Kick-Off Workshop
SFB Sparsity and Singular Structures
September 21–23, 2022, RWTH Aachen University

All lectures take place in Hörsaal II, 2nd floor Main Building RWTH, Templergraben 55

Coffee Breaks take place in room 149, 1st floor Main Building RWTH, Templergraben 55

Wednesday, September 21

14:00 - 14:10 Opening
Holger Rauhut & Carsten Honercamp (Dean of the Faculty of Mathematics, Computer Science and Natural Sciences)

14:10 - 14:30 Holger Rauhut
Overview on SFB 1481 Sparsity and Singular Structures

14:30 - 15:20 Gero Frieseke (TU München)
Sparsity in optimal transport: From theory to new algorithms that break the curse of dimension

15:30 - 16:15 *Coffee* (Room 149)

16:15 - 17:05 Semih Cayci (University of Illinois Urbana Champaign)
Finite-time analysis of entropy-regularized natural actor-critic algorithm with neural network approximation

17:15 - 18:05 Olga Mula (TU Eindhoven)
Sparsity in some optimal transport problems

18:15 - 20:00 Dinner (not organized; ask local participants for options)

20:00 *Reception and Jam Session*
Mathe-Lounge, Pontdriesch 14-16, 3rd floor
Bring your instrument if you would like to play! (Piano, Bass, Drums will be there)

SFB 1481 Sparsity and Singular Structures

RWTH AACHEN UNIVERSITY

Thursday, September 22

9:00 - 9:50 Reinhold Schneider (TU Berlin)
An Eulerian approach for numerical Solution of potential mean field games

10:00 - 10:40 Coffee Break (Room 149)

10:40 - 11:30 Tim Laux (University of Bonn)
The large-data limit of the MBO scheme for data clustering

11:40 - 12:30 Discussion
Sparsity concepts for Singular Structures: promising opportunities?

12:30 -14:30 *Lunch* (Sandwiches in Room 149; please feel to go to a restaurant if you prefer a "real" lunch)

14:30 - 15:20 Gabriel Peyré (ENS, CNRS, Paris)
Scaling optimal transport for high dimensional learning

15:30 - 16:15 *Coffee* (Room 149)

16:15 - 17:05 Sören Bartels (University of Freiburg)
Modeling and simulation of thin sheet folding

20:00 *Conference Dinner*
Restaurant Elisenbrunnen, Friedrich-Wilhelm-Platz 14, 52062 Aachen

Friday, September 23

9:00 - 9:50 Massimo Fornasier (TU München)
Breaking nonconvexity: Consensus based optimization

10:00 - 10:40 Coffee (Room 149)

10:40 - 11:30 Discussion
Deep Neural Networks in Machine Learning and Simulation: Where do we go?

11:40 - 12:30 Ingrid Daubechies (Duke University)
Sparsity and Machine Learning: Looking Back & Forward

12:40 Closing

12:45 Coffee (Room 149)

SFB 1481
Sparsity and
Singular Structures

RWTH AACHEN UNIVERSITY

Abstracts

Gero Frieseke (TU München), Wednesday, 14:30 - 15:20
Sparsity in optimal transport: From theory to new algorithms that break the curse of dimension

The mother of all sparsity theorems in optimal transport is the celebrated Brenier's theorem from 1989: optimal transport plans between two absolutely continuous probability measures on $\mathbb{R}^d$ yielding the Wasserstein-2 distance are supported on the graph of a map from $\mathbb{R}^d$ to $\mathbb{R}^d$. Thus their local support dimension is only d, whereas that of general transport plans (which are probability measures on $\mathbb{R}^d \times \mathbb{R}^d$) is $2d$. I will discuss (1) even more striking dimension reduction results for multi-marginal problems, (2) analoga for discretized optimal transport including recent versions in the multi-marginal case, and (3) an ensuing novel algorithm which – unlike the well known Sinkhorn algorithm – preserves and exploits the sparsity. The algorithm makes many-marginal problems computationally accessible for the first time, overcoming the curse of dimension and allowing to solve problems with 10^{30} unknowns with Matlab on a laptop. It was developed jointly with Andreas Schulz and Daniela Voegler for symmetric problems from electronic structure (SIAM JSC 2022) and extended with Max Penka to general (nonsymmetric) problems as arising in data science (arXiv 2022).

Semih Cayci (University of Illinois Urbana Champaign), Wednesday, 16:15 - 17:05
Finite-time analysis of entropy-regularized natural actor-critic algorithm with neural network approximation

Natural actor-critic (NAC) and its variants, equipped with the representation power of neural networks, have demonstrated impressive empirical success in solving reinforcement learning problems with large state spaces. In this talk, we present a finite-time analysis of NAC with neural network approximation, and identify the roles of neural networks, regularization and optimization techniques (e.g., gradient clipping and averaging) to achieve provably good performance in terms of sample complexity, iteration complexity and overparametrization bounds for the actor and the critic. In particular, we prove that (i) entropy regularization and averaging ensure stability by providing sufficient exploration to avoid near-deterministic and strictly suboptimal policies and (ii) regularization leads to sharp sample complexity and network width bounds in the regularized Markov decision processes, yielding a favorable bias-variance tradeoff in policy optimization. In the process, we identify the importance of uniform approximation power of the actor neural network to achieve near-optimality in policy optimization due to distributional shift.

