Nonlinear Approximation, PDEs and High-Dimensional Problems

Workshop in honor of Wolfgang Dahmen

October 21st to 23rd, 2024 at RWTH Aachen

Super C, 6th floor, Ford Saal and Generali Saal, Templergraben 57

Monday, Oct 21, 2024

14:00 Welcome

14:20 **Jonathan Siegel**,
Convergence and error control of consistent PINNs for elliptic PDEs

15:00 **Peter Binev**,
Computational electron microscopy

15:40 Coffee break

16:20 **Rob Stevenson**,
Ultra-weak least squares discretizations for unique continuation and Cauchy problems

17:00 **Leszek Demkowicz** (online),
DPG discretization of nonlinear elasticity problems consistent with John Ball's theory

18:00 Reception (Super C)

Tuesday, Oct 22, 2024

09:00 **Gerrit Welper**,
Approximation and optimization theory for neural networks

09:40 **Karsten Urban**,
When offline cost hurts: A parallel batch greedy algorithm for reduced bases

10:20 Coffee break

11:00 **Ron DeVore**,
Optimal recovery meets mini-max

12:00 Lunch break (catering, Super C)

Tuesday, Oct 22, 2024 – part 2

14:00 Special session

14:30 **Endre Süli** (online),
Finite element approximation of the fractional porous medium equation

15:10 **Helmut Harbrecht**,
Samplets: Wavelet concepts for scattered data

15:50 Coffee break

16:30 **Claudio Canuto**,
Adaptive finite element methods based on virtual elements

17:10 **Gitta Kutyniok**,
An applied harmonic analysis tour of reliable and sustainable AI

19:30 Dinner (Elisenbrunnen)

Wednesday, Oct 23, 2024

09:00 **Olga Mula**,
Accuracy controlled schemes for the eigenvalue problem of the radiative transfer equation

09:40 **Alexei Shadrin**,
Some remarks on the B-spline basis condition number

10:20 Coffee break

11:00 **Matthieu Dolbeault**,
Approximation to elliptic PDEs with high contrast diffusion coefficients

11:40 Closing remarks

Abstracts

Peter Binev: *Computational electron microscopy*

Effective extraction of information from electron microscopy data requires improvement of the mathematical models of the data acquisition and effective processing algorithms. The talk features two topics from Computational Electron Microscopy. The first one is about algorithms for estimating the position distortion in the scanning transmission electron microscopy (STEM) data and the approximate inversion of the corresponding data distortion operator. The second one introduces a new approach in simulation of the electron microscope, the Lattice Multislice Algorithm (LMA) that is very effective in calculating error-controlled simulations of STEM data.

Claudio Canuto: *Adaptive finite element methods based on virtual elements*

The design and understanding of adaptive finite element methods still pose interesting mathematical challenges (see e.g. the recent review paper [1]). In this vein, we focus on 2D or 3D non-conforming meshes obtained by successive newest vertex bisections, without completion (i.e., without removal of the newly added hanging nodes).

We see such meshes as made of simplicial Virtual Elements, and we first show that in the a posteriori error estimator proposed by Cangiani *et al.* for VEM discretizations of elliptic problems, the stabilization term can be removed, under reasonable conditions. This is crucial, since the stabilization term is of the same order as the residual-type error estimator, but its presence would prevent the equivalence of the latter with the energy error.

Based on this result, we design a 2-loop adaptive algorithm, which alternates a data approximation step and a Galerkin step, and we prove its convergence via a contraction argument. In addition, we introduce approximation classes relative to VEM spaces, and we investigate the complexity of the proposed algorithm, establishing its quasi-optimality with respect to these classes.

Numerical results provide a quantitative insight into the performance of the algorithm, which in 3D is able to reduce the cardinality of the output mesh by up to 30% compared to a standard AFEM, for a fixed target accuracy.

The results we present are based on the papers [2-5] listed below.

