Spherical configurations and quadrature methods for integral equations of the second kind

Congpei An, Guizhou University, China

We propose and analyze a product integration method for the second-kind integral equation with weakly singular and continuous kernels on the unit sphere S^2. We employ quadrature rules that satisfy the Marcinkiewicz--Zygmund property to construct hyperinterpolation for approximating the product of the continuous kernel and the solution, in terms of spherical harmonics. By leveraging this property, we significantly expand the family of candidate quadrature rules and establish a connection between the geometrical information of the quadrature points and the error analysis of the method. We then utilize product integral rules to evaluate the singular integral with the integrand being the product of the singular kernel and each spherical harmonic. We derive a practical L^∞error bound, which consists of two terms: one controlled by the best approximation of the product of the continuous kernel and the solution, and the other characterized by the Marcinkiewicz--Zygmund property and the best approximation polynomial of this product. Numerical examples validate our numerical analysis. This is a joint work with Dr. Haoning Wu.

Digital Twins and Optimization

Harbir Antil, George Mason University

With recent advancements in computing resources and interdisciplinary collaborations, a new research field called Digital Twins (DTs) is starting to emerge. Data from sensors located on a physical system is fed into its DT, the DT in turn help make decisions about the physical system. This cycle then continues for the life-time of the physical system. A typical example is a bridge or a human heart.

In many cases, these problems can be cast as optimization problems with finite or infinite dimensional (partial differential equations) constraints. This talk will provide an introduction to this topic. Special attention will be given to: 1) Optimization algorithms that are adaptive and can handle inexactness, e.g., Trust-Regions and ALESQP; 2) Optimization under uncertainty and tensor train decomposition to overcome the curse of dimensionality; 3) Reduced order modeling for dynamic optimization using randomized compression.

Additionally, the DT framework may require coupling mutiphysics / systems / data with very different time scales. Keeping this in mind, a newly introduced notion of barely coupled problems will be discussed.

Realistic examples of DTs to identify weakness in structures such as bridges, wind turbines, electric motors, and neuromorphic imaging will be considered.

Learning a generalized multiscale prolongation operator

Eric T. Chung, CUHK

Multigrid preconditioners are one of the most efficient techniques for solving large sparse linear systems. In this research, we address Darcy flow problems with random permeability using the conjugate gradient method, enhanced by a two-grid preconditioner based on a generalized multiscale prolongation operator, which has been demonstrated to be stable for high contrast profiles. To circumvent the need for repeatedly solving spectral problems with varying coefficients, we harness deep learning techniques to expedite the construction of the generalized multiscale prolongation operator. Considering linear transformations on multiscale basis have no impact on the performance of the preconditioner, we devise a loss function by the coefficient-based distance between subspaces instead of L^2-norm of the difference of the corresponding multiscale bases. We discover that leveraging the inherent symmetry in the local spectral problem can effectively accelerate the neural network training process. In scenarios where training data are limited, we utilize the Karhunen-Loève expansion to augment the dataset. Extensive numerical experiments with various types of random coefficient models are exhibited, showing that the proposed method can significantly reduce the time required to generate the prolongation operator while maintaining the original efficiency of the two-grid preconditioner.

The research is partially supported by the Hong Kong RGC General Research Fund (Projects: 14304021 and 14302620).

Multicontinuum homogenization and applications

Yalchin Efendiev, Texas A&M University, College Station

In this talk, I will discuss a general approach for homogenization that uses multiple macroscopic continua. I will discuss its relation to existing methods and show some applications.

Neural Inverse Operators for Solving PDE Inverse Problems

Bjorn Engquist, UT Austin

A large class of inverse problems for PDEs are only well-defined as mappings from operators to functions. Existing operator learning frameworks map functions to functions and need to be modified to learn inverse maps from data. We propose an architecture termed Neural Inverse Operators (NIOs) to solve these PDE inverse problems. Motivated by the underlying mathematical structure and PDE constrained optimization techniques, NIO is based on a composition of DeepONets and Fourier Neural Operators to approximate mappings from operators to functions. Experiments will be presented to demonstrate the performance of the NIOs. They do very well compared to existing neural network baselines in solving PDE inverse problems robustly and accurately. The examples include the classical Calderon problem and optical and seismic imaging. PDE-constrained optimization methods currently can address more challenging problems, but the advantage of NIOs is that they are orders of magnitude faster.

