Workshop Novel Numerical Methods Shifting the Borders of Computability

Program and Book of Abstracts

	Montag (Chair: Markus Hegland)	**Dienstag** (Chair: Patrick Dewilde)	**Mittwoch** (Chair: Miriam Mehl)
08:15-08:30	Welcome Note		
08:30-09:10	Gilbert Strang	Klaus Diepold	Irad Yavneh
09:10-09:50	Shiv Chandrasekaran	Uwe Helmke	Ulrike Meier Yang
09:50-10:30	Eugene Tyrtyshnikov	Paul Van Dooren	Bart Vandereycken
Coffee Break			
11:00-11:40	Bart De Moor	Joos Vandewalle	Michel Verhaegen
11:40-12:20	Ivan Oseledets	Alle-Jan van der Veen	Justin Rice
12:20-13:00	Daniel Kressner	Samarjit Chakraborty	Yuli Eidelmann
Mittagessen			
14:00-14:40	Phd: Marco Artina	Phd: Juliane Siegl	Marc Van Barel
	Phd: Steffen Peter	Phd: Michael Lubasch	
14:40-15:20	Phd: David Sattlegger	Phd: Matthias Wong	Patrick Dewilde
	Phd: Benjamin Scharf	Phd: Christoph Kowitz	
Coffee Break			
15:50-16:30	Lars Grasedyck	Mansoor I. Yousefi	Open Problems Discussion
16:30-17:10	Thomas Huckle	Markus Hegland	
17:10-17:50	Massimo Fornasier	Miriam Mehl	
	Welcome Reception		**Dinner**

"Good" points for multivariate polynomial interpolation and approximation

Marc Van Barel
Computer Science, KU Leuven, Belgium

For the interval $[-1, 1]$, it is well-known that interpolating in Chebyshev points is much better than using equidistant points in the same interval (Runge phenomenon). This fact forms the basis of Chebfun, a Matlab-toolbox that uses Chebyshev points for the interpolation and Chebyshev polynomials for the representation of the interpolant. Quantitatively, the fact that the Chebyshev points are "good" points for interpolation with a polynomial of degree δ corresponds to the fact that the Lebesgue constant grows as $\log \delta$ while this Lebesgue constant grows much faster in case of equidistant interpolation points. The Lebesgue constant is the maximum of the Lebesgue-function on the geometry considered, in this case the interval $[-1, 1]$.

Also for the multivariate case and for different geometries, sets of "good" points were investigated (e.g., Padua points on the unit square) and other "good" point configurations were computed by optimization algorithms. In this talk we will describe an alternative optimization method to compute point configurations with a small Lebesgue constant for different geometries. This method consists of several smaller optimization procedures, taking each more and more computational effort but leading to smaller and smaller Lebesgue constants. It will turn out that the choice of a good basis for a specific geometry is essential to be able to solve the polynomial interpolation problem over that geometry. We will use an orthonormal basis with respect to a discrete inner product where the points of the inner product are lying in the geometry that we are considering at that moment. No explicit representation for these basis polynomials will be computed but we will evaluate them using a recurrence relation, generalizing the three-term recurrence relation on the real line and the Szegö recurrence relation on the complex unit circle.

A Minimum Sobolev Norm Numerical Technique for PDEs

Shivkumar Chandrasekaran
Electrical and Computer Engineering, University of California, Santa Barbara, USA,
joint work with H. Mhaskar

We present a method for the numerical solution of PDEs based on finding solutions that minimize a certain Sobolev norm. Fairly standard compactness arguments establish convergence. The method prefers that the PDE is presented in first order form. A single short Octave code is used to solve problems that range from first-order Maxwell's equations to fourth-order biharmonic problems on complicated geometries. The method is high-order convergent even on complex curved geometries.

Our method has its roots in generalized Birkhoff interpolation. Let x_i denote N points in R^d. Let f be an unknown function from R^d to R^q. Given N point-wise (possibly vector-valued) linear observations of f:

$$g(x_i) = \sum_{j \in \mathbb{N}^d}^{\|j\|_1 \leq M} A_j(x_i)\, \partial_j^{\|j\|_1} f(x_i), \tag{1}$$

the problem is to compute an approximation to f.

