

NUmerical methods for Compression and LEarning



Report of Contributions

Contribution ID: 2

Type: **Lecture Talk**

Hierarchical adaptive low-rank format with applications to discretized PDEs

Thursday, May 12, 2022 5:00 PM (1 hour)

When solving PDEs over tensorized 2D domains, the regularity in the solution often appears in form of an approximate low-rank structure in the solution vector, if properly reshaped in matrix form. This enables the use of low-rank methods such as Sylvester solvers (namely, Rational Krylov methods and/or ADI) which allow to treat separable differential operators. We consider the setting where this global smoothness is absent, but still locally exists almost everywhere. We show that the solution can still be efficiently stored by replacing low-rank matrices with appropriate hierarchical low-rank structures, and Sylvester solvers can be generalized to this setting. The structure can be determined with a black-box approximation scheme, that finds it adaptively. In addition, theoretical results that guarantee the structure preservation hold for these more general structure as well, and the computational complexities of the proposed method nicely interpolate between the low-rank and the completely unstructured case. We discuss how to effectively evolve the structure in time when approximating the solution of the PDE at different time steps, in the hypothesis of moving (but isolated) singularities.

Primary author: ROBOL, Leonardo (University of Pisa)

Presenter: ROBOL, Leonardo (University of Pisa)

Session Classification: Lecture talk

Contribution ID: 3

Type: **Lecture Talk**

The Reduced Basis Method in Space and Time: Challenges, Limits and Perspectives — Part 1

Thursday, May 12, 2022 9:00 AM (1 hour)

In many engineering applications, a partial differential equation (PDE) has to be solved very often (“multi-query”) and/or extremely fast (“realtime”) and/or using restricted memory/CPU (“cold computing”). Moreover, the mathematical modeling yields complex systems in the sense that:

- (i) each simulation is extremely costly, its CPU time may be in the order of several weeks;
- (ii) we are confronted with evolutionary, time-dependent processes with long time horizons or time-periodic behaviors (which often requires long-time horizons in order to find the time-periodic solution). All problems rely on time-dependent parameterized partial differential equations (PPDEs);
- (iii) the processes often involve transport and wave-type phenomena as well as complex coupling and nonlinearities.

Without significant model reduction, one will not be able to tackle such problems. Moreover, there is a requirement in each of the above problems to ensure that the reduced simulations are certified in the sense that a reduced output comes along with a computable indicator which is a sharp upper bound of the error.

The Reduced Basis Method (RBM) is a well-established method for Model Order Reduction of PPDEs. We recall the classical framework for well-posed linear problems and extend this setting towards time-dependent problems of heat, transport, wave and Schrödinger type. The question of optimal approximation rates is discussed and possible benefits of ultraweak variational space-time methods are described.

Primary author: URBAN, Karsten (University of Ulm)

Presenter: URBAN, Karsten (University of Ulm)

Session Classification: Lecture talk

Contribution ID: 4

Type: **Poster**

DCT-Former: Efficient Self-Attention with Discrete Cosine Transform

The **Trasformer** family of Deep-Learning models is emerging as the dominating paradigm for both natural language processing and, more recently, computer vision applications.

An intrinsic limitation of this family of “fully-attentive” architectures arises from the computation of the dot-product attention, which grows both in memory consumption and number of operations as $O(n^2)$ where n stands for the input sequence length, thus limiting the applications that require modeling very long sequences. Several approaches have been proposed so far in the literature to mitigate this issue, with varying degrees of success. Our idea takes inspiration from the world of *lossy* data compression to derive an approximation of the attention module by leveraging the properties of the **Discrete Cosine Transform**. An extensive experimental analysis shows that our method takes up less memory and computation for similar performance, drastically reducing inference times.

We aim that the results of our research might serve as a starting point for a broader class of deep neural models with reduced memory footprint.

The implementation is publicly available at <https://github.com/cscribano/DCT-Former-Public>.

