HDA2023 Talk Abstracts

Name: Helmut Harbrecht
University/Institution: Department of Mathematics and Computer Science, University of Basel, Switzerland
Talk title: On tensor approximation of continuous functions
Time: 13:30 Monday 20th
Abstract: In this talk, we analyze tensor approximation schemes for continuous functions. We assume that the function to be approximated lies in either a standard Sobolev space or in a Sobolev space with dominating mixed smoothness. Then, we discuss the cost when approximating this function in a tensor format such as the tensor train format or the H-Tucker format.

Name: Marcin Wnuk
University/Institution: Osnabrueck University
Talk title: Why the ℓ_2/ℓ_∞ embedding cannot be solved by randomized algorithms
Time: 14:00 Monday 20th
Abstract: We are working in the framework of information-based complexity, the basic notions will be explained during the talk. Let S be a bounded linear operator between some separable Banach spaces. We say that S is solvable by a class of algorithms A if there exists a sequence $(A_n)_n$ of algorithms from A such that the errors made by $(A_n)_n$ when approximating S converge to 0. The definition of the error may, and usually does, depend on the class A under consideration.

It is well-known that the class of operators solvable by deterministic algorithms coincides with the compact operators. The main question we ask is: which operators are solvable by randomized algorithms? When proving negative results we rely on the Bakhvalov technique. In particular the results of S. Heinrich from "Lower bounds for the complexity of Monte Carlo function approximation" turn out to be very useful. Applying them we were able to obtain some elementary results on non-solvability of certain operators. The introductory part of the talk will be devoted to presenting those results.

In the main part of the talk I would like to focus on one of the simplest operators for which the method developed by Heinrich fails, namely when S is the indentical embedding of the sequence space ℓ_2 into ℓ_∞. Careful analysis of the finite-dimensional subproblems shows that for some constant $C > 0$ obtaining in the m-dimensional subproblems error smaller than C requires evaluating roughly $\log(m)$ (for the non-adaptive randomized algorithms) or $\log(\log(m))$ (for the adaptive randomized algorithms) functionals. Thus we cannot approximate S with error smaller than C using only finitely many functionals. If time permits I would like to go into some details of the proof drawing upon the classical information theory.

The talk is based on the joint work with Robert Kunsch (RWTH Aachen) and Erich Novak (University of Jena).

Name: Robin Rüßmann
University/Institution: TU Kaiserslautern
Time: 14:30 Monday 20th

Talk title: Equivalence of L^2-Approximation on Gaussian Spaces and Hermite Spaces

Abstract: We consider two types of reproducing kernel Hilbert spaces $H(L_\sigma)$ and $H(K_\beta)$ of d variables, where $d \in \mathbb{N}$ or $d = \infty$. Here, L_σ is given as the tensor product of univariate Gaussian kernels, i.e., $L_\sigma(x, y) := \prod_{j=1}^{d} \exp(-\sigma_j^2 \cdot (x_j - y_j)^2)$, and K_β is given as the tensor product of certain univariate Hermite kernels, i.e., $K_\beta(x, y) := \prod_{j=1}^{d} \sum_{\nu=0}^{\infty} \beta_j^\nu \cdot h_\nu(x_j) \cdot h_\nu(y_j)$, where h_v is the νth Hermite polynomial.

If the parameters σ and β are related in a certain way, we give a constructive one-to-one correspondence between algorithms for L^2-Approximation on $H(L_\sigma)$ and $H(K_\beta)$, preserving worst-case-error and cost. Using this, we investigate L^2-Approximation on $H(L_\sigma)$ in the case $d = \infty$. In the case $d \in \mathbb{N}$, we are able to apply known results form $H(L_\sigma)$ to $H(K_\beta)$ and vice versa in a constructive way. Similar results hold for the integration problem.