The above problem is well-posed for some special choices of the matrix coefficients A_j and points x_i in the sense that, as N approaches infinity, only the true solution can satisfy all the observations. In particular the numerical solution of linear PDEs can be posed as generalized Birkhoff interpolation problems. The classical approach to the above problem is to expand the unknown solution as a finite linear combination of basis functions, such that the constraints (1) become a system of square (or skinny) equations for the unknown coefficients. Then the (least-squares) solution of these equations is taken to be the computed solution for the unknown function. Our approach is almost the same, except that we pick more expansion coefficients so that we obtain a *fat* system of linear equations. As our solution we pick the one that minimizes a certain Sobolev norm. Assuming the true solution satisfies all the interpolation constraints and has a finite Sobolev norm, we can establish (see [1]) for the special case of classical interpolation) that

there is a uniform bound on the Sobolev norm of our computed solution independent of the number of constraints. It then follows from standard compactness arguments that our computed solution will converge to the true solution as the number of interpolation conditions increase.

[1] S. Chandrasekaran, H. N. Mhaskar and K. R. Jayaraman, "Minimum Sobolev norm interpolation with trigonometric polynomials on the torus," accepted for publication in *Journal of Computational Physics*, 2013.

Tensor Decompositions and Optimization Problems

Eugene Tyrtyshnikov
Numerical Mathematics, Russian Acamdemy of Sciences, Moscow, Russia

The main problem with data arrays in many dimensions is that we cannot use them to represent the data in numerical algorithms. Structure and generators become the key issue, and make the construction of algorithms to be kin to many works on structured matrices. The choice of generators is crucial for solving optimization and approximation problems.

We consider how matrix approximation results [3] transform into efficient optimization techniques for multi-index arrays with the tensor-train generators [1,2] and virtual dimensions [4]. A new approach based on the TT-CROSS algorithm [2] will be presented for the global optimization task arising in the docking problem.

[1] I.Oseledets, E.Tyrtyshnikov, Recursive decomposition of multidimensional tensors, Doklady Mathematics, vol. 80, no. 1 (2009), pp. 460-462.

[2] I.Oseledets, E. Tyrtyshnikov, TT-cross approximation for multidimensional arrays, Linear Algebra Appl., 432 (2010), pp. 70-88.

[3] S.Goreinov, E.Tyrtyshnikov, Quasioptimality of Skeleton Approximation of a Matrix in the Chebyshev Norm, Doklady Mathematics, vol. 83, no. 3 (2011), pp. 1-2.

[4] E.Tyrtyshnikov, Tensor approximations of matrices generated by asymptotically smooth functions, Sbornik Mathematics, vol. 194, no. 5-6 (2003), pp. 941-954.

6

Back to the Roots: Numerically Solving Polynomial Systems

Bart De Moor
ESAT-SISTA, KU Leuven, Belgium

We will demonstrate that the common roots of sets of multivariate polynomials can be calculated by solving (generalized) eigenvalue problems. The required insights derive from simple linear algebra (dependent and independent variables in sets of linear homogeneous equations), modelling the null space of the so-called Macaulay matrix using insights from nD-realization theory, and exploiting, just like in realization theory, the so-called shift structure of that null space, to obtain the (generalized) eigenvalue problems.

These insights open a wide avenue for numerical linear algebra algorithms to calculate the roots of multivariate polynomials (instead of the symbolic algebraic geometry algorithms that are widely used now), by exploiting the sparsity of the Macaulay matrix and the shift structure of the null space. In particular, when solving multivariate polynomial optimization problems (polynomial objective function with polynomial constraints), one only needs to find the mminimizing root. In addition, the solution to many other issues (such as conditioning, sensitivity, etc...) comes within reach.

