Workshop

Sparse Representation of Functions: Analytic and Computational Aspects

Berlin, December 10–14, 2012

Technische Universität Berlin
DFG Research Center MATHEON

Organized by:
Gitta Kutyniok (TU Berlin)
Volker Mehrmann (TU Berlin)
Marc Pfetsch (TU Darmstadt)

Schedule

<table>
<tr><th>Time</th><th>Monday 10</th><th>Tuesday 11</th><th>Wednesday 12</th><th>Thursday 13</th><th>Friday 14</th></tr>
<tr><td>8:00-8:45</td><td>Registration</td><td>—</td><td>—</td><td>—</td><td>—</td></tr>
<tr><td>8:45-9:00</td><td>Welcome</td><td>—</td><td>—</td><td>—</td><td>—</td></tr>
<tr><td>9:00-9:30</td><td>Dahmen</td><td>Daubechies</td><td>Vandenberghe</td><td>Calderbank</td><td>Kressner</td></tr>
<tr><td>9:40-10:10</td><td>Ward</td><td>Rauhut</td><td>Gribonval</td><td>Krahmer</td><td>Saab</td></tr>
<tr><td>10:20-10:50</td><td colspan="5">Coffee break</td></tr>
<tr><td>10:50-11:20</td><td>Hansen</td><td>Miedlar</td><td>Tillmann</td><td>Vybiral</td><td>Schneider</td></tr>
<tr><td>11:30-12:00</td><td>Grasedyck</td><td>Teschke</td><td>Cevher</td><td>Schröder</td><td>Final remarks</td></tr>
<tr><td>12:10-14:00</td><td colspan="4">Lunch break</td><td rowspan="7"></td></tr>
<tr><td>14:00-14:30</td><td>Saunders</td><td>Toledo</td><td rowspan="5">Excursion</td><td>Vavasis</td></tr>
<tr><td>14:40-15:10</td><td>Fornasier</td><td>Ullrich</td><td>Philipp</td></tr>
<tr><td>15:20-16:10</td><td colspan="2">Coffee break</td><td>Coffee break</td></tr>
<tr><td>16:10-16:40</td><td>Lemvig</td><td>King</td><td>Jokar</td></tr>
<tr><td>16:50-17:20</td><td>Mehrmann</td><td>Pfetsch</td><td>Kutyniok</td></tr>
<tr><td>18:00-24:00</td><td>—</td><td>—</td><td>Conference dinner</td><td>—</td></tr>
</table>

Program: Workshop on Sparse Representation of Functions: Analytic and Computational Aspects

Monday, December 10, 2012

08:00–08:45	**Registration**
08:45–09:00	**Welcome address**
09:00–09:30	*Wolfgang Dahmen*: Tensor Sparsity of Solutions to High Dimensional PDEs
09:40–10:10	*Rachel Ward*: Symbology-based Algorithms for Robust Bar Code Recovery
10:10–10:50	**Coffee break**
10:50–11:20	*Anders Hansen*: Breaking the Coherence Barrier: Semi-random Sampling in Compressed Sensing
11:30–12:00	*Lars Grasedyck*: Sparse Representation in Hierarchical Tensor Formats
12:10–14:00	**Lunch break**
14:00–14:30	*Michael Saunders*: BPdual: Another Solver for BP, BPDN, NNLS, and More
14:40–15:10	*Massimo Fornasier*: Sparse Stabilization and Optimal Control of the Cucker and Smale System
15:20–16:10	**Coffee break**
16:10–16:40	*Jakob Lemvig*: Sparse dual frames
16:40–17:20	*Volker Mehrmann*: Adaptivity in the numerical solution of eigenvalue problems for differential equations

Tuesday, December 11, 2012

09:00–09:30 *Ingrid Daubechies*:
TBA

09:40–10:10 *Holger Rauhut*:
Function interpolation via weighted ℓ_1-minimization

10:10–10:50 **Coffee break**

10:50–11:20 *Agnieszka Miedlar*:
Adaptive Finite Element Method finds "Sparse" Solution of PDE Eigenvalue Problem

11:30–12:00 *Gerd Teschke*:
On Sampling and Sparse Recovery in Inverse Problems

12:10–14:00 **Lunch break**

14:00–14:30 *Sivan Toledo*:
Random Sampling in Numerical Linear Algebra

14:40–15:10 *Tino Ullrich*:
Hammersley and Fibonacci QMC beat Sparse Grids

15:20–16:10 **Coffee break**

16:10–16:40 *Emily J. King*:
Image Inpainting via Analysis-Side ℓ_1-Minimization

16:40–17:20 *Marc Pfetsch*:
Computational Solver Comparison for Basis Pursuit

Wednesday, December 12, 2012

09:00–09:30 *Lieven Vandenberghe*:
Nuclear norm optimization in system identification