Primary authors: SCRIBANO, Carmelo (University of Modena and Reggio Emilia); Dr FRANCHINI, Giorgia (University of Modena and Reggio Emilia)

Co-authors: Prof. PRATO, Marco (University of Modena and Reggio Emilia); Prof. BERTOGNA, Marko (University of Modena and Reggio Emilia)

Presenter: SCRIBANO, Carmelo (University of Modena and Reggio Emilia)

Session Classification: Poster

Contribution ID: 5

Type: **Lecture Talk**

Koopman operators and a programme on the foundations of infinite-dimensional spectral computations

Friday, May 13, 2022 9:30 AM (1 hour)

Koopman operators are infinite-dimensional operators that globally linearise nonlinear dynamical systems, making their spectral information valuable for understanding dynamics. Their increasing popularity, dubbed “Koopmania”, includes 10,000s of articles over the last decade. However, Koopman operators can have continuous spectra and lack finite-dimensional invariant subspaces, making computing their spectral properties a considerable challenge. This talk describes data-driven algorithms with rigorous convergence guarantees for computing spectral properties of Koopman operators from trajectory data. We introduce residual dynamic mode decomposition (ResDMD), the first scheme for computing the spectra and pseudospectra of general Koopman operators from trajectory data without spectral pollution. By combining ResDMD and the resolvent, we compute smoothed approximations of spectral measures associated with measure-preserving dynamical systems. When computing the continuous and discrete spectrum, explicit convergence theorems provide high-order convergence, even for chaotic systems. Kernelized variants of our algorithms allow for dynamical systems with a high-dimensional state-space. We compute the spectral measures of a protein molecule (20,046-dimensional state-space) and computing nonlinear Koopman modes with error bounds for chaotic turbulent flow past aerofoils with Reynolds number greater than 100,000 (295,122-dimensional state-space). Finally, these algorithms are placed within a broader programme on the foundations of infinite-dimensional spectral computations, which extends beyond spectra to numerical solutions of PDEs, optimisation, computer-assisted proofs, and the foundations of AI.

Primary author: COLBROOK, Matthew**Presenter:** COLBROOK, Matthew**Session Classification:** Lecture talk

Contribution ID: 6

Type: **Lecture Talk**

Compression of partial differential operators by numerical homogenization

Thursday, May 12, 2022 11:30 AM (1 hour)

Numerical homogenization is a methodology for the computational solution of multiscale partial differential equations. It aims at the compression of the corresponding partial differential operators to finite-dimensional sparse surrogate models. The surrogates are valid on a given target scale of interest, thereby accounting for the impact of features on under-resolved scales. This talk shows how to construct such surrogates by localized orthogonal decompositions and discusses the underlying mathematics as well as applications to random diffusion and Schrödinger operators.

Primary author: PETERSEIM, Daniel (University of Augsburg)

Presenter: PETERSEIM, Daniel (University of Augsburg)

Session Classification: Lecture talk

Contribution ID: 7

Type: **Poster**

Penalty Hyperparameters Optimization in Non-negative Matrix Factorization problems

The problem of hyperparameter optimization (HPO) in learning algorithms represents an open issue of great interest, since it has a direct impact on the performance of the algorithms as well as on the reproducibility of the same, especially in the context of unsupervised learning.

In this scenario are placed the well-known Matrix Decompositions (MDs), which are gaining attention in Data Science as mathematical techniques capable to capture latent information embedded in large datasets. Among the low-rank MDs, Nonnegative Matrix Factorization (NMF) is one of the most effective methods for analyzing real-life nonnegative data. To emphasize the useful properties of the data matrix, we often use the penalized NMF.

How to automatically choose optimal penalization hyperparameters is an open question in this context. To the best of our knowledge, the literature panorama lacks a general non-black-box framework that addresses this problem.

In this work, we consider the hyperparameter selection problem using a bi-level approach: the selection of hyperparameters is incorporated directly into the algorithm as part of the updating process. This problem is approached from two perspectives: the existence and convergence theorems of numerical solutions, under appropriate assumptions, are presented together with the proposal of a new algorithm for tuning hyperparameters in NMF problems. The proposed approach provides competitive results for controlling sparsity on synthetic and real datasets.

