Conference on Tensor Methods in Mathematics and Data Science

Time: 18th~20th, November, 2024

Address: 9th floor, Library, Shenzhen MSU-BIT University

INM RAS, MSU-BIT, MSU, and Moscow RTT Algorithm Lab

AGENDA

LIST OF SCHEDULE EVENTS

DATE	TIME	EVENTS	LOCATION
11.18	08:30-09:00	Conference - Opening Ceremony	9th floor, Library
	09:00-09:30	Group Photo	1st floor, Near the Library
	09:30-15:30	Conference - Reports and Discussions	9th floor, Library
	15:30-17:00	Student Posters	9th floor, Library
	16:00-17:30	Huawei Galileo Hall Tour (all interested experts)	Huawei
11.19	09:00-17:00	Conference - Reports and Discussions	9th floor, Library
	14:00-16:00	Round Table: How Mathematics Revolutionizes Computing	1st floor, Library
11.20	09:00-15:00	Conference - Reports and Discussions	9th floor, Library
	15:00-15:30	Conference - Closing Remarks	9th floor, Library

If there are any changes to the agenda, the organizing committee will inform all the experts and participants in a timely manner.

MONDAY

2024.11.18

9th FLOOR OF THE LIBRARY

TIME	NAME	TOPIC
08:30 - 08:40	WELCOME SPEECH BY LI HEZHANG	
08:40 - 08:50	WELCOME SPEECH BY IVANCHENKO SERGEY NIKOLAEVICH	
08:50 - 09:00	WELCOME SPEECH BY CHEN BING	
09:00 - 09:30	GROUP PHOTO	
09:30 - 10:00	EUGENE TYRTYSHNIKOV	Tensors and optimization
10:00 - 10:30	MICHAEL KWOK-PO NG	Tensor Representations in Data Science
10:30 - 11:00	IVAN OSELEDETS	Matrix and tensor methods in machine learning
11:00 - 11:30	COFFEE BREAK	
11:30 - 12:00	ANDREI KRYLOV	Using Mathematical Models in Deep Learning
12:00 - 12:30	YIMIN WEI	Tensor-based model reduction and identification for generalized memory polynomial
12:30 - 14:00	LUNCH	
14:00 - 14:30	SERGEY KABANIKHIN	Regularization: from Algebraic Equations to Neural Networks
14:30 - 15:00	CARMINE DI FIORE	Insights on some theoretical results related to the convergence of power and Gauss-Seidel iterations
15:00 - 15:30	SERGEY DOLGOV	Tensor train approximation of deep transport maps for Bayesian inverse problems
15:30 - 17:30	Student posters / Huawei Galileo Hall tour. All interested professors.	

TUESDAY
2024.11.19

9th FLOOR OF THE LIBRARY

TIME	NAME	TOPIC
09:00 - 09:30	WELCOME COFFEE	
09:30 - 10:00	ALEXANDER OSINSKY	Speed-accuracy trade-off in cross matrix approximations
10:00 - 10:30	DMITRY YAROTSKY	SGD with memory: theoretical acceleration of mini-batch SGD for overparameterized models
10:30 - 11:00	ALEKSANDR MIKHALEV	NNTile: towards dynamic task-based parallelism for training and inference of large neural networks
11:00 - 11:30	COFFEE BREAK	
11:30 - 12:00	GIANLUCA CERUTI	Recent Developments in Tensor Techniques for Dynamical Low-Rank Approximation
12:00 - 12:30	DARIO FASINO	Computing the stationary probability vector of stochastic tensors and higher-order Markov chains
12:30 - 14:00	LUNCH	
14:00 - 14:30	SERGEY MATVEEV	Nonnegative low-rank tensor approximations: alternating projections and fast corrections
14:30 - 15:00	XIAOMAN HU	The architecture and Ecosystem of openMind AI platform
15:00 - 15:30	MAXIM SHISHLENIN	Matrix methods, block-Toeplitz matrices in inverse problems of wave tomography
15:30 - 16:00	COFFEE BREAK	
16:00 - 16:30	CHAO WANG	A Scale-Invariant Relaxation in Low-Rank Tensor Recovery
16:30 - 17:30	FREE DISCUSSIONS	