Name: Sifan Liu
University/Institution: Stanford University
Talk title: Pre-integration and Dimension Reduction with Generalized Active Subspaces
Time: 15:30 Monday 20th
Abstract:

Pre-integration and dimension reduction are useful techniques for improving randomized quasi-Monte Carlo (RQMC) accuracy. To choose the variable to pre-integrate with, one needs to consider both the variable importance and the tractability of the pre-integration step. For integrals over a Gaussian distribution, one can potentially pre-integrate over any linear combination of variables. I will talk about how to find the important direction to pre-integrate with by active subspaces while incorporating certain constraints so that the pre-integration step has a closed form. Meanwhile, this generalized active subspaces method can reduce the effective dimension of the integrand. The proposed algorithm is shown to work efficiently in some financial applications, including pricing exotic options and computing Greeks. In particular, the proposed method brings a big variance reduction for pricing Asian options under stochastic volatility models.

Name: Peter Kritzer
University/Institution: RICAM, Austrian Academy of Sciences
Talk title: Fast computation of matrix-vector products using reduced lattice points
Time: 16:00 Monday 20th
Abstract:

In various applications (e.g., mathematical finance or PDEs with random coefficients) one is interested in approximating integrals of the form

$$\int_D f(\boldsymbol{x}^\top A) \mathrm{d}\mu(\boldsymbol{x}),$$

for a domain $D \subseteq \mathbb{R}^s$, an $s \times t$ matrix $A \in \mathbb{R}^{s \times t}$, and a function $f : D \to \mathbb{R}$, by quasi-Monte Carlo integration rules of the form

$$Q_N(f) = \frac{1}{N} \sum_{k=0}^{N-1} f(\boldsymbol{x}_k^\top A).$$

We are interested in situations where the main computational cost of computing $Q_N(f)$ arises from the vector-matrix multiplication $\boldsymbol{x}_k^\top A$ for all N points, which requires $\mathcal{O}(N s\, t)$ operations.

It was shown by Dick, Kuo, Le Gia, and Schwab that, when using particular types of QMC rules, the cost to evaluate $Q_N(f)$ can be reduced to only $\mathcal{O}(t\, N \log N)$ operations. This drastic reduction in computational cost (as long as $\log N \ll s$) is achieved by a fast matrix-matrix multiplication that exploits the fact that for the chosen point sets the matrix X can be re-ordered to be of circulant structure. The fast multiplication is then realized by the use of the fast Fourier transformation (FFT).

In our talk, we will outline a different method which can also drastically reduce the computation cost of evaluating $Q_N(f)$. The reduction in computational complexity is achieved by using point sets which possess a certain repetitiveness in their components.

Name: Weiwen Mo
University/Institution: KU Leuven
Talk title: Constructing Embedded Lattice-based Algorithms to Approximate Multivariate Function with a Composite Number of Points
Time: 16:30 Monday 20th
Abstract:
Rank-1 lattice rules, a main family of QMC rules, are characterised by generating vectors, and the component-by-component (CBC) construction is an efficient way to construct the generating vector. In this talk, we introduce some new results in approximating multivariate one-periodic functions using function values at rank-1 lattice points. The number of points is not limited to a prime number as in currently available literature, but can take any composite value. The new results cannot be trivially generalised from existing results for prime number and a new proof technique is required. We also provide new CBC algorithm to construct embedded lattice sequences for a range of number of points. We theoretically show that embedded lattice sequences are essentially as good as lattice rules with the same number of points. With some modifications, the search criterion in the CBC construction under L_2 norm can also be applied to L_∞ norm, which allows fast CBC imp lementations for both L_2 and L_∞ norms.

Name: Tiangang Cui
University/Institution: Monash
Talk title: Self-reinforced Knothe–Rosenblatt rearrangements for high-dimensional stochastic computation
Time: 9:30 Tuesday 21st
Abstract:
Characterising intractable high-dimensional random variables is a fundamental task in stochastic computation. It has broad applications in statistical physics, machine learning, uncertainty quantification, and beyond. The recent surge of transport maps offers new insights into this task by constructing variable transformations that couple intractable random variables with tractable reference random variables. In this talk, we will present numerical methods that build the Knothe–Rosenblatt (KR) rearrangement— a family of transport maps in a triangular form—in high dimensions. We first design function approximation tools to realise the KR rearrangement that ensures the order-preserving property with controlled statistical errors. We then introduce a self-reinforced procedure to adaptively precondition the construction of KR rearrangements to significantly expand their capability of handling random variables with complicated nonlinear interactions and concentrated density functions. We demonstrate the efficiency of the resulting self-reinforced KR rearrangements on applications in statistical learning and uncertainty quantification, including parameter estimation for dynamical systems, PDE-constrained inverse problems, and rare event estimation.