Primary authors: SELICATO, Laura (Università degli Studi di Bari); Prof. DEL BUONO, Nicoletta (Università degli Studi di Bari); Dr ESPOSITO, Flavia (Università degli Studi di Bari); Prof. ZDUNEK, Rafal (Politechnika Wroclawska)

Presenter: SELICATO, Laura (Università degli Studi di Bari)

Session Classification: Poster

Contribution ID: 8

Type: **Lecture Talk**

An SDP approach for tensor product approximation of linear operators on matrix spaces

Friday, May 13, 2022 8:30 AM (1 hour)

Tensor structured linear operators play an important role in matrix equations and low-rank modelling. Motivated by this we consider the problem of approximating a matrix by a sum of Kronecker products. It is known that an optimal approximation in Frobenius norm can be obtained from the singular value decomposition of a rearranged matrix, but when the goal is to approximate the matrix as a linear map, an operator norm would be a more appropriate error measure. We present an alternating optimization approach for the corresponding approximation problem in spectral norm that is based on semidefinite programming, and report on its practical performance for small examples. This is joint work with Venkat Chandrasekaran and Mareike Dressler.

Primary author: USCHMAJEW, André (MPI Leipzig)

Presenter: USCHMAJEW, André (MPI Leipzig)

Session Classification: Lecture talk

Contribution ID: 9

Type: **Lecture Talk**

Neural and operator network approximations for elliptic PDEs

Friday, May 13, 2022 11:00 AM (1 hour)

The application of neural networks (NNs) to the numerical solution of PDEs has seen growing popularity in the last five years: NNs have been used as an ansatz space for the solutions, with different training approaches (PINNs, deep Ritz methods, etc.); they have also been used to infer discretization parameters and strategies.

In this talk, I will focus on deep ReLU NN approximation theory. I will first show how NNs accurately approximate functions with isolated singularities, for example the solutions to elliptic problems in polygons and polyhedra, or eigenfunctions of problems with singular potentials that arise in quantum chemistry. I will then introduce operator networks, which approximate the solution operator of PDEs. I will, in particular, consider operator networks that, given a fixed right-hand side, map sets of diffusion-reaction coefficients into the space of solutions (coefficient-to-solution map). When the coefficients are smooth, the size of the networks can then be bounded with respect to the H^1 norm of the error, uniformly over the parameter set. The proofs of our approximation rates combine elliptic regularity, classical and recent results in numerical analysis, and tools from NN approximation theory.

Primary author: MARCATI, Carlo (University of Pavia)

Presenter: MARCATI, Carlo (University of Pavia)

Session Classification: Lecture talk

Contribution ID: 10

Type: **Poster**

Semi-supervised Learning for Aggregated Multilayer Graphs Using Diffuse Interface Methods and Fast Matrix Vector Products

We generalize a graph-based multiclass semi-supervised classification technique based on diffuse interface methods to multilayer graphs. Besides the treatment of various applications with an inherent multilayer structure, we present a very flexible approach that interprets high-dimensional data in a low-dimensional multilayer graph representation. Highly efficient numerical methods involving the spectral decomposition of the corresponding differential graph operators as well as fast matrix-vector products based on the nonequispaced fast Fourier transform (NFFT) enable the rapid treatment of large and high-dimensional data sets. In fact, the computational complexity depends linearly on the network size in each iteration step in all stages of our algorithm.

Primary author: BERGERMANN, Kai (TU Chemnitz)

Co-authors: Prof. STOLL, Martin (TU Chemnitz); Dr VOLKMER, Toni (TU Chemnitz)

Presenter: BERGERMANN, Kai (TU Chemnitz)

Session Classification: Poster

Contribution ID: 11

Type: **Lecture Talk**

Computing Means of SPD matrices – Part 1

Thursday, May 12, 2022 10:00 AM (1 hour)

One of the most fruitful tasks in data processing is to identify structures in the set where data lie and exploit them to design better models and reliable algorithms.

As a paradigm of this process we show how the cone of positive definite matrices can be endowed with Riemannian geometries alternative to the customary Euclidean geometry. This can provide new tools for data scientists, in terms of averaging and clustering techniques.

These geometries have been used to give a definition of the geometric and the power mean of positive definite matrices. We describe the way in which these objects have been understood and the matrix analytic, geometric and computational tools needed to describe and compute them.

In particular, we will use computational techniques related to primary matrix functions, rational Krylov subspaces and Riemannian optimization.