09:40–10:10 *Rémi Gribonval*:
Sparse dictionary learning in the presence of noise & outliers

10:10–10:50 **Coffee break**

10:50–11:20 *Andreas Tillmann*:
Branch & Cut for L0-Minimization

11:30–12:00 *Volkan Cevher*:
Learning non-parametric basis independent modelsÂăfrom point queries via low-rank methods

12:10–14:00 **Lunch break**

14:00–18:00 **Excursion**

18:00–24:00 **Conference dinner**

Thursday, December 12, 2012

09:00–09:30	*Robert Calderbank*: Communications Inspired Linear Discriminant Analysis
09:40–10:10	*Felix Krahmer*: Compressive Imaging: Sampling Strategies and Reconstruction Guarantees
10:10–10:50	**Coffee break**
10:50–11:20	*Jan Vybiral*: Learning Functions of Few Arbitrary Linear Parameters in High Dimensions
11:30–12:00	*Christian Schröder*: Spectral error bounds for an inexact Arnoldi method
12:10–14:00	**Lunch break**
14:00–14:30	*Steve Vavasis*: Algorithms and complexity for nonnegative matrix factorization
14:40–15:10	*Friedrich Philipp*: Scalable Frames
15:20–16:10	**Coffee break**
16:10–16:40	*Sadegh Jokar*: A Multilevel Sparse Approximation Approach for the Solution of Partial Differential Equations via Customized Dictionaries
16:40–17:20	*Gitta Kutyniok*: Clustered Sparsity and Separation of Cartoon and Texture

Friday, December 13, 2012

09:00–09:30	*Daniel Kressner*: Low-rank tensor completion by Riemannian optimization
09:40–10:10	*Rayan Saab*: High accuracy finite frame quantization using Sigma-Delta schemes
10:10–10:50	**Coffee break**
10:50–11:20	*Reinhold Schneider*: Tensor completion and tensor recovery in recent tensor formats
11:30–12:00	**Final remarks**

List of Abstracts

Abstracts

Communications Inspired Linear Discriminant Analysis

Robert Calderbank

Duke University, USA

We study the problem of supervised linear dimensionality reduction, taking an information-theoretic viewpoint. The linear projection matrix is designed by maximizing the mutual information between the projected signal and the class label. By harnessing a recent theoretical result on the gradient of mutual information, the above optimization problem can be solved directly using gradient descent, without requiring simplification of the objective function. Theoretical analysis and empirical comparison are made between the proposed method and two closely related methods, and comparisons are also made with a method in which Renyi entropy is used to define the mutual information (in this case the gradient may be computed simply, under a special parameter setting). Relative to these alternative approaches, the proposed method achieves promising results on real datasets.

Learning non-parametric basis independent modelsÂăfrom point queries via low-rank methods

Volkan Cevher

EPFL Lausanne, Switzerland

We consider the problem of learning multi-ridge functions of the form $f(x) = g(Ax)$ from point evaluations of f. We assume that the function f is defined on an ℓ_2-ball in $\mathbb{R}^d$, g is twice continuously differentiable almost everywhere, and $A \in \mathbb{R}^{k \times d}$ is a rank k matrix, where k is much greater than d. We propose a randomized, polynomial-complexity sampling scheme for estimating such functions. Our theoretical developments leverage recent techniques from low rank matrix recovery, which enables us to derive a polynomial time estimator of the function f along with uniform approximation guarantees. We prove that our scheme can also be applied for learning functions of the form: $f(x) = \sum_{i=1}^{k} g_i(a_i^T x)$, provided f satisfies certain smoothness conditions in a neighborhood around the origin. We also characterize the noise robustness of the scheme. Finally, we present numerical examples to illustrate the theoretical bounds in action.

Tensor Sparsity of Solutions to High Dimensional PDEs

Wolfgang Dahmen

RWTH Aachen, Germany

Functions of a possibly very large number of variables arise for instance as solutions to parametric families of PDEs or of PDEs with a large number differentiable variables such as the SchrÃűdinger equation in quantum chemistry or Fokker-Planck equations modeling dilute polymers. In this talk we discuss the (near) tensor sparsity of solutions to certain high dimensional diffusion equations resulting from operator splitting applied to Fokker-Planck equations. We report on results obtained jointly with L. Grasedyck, E. Süli, and partly with R. DeVore concerning the question: given (nearly) tensor sparse data for such equations, are the solutions (nearly) tensor sparse and if so, how can this be quantified?