In particular, we will show that many identification problems (e.g. prediction error methods) boil down to solving a large scale eigenvalue problem. We will also elaborate on many other matrix approximation problems, which in essence are polynomial optimization problems (e.g. noisy realization, dynamic total least squares, misfit versus latency models, etc.)

Numerical Tensor Methods in Higher Dimensions

Ivan Oseledets
Numerical Mathematics, Russian Academy of Sciences, Moscow, Russia

I will give an overview of the recent developments in numerical multilinear algebra. Novel tensor formats (Tensor Train and Hierarchical Tucker formats) allow for stable algorithms and enormous complexity reduction in different applications in physics, chemistry and data mining. I will discuss, how tensor-based models emerge in these applications, present algorithms for solving linear systems, interpolation problems, eigenvalue problems and non-stationary problems, as well as some applications (including global optimization).

8

A Low-Rank Tensor Method for Large-Scale Markov Chains

Daniel Kressner
Mathematics, ETH Zurich, Switzerland
joint work with Francisco Macedo

A number of practical applications lead to Markov Chains with extremely large state spaces. Such an instance arises from models for calcium channels, which are structures in the body that allow cells to transmit electrical charges to each other. These charges are carried on a calcium ion which can travel freely back and forth through the calcium channel. The state space of a Markov process describing these interactions typically grows exponentially with the number of cells. More generally, Stochastic Automata Networks (SANs) are networks of interacting stochastic automata and the dimension of the resulting state space grows exponentially with the number of involved automata. Several techniques have been established to arrive at a formulation such that the transition matrix has Kronecker product structure. This allows, for example, for efficient matrix-vector multiplications. However, the number of possible automata is still severely limited by the need of representing a single vector (e.g., the stationary vector) explicitly. We propose the use of low-rank tensor techniques to avoid this barrier. More specifically, an algorithm will be presented that allows to approximate the solution of certain SANs very efficiently in a low-rank tensor format.

Low Rank Tensor Reconstruction by Sampling

Lars Grasedyck
Mathematics, RWTH Aachen, Germany
joint work with Markus Bachmayr, Jonas Ballani, Wolfgang Dahmen, Ron DeVore, Melanie Kluge

We consider the problem to reconstruct (resp. approximate) a tensor from just a few samples (pointwise entries) under the assumption (or constraint) that the tensor has small tensor rank in the sense of the hierarchical or MPS ranks. The situations we consider are threefold: 1) If the samples are given and more or less random, then an alternating directions fit is appropriate, but the reliability and convergence of the iterative scheme is not satisfactorial. Nonetheless, it typically dominates other more technically advanced methods. 2) If we are allowed to provide some extra rules on how the samples should be taken (quasi-random), then the situation changes entirely and the reconstruction is often close to perfect and fast. 3) If the samples are taken consecutively and adaptively, i.e. each sample position is allowed to depend on the previous samples, then the situation is as good as it gets and we can nicely reconstruct tensors. In both cases 2) and 3) we are able to prove optimality of the method in the rank one case, and we provide examples and counterexamples that underline the sharpness of the results for rank one and higher rank.

All results are valid for the case that the network is a tree, and we provide a method that aims at finding this tree structure.

10

Tensor Methods for Structured Vectors and Matrices

Thomas Huckle
Computer Science / Mathematics, Technische Universität München, Garching, Germany

We describe how symmetries can be used in quantic tensor train (or matrix product state) representations of vectors and matrices. Besides the well-known physical symmetries we also consider mathematical properties related to the underlying matrix. Furthermore, in related numerical methods like eigenvector computations or vector approximations these symmetries can be used to derive more efficient numerical algorithms that lead to faster convergence and better representations. By a well-known similar approach the numerical treatment of the periodic boundary case can be reduced to the open boundary case. Addiditonally, we discus the use of tensor methods for approximating sparse data. This approach might be useful for application like compressed sensing in order to recover sparse signals.