Primary author: IANNAZZO, Bruno (University of Perugia)

Presenter: IANNAZZO, Bruno (University of Perugia)

Session Classification: Lecture talk

Contribution ID: 12

Type: **Lecture Talk**

The Reduced Basis Method in Space and Time: Challenges, Limits and Perspectives — Part 2

Thursday, May 12, 2022 2:00 PM (1 hour)

In many engineering applications, a partial differential equation (PDE) has to be solved very often (“multi-query”) and/or extremely fast (“realtime”) and/or using restricted memory/CPU (“cold computing”). Moreover, the mathematical modeling yields complex systems in the sense that:

- (i) each simulation is extremely costly, its CPU time may be in the order of several weeks;
- (ii) we are confronted with evolutionary, time-dependent processes with long time horizons or time-periodic behaviors (which often requires long-time horizons in order to find the time-periodic solution). All problems rely on time-dependent parameterized partial differential equations (PPDEs);
- (iii) the processes often involve transport and wave-type phenomena as well as complex coupling and nonlinearities.

Without significant model reduction, one will not be able to tackle such problems. Moreover, there is a requirement in each of the above problems to ensure that the reduced simulations are certified in the sense that a reduced output comes along with a computable indicator which is a sharp upper bound of the error.

The Reduced Basis Method (RBM) is a well-established method for Model Order Reduction of PPDEs. We recall the classical framework for well-posed linear problems and extend this setting towards time-dependent problems of heat, transport, wave and Schrödinger type. The question of optimal approximation rates is discussed and possible benefits of ultraweak variational space-time methods are described.

Primary author: URBAN, Karsten (University of Ulm)

Presenter: URBAN, Karsten (University of Ulm)

Session Classification: Lecture talk

Contribution ID: 13

Type: **Lecture Talk**

Computing Means of SPD matrices — Part 2

Thursday, May 12, 2022 3:00 PM (1 hour)

One of the most fruitful tasks in data processing is to identify structures in the set where data lie and exploit them to design better models and reliable algorithms.

As a paradigm of this process we show how the cone of positive definite matrices can be endowed with Riemannian geometries alternative to the customary Euclidean geometry. This can provide new tools for data scientists, in terms of averaging and clustering techniques.

These geometries have been used to give a definition of the geometric and the power mean of positive definite matrices. We describe the way in which these objects have been understood and the matrix analytic, geometric and computational tools needed to describe and compute them.

In particular, we will use computational techniques related to primary matrix functions, rational Krylov subspaces and Riemannian optimization.

Primary author: IANNAZZO, Bruno (University of Perugia)

Presenter: IANNAZZO, Bruno (University of Perugia)

Session Classification: Lecture talk

Contribution ID: 14

Type: **Lecture Talk**

Neural networks, flexible activation functions and tensor decompositions

Thursday, May 12, 2022 6:00 PM (1 hour)

Neural networks are a fundamental tool for solving various machine learning tasks, such as supervised and unsupervised classification.

Despite this success, they still have a number of drawbacks, including lack of interpretability and large number of parameters.

In this work, we are particularly interested in learning neural network architectures with flexible activation functions (contrary to fixed activation functions commonly used).

Our approach relies on a tensor-based framework for decomposition of multivariate maps, developed in the context of nonlinear system identification.

We propose a new compression algorithm which is based on a constrained coupled matrix-tensor factorization (CMTF) of the Jacobian tensor and the matrix of function evaluations.

Primary author: USEVICH, Konstantin (CNRS Nancy)

Presenter: USEVICH, Konstantin (CNRS Nancy)

Session Classification: Lecture talk

Contribution ID: 15

Type: **Lecture Talk**

Sparse optimization methods for infinite-dimensional variational problems

Friday, May 13, 2022 12:00 PM (1 hour)

In this talk, I will review several recent results about the sparse optimization of infinite-dimensional variational problems. First, I will focus on the so-called representer theorems that allow to prove, in the case of finite-dimensional data, the existence of a solution given by the linear combination of suitably chosen atoms. In particular, I will try to convey the importance of such statements in understanding the sparsity in infinite-dimensional settings, describing several possible applications for various relevant problems.

In the second part of the talk, I will focus on a sparse optimization algorithm, named generalized conditional gradient method, that is built on the characterization of sparse objects for infinite-dimensional variational problems. This algorithm is a variant of the classical Frank-Wolfe algorithm, and it does not require an a priori discretization of the domain. I will show convergence results under general assumptions on the variational problem and finally, I will present some numerical examples in the context of dynamic inverse problems regularized with optimal transport energies.