TBA

Ingrid Daubechies

Duke University, USA

Sparse Stabilization and Optimal Control of the Cucker and Smale System

Massimo Fornasier

TU München, Germany

>From a mathematical point of view *self-organization* can be described as steady patterns, to which certain dynamical systems modeling social dynamics tend autonomously to be attracted. In this talk we explore situations beyond self-organization, in particular how to externally control such dynamical systems in order to eventually enforce pattern formation also in those situations where this wished phenomenon does not result from spontaneous and autonomous convergence. Our focus is on dynamical systems of Cucker-Smale type, modeling consensus emergence, and we question the existence of stabilization and optimal control strategies which require the minimal amount of external intervention for nevertheless inducing consensus in a group of interacting agents. In mathematical terms, our main result realizes the connection between certain variational problems involving ℓ_1-norm terms and optimally sparse controls. Our findings can also be informally stated in terms of the general principle for which a policy maker should always consider more favorable to intervene with stronger actions on the fewest possible instantaneous optimal leaders than try to control more agents with minor strength, in order to achieve group consensus.

Sparse Representation in Hierarchical Tensor Formats

Lars Grasedyck

RWTH Aachen, Germany

The (numerical) linear algebra related to tensor computations plays an immensely powerful role in theoretical and applied sciences: applied mathematics, biology, chemistry, information sciences, physics and many other areas. In this talk we consider high-order or high-dimensional tensors, respectively functions, of the form

$$A \in \mathbb{R}^{n\times\cdots\times n}, \qquad f : [0,1]^d \to \mathbb{R}$$

as they appear in multivariate (discrete or continuous) approximation problems. These require special techniques in order to allow a representation and approximation for high dimension d (cf. the curse of dimension). A typical example is the canonical polyadic (CP) format

$$A_{i_1,\ldots,i_d} = \sum_{j=1}^{r}\prod_{\mu=1}^{d} a_{\mu,j}(i_\mu), \qquad f(x_1,\ldots,x_d) = \sum_{j=1}^{r}\prod_{\mu=1}^{d} f_{\mu,j}(x_\mu)$$

with discrete, respectively continuous, functions

$$a_{\mu,j} : \{1,\ldots,n\} \to \mathbb{R}, \qquad f_{\mu,j} : [0,1] \to \mathbb{R}$$

that require $\mathcal{O}(d\cdot n\cdot r)$ degrees of freedom or the representation of dr one-dimensional functions $f_{\mu,j}$. However, this simple format comes along with several numerical obstructions, because it is highly non-linear.

In the last four years some new hierarchical formats (namely Tensor Train and Hierarchical Tucker) have been developed and analysed. These include as a subset all tensors of CP format (rank parameter r fixed), but in addition they allow for some nice linear algebra approaches, e.g. a hierarchical SVD in complexity $\mathcal{O}(d\cdot n\cdot r^2 + d\cdot r^4)$. The hierarchical SVD is similar to the SVD for matrices in the sense that it allows us to reliably compute a quasi-best approximation in the hierarchical format.

We will outline the basic linear algebra behind the new approaches as well as open questions and possible applications. In particular, we will consider a standard two- or three-dimensional fine-scale discretization (like in finite elements) and use the tensorization technique in order to reduce the problem to coarse-scale. We will show that this can be considered as a generalization of periodicity, or alternatively as a kind of upscaling.

Sparse dictionary learning in the presence of noise & outliers

Rémi Gribonval

Projet METISS, IRISA, Rennes, France

A popular approach within the signal processing and machine learning communities consists in modelling signals as sparse linear combinations of atoms selected from a *learned* dictionary. While this paradigm has led to numerous empirical successes in various fields ranging from image to audio processing, there have only been a few theoretical arguments supporting these evidences. In particular, sparse coding, or sparse dictionary learning, relies on a non-convex procedure whose local minima have not been fully analyzed yet. Considering a probabilistic model of sparse signals, we show that, with high probability, sparse coding admits a local minimum around the reference dictionary generating the signals. Our study takes into account the case of over-complete dictionaries and noisy signals, thus extending previous work limited to noiseless settings and/or under-complete dictionaries. The analysis we conduct is non-asymptotic and makes it possible to understand how the key quantities of the problem, such as the coherence or the level of noise, can scale with respect to the dimension of the signals, the number of atoms, the sparsity and the number of observations.