Sparse stabilization and control of consensus models

Massimo Fornasier
Mathematics, Technische Universität München, Germany

From a mathematical point of view self-organization can be described as patterns to which certain dynamical systems modeling social dynamics tend spontaneously 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 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 interven- tion for nevertheless inducing consensus in a group of interacting agents. First we follow a greedy approach, by designing instantaneous feedback controls with two different sparsity properties: com- ponentwise sparsity, meaning that the controls have at most one nonzero component at every instant of time and their implementation is based on a variational criterion involving l1-norm penalization terms; time sparsity, meaning that the number of switchings is bounded on every compact interval of time, and such controls are realized by means of a sample-and-hold procedure. Controls sharing these two sparsity features are very realistic and convenient for practical issues. Moreover we show that among the controls built out of the mentioned variational principle, the maximally sparse ones are instantaneously optimal in terms of the decay rate of a suitably designed Lyapunov functional, measuring the distance from consensus. As a consequence we provide a mathematical justification to the general principle according to which ”sparse is better” in the sense that a policy maker, who is not allowed to predict future developments, should always consider more favorable to intervene with stronger action on the fewest possible instantaneous optimal leaders rather than trying to control more agents with minor strength in order to achieve group consensus. We then establish local and global sparse controllability properties to consensus. Finally, we analyze the sparsity of solutions of the finite time optimal control problem where the minimization criterion is a combina- tion of the distance from consensus and of the l1-norm of the control. Such an optimization models the situation where the policy maker is actually allowed to observe future developments. We show that the lacunarity of sparsity is related to the codimension of certain manifolds in the space of cotangent vectors.

12

Video is a Cube or Even More

Klaus Diepold
Electrical Engineering and Information Technology, Technische Universität München, Germany

In this contribution we are concerned with the development of video quality metrics using multiway data analysis tools aka as tensor based data analysis. Video quality metrics have the purpose to predict the subjectively perceived quality of a video sequence if it is judged by a human observer using only technically measurable quantities extracted form the video. Video is often considered to be a simple extension of still images, i.e. video consists of a sequence of individual images. However, the temporal dimension makes all the difference between still images and video. In many video processing tasks and in video qualimetrics in particular, the current state of the art of algorithms extract features individual image or frames of the video. Each such image is represented as a two-dimensional matrix. This way, effects which happen along the time axis are neglected and the temporal dimension is managed by temporal pooling. Temporal pooling means that the features throughout the video are mapped into one feature value, which is taken to represent the whole video sequence. The metric then determines its values on the basis of such pooled features. Alternatively, the values of the metric are determined for each frame of the video individually and then these values are pooled temporally over all frames to gain one metric value for the whole sequence. Any type of pooling strategy obscures the influence of temporal effects on the human perception of quality, as intrinsic dependencies and structures in the temporal dimension are disregarded.

In this context we discuss the introduction of an additional temporal dimension directly in the design of the video quality metrics an show that this extension can improve the prediction performance for the quality metric. We propose therefore to consider video in its natural 3-D structure as a cube. Extending the data analysis approach, we add an additional dimension to our data set and thus arrive at multidimensional data analysis, an extension of the two dimensional data analysis. In doing so, we gain a better understanding of the video's properties and will thus be able to interpret the extracted features better. We no longer employ an a priori temporal pooling step but use the whole video cube to generate the prediction model for the visual quality, and thus consider the temporal dimension of video

more appropriately. If we extend this point of view to stereo video or even multi view video, we end up at more than a cube - a hyper cube

14

Computational Challenges in the Control of Interconnected Systems

Uwe Helmke
Mathematics, Universität Würzburg, Germany

Abstract: The control of interconnected dynamical systems is an exciting research area that has widespread applications to e.g. biological systems, quantum control, robotics and computer science. The need for constructing distributed control laws for large scale networks poses new computational challenges to the theory, requiring a combination of methods from numerical linear algebra, graph theory and control theory. This talk addresses some of these system-theoretic and computational issues.