Primary author: CARIONI, Marcello (University of Twente)

Presenter: CARIONI, Marcello (University of Twente)

Session Classification: Lecture talk

Contribution ID: 16

Type: **Poster**

Computing graph p -Laplacian eigenpairs by a dynamical method

Graph p -Laplacian eigenpairs, and in particular the two limit cases $p = 1$ and $p = \infty$, reveal important information about the topology of the graph. Indeed, the 1-Laplacian eigenvalues approximate the Cheeger constants of the graph, while the ∞ -eigenvalues can be related to distances among nodes, to the diameter of the graph, and more generally to the maximum radius that allows to inscribe a given number of disjoint balls in the graph. We provide a characterization of the p -Laplacian eigenpairs in terms of constrained weighted linear Laplacian eigenpairs that can be computed by gradient flows for a family of energy functions. Moreover, we show that this approach is suitable to deal also with the degenerate case $p = \infty$.

Primary authors: DEIDDA, Piero (Department of Mathematics "Tullio Levi-Civita", University of Padova); SEGALA, Nicola (University of Padova); PUTTI, Mario (University of Padova)

Co-authors: BURGER, Martin (Friedrich Alexander University of Erlangen-Nurnberg); TUDISCO, Francesco (GSSI)

Presenter: DEIDDA, Piero (Department of Mathematics "Tullio Levi-Civita", University of Padova)

Session Classification: Poster

Contribution ID: 17

Type: **Lecture Talk**

Asymptotic control theory for affine switching systems of oscillators

The talk will be devoted to continuous-time affine control systems and their reachable sets. I will focus on the case when all eigenvalues of the linear part of the system have zero real part. In this case, the reachable sets usually have a non-exponential growth rate as $T \rightarrow \infty$, and it is usually polynomial. The simplest non-trivial example is the problem of stabilisation (or, conversely, destabilisation) of two pendulums by the same common control. An exact description of reachable sets is most often impossible here, but their asymptotic behaviour as $T \rightarrow \infty$ can be found with high accuracy. In the talk, I will present the asymptotic behaviour of reachable sets in the problem of controlling a system of N independent oscillators, and in the problem of controlling the wave equation for a closed string. In particular, in these problems the corresponding analog of the Lyapunov function can be found explicitly, and, consequently, the optimal behaviour at high energies can be found very accurately.

Presenter: LOKUTSIEVSKIY, Lev (Steklov Institute, Moscow, Russia)

Session Classification: Lecture talk

Contribution ID: 18

Type: **Poster**

Distance to singularity for matrix polynomials

Given a regular matrix polynomial, an interesting problem consists in the computation of the nearest singular matrix polynomial, which determines its distance to singularity. We consider - only for simplicity - the quadratic case $\lambda^2 A_2 + \lambda A_1 + A_0$ with $A_2, A_1, A_0 \in \mathbb{C}^{n \times n}$ and look for the nearest singular quadratic matrix polynomial $\lambda^2(A_2 + \Delta A_2) + \lambda(A_1 + \Delta A_1) + (A_0 + \Delta A_0)$. Whenever the singularity of the polynomial is determined by the property that the perturbed matrices $(A_2 + \Delta A_2), (A_1 + \Delta A_1), (A_0 + \Delta A_0)$ have a common null (right/left) kernel, it can be shown that the perturbations have a low-rank property, which can be exploited in the computation. The algorithm we propose is a two-level procedure for a matrix nearness problem, where in an inner iteration a gradient flow drives perturbations to stationary points and in an outer iteration the perturbation size is optimized.