References

[1] A. Bonito, C. Canuto, R.H. Nochetto, and A. Veeser. *Adaptive finite element methods,* Acta Numerica 33 (2024), 163-485

[2] L.Beirao da Veiga, C. Canuto, R.H. Nochetto, G. Vacca, and M. Verani*, Adaptive VEM: stabilization-free a posteriori error analysis and contraction property,* SIAM J. Numer. Anal. 61 (2023), 457-494

[3] Beirao da Veiga, C. Canuto, R.H. Nochetto, G. Vacca, and M. Verani*, Adaptive VEM for variable data: convergence and optimality,* IMA J. Numer. Anal., (2023), 1-50

[4] C. Canuto and D. Fassino*, Higher-order adaptive virtual element methods with contraction properties,* Mathematics in Engineering 5:6 (2023)*, 1-33*

[5] S. Berrone, D. Fassino, and F. Vicini*, 3D adaptive VEM with stabilization-free a posteriori error bounds,* arXiv:2407.17858v1 (2024)

Leszek Demkowicz: *DPG discretization of nonlinear elasticity problems consistent with John Ball's theory*

There are very few so well known and appreciated mathematical results as John Ball's theory of polyconvexity and existence of energy minimizers for the nonlinear elasticity. And yet, the absolute majority of existing engineering codes is based on the standard Principle of Virtual Work (PVW) ignoring the polyconvexity theory. We will present a novel variational formulation for 2D nonlinear elasticity problems based on identifying the determinant of the deformation gradient as an additional unknown and reimposing the relation between the new unknown and the determinant through Lagrange multipliers AND relaxation of the constraint consistent with Ball's theory. The resulting variational formulation is analysed formally to assure that all domain and boundary integrals are well defined.

We discretize then the formulation with the Discontinuous Petrov Galerkin (DPG) technology - a minimum residual technique. The resulting nonlinear problem is solved by combining a continuation in load with the standard Newton-Rapshon iterations which can be interpreted as the Gauss-Newton method for minimizing the residual.

We test the methodology with a number of manufactured solutions and the classical Cook's membrane problem. A comparison with the classical PVW-based methodology shows that we can solve problems with several orders of magnitude larger deformation [1].

[1] J. Zhang and L. Demkowicz, *DPG Method for Nonlinear Elasticity. I. Comparison of Various Variational Formulations,* In preparation (2024)

Ron DeVore: *Optimal recovery meets mini-max*

We consider the problem of numerically recovering an unknown function f from m point samples of f with error to be measured in some Banach space norm in X. Bounds on the error of recovery can only be proved if there is additional information in the form that f$\in K$ where K is a compact subset of X. Two theories have emerged to define optimal performance of such a numerical algorithm. Optimal recovery assumes the point samples have no noise. Mini-max estimates assume the measurements are corrupted by additive i.i.d. Gaussian noise of mean zero and variance σ^2. One would expect that the minimax bounds (claimed to be optimal) would converge to the Optimal Recovery bounds as $\sigma \to 0$. However, the existing mini-max bounds in the literature do not provide such estimates. The goal of this talk is to understand what is going on.

We restrict our attention to the case f is defined on a nice domain $\Omega \in R^d$ and the model class K is the unit ball of a Besov space $B_\tau^s(L_q(\Omega))$ and the error is to be measured in an $L_q(\Omega)$ norm. We show that the existing mini-max rates in the literature are not clearly stated in terms of their dependence on σ. We go on to establish the true minimax rates as a function of σ and show that these rates converge to the optimal recovery rate when σ converges to zero. Another important aspect of our analysis is that it does not depend on wavelet decompositions which are somewhat opaque when the support of the wavelet intersects the bounday. This is joint work in collaboration with Robert Nowak, Rahul Parhi, Guergana Petrova, and Jonathan Siegel.

Matthieu Dolbeault: *Approximation to elliptic PDEs with high contrast diffusion coefficients*

We consider an elliptic diffusion equation with a diffusion coefficient that is piecewise constant, but can take arbitrary positive values on each piece. Although the problem is linear, it becomes degenerate when these values tend to zero or infinity. We thereby present specific constructions of reduced models as well as preconditioners to approximate the solutions.