A New Indefinite Matrix Decomposition and its Applications

Paul Van Dooren
Mathematical Engineering, Catholic University of Louvain, Belgium,
joint work with N. Mastronardi

Indefinite symmetric matrices occur in many applications, such as optimization, least squares problems, partial differential equations, and variational problems. In these applications one is often interested in computing a factorization of the indefinite matrix that puts into evidence the inertia of the matrix or possibly provides an estimate of its eigenvalues. We propose an algorithm that provides this information for any symmetric indefinite matrix by transforming it to a block antitriangular form using orthogonal similarity transformations. We show that the algorithm is backward stable and has a complexity that is comparable to existing matrix decompositions for dense indefinite matrices. On the other hand, for sparse matrices that typically come from saddle point problems or from model predictive control problems, this factorization, implemented in a suitable way, can be very efficient. References : [1] N. Mastronardi, P. Van Dooren, The anti-triangular factorization of symmetric matrices, SIAM Journal on Matrix Analysis and Applications, to appear. [2] C. Kirches, H. Bock, J.P. Schloder, S. Sager, A factorization with update procedures for a KKT matrix arising in direct optimal control, Mathematical Programming Computation, 3(4), pp.319348, 2011 [3] Y. Wang, S. Boyd, Fast Model Predictive Control Using Online Optimization, IEEE Transactions on Control Systems Technology, 18(2), pp.267-278, 2010

Can and should tensor algebra, geometry and visualisation methods meet each other for data, signals and systems?

Joos Vandewalle
ESAT/SCD, KU Leuven, Belgium

It is the intention of the talk to reflect about the evolution in the past 40 years in our research field where several tendencies have diverged rather than converged. Algebra, geometry and visualizations have indeed diverged for many pragmatic reasons. Widespread use of computers and progress in computer power have been pushing for impressive algorithmic developments based on strong algebraic approaches, thereby often neglecting the geometry and the visualisation. Higher dimensional data sets have naturally moved users away from any visualization, and left them with a black box feeling. It is however for the researchers, practictioners, educators and users reconciling, and beneficiary if algebra, geometry and visualization can meet. In the talk we will highlight a couple of efforts that work on this reconciliation: the parallel coordinates ([1]), multilinear singular value decomposition ([2]), tensor decompositions, their generalizations, data fusion and signal separation techniques ([3-5]).

References:

[1] A. Inselberg, Parallel Coordinates: Visual Multidimensional Geometry and its Applications. Springer, 2009.

[2] Lieven De Lathauwer, Bart De Moor and Joos Vandewalle ”A multilinear Singular Value Decomposition”. SIAM Journal on Matrix Analysis and Applications 21, April 2000.

[3] M. Vetterli and V. Goyal ”Teaching signal processing with geometry”, http://www.eurasip.org/Proceedings/Eusipco/Eusipco2012/Conference/Tutorials/Tutorial% 1201/T1_SlideHandout.pdf

[4] J. Suykens, J. Vandewalle, ”Least squares support vector machine classifiers”, Neural processing letters, Vol. 9, 1999.

[5] L Omberg, GH Golub, O Alter, ”A tensor higher-order singular value decomposition for integrative analysis of DNA microarray data from different studies”, Proceedings of the National Acad Sciences PNAS, 2007.

Subspace Intersection Tracking Using GSVD and the Signed URV Algorithm

Alle-Jan van der Veen
Electrical Engineering, Delft University of Technology, Netherlands
joint work with Mu Zhou

We consider the separation of partially overlapping data packets by an antenna array in narrowband communication systems. This problem occurs in asynchronous communication systems and several transponder systems such as RFID, AIS and ADS-B. Arbitrary arrival times of interfering data packets cause nonstationary scenarios and makes it difficult to identify the interfering signals using standard blind beamforming techniques.

We propose subspace-based algorithms to suppress the intermittent interference. The algorithms are based on subspace intersection and oblique projections, and computed using generalized SVD (GSVD) and generalized eigenvalue (GEV) decompositions.