WEDNESDAY 2024.11.20

9th FLOOR OF THE LIBRARY

TIME	NAME	TOPIC
09:00 - 09:30	WELCOME COFFEE	
09:30 - 10:00	ALEXANDER BEZNOSIKOV	Various aspects of efficient optimization:stochasticity, adaptivity, distributedness
10:00 - 10:30	TIANGANG CUI	Tensor-Train Methods for Sequential State, Parameter Estimation in State-Space Models
10:30 - 11:00	SEMYON DOROKHIN	Tensor-Based XL-MIMO Framework Design and Practical Challenges
11:00 - 11:30	COFFEE BREAK	
11:30 - 12:00	RAYMOND CHAN	Selecting Regularization Parameters for Minimization Problems in Imaging
12:00 - 12:30	LIEVEN DE LATHAUWER	From tensor-based blind source separation to tensor-based blind source matching
12:30 - 14:00	LUNCH	
14:00 - 14:30	VLADIMIR KAZEEV	Low-rank tensors for leveraging hidden structure in PDE problems
14:30 - 15:00	ALEXANDER HVATOV	Nonlinear dimensions and how to discover them automatically
15:00 - 15:30	CLOSING REMARKS	

MONDAY
2024.11.18

EUGENE TYRTYSHNIKOV

Institute of numerical mathematics
of the russian academy of sciences (INM RAS),
Moscow, Russia

09:30 - 10:00

Tensors and optimization

We consider optimization problems in connection with tensor decompositions and a general global optimization algorithm based on Tensor Train decompositions

MONDAY
2024.11.18

MICHAEL KWOK-PO NG

Lingnan University, Hongkong, China

10:00 - 10:30

Tensor Representations in Data Science

Higher-order tensors are suitable for representing multi-dimensional data in real-world, e.g., color images and videos, low-rank tensor representation has become one of the emerging areas in machine learning and computer vision. However, classical low-rank tensor representations can solely represent multi-dimensional discrete data on meshgrid, which hinders their potential applicability in many scenarios beyond meshgrid. In this talk, we discuss the recent development of tensor representations in data science. Both theoretical results and numerical examples are presented to demonstrate the usefulness of tensor representations.

MONDAY
2024.11.18

IVAN OSELEDETS

Laboratory of Computational Intelligence,
Center for Artificial Intelligence Technology,
Moscow, Russia

10:30 - 11:00

Matrix and tensor methods in machine learning.

In this talk, I will describe the questions related to the efficient usage of linear and multilinear algebra to different topics in machine learning, including generative models, compression of neural networks.
I will also discuss existing problems and possible future research directions.

MONDAY
2024.11.18

ANDREY KRYLOV

Moscow State University, Moscow, Russia

11:30-12:00

Using Mathematical Models in Deep Learning

The last few years have been characterized by a very rapid development of neural network methods, using, among other things, mathematical models built to describe the analyzed process or object. An example is neural networks that take into account the physics of the process is Physics-Informed Neural Networks (PINN), which allow you to build a neural network solution for non-linear partial differential equations. At the same time, during training, the loss function takes into account the requirement to satisfy the differential equation on the grid and its initial and boundary conditions. The report analyzes this approach, its advantages and disadvantages.

More promising, according to many authors, is a new approach based on the use of so-called neural operators. The most commonly used in practice are the Fourier neural operator (FNO) and the neural operator DeepONet. These neural operators are based on the use of the theorem on the approximation of an arbitrary continuous nonlinear operator by a perceptron with any given accuracy. Of great interest are also methods based on the use of Kolmogorov-Arnold neural networks (KAN), which arose this year and are based on the use of the representation theorem, which states that each multidimensional continuous function can be represented as a superposition of continuous functions of one variable.