Name: Abi Srikumar
University/Institution: UNSW Sydney
Time: 10:00 am Tuesday 21st
Talk title: Approximating distribution functions in uncertainty quantification using quasi-Monte Carlo methods
Abstract: As high-dimensional problems become increasingly prevalent in many applications, the effective evaluation of these problems within the limits of our current technology poses a great

hurdle due to the exponential increase in computational cost as dimensionality increases. One class of strategies for evaluating such problems efficiently are quasi-Monte Carlo (QMC) methods. Recently the application of quasi-Monte Carlo methods to approximate expected values associated with solutions to elliptic partial differential equations with random coefficients in uncertainty quantification has been of great interest. In this talk, we look into extending this from the computation of expected values of functionals of the PDE solution to the approximation of distribution functions. This done by reformulating these functions as expectations of an indicator function. However due to the discontinuous nature of the indicator functions, we do not have an integrand that is conducive to obtaining the optima l rate of error convergence. We seek to alleviate this issue using preintegration, whereby we integrate out a single variable of the discontinuous function in order to obtain a function of one dimension less with a sufficient level of smoothness to apply QMC methods. Some theoretical results regarding the error bounds associated with such approximations and the results of numerical experiments will be presented.

Name: Andreas Rupp
University/Institution: LUT University
Talk title: Quasi-Monte Carlo methods and discontinuous Galerkin
Time: 11:00 Tuesday 21st
Abstract: In this talk, we design and develop Quasi-Monte Carlo (QMC) cubatures for non-conforming discontinuous Galerkin approximations of elliptic PDE problems with random coefficients. In particular, we are interested in using QMC cubatures to compute the response statistics (expectation and variance) of the discretized PDE problem and we derive rigorous QMC convergence rates for this problem.

Name: Ian H. Sloan
University/Institution: University of New South Wales, Sydney Australia
Talk title: Periodic kernel-based high-dimensional approximation
Time: 11:30 Tuesday 21st
Abstract:
This talk describes a recent periodic-kernel-based QMC method for high-dimensional approximation. In this work, jointly with Vesa Kaarnioja, Yoshihito Kazashi, Frances Kuo and Fabio Nobile, an elliptic partial differential equation with a random input field is modelled in a non-standard way with periodic random variables together with kernel-based interpolation, giving a computational cost that grows merely linearly with dimensionality. The method is feasible even with hundreds of parameters in an input random field.

Name: Alexander Gilbert
University/Institution: UNSW Sydney
Talk title: Theory and construction of lattice rules after preintegration for option pricing
Time: 12:00 Tuesday 21st
Abstract:
Since the 1990's it has been known that quasi-Monte Carlo (QMC) methods are extremely effective at computing the integrals required for pricing Asian options, despite the fact that the kink in the payoff function means that such problems are not smooth enough for the QMC theory to apply. A popular remedy is to first smooth the payoff by preintegration, which is a specific case of the more general method of conditional sampling. Here one first integrates the non-smooth payoff with respect a specially chosen variable (or equivalently conditions on partial information), with the result being a smooth function, but now in one dimension less. In this talk we look at pricing an Asian option by performing preintegration then applying a randomly shifted lattice rule to the result. By studying the

regularity of the payoff after preintegration, we can choose appropriate weights for each specific Asian option problem, which in turn allows tailored randomly shifted lattice rules to be constructed using the component-by-component algorithm. We give rigorous error bounds with first-order convergence and which are explicit in the dependence on dimension. Along the way I will also highlight a key step in the analysis: an equivalence between the Sobolev spaces used to analyse preintegration and the weighted spaces used to analyse randomly shifted lattice rules.

Name: Rüdiger Kempf
University/Institution: University of Bayreuth
Talk title: The Double-Adaptive Tensor Product Multilevel Method
Time: 13:30 Tuesday 21st
Abstract:

The *Tensor Product Multilevel Method (TPML)* is a combination of Smolyak's construction for a high-directional approximation operator with low-dimensional kernel-based multilevel schemes. This yields a high-order method which allows scattered data approximation on relatively arbitrary high-dimensional domains.

