Germany
[2]Center for Free-Electron Laser Science (CFEL), Deutsches Elektronen-Sychrotron DESY, Hamburg, Germany
[3]Department of Physics, Universität Hamburg, Hamburg, Germany
[4]Center for Ultrafast Imaging, Universität Hamburg, Hamburg, Germany

Recently, there has been a significant research interest in using neural networks for solving partial differential equations (PDEs) in general [2], and Schrödinger equations in particular [6]. The use of neural networks was shown to mitigate, or even break the curse of dimensionality [4] encountered in standard numerical methods, such as finite-volume or spectral methods. However, standard neural networks for solving PDEs seem to be fragile [3], since they require a lot of engineering and show high sensitivity to hyperparameters. In the context of quantum mechanics, neural networks were shown to accurately approximate ground-states, i.e., eigenfunctions corresponding to smallest eigenvalues, of molecular systems, while scaling moderately with the dimension of the problem [5]. However, extensions to computing many excited states, i.e., eigenfunctions corresponding to larger eigenvalues, suffer from convergence issues and remain challenging [1].

In this talk I introduce a neural-network based paradigm, where complex and rich families in L^2 are produced by pushing forward standard basis sets through non-singular measurable mappings. I show that a bijectivity assumption on the mapping is a necessary and sufficient conditions for the resulting families to be dense in L^2 [8]. This allows us to model these mappings using normalizing flows, an important tool from generative probabilistic modeling. I present a nonlinear variational framework to approximate molecular wavefunctions in the linear span of these flow-induced families. The framework allowed to compute many eigenstates of various molecular systems with orders-of-magnitude improved accuracy over standard linear methods [7].

REFERENCES

[1] Cuzzocrea, A., Scemama, A., Briels, W. J., Moroni, S., and Filippi, C. (2020). Variational principles in quantum Monte Carlo: The troubled story of variance minimization. 16(7):4203-4212.

[2] E, W. and Yu, B. (2018). The deep Ritz method: a deep learning-based numerical algorithm for solving variational problems. Commun. Math. Stat., 6(1):1-12.

[3] E, W., Ma, C., Wojtowytsch, S., and Wu, L. (2020). Towards a mathematical understanding of neural network-based machine learning: what we know and what we don't.

[4] Grohs, P., Hornung, F., Jentzen, A., and Von Wurstemberger, P. (2018). A proof that artificial neural networks overcome the curse of dimensionality in the numerical approximation of Black-Scholes partial differential equations.

[5] Hermann, J., Schätzle, Z., and Noé, F. (2020). Deep-neural-network solution of the electronic Schrödinger equation.

[6] Hermann, J., Spencer, J., Choo, K., Mezzacapo, A., Foulkes, W., Pfau, D., Carleo, G., and Noé, F. (2022). Ab-initio quantum chemistry with neural-network wavefunctions.

[7] Saleh, Y., Corral, Á. F., Iske, A., Küupper, J., and Yachmenev, A. (2023). Computing excited states of molecules using normalizing flows.

[8] Saleh, Y., Iske, A., Yachmenev, A., and Küpper, J. (2023). Augmenting basis sets by normalizing flows. Proc. Appl. Math. Mech., 23(1):e202200239.

A neural network approach to learn delay differential equations via pseudospectral collocation

Breda Dimitri and Tanveer Muhammad

CDLab Computational Dynamics Laboratory
Department of Mathematics, Computer Science and Physics
University of Udine

The use of neural networks to learn the dynamics of Delay Differential Equations (DDEs) has been recently considered in [1]. The methodology proposed therein relies on reducing to finite dimension the infinite-dimensional dynamical system associated to the DDE by following the approach presented in [2]. In particular, the approximating Ordinary Differential Equation (ODE) is obtained by using a discretization based on the standard Euler's method. In this talk we illustrate a recent extension based on pseudospectral collocation, which leads in general to a more efficient approximation of the underlying DDE. We focus our interest on analyzing the role of the degree of the collocation polynomial in terms of the effectiveness of the learning algorithm and in view of understanding its interplay with the newtwork hyperparameters. Experimental results refer both to the celebrated Mackey-Glass equation as well as to a delay model for a supply chains.

REFERENCES

[1] Ji, X. A., and Gbor O. (2022). Learning Time Delay Systems with Neural Ordinary Differential Equations. IFAC-PapersOnLine 55.36, pp. 79-84.

[2] Breda, D., Diekmann, O., Gyllenberg, M., Scarabel, F., and Vermiglio, R. (2016). Pseudospectral discretization of nonlinear delay equations: new prospects for numerical bifurcation analysis. SIAM Journal on applied dynamical systems 15.1, pp. 1-23.