Joint work with Rodolphe Jenatton & Francis Bach

Breaking the Coherence Barrier: Semi-random Sampling in Compressed Sensing

Anders Hansen

University of Cambridge, United Kingdom

It is well known that compressed sensing relies on two crucial assumptions: incoherence and sparsity. Although there are examples where one has both of these features, there are surprisingly many cases in applications where these criteria are not satisfied. In particular, Fourier sampling and wavelet or polynomial reconstruction (both essential in Magnetic Resonance Imaging (MRI)). To overcome this obstacle we introduce a new theory where the assumptions of incoherence and sparsity are replaced by two new features: asymptotic incoherence and asymptotic sparsity. These two ideas combined with different semi-random sampling schemes allow for dramatic subsampling in cases that previously were limited by the lack of incoherence/sparsity. Moreover, we demonstrate that asymptotic incoherence and asymptotic sparsity + semi-random sampling yield a rather intriguing effect: higher resolution allows for substantially better subsampling than with lower resolution. In particular, this new technique is excellent for high resolution problems.

Joint work with Ben Adcock

A Multilevel Sparse Approximation Approach for the Solution of Partial Differential Equations via Customized Dictionaries

Sadegh Jokar

TU Berlin, Germany

We present a framework for sparse representation solutions of PDEs using the notion of multilevel spline dictionaries and Galerkin approach. We use the information in the PDE operator, the right hand side and the problem geometry to design a dictionary combining a union of weighted and extended B-splines and problem specific enrichment functions which gives us a sparse representation of the solution. The PDE is discretized in the Galerkin framework and solved using a recursive frame refinement procedure, that uses so called Orthogonal matching pursuit with replacement to find a good approximation to the sparse solution on a given refinement level. We consider some ODEs and PDEs with sharp gradient and complicated geometry to demonstrate the applicability of our method to a range of problems.

Joint work with Volker Mehrmann and Sarosh Quraishi

Image Inpainting via Analysis-Side ℓ_1-Minimization

Emily J. King

TU Berlin, Germany

An issue in data analysis is that of incomplete data, for example a photograph with scratches or seismic data collected with fewer than necessary sensors. There exists a unified approach to solving this problem and that of data separation: namely, minimizing the norm of the analysis (rather than synthesis) coefficients with respect to particular frame(s). There have been a number of successful applications of this method recently. Analyzing this method using the concept of clustered sparsity leads to theoretical bounds and results, which will be presented. Furthermore, necessary conditions for the frames to lead to sufficiently good solutions will be shown. Finally, this theoretical framework will be use to show that shearlets are able to inpaint larger gaps than wavelets.

Joint work with Gitta Kutyniok Xiaosheng Zhuang

Compressive Imaging: Sampling Strategies and Reconstruction Guarantees

Felix Krahmer

U Göttingen, Germany

In many applications such as Magnetic Resonance Imaging, images are acquired using Fourier transform measurements. Such measurements can be expensive and so it is of interest to exploit the wavelet domain sparsity of natural images to reduce the number of measurements without destroying the quality of image reconstruction. Much work in compressed sensing has been devoted to this and related problems in recent years. However, a rigorous theory for sampling with compressive frequency measurements has to date only been developed for bases that, unlike wavelet bases, are incoherent to the Fourier basis. Nevertheless, it has been shown empirically that variable density sampling strategies seem to overcome this obstacle. We introduce a theory which reveals suitable variable density sampling strategies and provides the first theoretical reconstruction results for compressive imaging via frequency measurements.

Joint work with Rachel Ward

Low-rank tensor completion by Riemannian optimization

Daniel Kressner

EPFL Lausanne

The aim of tensor completion is to fill in missing entries of a partially known tensor under a low-rank constraint. We propose a new algorithm that performs Riemannian optimization techniques on the manifolds of tensors of fixed multilinear rank. More specifically, a variant of the nonlinear conjugate gradient method is developed. Particular attention is paid to the efficient computation of the different ingredients, including the gradient, retraction and vector transport. Examples with synthetic data demonstrate good recovery even if the vast majority of the entries is unknown. Finally, we illustrate the use of the developed algorithm for the approximation of functions with singularities.

Joint work with Michael Steinlechner (EPFL) and Bart Vanderecken (U Princeton)

Clustered Sparsity and Separation of Cartoon and Texture

Gitta Kutyniok

TU Berlin, Germany

Natural images are typically a composition of cartoon and texture structures. A medical image might, for instance, show a mixture of gray matter and the skull cap. One common task is to separate such an image into two single images, one containing the cartoon part and the other containing the texture part. Using compressed sensing techniques, numerous inspiring empirical results have already been obtained.