Helmut Harbrecht: *Samplets: Wavelet concepts for scattered data*

This talk is dedicated to recent developments in the field of wavelet analysis for scattered data. We introduce the concept of samplets, which are signed measures of wavelet type and may be defined on sets of arbitrarily distributed data sites in possibly high dimension. By employing samplets, we transfer well-known concepts known from wavelet analysis, namely the fast basis transform, data compression, operator compression and operator arithmetics to scattered data problems. Especially, samplet matrix compression facilitates the rapid solution of scattered data interpolation problems, even for kernel functions with nonlocal support. Finally, we demonstrate that sparsity constraints for scattered data approximation problems become meaningful and can efficiently be solved in samplet coordinates.

Gitta Kutyniok: *An applied harmonic analysis tour of reliable and sustainable AI*

The new wave of artificial intelligence is impacting industry, public life, and the sciences in an unprecedented manner. However, one current major drawback is the lack of reliability as well as the enormous energy problem of AI.

The goal of this lecture is to first provide an introduction into this new vibrant research area, and showcase some recent successes from an applied harmonic analysis viewpoint to ensure reliability. We will then discuss the necessity of an analog approach for overcoming the energy problem, leading us naturally to the mathematical model of spiking neural networks.

Olga Mula: *Accuracy controlled schemes for the eigenvalue problem of the radiative transfer equation*

This talk is about a collaboration with Wolfgang Dahmen on the radiative transfer equation which is a linear Boltzmann-type PDE that models the propagation of light and neutrons in a collisional medium. The computation of the largest eigenvalue of this Boltzmann operator is crucial in nuclear safety studies but it has classically been formulated only at a discretized level, so the predictive capabilities of such computations are fairly limited. In this talk, I will give an overview of the modeling for this equation, as well as recent analysis that leads to an infinite dimensional formulation of the eigenvalue problem. We leverage this point of view to give a conceptual path to build numerical schemes that come with a rigorous, a posteriori estimation of the error between the exact, infinite-dimensional solution, and the computed one.

Alexei Shadrin: *Some remarks on the B-spline basis condition number*

Uniform boundedness of the B-spline basis condition number $\kappa_{k,p}$ (of order k, for the L_p-norm) is one of the key features in spline theory. It guarantees stability of numerical calculations with splines, provides a good local spline approximation, ensures existence of a bounded interpolating spline projector for any partition, shows how small the kth derivative of any interpolant could be, and many other things. In this talk, we discuss three conjectures of de Boor regarding this number made in mid 70s and in the 90s.

(1) The first one was that, for all k and p, this number grows as 2^k. This seems to be correct for the max-norm, but most likely we have a bit slower growth $k^{-1/2p}2^k$ for $p < \infty$. (The current upper bound is $k\ 2^k$ for all p.)
(2) The second one was that the extreme case occurs for the partition with no interior knots (the so-called Bernstein knots). This was shown to be wrong for the max-norm by de Boor himself, we show that it is not the case also for large $p < \infty$.
(3) And the third conjecture by de Boor was that "the exact condition of the B-spline basis may be hard to determine". For this one, our correcting guess is that "the exact condition of the B-spline basis will never be determined".

Jonathan Siegel: *Convergence and error control of consistent PINNs for elliptic PDEs*

We study the convergence rate, in terms of the number of collocation points, of Physics-Informed Neural Networks (PINNs) for the solution of elliptic PDEs. Specifically, given Sobolev or Besov space assumptions on the right-hand side of the PDE and on the boundary values, we determine the minimal number of collocation points required to achieve a given accuracy. These results apply more generally to any collocation method which only makes use of point values. Based upon this theory we introduce novel PINNs loss functions which we call consistent PINNs. We derive an a posteriori error estimator based upon the consistent PINNs loss functions. Finally, we present numerical experiments which demonstrate that consistent PINNs result in significantly improved error compared with the original PINNs loss function.

