2021 年科学与工程计算青年研讨会报告题目摘要

Approximation Problems Related to Non-smooth Solutions of a Class of Nonlinear Integrable PDEs

常向科
（中国科学院数学与系统科学研究院）

Abstract: A class of nonlinear integrable PDEs admit non-smooth solitary wave solutions as certain special weak solutions, which can be used to model non-linear phenomena such as the breakdown of regularity and the onset of shocks. Due to the Lax integrability, the inverse spectral method is a powerful tool in the construction of these solutions. In this talk, I'll stress the approximation theory, such as continued fractions, Hermite-Pade approximations, involved in the inverse spectral problems.

A Block-Randmoized Stochastic Algorithm for Low-Rank Tensor CP Decomposition

崔春风
（北京航空航天大学）

Abstract: The block-randomized stochastic algorithm has shown its power in handling high-dimensional low-tank tensor canonical polyadic decomposition (CPD). Since computing CPD is computationally expensive, there is great interest in speeding up the convergence. In this talk, we introduce several accelerated version of the block-randomized stochastic gradient descent (SGD) algorithm for low-rank tensor CPD. Under some mild conditions, we also show the global convergence. Our preliminary numerical experiments for the synthetic and real data demonstrated that our accelerated algorithms are efficient, and can achieve better performance in terms of objection function value, mean squared error, and structural similarity value. This talk is based on the joint work with Qingsong Wang and Prof. Deren Han.

Explaining Neural Networks with Representation Theory

范凤磊
（康奈尔大学）

Abstract: Deep learning has recently achieved huge successes in many applications. It is widely recognized that machine learning, especially deep learning, is a paradigm shift in many fields. However, the success of deep learning is not well underpinned by the effective theory. Lacking interpretability has become a primary obstacle to the widespread translation and further development of deep learning techniques. In this talk, we attempt to understand the inner working of neural networks with representation learning theory. Particularly, we shed light on the following two interesting questions: why shortcuts are a powerful type of machinery in neural network design? Do neural networks have a directional preference, or could network

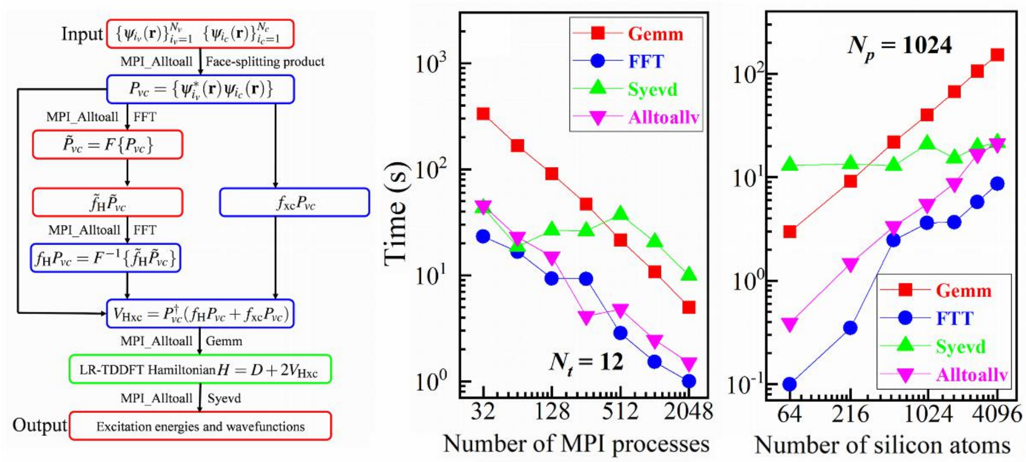
width & depth be exchangeable in some way? For the former question, we show that a network with shortcuts can approximate a much more complicated function than a network without shortcuts can, while for the latter question, we prove that the width and depth of a network can be converted to each other up to an arbitrarily small error, thereby establishing the quasi-equivalency of width and depth.