In this paper we provide the first thorough theoretical study of the separation of a combination of cartoon and texture structures in a model situation. The methodology we consider expands the image in a combined dictionary consisting of a shearlet tight frame and a Gabor tight frame and minimizes the ℓ_1 norm on the analysis side. Sparse approximation properties then force the cartoon components into the shearlet coefficients and the texture components into the Gabor coefficients, thereby separating the image. Utilizing the fact that the coefficients are clustered geometrically and endowing a Gabor system with a scale, we prove that at sufficiently fine scales arbitrarily precise separation is possible.

Sparse dual frames

Jakob Lemvig

TU Danmark, Danmark

Frames are generalizations of bases which lead to redundant signal expansions, and they play an important role in many applications, e.g., in the theory of nonuniform sampling, wireless communications, and Sigma-Delta quantization. Sparsity of frame vectors (in some fixed orthonormal basis) is a new paradigm in frame theory that among other things allows for simple representation of the frame vectors and fast analysis and reconstruction procedures. Recently, a general construction method for sparse tight frames was considered in [Casazza, Heinecke, Krahmer, Kutyniok, *Optimally sparse frames*, IEEE Trans. IT]. In this talk, we study sparse *dual* frames of a given (not necessarily tight nor sparse) frame. We characterize the optimal sparsity level in the set of all duals and present numerical algorithms for finding sparse dual with desirable properties.

Joint work with Felix Krahmer and Gitta Kutyniok

Adaptivity in the numerical solution of eigenvalue problems for differential equations

Volker Mehrmann

TU Berlin, Germany

We discuss multi-way adaptive algorithms for eigenvalue problems associated with non-selfadjoint partial differential operators. Apart from the adaptive grid refinement, also adaptivity for the associated linear algebra problems and adaptivity in the context of homotopy is discussed and all errors are measured and balanced in the same framework.

Joint work with C. Carstensen, J. Gedicke, and A. Miedlar

Adaptive Finite Element Method finds "Sparse" Solution of PDE Eigenvalue Problem

Agnieszka Miedlar

TU Berlin, Germany

PDE eigenvalue problems of a general form $L(\lambda, u) = 0$ with partial differential operator L, arise in many modern technological applications, e.g. vibration of structures or quantum phase transitions. Recently, a particular interest has been concerned with respect to so-called Adaptive Finite Element Methods (AFEM) which, based on the quality of the numerical approximation (a posteriori error estimator), automatically adjust the finite element space in order to determine the sufficiently accurate final solution. A typical loop of the AFEM consists of the four steps Solve $\rightarrow$ Estimate $\rightarrow$ Mark $\rightarrow$ Refine. We will first introduce an extended approach AFEMLA for the adaptive finite element solution of selfadjoint elliptic PDE eigenvalue problems that incorporates the solution (in finite precision arithmetic) of the algebraic problem into the adaptation process and uses an early terminated iterative Krylov subspace method to compute a few smallest eigenvalues of a selfadjoint elliptic PDE eigenvalue problems. An additional advantage of the proposed algorithm over the standard AFEM is that this adaptivity technique can be applied even without explicit knowledge of the underlying PDE. Let us assume that we would like to obtain the solution of some physical problem, e.g., compute the noise level inside the car, but the only available information are the matrix representation of the problem and the corresponding finite element grid. Although we do not have a PDE equation describing the problem and we are not able to construct appropriate an a posteriori error estimator, we still can construct an adaptive algorithm which allows us to obtain a good approximation of the exact solution at a reasonable cost. Since in the AFEMLA algorithm adaptivity is governed by the algebraic residual it can be used also when the problem come in discretized form, e.c. from the finite element modeling. In some sense an Adaptive Finite Element Method can be viewed as a computational method trying to find the "sparse" solution of the underlying algebraic eigenvalue problem. Sparsity not only means that very few components of the residual vector or eigenvector are nonzero, 1moreover these nonzero components directly correspond to particular degrees of freedom and basis functions support. In other words "sparse" solution obtain to the solution corresponding to some very coarse grid.

Joint work with V. Mehrmann

Computational Solver Comparison for Basis Pursuit

Marc Pfetsch

TU Darmstadt, Germany

The problem of finding a minimum ℓ_1-norm solution to an underdetermined linear system (basis pursuit) is undoubtedly one of the central problems in compressed sensing. Many specialized solution approaches have been proposed and (matlab) implementations are available. This talks presents an extensive numerical comparison of seven such solvers and discusses the benchmarking methodology. Moreover, we propose a heuristic optimality check (HOC) as a general tool for ℓ_1-minimization, which often allows for early termination by "guessing" a primal-dual optimal pair based on an approximate support.