Primary authors: GNAZZO, Miryam (Gran Sasso Science Institute); Prof. GUGLIELMI, Nicola (Gran Sasso Science Institute)

Presenter: GNAZZO, Miryam (Gran Sasso Science Institute)

Session Classification: Poster

Contribution ID: 19

Type: **Poster**

Rank-1 ODE for structured eigenvalue optimization

A new approach to solve eigenvalue optimization problems for large structured matrices is proposed and studied. The class of optimization problems considered is related to compute structured pseudospectra and their extremal points, and to structured matrix nearness problems such as computing the structured distance to instability. The structure can be a general linear structure and includes, for example, large matrices with a given sparsity pattern, matrices with given range and co-range, and Hamiltonian matrices. Remarkably, the eigenvalue optimization can be performed on the manifold of complex rank-1 matrices, which yields a significant reduction of storage and computational cost. The method relies on a constrained gradient system and the projection of the gradient onto the tangent space of the manifold of complex rank-1 matrices. It is shown that near a local minimizer this projection is very close to the identity map, and so the computationally favorable rank-1 projected system behaves locally like the gradient system.

Primary authors: Prof. LUBICH, Christian (University of Tübingen); Prof. GUGLIELMI, Nicola (Gran Sasso Science Institute); SICILIA, Stefano (Gran Sasso Science Institute)

Presenter: SICILIA, Stefano (Gran Sasso Science Institute)

Session Classification: Poster

Contribution ID: 20

Type: **Poster**

Contour Integral Methods and Reduced Basis for parametric convection-diffusion PDEs

We discuss a reduced basis method for linear evolution PDEs, which is based on the application of the Laplace transform. The main advantage of this approach consists in the fact that, differently from time stepping methods, like Runge-Kutta integrators, the Laplace transform allows to compute the solution directly at a given instant, which can be done by approximating the contour integral associated to the inverse Laplace transform by a suitable quadrature formula. In terms of the reduced basis methodology, this determines a significant improvement in the reduction phase, like the one based on the classical proper orthogonal decomposition (POD), since the number of vectors to which the decomposition applies is drastically reduced as it does not contain all intermediate solutions generated along an integration grid by a time stepping method. We show by some illustrative parabolic PDEs arising from finance the effectiveness of the method and also provide some evidence that the method we propose, when applied to a simple advection equation, does not suffer the problem of slow decay of singular values which instead affects methods based on time integration of the Cauchy problem.

Primary authors: MANUCCI, Mattia (Gran Sasso Science Institute); Prof. GUGLIELMI, Nicola (Gran Sasso Science Institute)

Presenter: MANUCCI, Mattia (Gran Sasso Science Institute)

Session Classification: Poster

Contribution ID: 21

Type: **Poster**

Nonlinear Label Spreading on Hypergraphs

Semi-supervised learning is the problem of finding clusters in a graph or a point-cloud dataset where we are given “few” initial input labels. Label Spreading (LS) is a standard technique for this problem, which can be interpreted as a random walk diffusion process of the labels on the graph [1].

Recent work in network science has shown that in many graph algorithms a great advantage can be obtained when accounting directly for higher-order interactions modeled as hypergraphs, where, for example, a hyperedge directly connects all the authors of a paper in a co-authorship network [2].

%

In this work we propose a variation of LS for hypergraphs where the standard random walk Laplacian matrix is replaced by a nonlinear Laplacian-inspired map which is defined in terms of the hypergraph incidence matrix [3].

%

We show the convergence of the new nonlinear diffusion process to the global minimum of a regularized objective function that aims at reducing the variance across the hyperedge nodes and we demonstrate the efficiency of our approach on a variety of point cloud and network datasets.

Primary authors: PROKOPCHIK, Konstantin (GSSI); TUDISCO, Francesco (GSSI)

Presenter: PROKOPCHIK, Konstantin (GSSI)

Session Classification: Poster

Contribution ID: 22

Type: **Poster**

Gradient Flow Optimisation for Graph Topological Stability

The topological structure of data is widely relevant in various applications, hence raising the question of the stability of topological features. In this talk we address the stability of 1-dimensional holes in a simplicial complex through the optimisation of a functional that combines the spectra of the classical graph Laplacian with the one of the higher-order Hodge Laplacian. The proposed procedure is based on a matrix ODE formulation and a constrained gradient flow approach; the method's performance is illustrated on synthetic quasi-triangulation datasets and transportation networks.

Primary authors: SAVOSTIANOV, Anton (Gran Sasso Science Institute); TUDISCO, Francesco (GSSI); Prof. GUGLIELMI, Nicola (Gran Sasso Science Institute)

Presenter: SAVOSTIANOV, Anton (Gran Sasso Science Institute)

Session Classification: Poster