In the second part of the talk, these algorithms are refined using a recently developed subspace estimation tool, the Signed URV algorithm, which is in the class of "Schur subspace estimators". This class provides a complete parametrization of all "principal subspace estimates", defined as the column spans of corresponding low-rank matrix approximants that lie within a specified 2-norm distance of a given matrix. The parametrization is in terms of a two-sided hyperbolic decomposition, which can be computed using hyperbolic rotations. Although such rotations are commonly associated with numerical instabilities, the proposed SURV algorithm implicitly imposes certain constraints such that important norm bounds are achieved that guarantee stability. The algorithm is also non-iterative and the decomposition can be updated efficiently (similar to QR updates), and downdated in exactly the same way. The subspace estimates are close to the principal subspace provided by the SVD (which is a special case within the class). As we will show, the decomposition can be used to implement a "truncated Generalized SVD".

18

Semantics-Preserving Implementations of Embedded Control Systems

Samarjit Chakraborty
Institute for Real-Time Computer Systems, Technische Universität München, Germany

Control systems design usually starts with high-level models where most of the implementation-level realities are abstracted away. Often, many of these assumptions do not hold true in software implementations of the designed controllers, thereby leading to instability or poor control performance. These might stem from computation times, delays in distributed implementations or the limited numerical accuracy available. We will discuss some of these problems and methods to address them.

Numerical Methods for the Computation of the Nonlinear Fourier Transform

Mansoor I. Yousefi
Institute for Advanced Study, Technische Universität München, Germany

The nonlinear Fourier transform (NFT), a powerful tool in soliton theory and exactly solvable models, is a method for solving integrable partial differential equations governing wave propagation in certain nonlinear media. The NFT decorrelates signal degrees-of-freedom in such models, in much the same way that the Fourier transform does for linear systems. Just as the (ordinary) Fourier transform converts a convolution into a multiplication operator in the frequency domain, the nonlinear Fourier transform converts a nonlinear dispersive equation described by a Lax convolution into a multiplication operator in the nonlinear spectral domain. In this talk, numerical methods are suggested to compute the discrete and the continuous spectrum of a signal with respect to the Zakharov-Shabat system, a Lax operator underlying numerous integrable equations including the nonlinear Schrödinger equation, modeling pulse propagation in optical fibers. These methods are subsequently tested and their ability to estimate the spectrum are compared against each other. These methods are used to compute the spectrum of various pulses commonly used in the optical fiber communications.

Biological signalling cascades and tensor factorisations

Markus Hegland
Mathematics, Australian National University, Canberra, Australia
joint work with Jochen Garcke (Uni Bonn and Frauenhofer Institute)

While molecular biological systems may contain large numbers of chemical reactions, the number of copies of each substance involved is often relatively small. Thus the randomness of chemical reactions - which manifests itself as noise - has an impact on the performance of biological systems. The chemical master equation describes the change of the probability distribution of states and reaction numbers over time. Here we consider structure and numerical solution of the chemical master equation. In particular we discuss how tensor factorisation methods may be used to simplify and compute the probability distribution in the case of biological signalling cascades.

Multiphysics Applications – Just One Plus One?

Miriam Mehl
Mathematics / Computer Science, Technische Universität München, Germany
joint work with Bernhard Gatzhammer and Benjamin Uekermann

Going to multiphysics is one important way to improve existing models for a broad variety of applications in engineering and science. We can only give a few examples here: taking into account the flapping of wind turbine blades instead of only modelling the airflow; adding the flow-induced noise in a numerical virtual wind tunnel for newly developed aircrafts; enhancing simulations of electromagnetic flields with additional simulation of the underlying airflow, ... We could arbitrarily continue with this list. All these applications pose similar challenges to numerical solvers. Although the involved single-physics (fluid flow, structural mechanics, electromagnetic fields, acoustics, . . .) are well-understood, combining them to a multiphysics simulation environment is not trivial. In fact the 'plus' in this one-plus-one procedure is the main difficulty for the following reasons: Discretizing the whole set of equations describing a multiphysics problem tends to lead to ill-conditioned system matrices that are hard to solve with sufficient accuracy; implementing a new code for every possible (and required) combination of single-physics phenomena would be an immense effort; reusing existing single-physics codes and just glueing them together requires a lot of numerics such as data mapping between non-matching grids and numerical iteration schemes to regain the solution of the fully coupled system and technical solutions for code-to-code communication; the high accuracy of multiphysics models can only be exploited with a very high resolution of the underlying computational grids which makes the use of massively parallel supercomputers mandatory.