Joint work with Dirk Lorenz and Andreas Tillmann

Scalable Frames

Friedrich Philipp

TU Berlin, Germany

Tight frames can be characterized as those frames which possess optimal numerical stability properties. Here, we consider the question of modifying a general frame to generate a tight frame by simply scaling its frame vectors; a process which can also be regarded as perfect preconditioning of a frame by a diagonal matrix. A frame is called scalable, if such a diagonal matrix exists. We derive various characterizations of scalable frames and provide a geometrical interpretation of scalability in terms of conical surfaces. Finally, it is shown that – loosely speaking – the set of scalable frames is thin in the set of frames.

Function interpolation via weighted ℓ_1-minimization

Holger Rauhut

University of Bonn, Germany

We consider the problem of interpolating a function from sample values. Building on insights from compressive sensing, we study weighted ℓ_1-minimization as recovery/interpolation method. Assuming that the function to be recovered has a sparse expansion in terms of the Fourier system or more general orthonormal systems including orthogonal polynomials, compressive sensing predicts that the function can be recovered from few samples at randomly chosen locations via (unweighted) ℓ_1-minimization. In practice, however, strict sparsity occurs rarely and has to be replaced with approximate sparsity. While also this setup is already well-understood, we take a step further and take into consideration that functions often possess a certain smoothness. Hence, low-order Fourier coefficients are more likely to appear in the best s-term approximation than high-order Fourier coefficients. Taking this observation into account, we use weighted lp-norms with $p \leq 1$ on the Fourier coefficients in order to model the functions to be reconstructed. The corresponding natural recovery method turns out to be weighted l1-minimization. We will present theoretical results and promising numerical experiments. This approach is able to overcome certain limitations in the context of recovering spherical harmonic expansions. If time permits also connections to numerically solving parametric (stochastic) partial differential equations will be discussed.

Joint work with Rachel Ward

High accuracy finite frame quantization using Sigma-Delta schemes

Rayan Saab

Duke University, USA

In this talk, we address the digitization of oversampled signals in the finite-dimensional setting. In particular, we show that by quantizing the N-dimensional frame coefficients of signals in $\mathbb{R}^d$ using Sigma-Delta quantization schemes, it is possible to achieve root- exponential accuracy in the oversampling rate $\lambda := N/d$ (even when one bit per measurement is used). These are currently the best known error rates in this context. In particular, such error rates holds for an "optimal" family of frames (the Sobolev self-dual frames), as well as to the well-known Harmonic frames, and surprisingly to random frames from appropriate distributions. We also discuss connections and applications to quantization of compressed sensing measurements as well as open problems.

Joint work with F. Krahmer, R. Ward, and O. Yilmaz

BPdual: Another Solver for BP, BPDN, NNLS, and More

Michael Saunders

Stanford University, USA

Many imaging and compressed sensing applications seek sparse solutions to under-determined least-squares problems. The basis pursuit (BP) approach minimizes the 1-norm of the solution, and BP denoising (BPDN) balances it against the least-squares fit. The duals of these problems are conventional linear and quadratic programs. We introduce the following parameterization of the BPDN problem and its dual:

$$\min_{x,\,y} \quad \|x\|_1 + \tfrac{1}{2}\lambda \|y\|_2^2 \quad \text{s.t.} \quad Ax + \lambda y = b, \qquad\qquad \max_{y} \quad b^T y - \tfrac{1}{2}\lambda \|y\|_2^2 \quad \text{s.t.} \quad -e \le A^T y \le e,$$

where e is a vectors of 1s. Exploring the effectiveness of active-set methods for solving the BPDN dual led us to the generalized problems

$$\mathrm{BP}_\lambda\colon \quad \min_{x,\,y} \quad c(x)^T x + \tfrac{1}{2}\lambda \|y\|_2^2 \quad \text{s.t.} \quad Ax + \lambda y = b, \qquad\qquad \mathrm{QP}_\lambda\colon \quad \max_{y} \quad b^T y - \tfrac{1}{2}\lambda \|y\|_2^2 \quad \text{s.t.} \quad -l \le A^T y \le u,$$

where

$$c_j(x) = \begin{cases} \ell_j & \text{if } x_j \le 0, \\ u_j & \text{if } x_j > 0. \end{cases}$$

Our algorithm for solving QP_λ unifies several existing algorithms and is applicable to large-scale examples.