Rob Stevenson: *Ultra-weak least squares discretizations for unique continuation and Cauchy problems*

We present conditional stability estimates for unique continuation and Cauchy problems associated to the Poisson equation in ultra-weak variational form. Numerical approximations are obtained as minima of regularized least squares functionals. The arising dual norms are replaced by discretized dual norms, which leads to a mixed formulation in terms of trial- and test-spaces. For stable pairs of such spaces, and a proper choice of the regularization parameter, the L_2-error on a subdomain in the obtained numerical approximation can be bounded by the best possible fractional power of the sum of the data error and the error of best approximation. Compared to the use of a standard variational formulation, the latter two errors are measured in weaker norms. To avoid the use of C^1-finite element test spaces, nonconforming finite element test spaces can be applied as well. They either lead to the qualitatively same error bound, or in a simplified version, to such an error bound modulo an additional data oscillation term. Numerical results illustrate our theoretical findings.

Endre Süli: *Finite element approximation of the fractional porous medium equation*

We construct a finite element method for the numerical solution of a fractional porous medium equation on a bounded open Lipschitz polytopal domain. The pressure in the model is defined as the solution of a fractional Poisson equation, involving the fractional Neumann Laplacian in terms of its spectral definition. We perform a rigorous passage to the limit as the spatial and temporal discretization parameters tend to zero and show that a subsequence of the sequence of finite element approximations defined by the proposed numerical method converges to a bounded and nonnegative weak solution of the initial-boundary-value problem under consideration. This result can be therefore viewed as a constructive proof of the existence of a nonnegative, energy-dissipative, weak solution to the initial-boundary-value problem for the fractional porous medium equation under consideration, based on the Neumann Laplacian. The convergence proof relies on results concerning the finite element approximation of the spectral fractional Laplacian and compactness techniques for nonlinear partial differential equations, together with properties of the equation, which are shown to be inherited by the numerical method. We also prove that the total energy associated with the problem under consideration exhibits exponential decay in time. These results are joint work with Jose Antonio Carrillo (Oxford) and Stefano Fronzoni (Oxford).

Karsten Urban*: When offline cost hurts: A parallel batch greedy algorithm for reduced* bases

The "classical" (weak) greedy algorithm is widely used within model order reduction in order to compute a reduced basis in the offline training phase: An a posteriori error estimator is maximized and the snapshot corresponding to the maximizer is added to the basis. Since these snapshots are determined by a sufficiently detailed discretization, the offline phase is often computationally extremely costly.

We suggest to replace the serial determination of one snapshot after the other by a parallel approach. In order to do so, we introduce a batch size b and add b snapshots to the current basis in every greedy iteration. These snapshots are computed in parallel.

Based upon earlier work by Wolfgang Dahmen and others, we prove convergence rates for this new batch greedy algorithm and compare them to those of the classical (weak) greedy algorithm. Moreover, we present numerical results where we apply a (parallel) implementation of the proposed algorithm to the linear elliptic thermal block problem. We analyze the convergence rate as well as the offline and online wall-clock times for different batch sizes. We show that the proposed variant can significantly speed-up the offline phase while the size of the reduced problem is only moderately increased. The benefit of the parallel batch greedy increases for more complicated problems.

This talk is based upon joint work with Niklas Reich and Jürgen Vorloeper (HS Ruhr-West).

Gerrit Welper: *Approximation and optimization theory for neural networks*

The literature has shown excellent performance of neural networks for function approximation. For example, they achieve dimension independent approximation rates for Barron smooth targets or super-convergence results for classical smoothness with about twice the rates of classical approximation methods. However, these results are of theoretical nature because the network weights are either hand picked or sampled from practically unknown distributions.

On the other hand, practical networks are trained by gradient descent. This is best understood in severely over-parametrized regimes, with more weights than samples, where the training dynamics has an almost linear behavior, governed by the neural tangent kernel (NTK). These results are not immediately compatible with approximation theory because the latter measures the error in L_2 norms, corresponding to infinite samples, and is therefore under-parametrized. We discuss new results, that reconcile NTK and approximation theory and provide approximation bounds for fully connected deep neural networks trained by gradient descent.

Since the NTK theory relies on linearization, it cannot capture the full nonlinear potential of the networks. Therefore, we also provide some preliminary nonlinear results, which show that all critical points of the loss of a sufficiently wide simple model network have approximation properties comparable to classical bounds of nonlinear approximation.