The report provides practical examples of using PINN, FNO and KAN in problems of analyzing the results of diagnostics of materials, solving the problem of numerical differentiation of noisy data and suppressing Gibbs oscillations in images that occur in medical MRI images.

MONDAY
2024.11.18

YIMIN WEI

Fudan University, Shanghai, China

12:00-12:30

Tensor-based model reduction and identification for generalized memory polynomial

Abstract: Power amplifiers (PAs) are essential components in wireless communication systems, and the design of their behavioral models has been an important research topic for many years. The widely used generalized memory polynomial (GMP) model suffers from rapid growth in the number of parameters with increasing memory depths and nonlinearity order, which leads to significant increase in model complexity and the risk of overfitting. In this study, we introduce tensor networks to compress the unknown coefficient tensor of the GMP model, resulting in three novel tensor-based GMP models. These models can achieve comparable performance as the GMP model, but with much fewer parameters and lower complexity. For the identification of these models, we derive alternating least-squares (ALS) method to ensure the rapid updates and convergence of model parameters in an iterative manner. In addition, we notice that the horizontal slices of the third-order data tensor constructed from the input signals are Vandermonde matrices, which have numerical low-rank structure. Hence, We further propose the RP-ALS algorithm, which first performs a truncated higher-order singular value decomposition on the data tensor to generate random projections, then conducts the ALS algorithm on downscaled multilinear least-squares problems, thus reducing the computational effort of the iterative process. The experimental results show that the proposed models outperform the full GMP model and sparse GMP model via LASSO regression in terms of the reduction in the number of parameters and running complexity.

MONDAY
2024.11.18

SERGEY KABANIKHIN

Sobolev Institute of Mathematics
Siberian Branch of Russian Academy of Science,
Novosibirsk, Russia

14:00 - 14:30

Regularization: from Algebraic Equations to Neural Networks

Regularization is one of the most effective way to construct approximation for the solution of inverse ill-posed problems of mathematical physics and neural networks. The fundamental instrument of regularization is linear algebra. Methods of classical regularization can help to construct the best possible matrix of weights of linear part of the neural networks in a sense of normal and\or pseudo-solutions.

MONDAY
2024.11.18

CARMINE DI FIORE

University of Rome Tor Vergata, Rome, Italy

14:30 - 15:00

Insights on some theoretical results related to the convergence of power and Gauss-Seidel iterations

While studying the rate of convergence of power method applied to A^{-1}, where A is a $n\times n$ not diagonalizable invertible matrix, one needs to consider the matrix Y which transforms A^{-1} into Jordan canonical form,and to express such Y in terms of the matrix X which transforms A into Jordan canonical form. The matrix Y turns out to be equal to XM with M block diagonal where the diagonal blocks have an interesting upper triangular Tartaglia-Toeplitz structure.
While searching for linear systems $Ax = b$ which are solvable via the Gauss-Seidel (G.-S.) method even if the standard sufficient conditions on A for G.-S. convergence are not satisfied, one meets an interesting class of 2×2-block matrices A,for which G.-S. converges for A every time it converges for its two diagonal blocks, of order r and $n-r$, and $|a_{r+1,1}a_{1,r+1}/a_{1,1}a_{r+1,r+1}|<1$.
The above are two examples of how in-depth studies of classical subjects of numerical mathematics can lead to new interesting remarks. E. Bozzo , P. Deidda, C. Di Fiore, The Jordan and Frobenius pairs of the inverse LINEAR & MULTILINEAR ALGEBRA, 2023

TUESDAY
2024.11.19

ALEXANDER OSINSKY

Marchuk Institute of Numerical Mathematics
of the Russian Academy of Sciences (INM RAS),
Moscow, Russia

09:30 - 10:00

Speed-accuracy trade-off in cross matrix approximations

When constructing cross matrix approximations, the key property, connected to the approximation accuracy, is the volume of the submatrix at the intersection of the selected rows and columns. Many methods of constructing cross approximations can be considered greedy volume maximization algorithms. However, to build such approximations as quickly as possible, approaches that do not reach maximum volume are often used. During this talk, we will examine exactly which simplifications are used in cross methods, that frequently lead to a loss of accuracy, and how to avoid them without significantly increasing the computational complexity of the algorithms. The possibility of using cross approximations as an approximate SVD will also be considered, and in which cases such use is acceptable.