Both building blocks of TPML have adaptive versions which can be combined in the same way as their non-adaptive ones. This allows us to iteratively build the index set the Smolyak algorithm uses and/or adaptively choose the sites the multilevel scheme uses. We call the resulting version of TPML the *Double-Adaptive Tensor Product Multilevel Method*.

In this talk we briefly repeat the basics of TPML and introduce our ideas for the double-adaptive TPML. If time permits we show the numerical performance of the methods with the help of an example.

Name: Yiran Zhao
University/Institution: Monash University
Talk title: Tensor-based methods for sequential state and parameter estimation in state space models
Time: 14:00 Tuesday 21st
Abstract:

Numerous real-world applications involve the filtering problem: one aims to sequentially estimate the states of a (stochastic) dynamical system from incomplete, indirect, and noisy observations over time to forecast and control the underlying system. In addition to the filtering problem, it is often of interest to estimate some parameters that govern the evolution of the system. Both the filtering and the parameter estimation can be naturally formalized under the Bayesian framework.

However, the Bayesian solution poses some significant challenges. For example, the most widely used particle filters can suffer from particle degeneracy and the more robust ensemble Kalman filters rely on the rather restrictive Gaussian assumptions. The solution to parameter and state estimation in high-dimensional model is rather unsolved. Exploiting the interplay between the low-rank tensor structure and Markov property of the filtering problem, we present a tensor-train-based approach for tackling Bayesian filtering and parameter estimation altogether.

Our approach aims at exact Bayesian solutions and does not suffer from particle degeneracy. Applications on several challenging numerical examples are presented.

Name: David Krieg
University/Institution: Johannes Kepler University, Linz, Austria
Talk title: Uniform recovery of periodic functions
Time: 14:30 Tuesday 21st
Abstract:

We consider the problem of function recovery for classes of high-dimensional periodic and continuous functions in the uniform norm. We show that algorithms based on function values (samples) are as powerful as algorithms based on general linear information (like Fourier coefficients). Moreover, there is a simple algorithm that achieves almost the optimal error with high probability. We also discuss a result for classes of non-periodic functions.

Name: Andreas Zeiser
University/Institution: HTW Berlin, Germany
Time: 15:30 Tuesday 21st
Talk title: Dynamical low-rank approximation of Vlasov-Poisson equations on polygonal spatial domains
Abstract: We consider dynamical low-rank approximation (DLRA) for the numerical simulation of Vlasov-Poisson equations based on separation of space and velocity variables, as proposed in several recent works. The standard approach for the time integration in the DLRA model uses a splitting of the tangent space projector for the low-rank manifold according to the separated variables. It can also be modified to allow for rank-adaptivity. A less studied aspect is the incorporation of boundary conditions in the DLRA model. We consider a variational formulation of the projector splitting which allows to handle inflow boundary conditions on piecewise polygonal spatial domains. Numerical experiments demonstrate the principle feasibility of this approach.

Name: Paolo Climaco
University/Institution: University of Bonn
Talk title: Investigating the effects of minimising the training set fill distance in machine learning regression
Time: 16:00 Tuesday 21st
Abstract:
Machine learning (ML) regression methods are powerful tools leveraging large datasets for nonlinear function approximation in high-dimensional domains. Unfortunately, learning from large datasets may be unfeasible due to computational limitations and the cost of producing labels for the data points, as in the case of quantum-chemistry applications. Therefore, an important task in scientific ML is sampling small training sets from large pools of unlabelled data points to maximise models performance while keeping the computational effort of the training and labelling processes low. In this talk, we analyse the advantages of a common approach for training set sampling aiming to minimise the fill distance of the selected set, which can be considered a measure of the data distribution. We provide theoretical results that minimising the training set fill distance reduces the worst-case approximation error of several ML models, which can be interpreted as an indicator of the robustness of a model's approximation. We support our theoretical results with experiments considering various ML techniques, such as Kernel methods and Neural Networks. The application context is quantum-chemistry, where different datasets have been generated to develop effective ML techniques and approximate the functions mapping high-dimensional representations of molecules to their physical and chemical properties.