Low-rank approximation of high-dimensional functions in isotropic and anisotropic Sobolev spaces

Helmut Harbrecht, University of Basel, Switzerland

Tensor approximation schemes provide a powerful tool to approximate high-dimensional problems. In order to clarify which problems can efficiently be approximated by tensor approximation schemes, we analyze in this talk the

approximation power of such schemes when applied to high-dimensional functions in the continuous setting. To this end, we assume that the function to be approximated lies either in an isotropic Sobolev space or an anisotropic Sobolev space, possibly equipped with dimension weights. We apply successively the truncated singular value decomposition in order to discuss the cost when approximating the function under consideration in the continuous analogues of tensor formats such as the Tucker tensor format or the tensor train format.

Analog quantum simulation for partial differential equations

Shi Jin, Shanghai Jiao Tong University

To simulation PDEs using quantum simulation, we introduce a method called Schrödingerisation, that directly maps D-dimensional linear PDEs onto a $(D+1)$-qumode quantum system of Schrodinger type, where analog or continuous-variable (CV) Hamiltonian simulation on $D+1$ qumodes can be used. This very simple methodology does not require one to discretise PDEs first, and it is not only applicable to linear PDEs but also to some nonlinear PDEs and systems of nonlinear ordinary differential equations. This also raises the possibility that some PDEs may be simulated directly on analog quantum systems by using Hamiltonians natural for those quantum systems.

Furthermore, we present a simplified analog quantum simulation protocol for preparing quantum states that embed solutions of parabolic partial differential equations, including the heat, Black-Scholes and Fokker-Planck equations. The key idea is to approximate the heat equations by a system of hyperbolic heat equations that involve only first-order differential operators. This scheme requires relatively simple interaction terms in the Hamiltonian, which are the electric and magnetic dipole moment-like interaction terms that would be present in a Jaynes-Cummings-like model. For a d-dimensional problem, we show that it is much more appropriate to use a single d-level quantum system- a qudit-instead of its qubit counterpart, and $d+1$ qumodes. The total resource cost is efficient in d and precision error, and has potential for realisability for instance in cavity and circuit QED systems.

Proximal random reshuffling under local Lipschitz continuity

Lexiao Lai, HKU

We study proximal random reshuffling for minimizing the sum of locally Lipschitz functions and a proper lower semicontinuous convex function without assuming coercivity or the existence of limit points. The algorithmic guarantees pertaining to near approximate stationarity rely on a new tracking lemma linking the iterates to trajectories of conservative fields. One of the novelties in the analysis consists in handling conservative fields with unbounded values.

Algorithmic Stability of Sharpness Aware Minimization

Yunwen Lei, HKU

Sharpness aware minimization (SAM) aims to find flat minima by introducing a perturbation step in gradient methods, which has achieved an impressive success in improving the generalization behavior. An important problem is to understand how the perturbation affects the generalization of SAM, which has not been well studied in the literature. In this talk, we leverage the concept of algorithmic stability to study the generalization performance of SAM. We first build the approximate nonexpansiveness of the gradient operator for SAM, based on which we derive stability bounds as measured by the norm of the model weights. Then, we develop convergence rates in terms of function values for minibatch SAM, which is stronger than the existing rates on gradient norms for convex problems. Our results clearly show how the batch size and the perturbation size affect both the optimization and generalization, and how they should be balanced to achieve the best performance in prediction.