TUESDAY

2024.11.19

DMITRY YAROTSKY

Laboratory of Mathematical Foundations
of Artificial Intelligence, Center for Artificial Intelligence
Technology, Moscow, Russia

10:00 - 10:30

SGD with memory: theoretical acceleration of mini-batch SGD for overparameterized models

Modern predictive models are usually overparametrized so that the spectrum of the loss minimization problem has a long tail of eigenvalues converging to 0. It is well-known both theoretically and empirically that Gradient Descent without noise in this setting can be substantially accelerated by introducing momentum and performing Heavy Ball with a suitable schedule. However, when using Stochastic GD, sampling noise causes this algorithm to diverge. We show theoretically and experimentally that SGD can nevertheless be accelerated by considering a broader class of SGD with memory.

TUESDAY
2024.11.19

ALEXANDER MIKHALEV

Laboratory “Multiscale Neurodynamics for Intelligent Systems”, Center for Artificial Intelligence Technology, Moscow, Russia

10:30 - 11:00

NNTile: towards dynamic task-based parallelism for training and inference of large neural networks

A will to train and apply extremely large neural networks requires humanity to construct extremely large supercomputers. The larger supercomputer is, the harder it is to keep hardware utilization ratio at an appropriate level. Fortunately, many groups of scientists and engineers work on the problem. Parallel processing decisions made by humans are based on lots of assumptions, but are nearly optimal. To overcome limits of all those assumptions we propose to switch to a task-based parallel programming model. It allows to rely on an automated decision-making heuristics. Once mastered, such automated tools, at the first glance, would allow us to get good performance automatically on any hardware, be it homogeneous or heterogeneous. But the main profit is yet to be uncovered in the future — in a case of hardware failure there will be no need to restart entire process of training from a checkpoint, but dynamic parallelism will just rebalance the load. In this talk we present our software, titled NNTile, capable of training GPT2 and LLaMa models on a single heterogeneous node within the task-based parallelism. The software is yet at its early stage, but it is in active development.».

TUESDAY
2024.11.19

GIANLUCA CERUTI

University of Innsbruck, Innsbruck, Austria

11:30 - 12:00

Recent Developments in Tensor Techniques for Dynamical Low-Rank Approximation

In recent years, tensor techniques have emerged as powerful tools for addressing the challenges posed by high-dimensional data, often referred to as the curse of dimensionality. This phenomenon involves a sudden and exponential increase in the number of parameters required to accurately describe a system's state. Fields that frequently encounter this issue include complex dynamical systems, such as molecular kinetics, where the degrees of freedom necessary to depict the system's state are extensive.

One of the most promising approaches to managing this data explosion is the use of Dynamical Low-Rank Approximation (DLRA). DLRA aims to minimize the residual error of the system within a predefined search space, providing an efficient way to handle large-scale data and reduce memory usage without significantly compromising accuracy. Over the past decade, significant efforts have been made to systematically explore and develop tensorial numerical integrators for DLRA. These efforts have led to numerous advancements in the field, resulting in improved methodologies and expanded applications.

This talk will delve into the latest advancements in this field, highlighting novel methodologies and applications. We will explore state-of-the-art algorithms that enhance the efficiency of using tensor decompositions, particularly in the context of time-dependent problems. The focus will be on how these techniques address the challenges of large-scale data by reducing computational complexity and memory requirements. Key topics include an innovative approach encompassing advanced tensorial decompositions, such as tensor train decompositions and hierarchical Tucker formats, with applications in molecular quantum dynamics, kinetic equations, and computer vision for machine learning.