Joint work with Michael Friedlander

Tensor completion and tensor recovery in recent tensor formats

Reinhold Schneider

TU Berlin, Germany

Hierarchical Tucker tensor format (Hackbusch) and a particular case Tensor Trains (TT) (Tyrtyshnikov) have been introduced recently offering stable and robust approximation by a low order cost, and generalizing the well established Tucker format.

Motivated by recent progress in matrix completion, we consider the problem of tensor reconstruction under an a priori assumptions of low ranks $\underline{r} = (r_1, \ldots)$. In contrast to matrix completion their does not exist a counterpart for l_1 or nuclear norm optimization. The computation of such a norm seems to be NP-hard (Lim).
We will discuss an iterative hard thresholding algorithm for tensor reconstruction based on the higher order SVD (HOSVD). The HOSVD provides only a quasi-optimal rank $\underline{r}$ approximation, whereas the computation of an optimal approximation is in general NP-hard, according to recent results of Lim et al. Under a natural restricted isometry condition (RIP) with sufficiently good constants, we can show convergence to a unique solution of the reconstruction problem.

Spectral error bounds for an inexact Arnoldi method

Christian Schröder

TU Berlin, Germany

Arnoldis method is among the most popular methods to solve the Hermitian eigenvalue problem

$$Ax = \lambda x$$

searching a Krylov subspace

$$\mathcal{K}_k(A, v_1) = \text{span}(v_1, Av_1, A^2 v_1, \ldots, A^{k-1} v_1)$$

for approximate eigenvectors of A. The two dominant operations in this Algorithms are i) matrix applications and ii) vector sums for the aim of orthogonalization as

$$w_k := Av_k, \quad l_k := w_k + \alpha_1 v_1 + \ldots + \alpha_k v_k.$$

Arnoldis method generates an orthonormal basis of the Krylov space $\mathcal{K}_k(A, v_1)$, approximate eigenvectors in that subspace and corresponding approximate eigenvalues.
In practice the two named operations often cannot be carried out exactly, but with some error, i.e., instead of w_k, l_k we end up with

$$\tilde{w}_k = w_k + g_k \quad \text{and} \quad \tilde{l}_k = l_k + f_k$$

with (assumed) small perturbations g_k, f_k.
The talk discusses how good an approximation the in such a way obtained eigenvalue and –vectors are. The presented error bounds generalize those known for the unperturbed Arnoldi method. Other questions that arise are the distance to orthonormality of the now-not-anymore-orthonormal basis $\tilde{v}_1, \ldots, \tilde{v}_k$ and whether all perturbations may be interpreted as backwards error for the matrix A.
A connection to compressed sensing is drawn when the perturbations $w_k \mapsto \tilde{w}_k$ and $l_k \mapsto \tilde{l}_k$ are carried out by sparsification, i.e., setting small vector elements to zero.

On Sampling and Sparse Recovery in Inverse Problems

Gerd Teschke

Institute for Computational Mathematics in Science and Technology, Neubrandenburg University of Applied Sciences

Generalized sampling is new framework for sampling and reconstruction in infinite-dimensional Hilbert spaces. Given measurements (inner products) of an element with respect to one basis, it allows one to reconstruct in another, arbitrary basis, in a way that is both convergent and numerically stable. However, generalized sampling is thus far only valid for sampling and reconstruction in systems that comprise bases. We extend this framework from bases to frames, and provide fundamental sampling theorems for this more general case and show how the developed theory can be applied to the solution of inverse problems. Moreover, we are concerned with extending the idea of generalized sampling to sparse recovery in the context of nonlinear approximation.

Joint work with Anders Hansen and Ben Adcock

Branch & Cut for L0-Minimization

Andreas Tillmann

TU Darmstadt, Germany

While there are many heuristics for finding a sparse solution to an underdetermined linear equation system, hardly any exact solution methods are known beside the trivial total enumeration procedure. Since finding the sparsest solution, i.e., minimizing the ℓ_0-quasinorm of a vector x satisfying $Ax = b$, is NP-hard, there is little hope of devising a polynomial-time solution algorithm. In this talk, we discuss work in progress on a novel Branch & Cut scheme for ℓ_0-minimization.

Random Sampling in Numerical Linear Algebra

Sivan Toledo

University Tel Aviv, Israel

The talk will focus on random sampling in numerical linear algebra. Traditionally, algorithms in numerical linear algebra strive for accuracy (small residual errors), which straightforward random-sampling methods do not always deliver. The talk will explain how the accuracy of randomly-sampled approximations can be improved using residual-correction methods, such as LSQR. The talk will describe cases where we do have an appropriate residual correction method and cases where we do not have such methods. We will discuss both mixing-based (random projection) methods, which are appropriate for dense problems, and leverage-score-based methods, which are appropriate for sparse problems.