Olga Mula (TU Eindhoven), Wednesday 17:15 - 18:05
Sparsity in some optimal transport problems

Probability measures are fundamental objects in machine learning, and their treatment often boils down to solving optimal transport problems. This talk discusses different notions of sparsity that one can introduce to i) approximate parametrized families of measures, and ii) to address computational bottlenecks arising in high-dimensional OT.

i) For the approximation of parametric families of measures, we introduce the notion of best n-term and sparse interpolation with Wasserstein barycenters from a dictionary of measures. We discuss optimality properties, and numerical strategies to compute the sparse barycenters in practice.

ii) To fight against the curse of dimensionality in OT problems, we discuss an ongoing work related to a moment approach to solve OT problems, and how to leverage sparse approximations of transport plans in high dimensions.

Reinhold Schneider (TU Berlin), Thursday 9:00 - 9:50
An Eulerian approach for numerical solution of potential mean field games

We consider Potential Mean Field Games and are mainly interested in the numerical solution of (deterministic/stochastic) HJB. We will consider control affine dynamical systems and quadratic cost for the control. In contrast to a Lagrangian approach, we considered earlier we want to follow the Eulerian perspective introduced by Benamou & Brennier for the optimal transport problem and extended by Benamou et al.. We are mainly interested in the approximation of the value function rather than the density and momentum. The value function can be obtained by solving a saddle pint problem and apply augmented Lagrangian iteration. Also the present approach can be performed by Neural Networks we focus on sparse polynomial approximation and low rank tensor product representation. For many high dimensional PDEs of practical interest, e.g. Backward Kolmogorov equations, HJB etc. the PDE operator cannot be easily expanded in tensor form. In this case, we propose machine learning approach for a modified gradient recovery confined to the manifold of tree based tensors with fixed multi-rank.

Tim Laux (University of Bonn), Thursday 10:40 - 11:30
The large-data limit of the MBO scheme for data clustering

The MBO scheme is an efficient scheme for data clustering, the task of partitioning a given dataset into several meaningful clusters. In this talk, I will present the first rigorous analysis of this scheme in the large-data limit. The starting point is that each iteration of the MBO scheme corresponds to one step of implicit gradient descent for the thresholding energy on the similarity graph of the dataset. For a subset of the nodes of the graph, the thresholding energy measures the amount of heat transferred from the subset to its complement. It is then natural to think that outcomes of the MBO scheme are (local) minimizers of this energy. We prove that the algorithm is consistent, in the sense that these (local) minimizers converge to (local) minimizers of a suitably weighted optimal partition problem. This is joint work with Jona Lelmi (U Bonn).

Gabriel Peyré (ENS, CNRS, Paris), Thursday 14:30 - 15:20
Scaling optimal transport for high dimensional learning

Optimal transport (OT) has recently gained a lot of interest in machine learning. It is a natural tool to compare in a geometrically faithful way probability distributions. It finds applications in both supervised learning (using geometric loss functions) and unsupervised learning (to perform generative model fitting). OT is however plagued by the curse of dimensionality, since it might require a number of samples which grows exponentially with the dimension. In this talk, I will explain how to leverage entropic regularization methods to define computationally efficient loss functions, approximating OT with a better sample complexity. More information and references can be found on the website of our book "Computational Optimal Transport" `https://optimaltransport.github.io/`

Sören Bartels (University of Freiburg), Thursday 16:15 - 17:05
Modeling and simulation of thin sheet folding

The folding of a thin elastic sheet along a curved arc has various applications including the construction of bistable devices. We discuss the derivation of a plate model from three-dimensional hyperelasticity and rigidity properties of admissible deformations and minimizers. The numerical solution is based on an isoparametric discontinuous Galerkin finite element method that provides a suitable geometric approximation of the folding arc. Error estimates are presented for a linearized version of the model problem.

Joint work with Andrea Bonito, Peter Hornung and Philipp Tscherner

Massimo Fornasier (TU München), Friday 9:00 - 9:50
Breaking nonconvexity: Consensus based optimization

Consensus-based optimization (CBO) is a multi-agent metaheuristic derivative-free optimization method that can globally minimize nonconvex nonsmooth functions and is amenable to theoretical analysis. In fact, optimizing agents (particles) move on the optimization domain driven by a drift towards an instantaneous consensus point, which is computed as a convex combination of particle locations, weighted by the cost function according to Laplace's principle, and it represents an approximation to a global minimizer. The dynamics is further perturbed by a random vector field to favor exploration, whose variance is a function of the distance of the particles to the consensus point. In particular, as soon as the consensus is reached the stochastic component vanishes. Based on an experimentally supported intuition that CBO always performs a gradient descent of the squared Euclidean distance to the global minimizer, we show a novel technique for proving the global convergence to the global minimizer in mean-field law for a rich class of objective functions. The result unveils internal mechanisms of CBO that are responsible for the success of the method. In particular, we present the proof that CBO performs a convexification of a very large class of optimization problems as the number of optimizing agents goes to infinity. We further present formulations of CBO over compact hypersurfaces and the proof of convergence to global minimizers for nonconvex nonsmooth optimizations on the hypersphere. We conclude the talk with several numerical experiments, which show that CBO scales well with the dimension and is extremely versatile. To quantify the performances of such a novel approach, we show that CBO is able to perform essentially as good as ad hoc state of the art methods using higher order information in challenging problems in signal processing and machine learning, namely the phase retrieval problem and the robust subspace detection.

Ingrid Daubechies (Duke University), Friday 11:40 - 12:30
Sparsity and Machine Learning: Looking Back & Forward

The presentation will give an overview of the role of sparsity in several directions of signal analysis and machine learning in which the speaker has been involved, from wavelets to compressed sensing and manifold learning.