Name: Jochen Garcke
University/Institution: University of Bonn
Talk title: Learning Embeddings of Deformation Patterns of Surface Meshes
Time: 16:30 Tuesday 21st
Abstract:
The underlying dynamics and patterns of 3D surface meshes deforming over time can be discovered by unsupervised learning, especially autoencoders, which calculate low-dimensional embeddings of the

surfaces. To study the deformation patterns of unseen shapes by transfer learning, we want to train an autoencoder that can analyze new surface meshes without training a new network. Here, most state-of-the-art autoencoders cannot handle meshes of different connectivity and therefore have limited to no generalization capacities to new meshes. Also, reconstruction errors strongly increase in comparison to the errors for the training shapes. To address this, we propose novel spatial and spectral CoSMA (Convolutional Semi-Regular Mesh Autoencoder) network, respectively.

Name: Dinh Dũng
University/Institution: Vietnam National University, Hanoi
Talk title: Numerical Gaussian-weighted integration of functions having mixed smoothness
Time: 9:00 Wednesday 22nd
Abstract:

We investigate the multivariate numerical approximation of Gaussian-weighted integrals over $\mathbb{R}^d$ for integrands in the unit ball of Sobolev spaces of functions having mixed smoothness bounded in the Gaussian-weighted p-th integral norm. For $1 < p < \infty$, we prove the right convergence rate of optimal quadrature with respect to the number of integration nodes and propose a novel method for constructing asymptotically optimal quadrature. For $p = 1$, we prove tight upper and lower bounds (with a difference in additional logarithm terms) of the error for optimal quadrature and show that the upper bound is performed by a sparse-grid quadrature with integration nodes on step hyperbolic crosses in the function domain. This talk is based on the following works:

[1]. D. Dũng and V. K. Nguyen. Optimal numerical integration and approximation of functions on $\mathbb{R}^d$ equipped with Gaussian measure, arXiv:2207.01155 [math.NA], 2022.

[2]. D. Dũng, Numerical weighted integration of functions having mixed smoothness, arXiv:2208.0908 [math.NA], 2022.

Name: Guanglian Li
University/Institution: The University of Hong Kong
Talk title: Quasi-Monte Carlo finite element approximation of the Navier–Stokes equations with initial data modeled by log-normal random fields
Time: 9:30 Wednesday 22nd
Abstract:

We analyze the numerical approximation of the Navier–Stokes problem over a polygonal domain in $\mathbb{R}^2$, where the initial condition is modeled by a log-normal random field. This problem usually arises in the area of uncertainty quantification. We aim to compute the expectation value of linear functionals of the solution to the Navier–Stokes equations and perform a rigorous error analysis for the problem. In particular, our method includes the finite element, fully-discrete discretizations, truncated Karhunen–Loéve expansion for the realizations of the initial condition, and lattice-based quasi-Monte Carlo (QMC) method to estimate the expected values over the parameter space. Our QMC analysis is based on randomly-shifted lattice rules for the integration over the domain in high-dimensional space, which guarantees the error decays with $\mathcal{O}(N^{-1+\delta})$, where N is the number of sampling points, $\delta > 0$ is an arbitrary small number, and the constant in the decay estimate is independent of the dimension of integration.

Name: Andrea Scaglioni
University/Institution: Vienna university of technology
Talk title: Sparse grid approximation of stochastic dynamic micromagnetics
Time: 10:00 Wednesday 22nd
Abstract:

We consider the stochastic Landau-Lifschitz-Gilbert problem, an SPDE model for dynamic micromagnetism. We first convert the problem to a (highly nonlinear) PDE with parametric coefficients using the Doss-Sussmann transform and the Lévy-Ciesielsky parametrization of the Brownian motion. We prove analytic regularity of the parameter-to-solution map and derive estimates of arbitrary derivatives of the parameter-to-solution map. We apply these estimates to prove dimension-independent algebraic convergence of a piecewise-polynomial sparse grid discretization. If time permits, we discuss space and time discretization and efficient multilevel strategies for large-scale simulation.