TUESDAY
2024.11.19

DARIO FASINO

University of Udine, Udine, Italy

12:00 - 12:30

Computing the stationary probability vector of stochastic tensors and higher-order Markov chains

Markov chains and random walks on graphs have long provided a solid theoretical foundation for modelling diffusion and navigation dynamics in complex networks. However, their limited expressiveness hinders understanding more complex dynamics involving memory terms and high-dimensional data. For these reasons, attention has shifted to more sophisticated stochastic models capable of handling long-term dependencies and nonlinearities. These models are described through stochastic cubic tensors, i.e. three-index arrays of non-negative numbers whose column sum equals 1. Still, fundamental questions such as the convergence of the process to a limiting distribution and the uniqueness of such a limit are the subject of rich recent literature.
Ergodicity coefficients for stochastic tensors can provide a partial but effective answer to these questions. They yield explicit formulas that guarantee the uniqueness of nonlinear Perron eigenvectors of stochastic tensors of order three, provide bounds on the sensitivity of such eigenvectors, and ensure the global convergence of various higher-order or non-linear stochastic processes to the stationary distribution. Time permitting, we also present a randomized Kaczmarz-type method to compute stationary probability vectors of stochastic cubic tensors. Compared with other iterative methods, this can reduce memory requirement and computational cost at the price of reduced accuracy, a worthwhile compromise in machine learning and data mining applications.

TUESDAY
2024.11.19

SERGEY MATVEEV

Marchuk Institute of Numerical Mathematics of the Russian Academy of Sciences (INM RAS), Moscow, Russia

14:00 - 14:30

Nonnegative low-rank tensor approximations: alternating projections and fast corrections

In this talk we will discuss a summary on recent progress in development of low-rank nonnegative matrix and tensor approximations. In particular, we will show that application of the quasi-optimal projection methods does not lead to significant loss of accuracy in the alternating projections method but allows to accelerate the whole computational process significantly. Further, we will show that straight-forward generalization of the alternating projections procedure allows also to deal with the Tucker and tensor train formats successfully. All in all, we will show that a simple heuristics based on the efficient search of the global maximal element within TT-format allows us to clean up the whole tensor from the artificial negative elements by application of simple rank-1 corrections without any operation having exponential complexity.

TUESDAY
2024.11.19

XIAOMAN HU

Huawei

14:30- 15:00

The architecture and Ecosystem of openMind AI platform

OpenMind is a full-process application enablement suite that can help enterprises and developers achieve end-to-end development training and inference deployment, saving time and costs. This presentation will show openMind technical architecture and ecosystem development in detail.

TUESDAY
2024.11.19

MAXIM SHISHLENIN

Sobolev Institute of Mathematics Siberian Branch of Russian Academy of Science, Novosibirsk, Russia

15:00 - 15:30

Matrix methods, block-Toeplitz matrices in inverse problems of wave tomography.

The development of direct methods for solving inverse problems is a very important problem in wave tomography applications. The main idea is to reduce inverse problems after discretization to systems of linear algebraic equations with ill-posed or block-toeplitz matrices of huge dimension. Two approaches are considered. One approach is to reduce the coefficient inverse problem to a system of integral equations, the solution of which is reduced to solving systems of linear algebraic equations with a block-Toeplitz matrix of huge dimension. Another approach is based on the method of lines: using spatial approximation, we proceed to a system of ordinary differential equations, which is solved by the matrix method. One of the methods for inverting huge matrices is low-rank approximation for solving real-time wave tomography problems.

TUESDAY
2024.11.19

CHAO WANG

Southern University of Science and Technology (SUSTech), Shenzhen, China

16:00 - 16:30

A Scale-Invariant Relaxation in Low-Rank Tensor Recovery

In this talk, we consider a low-rank tensor recovery problem. Based on the tensor singular value decomposition (t-SVD), we propose the ratio of the tensor nuclear norm and the tensor Frobenius norm (TNF) as a novel nonconvex surrogate of tensor's tubal rank. The rationale of the proposed model for enforcing a low-rank structure is analyzed as its theoretical properties. Specifically, we introduce a null space property (NSP) type condition, under which a low-rank tensor is a local minimum for the proposed TNF recovery model. Numerically, we consider a low-rank tensor recovery problem and employ the alternating direction method of multipliers (ADMM) to secure a model solution with guaranteed subsequential convergence under mild conditions. Extensive experiments demonstrate the superiority of our proposed model over state-of-the-art methods

TUESDAY
2024.11.19

SERGEY DOLGOV

University of Bath, Bath, England

16:30 - 17:00

Tensor train approximation of deep transport maps for Bayesian inverse problems.