基于机器学习的低秩分解算法加速第一性原理高精度电子结构低标度高性能计算

胡伟
（中国科学技术大学）

Abstract: 分子固体材料体系的光电性质严重地依赖于激发态电子-空穴对（激子）的相互作用。通过基于含时密度泛函线性响应理论LR-TDDFT激发态电子结构计算方法，可以准确地计算体系的激子效应。但是这种激发态电子结构计算速度非常慢，尤其是在高精度的周期性平面波基组下，由于昂贵的多中心积分计算，计算标度更是高达O(N^5~N^6)，计算复杂度大，内存需求多，传统的第一性原理材料模拟激发态电子计算只能处理数百原子的体系。

为了解决这些问题，我们提出了针对大尺度分子固体材料体系的LR-TDDFT激发态电子结构低标度计算方法，通过插值可分离密度拟合ISDF低秩分解算法，结合机器学习聚类降维算法和隐式迭代对角化方法，可以将LR-TDDFT激发态电子结构的计算复杂度从高标度O(N^5~N^6)降低至立方标度O(N^3)（与LDA/GGA类似）。而且，这种低标度计算方法有利于在超级计算机上实现大规模高性能并行计算（万核以上）。在自主开发的标准平面波计算软件PWDFT中，通过混合CPU-MPI/OpenMP + GPU-CUDA异构并行计算模式，我们计算了包含4,096原子体相硅的激发态电子结构性质，并行计算的规模高达12,288CPU核。



[1] Hu, W.; Liu, J.; Li, Y.; Ding, Z.; Yang, C.; Yang J. *J. Chem. Theory Comput.* 16, 964 (2020).
[2] Liu, J.; Hu, W.; Yang, J. *J. Chem. Phys.* 154, 064101 (2021).
[3] Wan, L.; Liu, X.; Liu, J.; Qin, X.; Hu, W.; Yang, J. *Electron. Struct.* 3, 024004 (2021).

A Riemannian Optimization Approach to Clustering Problems

黄文
（厦门大学）

Abstract: The task of clustering problems is to group a set of objects such that the objects in the same group are more similar under certain criterion to each other than to those in other groups. The applications include but not limited to K-means, spectral clustering, normalized cuts, and community detection. In this presentation, we formulate the clustering problems as nonsmooth optimization problems on a subset of the Stiefel manifold. It is shown that the subset forms an embedded submanifold. The geometry structures are derived. A proximal gradient method is proposed and used. Numerical experiments are used to demonstrate the performance of the proposed model and algorithm. This is joint work with Meng Wei, Kyle A. Gallivan, and Paul Van Dooren.

智能科学计算——HPC+AI 在第一性原理分子动力学中的应用

贾伟乐
（中国科学院计算技术研究所）

Abstract: 智能超算为超算应用的发展提出了新的挑战和方向。如何融合传统的”HPC+物理模型”的计算模式与新的智能超算成为新的课题。本报告从典型的科学计算出发，以第一性原理分子动力学为例，展示一种全新的智能科学计算（HPC+AI+物理模型）的计算模式。相比传统的高性能计算，智能科学计算在计算能效上有有成量级的提升，同时也对传统的计算机体系架构提出了新的挑战。本报告从深度学习分子动力学软件出发，分享智能科学计算的一点进展和思考。

Nonlocal Cahn-Hilliard Equation and its Applications in 3D Printing and Fingerprint Restoration

李义宝
（西安交通大学）

Abstract: In this talk, we will introduce a nonlocal Cahn-Hilliard equation and its applications in 3D printing and fingerprint restoration. A Crank-Nicolson method is proposed to discrete the nonlocal Cahn-Hilliard equation, which was developed for modeling microphase separation of diblock copolymers. We prove that our proposed scheme is unconditionally energy stable. Then we will present a robust and efficient fingerprint image restoration algorithm and surface pattern generation method for 3D printing. The proposed method has a merit that the pixel values in the damaged fingerprint domain can be obtained using the image information from the outside of the damaged fingerprint region. Restoration of fingerprint based on the adjacent pixel information can ensure the accuracy of fingerprint information with low computational cost. Computational experiments are presented to demonstrate the efficiency of the proposed method.

Solving High-speed Flows Using PINNs and DeepMMnets

毛志平
（厦门大学）

Abstract: Recently, neural network-based deep learning methods, which are different from the classical numerical methods, have attracted lots of attention not only in the traditional artificial intelligence community but also the scientific computing community. In this talk, I will introduce my work using physics-informed neural networks (PINNs) and deep multi-scale multi-physics nets (DeepMMnet) for high-speed flows. In particular, I shall solve the inverse problems of the shock wave problems in supersonic flow by using PINNs based on the information of density gradient $\nabla\rho$ and limited data of pressure and inflow conditions instead of using boundary conditions. Then I will introduce the inference of the flow past a normal shock in hypersonic flow by using the DeepMMnets with the help of DeepOnets.