A discretization-invariant extension of deep operator networks and its application in multiscale problems

Wing Tat Leung, City University of Hong Kong

Operator learning trains a neural network to map functions to functions. An ideal operator learning framework should be mesh-free, meaning it doesn't require a specific discretization of the input functions during training, can handle input and output functions on different domains, and supports varying grids between samples. We present a mesh-free neural operator for solving parametric partial differential equations. Our Basis Enhanced Learning Network (BelNet) projects input functions into a latent space and reconstructs output functions with a unique feature that allows the network to learn "basis" functions during training. This approach generalizes the universal approximation theory of nonlinear operators by Chen and Chen to accommodate differences in input and output meshes. Through various challenging high-contrast and multiscale problems, we show that our method outperforms existing operator learning techniques and provides greater flexibility in the sampling and discretization process.

Wavelet-based Edge Multiscale Finite Element Methods

Guanglian Li, HKU

We propose a novel efficient and robust Wavelet-based Edge Multiscale Finite Element Method (WEMsFEM) to solve PDEs with heterogeneous coefficients. The main idea is to first establish a local splitting of the solution over a local region by a local bubble part and local Harmonic extension part, and then derive a global splitting by means of Partition of Unity. This facilitates a representation of the solution as a summation of a global bubble part and a global Harmonic extension part, where the first part can be computed locally in parallel. To approximate the second part, we construct an edge multiscale ansatz space locally with hierarchical bases or Haar wavelets as the local boundary data that has a guaranteed approximation rate without higher regularity requirement on the solution. The key

innovation of this proposed method lies in a provable convergence rate with little restriction on the mesh size or the regularity of the solution. Its convergence rate with respect to the computational degree of freedom is rigorously analyzed for many multiscale problems, including the elliptic problems with heterogeneous coefficients, convection dominated diffusion problems and Maxwell equations with heterogeneous permittivity. Moreover, they are verified by extensive 2-d and 3-d numerical tests. This is a joint work with Eric Chung (CUHK, Hong Kong), Yalchin Efendiev (Texas A&M University, College Station), Shubin Fu (Eastern Institute of Technology, P.R. China) and Yueqi Wang (HKU).

Random Batch Methods and Kinetic Monte Carlo for Interacting Particle Systems with Lévy Noise

Jian-Guo Liu, Duke University

The random batch method (RBM) has significantly reduced computational costs across various applications in machine learning and data science. Building on this concept, we have developed efficient algorithms for simulating large-scale interacting particle systems comprising N indistinguishable particles. These systems model diverse phenomena in natural and social sciences, ranging from molecular dynamics to collective behavior in biological and social contexts.

The central idea of RBM involves randomly selecting small batches of size $p \ll N$, where particles within a batch interact with each other over a short time. To preserve the statistical properties of the original system, the weak interactions (of strength $\frac{1}{N-1}$) are replaced by stronger interactions (of strength $\frac{1}{p-1}$) within each batch. This innovative replacement reduces the computational complexity per time step from $O(N^2)$ to$O(pN)$, while maintaining an unbiased approximation of the force or velocity field in the original system. Remarkably, RBM with replacement has a fundamental equivalence with the well-established Kinetic Monte Carlo (KMC) commonly used in particle simulations.

In many real-world scenarios, the underlying random fluctuations are non-Gaussian, particularly in contexts where heavy-tailed data distributions arise, such as COVID-19 mortality statistics and hurricane-related economic losses. Such non-Gaussian phenomena call for Lévy noise, which accommodates jumps and extreme variations.

Recent research has also employed Lévy-type diffusion approximations for stochastic gradient descent (SGD) to better address heavy-tailed distributions in data. The RBM framework seamlessly extends to handle systems driven by Lévy noise, offering robust theoretical guarantees for accuracy and convergence.

In this talk, I will discuss the theoretical foundations and practical applications of random batch methods, with a focus on their integration into stochastic gradient descent and interacting particle systems driven by Lévy noise.