We develop a deep transport map for sampling concentrated distributions defined by an unnormalised density function. We approximate the target distribution as the pushforward of a reference distribution under a composition of transport maps formed by tensor-train approximations of bridging densities. We propose two bridging strategies: tempering the target density, and smoothing of an indicator function with a sigmoid. The latter opens the door to efficient computation of rare event probabilities in Bayesian inference problems.

WEDNESDAY

2024.11.20

ALEXANDER BEZNOSIKOV

Ivannikov Institute for System Programming (ISP RAS), Moscow, Russia

09:30 - 10:00

Various aspects of efficient optimization: stochasticity, adaptivity, distributedness

Finding a minima or maxima is an integral part of many application problems from resource allocation to neural network training. Therefore, optimization methods that solve these kinds of problems also play an important role in modern science. In the course, we will look at several stories around the state of the art in this field. First, we will talk about classical methods on the basis of which all new results are being developed. We will also discuss stochastic optimization algorithms and understand the intuition behind the SGD and Adam methods. Finally, we will consider stories from distributed optimization and federated learning.

WEDNESDAY
2024.11.20

TIANGANG CUI

University of Sydney, Sydney, Austarlia

10:00 - 10:30

Tensor-Train Methods for Sequential State and Parameter Estimation in State-Space Models

Numerous real-world applications require the estimation, forecasting, and control of dynamic systems using incomplete and indirect observations. These problems can be formulated as state-space models, where the challenge lies in learning the model states and parameters from observed data. We present new tensor-based sequential Bayesian learning methods that jointly estimate parameters and states. Our methods provide manageable error analysis and potentially mitigate the particle degeneracy encountered in many particle-based approaches. Besides offering new insights into algorithmic design, our methods naturally incorporate conditional transports, enabling filtering, smoothing, and parameter estimation within a unified framework.

WEDNESDAY
2024.11.20

SEMYON DOROKHIN

Huawei

10:30 - 11:00

Tensor-Based XL-MIMO Framework Design and Practical Challenges

Current trend of MIMO evolution is the increase both in bandwidth and antenna number, turning it into Extra Large MIMO (XL-MIMO) and making tensor processing a possible candidate. Although such evolution potentially offers better performance, channel estimation overhead with traditional methods becomes prohibitively large. Traditionally, Base Station needs to estimate channel matrices for each frequency "subband" and calculate singular vectors of these matrices. Alternatively, we can unite these matrices into one tensor and choose some low-parametric decomposition. Channel estimation can be replaced by parameters retrieval and singular vectors can be calculated directly from the estimated parameters.
Being a hot topic in current research, this idea is usually implemented by choosing CPD as low-parametric decomposition and merging reference signal transmission, antenna array scanning configuration and CPD factors into one joint optimization problem. As a result, existing solution impose rather tough restrictions on the reference signal allocation in time, antenna and frequency domains.
This talk will introduce a general framework for wireless tensor-based processing, featuring separation of channel elements estimation stage and tensor completion stage. This separation allows to use slightly modified reference signals from 3GPP standard to measure elements of MIMO OFDM channel tensor. Next step, we can reuse tensor completion theory and methods from other fields like fMRI scanning, image and video reconstruction etc. Simulation results based on geometry-based stochastic models demonstrate the potential of this approach. Finally, further challenges of the proposed framework will be outlined.

WEDNESDAY
2024.11.20

RAYMOND CHAN

Lingnan University, Hongkong, China

11:30 - 12:00

Selecting Regularization Parameters for Minimization Problems in Imaging.