We present methods for an efficient parallel simulation of an important class of multiphysics simulations, fluid-structure interactions. This includes a technical realization for the coupling of several single-physics codes, conservative and constistent data-mapping strategies, and numerical iteration methods allowing for the parallel execution of the involved solvers. The latter is different in todays standard approaches where mostly fluid and structure are executed in an alternating way which leads to a poor scaling behaviour on high performance computing architectures. All these methods have been developed and tested for fluid-structure interactions but can be extended to more general multiphysics applications in large parts.

A Multilevel Algorithm for L_1 Minimization with Application to Sparse Representation of Signals

Irad Yavneh
Computer Science, Israel Institute of Technology, Haifa, Israel
joint work with Eran Treister

The area of sparse representation of signals is drawing tremendous attention in recent years in diverse fields of science and engineering. The idea behind the model is that a signal can be approximated as a linear combination of a few "atoms" from a pre-specified and over-complete "dictionary" (typically represented by columns from a matrix with more columns than rows). The sparse representation of a signal is often achieved by minimizing an L_1 penalized least squares functional. Various iterative-shrinkage algorithms have recently been developed and are quite effective for handling these problems, often surpassing traditional optimization techniques. Here we suggest a new iterative multilevel approach that reduces the computational cost of existing solvers for these inverse problems. Our method takes advantage of the typically sparse representation of the signal, and, at each iteration, it adaptively creates and processes a hierarchy of lower-dimensional problems employing well-known iterated shrinkage methods. Analytical observations suggest, and numerical results confirm, that this new approach may significantly enhance the performance of existing iterative shrinkage algorithms in cases where the dictionary is an explicit matrix.

References:

[1] Michael Elad, *Sparse and Redundant Representations, From Theory to Applications in Signal and Image Processing*, Springer, 2010.

[2] Alfred M. Bruckstein, David L. Donoho, and Michael Elad, "From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images", *SIAM Review*, **51** (1), 34–81, 2009.

[3] Eran Treister and Irad Yavneh, "A multilevel iterated-shrinkage approach to L_1 penalized least-squares minimization", *IEEE Trans. Sig. Proc.*, **60** (12) , 6319–6329, 2012.

Reducing Communication in Parallel Algebraic Multigrid

Ulrike Meier Yang
Lawrence Livermore National Laboratory, USA

Algebraic multigrid (AMG) is a popular solver for large-scale scientific computing and an essential component of many simulation codes. AMG has shown to be extremely efficient on distributed memory architectures, particularly on BlueGene/P and BlueGene/Q. (On BlueGene/Q , it has proven to scale to over a million cores using almost 4.5 million threads.)

However, with single-core speeds plateauing, future increases in computing performance have to rely on increased concurrency provided by the architecture, leading to potentially billions of cores or threads. Applications have to match this increased level of concurrency to exploit the performance potential and hence face additional communication requirements. On the counter side, future systems will be subject to strict power limitations for overall system power and communication is known to be one of the most significant contributors to power consumption. These two trends are turning data movements into one of the most severe bottlenecks for the next generation of parallel systems and applications. To address this challenge and to successfully exploit future architectures, it is important to target the reduction of communication.

AMG obtains its optimal computation complexity by using smaller ”coarse grid” problems to accelerate the solution of the original “fine grid” problem. Since the number of nonzeros per row for the coarse grid operators grows, communication complexity also increases significantly, leading to a large number of messages. Contention at these levels can lead to a significant decrease in performance and scalability on architectures with slower networks or possibly on future exascale machines. To counter the high communication complexities at the coarse levels, new variants with reduced communication complexity and improved communication-computation overlap are needed.