Autonomous robotic interaction with flow fields

Jia Pan, HKU

In this talk, we will introduce our recent work on how to enable the robot to intelligently interact and manipulate flow fields, using ocean flow, human crowds, and transportation flow as typical examples. We will demonstrate that autonomous robots with sophisticated algorithms can efficiently recover the global properties of the flow field, interactively query local details of the field, and effectively manipulate the entire flow to achieve optimal flow behaviors. We will discuss interesting applications such as autonomous underwater robot, human-robot interactions, and intelligent transportation.

Executable Quantum Algorithms for Efficient PDE Simulations

Daniel Peterseim, University of Augsburg

In this talk, we present an executable quantum algorithm designed to solve second-order linear elliptic partial differential equations (PDEs) discretized by d-linear finite elements on Cartesian grids within bounded d-dimensional domains. A key feature of this approach is the BPX preconditioner, which transforms the linear system into a well-conditioned form optimized for quantum computation. The algorithm achieves a complexity that scales linearly with 1/tol for fixed dimension d, allowing the computation of relevant functionals of the solution to a given tolerance tol > 0.

We discuss the practical implementation of the quantum circuit and present numerical results from both quantum simulators and real quantum computers, demonstrating the applicability of the algorithm on currently accessible quantum hardware. In addition, we briefly introduce a novel framework for addressing nonlinear problems using quantum computing.

Enhancing CFD simulations combining model order reduction and scientific machine learning

Gianluigi Rozza, SISSA

Partial differential equations (PDEs) are invaluable tools for modeling complex physical phenomena. However, only a limited number of PDEs can be solved analytically, leaving the majority of them requiring computationally expensive numerical approximations. To address this challenge, reduced order models (ROMs) have emerged as a promising field in computational sciences, offering efficient computational tools for real-time simulations. In recent years, deep learning techniques have played a pivotal role in advancing efficient ROM methods with exceptional generalization capabilities and reduced computational costs [1, 2, 3]. In this talk we explore how classical ROM techniques can be elevated through the integration of some deep learning models. We will introduce hybrid approaches, which consider both physics-based and purely data-driven techniques [4, 5, 6, 7], as well as aggregated ones, where the model is built as the combination of different pre-trained models [8].

Our discussion encompasses a review of existing (intrusive and data driven) approaches to enhancing ROM by means of neural operators with applications in Computational Fluid Dynamics, also in presence of turbulence and compressibility.

[1] Benner, P., Grivet-Talocia, S., Quarteroni, A., Rozza, G., Schilders, W., & Silveira, L. M. (2021). System-and data-driven methods and algorithms (p. 378). De Gruyter.
[2] Benner, P., Schilders, W., Grivet-Talocia, S., Quarteroni, A., Rozza, G., & Miguel Silveira, L. (2020). Model Order Reduction: Volume 2: Snapshot-Based Methods and Algorithms (p. 348). De Gruyter.
[3] Benner, P., Schilders, W., Grivet-Talocia, S., Quarteroni, A., Rozza, G., & Miguel Silveira, L. (2020). Model order reduction: Volume 3: Applications (p. 466). De Gruyter.

[4] Ivagnes, A., Stabile, G., Mola, A., Iliescu, T., & Rozza, G. (2023). Hybrid data-driven closure strategies for reduced order modeling. Applied Mathematics and Computation, 448, 127920
[5] Ivagnes, A., Stabile, G., & Rozza, G. (2024). Parametric Intrusive Reduced Order Models enhanced with Machine Learning Correction Terms. arXiv preprint arXiv:2406.04169.
[6] Zancanaro, M., Mrosek, M., Stabile, G., Othmer, C., & Rozza, G. (2021). Hybrid neural network reduced order modelling for turbulent flows with geometric parameters. Fluids, 6(8), 296.
[7] Stabile, G., Zancanaro, M., & Rozza, G. (2020). Efficient geometrical parametrization for finite-volume-based reduced order methods. International journal for numerical methods in engineering, 121(12), 2655-2682.
[8] Ivagnes, A., Tonicello, N., Cinnella, P., & Rozza, G. (2024). Enhancing non-intrusive reduced-order models with space-dependent aggregation methods. Acta Mechanica, 1-30.