Name: Erika Hausenblas
University/Institution: Montanuniversity Leoben, Leoben, Austria
Talk title: reduced dynamic of the inhibitor-activator system
Time: 11:00 Wednesday 22nd
Abstract:

Chemical and biochemical reactions can exhibit surprisingly different behaviors from multiple steady-state solutions to oscillatory solutions and chaotic behaviors. Such behavior has been of great interest to researchers for many decades. The Briggs-Rauscher, Belousov-Zhabotinskii, and Bray-Liebhafsky reactions, for which periodic variations in concentrations can be visualized by changes in color, are experimental examples of oscillating behavior in chemical systems. These types of systems are modeled by a system of partial differential equations coupled with a nonlinearity.

However, analyzing the pattern, one may suspect that the dynamic is only generated by a finite number of spatial Fourier modes. In fluid dynamics, it is shown that for large times, the solution is determined by a finite number of spatial Fourier modes, called determining modes. In the article, we first introduce the concept of determining modes and show that, indeed, it is sufficient to characterize the dynamic by only a finite number of spatial Fourier modes.

Name: Vesa Kaarnioja
University/Institution: Free University of Berlin
Talk title: Quasi-Monte Carlo approach to Bayesian optimal experimental design problems governed by PDEs
Time: 11:30 Wednesday 22nd
Abstract: Contemporary quasi-Monte Carlo (QMC) methods are based on tailoring specially designed cubature rules for high-dimensional integration problems. By exploiting the smoothness and anisotropy of an integrand, it is possible to achieve faster-than-Monte Carlo convergence rates. To this end, QMC methods have become a popular tool for numerical treatment of partial differential equations involving random coefficients. Meanwhile, the goal in Bayesian optimal experimental design is to maximize the expected information gain for the reconstruction of unknown quantities when there is a limited budget for collecting measurement data. In this talk, we derive tailored QMC cubature rules to alleviate the computational burden associated with Bayesian optimal experimental design problems governed by partial differential equations.

Name: Laurence Wilkes
University/Institution: KU Leuven
Talk title: The fixed vector randomised lattice algorithm for high-dimensional integration
Time: 12:00 Wednesday 22nd
Abstract:

The recently developed fixed vector algorithm offers a very simple solution to the problem of producing a lattice-based randomised algorithm for numerical integration with the optimal randomised error. The essential idea of this algorithm is to shift the construction of the generating vector of the lattice to a precomputation so that the only randomised element is to choose the number of function

evaluations from a suitable range. In the talk, we will look at the existence result for such a fixed generating vector, a method to construct the vector and finally look at how the algorithm can be generalised to work in the half-period cosine space and a Sobolev space with a lower smoothness parameter.

Name: Leo Lebrat
University/Institution: CSIRO
Talk title: Sparse functional representations for Neuro-medical imaging
Time: 9:30 Thursday 23rd
Abstract:
The study of numerous neurodegenerative diseases relies on the reconstruction and analysis of the brain cortices from magnetic resonance imaging (MRI). Traditional volumetric representations for these surfaces fail to capture fine-grained cortical foldings. In this talk, we will present memory-efficient functional representations for those surfaces and exhibit empirical evidence for their robustness to dataset shift. Leveraging analogous sparse mathematical objects, we will also present a novel explainability method allowing us to produce plausible counterfactual examples.

[1] Lebrat, Leo, Santa Cruz, Rodrigo et al. "Corticalflow: a diffeomorphic mesh transformer network for cortical surface reconstruction." Advances in Neural Information Processing Systems 34 (2021)

[2] Santa Cruz, Rodrigo, Lebrat, Leo et al. "Deepcsr: A 3d deep learning approach for cortical surface reconstruction." Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2021.

[3] Peters, Joshua, Lebrat, Leo et al. "DBCE: A Saliency Method for Medical Deep Learning Through Anatomically-Consistent Free-Form Deformations." Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2023.