Hammersley and Fibonacci QMC beat Sparse Grids

Tino Ullrich

Hausdorff-Center for Mathematics/Institute for Numerical Simulation, Bonn, Germany

We constructively prove new asymptotically optimal error bounds for numerical integration in bivariate periodic Besov spaces with dominating mixed smoothness $S^r_{p,q}B(\mathbb{T}^2)$ where $1 \le p, q \le \infty$ and $r > 1/p$. Our first result uses Quasi-Monte Carlo integration on Fibonacci lattice rules and improves on the so far best known upper bound achieved by using cubature formula taking function values from a Sparse Grid. It is well known that there is no proper counterpart for Fibonacci lattice rules in higher dimensions. To this end, our second result is based on Hammersley (or Van der Corput) type point grids. Instead of exploiting a Hlawka-Zaremba type discrepancy duality, which is limited to $1/p < r \le 1$, we extend Hinrichs' recent results to larger orders r, namely $1/p < r < 2$. This direct approach is strongly conjectured to have a proper counterpart for higher orders r and, in addition, for functions on the d-torus $\mathbb{T}^d$. Last, but not least, we prove that any cubature rule based an a sparse grid in d dimensions has a significantly worse error order than the previously described methods. These results are a first step to understand the problem of optimal recovery of functions from a discrete set of function values from a completely new direction.

Nuclear norm optimization in system identification

Lieven Vandenberghe
UCLA, USA

The nuclear norm (sum of singular values) of a matrix plays an important role in extensions of 1-norm techniques for sparse optimization to optimization problems involving matrix rank minimization. It has been used successfully for applications in machine learning, signal and image processing, and statistics. Formulations based on the nuclear norm penalty are also increasingly used in dynamical system identification, as an alternative to the singular value decomposition commonly used for low-rank matrix approximations in subspace identification algorithms. The nuclear norm approach is attractive for several reasons. It preserves linear structure (such as block Hankel structure) in the low-rank matrix approximation, additional constraints or regularization terms are easily included in the formulation, and missing measurement data can be handled in a straightforward manner. In the talk we will discuss several system identification algorithms based on the nuclear norm penalty, as well as first-order algorithms for solving the resulting convex optimization problems.

Algorithms and complexity for nonnegative matrix factorization

Steve Vavasis
University of Waterloo, Canada

Nonnegative matrix factorization (NMF) is a tool for approximating a large nonnegative matrix with much smaller nonnegative sparse factors. Its power stems from its ability to automatically find features in the large matrix. It is widely used in machine learning applications including classifying text, finding features in images, interpreting the results of bioarray experiments, and even analysis of musical compositions. Recently several results have emerged concerning the theoretical complexity of NMF: it is NP-hard in general, but some interesting special cases are solvable in polynomial time. Convex relaxation in particular has proven to be a powerful solution method. Guaranteed polynomial-time algorithms and corresponding complexity bounds are a very promising line of attack in understanding NMF.

Parts of this talk represent joint work with N. Gillis, X. V. Doan and K.-C. Toh.

Learning Functions of Few Arbitrary Linear Parameters in High Dimensions

Jan Vybiral
TU Berlin, Germany

We study the uniform approximation of functions of many variables with the following inner structure. We assume, that $f(x) = g(Ax)$, where $x \in \mathbb{R}^d$, A is a $k \times d$ matrix and g is a (smooth) function on $\mathbb{R}^k$. Both g and A are unknown and their recovery is a part of the problem.
Under certain smoothness and variation assumptions on the function g, and an arbitrary choice of the matrix A, we present a sampling choice of the points drawn at random for each function approximation and algorithms for computing the approximating function. Due to the arbitrariness of A, the choice of the sampling points will be according to suitable random distributions and our results hold with overwhelming probability. Our approach uses tools taken from the compressed sensing framework, recent Chernoff bounds for sums of positive-semidefinite matrices, and classical stability bounds for invariant subspaces of singular value decompositions.

Symbology-based Algorithms for Robust Bar Code Recovery

Rachel Ward

University of Texas, USA

UPC bar codes can be characterized as sparse representations with respect to a certain symbology basis. We exploit this low-dimensional structure and introduce a greedy bar code reconstruction algorithm which can recover UPC bar codes from very noisy measurements and inaccurate parameter information. Extensions to general bar codes, radio-frequency identification, and text denoising will be discussed.