Convection-Diffusion Equation: An axiomatized Framework for Neural Networks

Zuoqiang Shi, Tsinghua

Bridging neural networks with partial differential equations holds significant importance, as it not only enhances the interpretability of neural networks but also sheds light on designing network architectures. In this talk, we establish convection-diffusion equation models based on rigorous theoretical analysis. The convection-diffusion equation model not only covers existing network structures, but also illuminates novel network design, COnvection dIffusion Networks (COIN). Numerical results demonstrate the effectiveness of COIN in various benchmarks, as well as its potential in novel tasks such as disease prediction.

Parametric PDE with Many Parameters– Kernel Methods and Deep Neural Networks

Ian H. Sloan, University of New South Wales, Sydney Australia

Parametric PDE with a large number of parameters, perhaps in the hundreds, face the “Curse of dimensionality”. After an introduction, the talk will describe recent work on a direct “kernel” method, and then describe work in progress on a deep

learning approach to a fast surrogate. Collabo rators on these projects include Frances Kuo, Vesa Kaarnioja, Fabio Nobile, Yoshihito Kazashi, Alec Gilbert, Alex Keller and Dirk Nuyens.

Operator Learning for Nonsmooth Optimal Control of PDEs

Yongcun Song, CityU

Optimal control problems with nonsmooth objectives and partial differential equation (PDE) constraints are challenging, mainly because of the underlying nonsmooth and nonconvex structures and the demanding computational cost for solving multiple high-dimensional and ill-conditioned systems after mesh-based discretization. To mitigate these challenges numerically, we propose an operator learning approach combined with an effective primal-dual optimization idea that can decouple the treatment of the control and state variables so that each resulting iteration only requires solving two PDEs. Our main purpose is to construct neural surrogate models for the involved PDEs by operator learning, allowing the solution of a PDE to be obtained with only a forward pass of the neural network. The resulting algorithmic framework offers a hybrid approach that combines the flexibility and generalization of operator learning with the model-based nature and structure-friendly efficiency of primal-dual-based algorithms. The primal-dual-based operator learning approach offers numerical methods that are mesh-free, easy to implement, and adaptable to various optimal control problems with nonlinear PDEs. Notably, the neural surrogate models can be reused across iterations and parameter settings hence, computational cost can be substantially alleviated. We validate the effectiveness and efficiency of the primal-dual-based operator learning approach across a range of typical optimal control problems with nonlinear PDEs.

Hessian estimates in the score based generative models

Zhongjian Wang, NTU

The Hessian estimates of the score potential, or Lipchitz estimate of score, play important roles in the well-posedness and convergence analysis of the score based generative models. In this talk, I will discuss the approaches of estimate in three different applicational scenarios: lower bound in singular case, polynomial in dimension upper bound for finite dimensional, trace type upper bound in the infinite

dimensional near Gaussian setup. Our approaches are application of PDE regularity theories on Fokker-Planck equation with OU drift and applies to both KL and Wasserstein bounds.

Small and Large Atomic/Language Models for Materials Science

Tongqi Wen, HKU

"AI for Materials" aims to facilitate the design of new materials with exceptional properties and enhance our understanding of composition-structure-property relationships by incorporating advanced AI techniques. This talk will be divided into two parts: the application of (i) small/large atomic models and (ii) large language models in materials science. In the first part, we will discuss the application of small AI atomic models in atomistic simulations, focusing on defect properties in structural materials including Ni and Ti-based alloys. We will then present our latest work on large atomic models for alloys covering 53 elements in the periodic table. In the second part, we will explore the development of materials-specific large language models built on general LLMs (GPT-4, Gemini-Pro, …) employing prompt engineering techniques for materials property classification and prediction. Our future research will focus on harnessing the power of AI, particularly language models and deep neural networks, to design advanced materials for structural, ferroelectric, and battery applications and understand relevant physical phenomena.