Name: Russell Tsuchida
University/Institution: Data61-CSIRO
Talk title: High dimensional inner products of SGD iterates in exponential family models resemble neural network kernels
Time: 10:00 Thursday 23rd
Abstract: We discuss a problem involving computing with point estimates of exponential family likelihoods in a high dimensional regime. We present background on exponential families, a general and mathematically convenient family of probability distributions studied by probabilists, statisticians and machine learners alike. In classical applications, a dependent variable follows an exponential family likelihood with an expectation that is a linear function of the independent variable. In an unsupervised setting, the independent variable is learnt, possibly together with the linear function. We consider the setup of updating the independent variable using stochastic gradient descent (SGD) in such an unsupervised setting. We find that in a high dimensional limit, the inner product between any two learnt independent variables can be updated exactly using a deterministic rule. This derived update rule resembles previously introduced infinitely wide neural network kernels. As part of our analysis, we are able to draw new connections between the statistician's inverse link function and the machine learner's activation function. We describe an interesting property of SGD in this high dimensional limit — even though individual iterates are random vectors, inner products of any two iterates are deterministic, and can converge to a unique fixed point as the number of iterates increases. Our update rule allows us to compute using infinite dimensional independent variables using a popular class of machine learning algorithms called kernel methods.

Name: Simon Foucart

University/Institution: Texas A&M University
Talk title: Recovery from corrupted data: recent results for various models
Time: 11:00 Thursday 23rd
Abstract:
The goal of Optimal Recovery is to uncover procedures that are worst-case optimal for the task of learning / recovering unknown functions from a priori information (the model) and a posteriori information (the observations). Optimal recovery procedures can sometimes be chosen as linear maps, at least when the observations are accurate. When they are corrupted, however, the problem becomes more complicated, especially if the focus is put on the practical construction of the (near) optimal procedure. This talk showcases positive results in three situations: deterministic observation errors bounded in ℓ_2; deterministic observation errors bounded in ℓ_1; and stochastic log-concave observation errors.

Name: Michael Gnewuch
University/Institution: Osnabrück University
Talk title: Integration and Function Recovery on Hermite Spaces
Time: 11:30 Thursday 23rd
Abstract: We consider spaces of functions of infinetely many variables that are defined with the help of all univariate Hermite polynomials, which (properly normalized) form an orthonormal basis of the space of all square-integrable functions over the real line endowed with the standard normal distribution. Those spaces belong to the class of reproducing kernel Hilbert spaces of increasing smoothness and their elements are defined on proper subsets of the sequence space (i.e., the space of all real-valued sequences). Their norms are induced by an underlying function space decomposition, namely the infinite-dimensional ANOVA decomposition. We discuss further properties of those spaces and present sharp results on numerical integration and on function recovery.

The Talk is based on joint work with Mario Hefter, Aicke Hinrichs, Klaus Ritter and Robin Rüssmann.

Name: Manfred Faldum
University/Institution: RWTH Aachen University
Talk title: A Space-Time Adaptive Low-Rank Method for High-Dimensional Parabolic PDEs
Time: 12:00 Thursday 22nd
Abstract: In this talk, we present the construction and analysis of a space-time adaptive method for parabolic partial differential equations that combines sparse wavelet expansions in time with adaptive low-rank approximations in the spatial variables. Similar to the existing adaptive low-rank method for elliptic problems [1], we use a perturbed Richardson iteration, where we apply two reduction operators to the iterates to keep the support as well as the arising ranks of the low-rank approximations near-optimal for a given error tolerance. This perturbed Richardson iteration is applied to a bi-infinite matrix-vector problem based on the space-time variational formulation [2] which is equivalent to the parabolic initial boundary value problem. For the analysis of the method, we propose a new approximation class for the temporal operator which is necessary due to the interaction between hierarchical tensor formats of different time indices. One of the main challenges is the fact that the parabolic operator is an isomorphism with respect to spaces not endowed with a cross norm. Therefore, as in [1], we use a method for preconditioning operators in low-rank format by exponential sum approximations. The method is shown to converge and satisfy similar complexity bounds as the existing adaptive low-rank methods for elliptic problems [1], establishing its suitability for parabolic problems on high-dimensional spatial domains and does not suffer from the curse of dimensionality. The construction also yields computable rigorous a posteriori error bounds for the total error depending on the activated basis functions and ranks in the approximation. The results are illustrated by numerical experiments for the heat equation in high dimensions, demonstrating practical efficiency.