Non-line-of-sight Reconstruction with Signal-object Collaborative Regularization

邱凌云
（清华大学）

Abstract: Non-line-of-sight imaging aims at recovering obscured objects from multiple scattered light. It has recently received widespread attention due to its potential applications such as autonomous driving, rescue operations, and remote sensing. However, in cases with high measurement noise, obtaining high-quality reconstructions remains a challenging task. In this work, we establish a unified regularization framework, which can be tailored for different scenarios, including indoor and outdoor scenes with substantial background noise under both confocal and non-confocal settings. The proposed regularization framework incorporates sparseness and non-local self-similarity of the hidden objects as well as smoothness of the measured signals. We show that the estimated signals, albedo, and surface normal of the hidden objects can be estimated reconstructed robustly even with high measurement noise under the proposed framework. Reconstruction results on synthetic and experimental data show that our approach recovers the hidden objects faithfully and outperforms state-of-the-art reconstruction algorithms in terms of both quantitative criteria and visual quality.

Exponential Convergence Theory of the Multipole and Local Expansions for the Three-dimensional Laplace Equations in Layered Media

汪波
（湖南师范大学）

Abstract: In this talk, we present the exponential convergence theory for the multipole and local expansions, shifting and translation operators for the Green's function of 3-dimensional Laplace equation in layered media. An immediate application of the

theory is to ensure the exponential convergence of the FMM which has been shown by the numerical results reported in our published paper. As the Green's function in layered media consists of free space and reaction field components and the theory for the free space components is well known, this paper will focus on the analysis for the reaction components. We first prove that the density functions in the integral representations of the reaction components are analytic and bounded in the right half complex wave number plane. Then, using the Cagniard-de Hoop transform and contour deformations, estimates for the remainder terms of the truncated expansions are given, and, as a result, the exponential convergence for the expansions and translation operators is proven.

Deep Learning Methods for PDEs and Reduced Order Model

王敏
（杜克大学）

Abstract: In this talk, we will be centred on the deep-learning-based methods for numerical solutions to high-dimensional PDEs. In particular, three related projects will be presented with details in correspondence to three major subtopics in this area: 1) the formulations of the optimization problems associated to the quadratic porous medium equation; 2) the generalization error analysis for deep Ritz method; and 3) a training scheme for ResNet. In addition, a brief discussion over the combination of deep learning techniques with the reduced order models (ROM) will also be included.

Implicit Regularization and Entrywise Convergence of Riemannian Optimization for Low Tucker-Rank Tensor Completion

魏轲
（复旦大学）

Abstract: This paper is concerned with the low Tucker-rank tensor completion problem, which is about reconstructing a tensor $\mathcal{T}\in\mathbb{R}^{n\times n\times n}$ of low multilinear rank from partially observed entries. We consider a manifold algorithm (i.e., Riemannian gradient method) for this problem and reveal an appealing implicit regularization phenomenon of non-convex optimization in low Tucker-rank tensor completion. More precisely, it has been rigorously proved that the iterates of the Riemannian gradient method stay in an incoherent region throughout all iterations provided the number of observed entries is essentially in the order of $O(n^{3/2})$. To the best of our knowledge, this is the first work that has shown the implicit regularization property of a non-convex method for low Tucker-rank completion under the nearly optimal sampling complexity. Additionally, the entrywise convergence of the method is further established. The analysis relies on the leave-one-out technique and the subspace projection structure within the algorithm. Some of technical results developed in the paper might be of broader interest in investigating the properties of other non-convex algorithms.

Towards Macroscopic Liquid-crystal Theory for General Rigid Molecules

徐劼
（中国科学院数学与系统科学研究院）

Abstract: Non-axisymmetric rigid molecules have proved to exhibit rich liquid-crystalline phases that are barely understood theoretically, especially for the connection between the molecular architecture and macroscopic phase behaviors. A promising route is to derive macroscopic theory from molecular theory, which has shown its success in the study of rod-like and bent-core molecules. We discuss several essential ingredients in this route to construct free energy for general rigid molecules: molecular symmetry; order parameters to describe mesoscopic states; expansion of interaction kernels; entropy. The molecular symmetry is characterized by a point group. We follow three steps below to arrive at the macroscopic free energy.

1) From the proper rotations in the point group, we identify, with explicit expressions, the invariant tensors, which are nonvanishing when averaged by the density function. It is from the averages of these invariant tensors that the order parameters are chosen.