Efficient Temporal GNN Training via Staleness-Aware Pipeline

Chuan Wu, HKU

Memory-based Temporal Graph Neural Networks (MTGNNs) are a class of temporal graph neural networks that utilize a node memory module to retain long-term temporal dependencies, leading to superior performance compared to memory-less counterparts. For MTGNN training, the iterative reading and updating process of the memory module to obtain up-to-date information needs to follow the temporal dependencies, incurring significant overhead and limiting training throughput. In this talk, I will share our recent work on an efficient learning framework for memory-based TGNNs that maximizes training throughput while maintaining model accuracy. Our design addresses challenges associated with fetching and updating

node memory states in MTGNNs by integrating staleness into the memory module. We introduce an online pipeline scheduling algorithm that strategically breaks temporal dependencies with minimal staleness and delays memory fetching to obtain fresher memory states. We also propose a staleness mitigation mechanism to enhance training convergence and model accuracy.

Multiscale computational method for radiative transfer equation and Landau-Lifschitz equation in heterogeneous media

Lei Zhang, Shanghai Jiaotong University

Multiscale computational methods are designed to capture fine-scale features while maintaining a manageable computational cost at coarser levels. In recent decades, methods such as numerical homogenization have matured, particularly for benchmark problems such as multiscale elliptic PDEs. In this talk, I will present two examples that extend beyond these benchmarks: the radiative transfer equation, which includes velocity dependence, and the Landau-Lifshitz equation in micromagnetics, a nonlinear vectorial equation—both exhibit spatial heterogeneity. We will formulate multiscale computational methods for these cases, present analyses and numerical experiments, and discuss the unique characteristics of each PDE.

DeepParticle: learning invariant measure by a deep neural network minimizing Wasserstein distance on data generated from an interacting particle method

Zhiwen Zhang, HKU

High-dimensional partial differential equations (PDEs) pose significant computational challenges, particularly when solutions exhibit large gradients or concentrations at unknown locations. Traditional mesh-based methods often struggle with such complexities. In this talk, we introduce

DeepParticle, an innovative approach that integrates deep learning (DL), optimal transport (OT), and interacting particle (IP) methods to address these challenges. Through a case study of Kolmogorov-Petrovsky-Piskunov (KPP) front speeds in incompressible flows, we demonstrate the effectiveness of our method. Our approach reduces the PDE problem to the computation of the principal eigenvalue of an advection-diffusion operator. By leveraging the Feynman-Kac formula, we derive a stochastic representation that enables the use of a genetic IP algorithm. This algorithm evolves particle distributions to a time-invariant measure, from which the front speed is extracted. To efficiently learn this family of invariant measures, we train a physically parameterized deep neural network using affordable data generated by IP computations at moderate Péclet numbers. This trained network then allows us to predict invariant measures and, consequently, front speeds at larger Péclet numbers, where direct IP computations are prohibitively expensive. Our methodology not only addresses the computational bottlenecks of traditional methods but also extends to more general stochastic particle dynamics problems. For example, we showcase its application in learning and generating aggregation patterns in Keller-Segel chemotaxis systems, highlighting the versatility and broad applicability of the DeepParticle approach.

Weak Generative Sampler to Solve High - Dimensional PDEs for Stochastic Models

Xiang Zhou, CityU

The solution of many typical high-dimensional PDEs (such as the Fokker-Planck, and McKean-Vlasov equations) is associated with a probability distribution. To solve such PDEs by deep learning techniques is usually to simply find a neural network for the density function itself, subject to certain positivity and normalization conditions. The further utilization of the solution requires random sampling again. We introduce a framework of Weak Generative Sampler (WGS) to both solve the PDE and generate

samples more efficiently than the PINN and the Ritz method. Our proposed loss function is based on the weak form and the generic probability interpretation of the loss function. The details of this talk will also explain why the efficiency and adaptivity are so easy to achieve in this WGS for high-dimensional PDEs.