[1] M. Bachmayr, W. Dahmen, Adaptive Low-Rank Methods: Problems on Sobolev Spaces, SIAM Journal on Numerical Analysis 54(2), pp.744-796 [2] C. Schwab and R. Stevenson, Space-Time Adaptive Wavelet Methods for Parabolic Evolution Problems, Mathematics of Computation 78(267), pp.1293-1318

Name: Ryan W Matzke
University/Institution: Vanderbilt University
Talk title: Energy and Discrepancy of Greedy Sequences on the Sphere
Time: 13:30 Thursday 23rd
Abstract:

Greedy algorithms are surprisingly effective in a wide range of optimization problems, though have only recently been considered as a possible way to find point configurations with low discrepancy or energy. On the sphere and other manifolds, greedy algorithms minimizing certain discrepancies or energies at each step have been shown to be uniformly distributed. In this talk, we further quantify some of this behavior on the sphere, in particular discussing the effectiveness of greedily generated point configurations as minimizers of Riesz energies and the quadratic spherical cap discrepancy. In particular, we show that they are optimal for singular Riesz energies. This is based on joint work with Dmitriy Bilyk, Michelle Mastrianni, and Stefan Steinerberger.

Name: Hoang Tran
University/Institution: Oak Ridge National Laboratory
Talk title: High-Dimensional Optimization with a Novel Nonlocal Gradient
Time: 14:00 Thursday 23rd
Abstract:

The problem of minimizing multi-modal loss functions with a large number of local optima frequently arises in machine learning and model calibration problems. Since the local gradient points to the direction of the steepest slope in an infinitesimal neighborhood, an optimizer guided by the local gradient is often trapped in a local minimum. To address this issue, we develop a novel nonlocal gradient to skip small local minima by capturing major structures of the loss's landscape in black-box optimization. The nonlocal gradient is defined by a directional Gaussian smoothing (DGS) approach. The key idea of DGS is to conducts 1D long-range exploration with a large smoothing radius along d orthogonal directions in $\mathbb{R}^d$, each of which defines a nonlocal directional derivative as a 1D integral. Such long-range exploration enables the nonlocal gradient to skip small local minima. The d directional derivatives are then assembled to form the nonlocal gradient. We use the Gauss -Hermite quadrature rule to approximate the d 1D integrals to obtain an accurate estimator. We provide a convergence theory in the scenario where the objective function is composed of a convex function perturbed by a highly oscillating, deterministic noise. We prove that our method exponentially converge to a tightened neighborhood of the solution, whose size is characterized by the noise wavelength. The superior performance of our method is demonstrated in several high-dimensional benchmark tests, machine learning and model calibration problems.

Name: Moody Chu
University/Institution: North Carolina State University
Talk title: Preparing Hamiltonians for Quantum Simulation: A Computational Framework for Cartan Decomposition via Lax Dynamics
Time: 14:30 Thursday 23rd
Abstract:

Simulating the time evolution of a Hamiltonian system on a classical computer is hard –The computational power required to even describe a quantum system scales exponentially with the number

of its constituents, let alone integrate its equations of motion. Hamiltonian simulation on a quantum machine is a possible solution to this challenge – Assuming that a quantum system composing of spin-$\frac{1}{2}$ particles can be manipulated at will, then it is tenable to engineer the interaction between those particles according to the one that is to be simulated, and thus predict the value of physical quantities by simply performing the appropriate measurements on the system. Establishing a linkage between the unitary operators described mathematically as a logic solution and the unitary operators recognizable as quantum circuits for execution is therefore essential for algorithm design and circuit implementation. Most current techniques are fallible because of truncation errors or t he stagnation at local solutions. This work offers an innovative avenue by tackling the Cartan decomposition with the notion of Lax dynamics. Within the integration errors that is controllable, this approach gives rise to a genuine unitary synthesis which not only is numerically feasible, but also can be utilized to gauge the quality of results produced by other means, and extend the knowledge to a wide range of applications. This paper aims at establishing the theoretic and algorithmic foundations by exploiting the geometric properties of Hamiltonian subalgebras and describing a common mechanism for deriving the Lax dynamics.