Many minimization problems in imaging consist of a data-fitting term and a regularization term. The weights of the two terms are balanced by a regularization parameter. In this talk we discuss how to choose this parameter when it has an L2 data-fitting term and the regularization term is either an L1-or a nuclear-norm term.

WEDNESDAY
2024.11.20

LIEVEN DE LATHAUWER

KU Leuven, Leuven, Belgium

12:00 - 12:30

From tensor-based blind source separation to tensor-based blind source matching

In this talk we take the step from tensor-based data analysis to tensor-based data comparison, and from the decomposition of a single tensor to the assessment of the similarity between components of different tensors.
It is well-known that Canonical Polyadic Decomposition (CPD) and Block Term Decomposition (BTD) of higher-order tensors are fundamental tools for blind signal separation and latent variable analysis, “beyond matrix techniques”.
On the other hand, assessing similarity is a key task in pattern recognition and machine learning. We will show that tensors also provide fundamentally new possibilities for blind similarity assessment and latent variable matching, “beyond matrix pair techniques”. Moreover, under mild conditions, the assessment of similarity can be done by conventional linear algebra. The number of terms and their multilinear rank (in the case of BTD) can be found as well. The results will be illustrated with applications.

WEDNESDAY
2024.11.20

VLADIMIR KAZEEV

University of Vienna, Vienna, Austria

14:00 - 14:30

Low-rank tensors for leveraging hidden structure in PDE problems

The approximation of complex data in suitable low-dimensional subspaces is fundamental to scientific computing and data science: many applications exhibit hidden low-dimensional structure in some form, allowing for storage and computations in terms of relatively few parameters. In the context of PDEs, much of the work of traditional numerical analysts consists in proving the presence of such structure in specific classes of PDE problems. That is typically done in a constructive way: a numerical method (approximation subspaces) for the chosen problem class is designed, and the properties of the method (of the subspaces) with respect to approximation, stability and convergence are analyzed.

In contrast, low-rank tensor decompositions, relying on well-established techniques of numerical linear algebra and optimization, realize the adaptive low-parametric approximation of solutions. That allows for discovering efficient approximation subspaces computationally and thereby for dramatically reducing the complexity of numerical solvers. One such a decomposition was proposed under the names of «matrix product states» (MPS) in computational quantum physics and «tensor train» (TT) in computational mathematics. In particular, the multilevel MPS-TT representation, building on the classical idea of Kronecker-product multilevel approximation, allows to handle generic but extravagantly large discretizations and leads to data-driven computations based on effective discretizations adapted to the problem class and to the data instead of problem-dependent discretizations (approximation spaces) designed analytically. This approach has been shown, both theoretically and experimentally, to lead to the efficient approximation of algebraic singularities, boundary layers and high-frequency oscillations arising in multiscale diffusion problems, achieving root-exponential convergence with respect to the total number of representation parameters.

In this talk, we will give an overview of the approach and discuss the stability of low-rank MPS-TT representations and the quasi-optimality of the associated approximation algorithms. As we will show, these points extend to the recently proposed low-rank multilevel frame representation of functions. We will elaborate on its application for the numerical solution of PDE problems, where the low-rank adaptivity is used to resolve the multiscale structure of the data and of the solution.

WEDNESDAY
2024.11.20

ALEXANDER HVATOV

University of Information Technologies,
Mechanics and Optics (ITMO),
Saint Petersburg, Russia

14:30 - 15:00

Nonlinear dimensions and how to discover them automatically

In data science, low-rank approximations provide linear projections for dimension reduction, but the manifold assumption in many datasets suggests that more flexible, nonlinear projections may better capture their structure. This talk will explore methods that go beyond traditional low-rank tensor decompositions to achieve nonlinear dimensionality reduction that optimizes for the manifold structures inherent in the data. I will discuss modern techniques for automated optimization of distance graphs, enabling adaptive, geometry-aware representations that optimize model performance in various applications.