2) By expanding the interaction kernels into coupling of tensors, the interaction free energy can be constructed. The improper rotations, although having no effect on order parameters, impose extra conditions on the expansion of interaction kernels.

3) The entropy term can be converted into a function of tensor order parameters by finding the maximum entropy state. We further propose a simple function as an approximation that maintains the essential properties of the original entropy, while avoiding the integrals in the original entropy.

ErrorCompensatedX: Error Compensation for Variance Reduced Algorithms

严明
（密歇根州立大学）

Abstract: Communication cost is one major bottleneck for the scalability for distributed learning. One approach to reduce the communication cost is to compress the gradient during communication. However, directly compressing the gradient decelerates the convergence speed, and the resulting algorithm may diverge for biased compression. Recent work addressed this problem for stochastic gradient descent by adding back the compression error from the previous step. This idea was further extended to one class of variance reduced algorithms, where the variance of the stochastic gradient is reduced by taking a moving average over all history gradients. However, our analysis shows that just adding the previous step's compression error, as done in existing work, does not fully compensate the compression error. So, we propose ErrorCompensatedX, which uses the compression error from the previous two steps. We show that ErrorCompensatedX can achieve the same asymptotic convergence rate with the training without compression. Moreover, we provide a unified theoretical analysis framework for this class of variance reduced algorithms, with or without error compensation.

带线型 Dirac 源项 Poisson 方程的自适应有限元方法

易年余
（湘潭大学）

Abstract: 我们针对二维带线型 Dirac 源项的 Poisson 方程，分析了解的正则性，并结合解在裂缝线上的连续性和法向导数的跳跃特征，构造了该问题有限元离散格式的后验误差估计子，得到了误差估计子的可靠性和有效性，数值结果说明了理论结果的正确性和自适应算法的有效性。

A Mathematical Theory of Computational Resolution Limit

张海
（香港科技大学）

Abstract: It is well-known that the resolution of optical imaging system is fundamentally limited by the optical wavelength. Based on this, Rayleigh proposed the Rayleigh criterion on the minimum resolvable distance between two point sources, the so called Rayleigh limit. Although widely used in the practice, this limit is not so useful for images that are subject to elaborated data processing. To remedy this, we develop a theory of computational resolution limit to characterize the fundamental resolution limit from the approximation theory point of view. The theory can be used to explain the phase transition phenomenon in the reconstruction problem. New efficient superresolution algorithm is also developed following the theory.

Towards Gradient-based Bilevel Optimization in Machine Learning

张进
（南方科技大学）

Abstract: Recently, Bi-Level Optimization (BLO) techniques have received extensive attentions from machine learning communities. In this talk, we will discuss some recent advances in the applications of BLO. First, we study a gradient-based bi-level optimization method for learning tasks with convex lower level. In particular, by formulating bi-level models from the optimistic viewpoint and aggregating hierarchical objective information, we establish Bi-level Descent Aggregation (BDA), a flexible and modularized algorithmic framework for bi-level programming. Second, we focus on a variety of BLO models in complex and practical tasks are of non-convex follower structure in nature. In particular, we propose a new algorithmic framework, named Initialization Auxiliary and Pessimistic Trajectory Truncated Gradient Method (IAPTT-GM), to partially address the lower level non-convexity. By introducing an auxiliary as initialization to guide the optimization dynamics and designing a pessimistic trajectory truncation operation, we construct a reliable approximation to the original BLO in the absence of lower level convexity hypothesis. Extensive experiments justify our theoretical results and demonstrate the superiority of the proposed BDA and IAPTT-GM for different tasks, including hyper-parameter optimization and meta learning.

Diffusion Coefficient Identification for Elliptic and Parabolic Problems: Finite Element Approximation and Error Analysis

周知
（香港理工大学）

Abstract: I will present a novel error analysis for recovering a spatially dependent diffusion coefficient in an elliptic or parabolic problem. It is based on the standard regularized output least-squares formulation with an $H^1(\Omega)$ seminorm penalty and then discretized using the Galerkin finite element method with conforming piecewise linear finite elements for both state and coefficient and backward Euler in time in the parabolic case. We derive a priori weighted $L^2(\Omega)$ estimates where the constants depend only on the given problem data for both elliptic and parabolic cases. Further, these estimates also allow deriving standard $L^2(\Omega)$ error estimates under a positivity condition that can be verified for certain problem data.