We will discuss two approaches that have shown to reduce communication. The use of redundancy and/or agglomeration on the coarse levels has shown to be effective to alleviate this communication complexity. We have developed a performance model that determines the optimal level at which to employ this technique. The second approach is an additive AMG variant.

While the classical additive approach increases parallelism and shows great potential for reduced communication, it also leads to severely reduced convergence. We investigate an additive variant, which shows convergence behavior comparable to the multiplicative version.

Differential geometry for rank-structured tensors from a numerical linear algebra perspective

Bart Vandereycken, Mathematics, Princeton University, USA

Many rank-structured tensors were very recently shown to admit smooth structures that turn them into differential manifolds, for example, tensor train (TT) or matrix product states (MPS) in [1,2], and hierarchal Tucker (HT) in [3]. In this talk, I will discuss why treating tensors as manifolds can be useful from a numerical linear algebra perspective.

In particular, I will first explain how the dynamical low-rank algorithm [4] can be used to approximate tensor differential equations in the HT or TT format. In this approach, the time derivative of the tensor to be approximated is projected onto the time-dependent tangent space of the approximation manifold along the solution trajectory. Next, I will explain how the framework of optimization on manifoldd [5] lends itself naturally for approximately solving high-dimensional tensor problems directly in the HT or TT format. In this case, the approach is conceptually simple: constrain the original problem to the manifold of fixed-rank HT/TT tensors and subsequently solve and analyze this problem using techniques from optimization on manifolds.

[1] S. Holtz, T. Rohwedder, and R. Schneider. On Manifolds Of Tensors Of Fixed TT-Rank. Num. Math., 2012.

[2] J. Haegeman, M. Marin, T. J. Osborne, and F. Verstraete. Geometry of Matrix Product States: metric, parallel transport and curvature. http://arxiv.org/abs/1210.7710v2, 2012.

[3] A. Uschmajew and B. Vandereycken. The geometry of algorithms using hierarchical tensors. Submitted, 2012.

[4] C. Lubich, T. Rohwedder, R. Schneider, B. Vandereycken, Dynamical approximation of hierarchical Tucker and tensor-train tensors. To appear in SIAM J. Matrix Anal. Appl., 2013.

[5] P.-A. Absil, R. Mahony, R. Sepulchre. Optimization Algorithms on Matrix Manifolds. Princeton University Press, 2008.

26

Infinite Matrices and Thoughts on Inverse Problems

Gilbert Strang
Mathematics, Massachusetts Institute of Technology, Boston, USA

This talk will describe two results on banded doubly infinite matrices. One is their factorization into A = LPU (lower-permutation-upper). The second concentrates on the permutation P, to find its "main diagonal" This may not be the zeroth diagonal (as for a shift)

Many applications lead to positive definite matrices of the form A'CA, where the diagonal matrix C contains the physical constants. How well can those be determined when A and A'CA are known ? This inverse problem is often well-conditioned and we open a discussion....

References for the first topics include recent papers by Patrick Dewilde and by Marko Lindner/GS in Linear Algebra and Its Applications – we also mention Three Steps on an Open Road (GS).

Sequentially Semi-Separable Matrices (and Cousins) in Distributed Systems

Justin Rice
Biomedical Engineering, City College of New York, USA

Distributed dynamical systems are all around us, from the flocks of birds in the sky and schools of fish in the sea to the traffic jams at rush hour to even the fluid dynamics of the air we breath and the synchronization of the cells in our hearts. The enormous dimension of such interconnected systems can lead to surprising and fascinating behavior, but also makes system identification, analysis, and controller synthesis prohibitively expensive.

In this presentation I'll provide an overview and explanation of how we can use the structure of such systems to drastically reduce the computationally complexity of these problems. The basic idea is to exploit the special matrix structure in such systems in iterative techniques for solving the relevant equations (Riccati and Lyapunov-type). Different system interconnection structures induce different matrix structures, but the common factor here is that they are all related to the "Matrices of Low Hankel Rank" AKA "Quasiseparable matrices", AKA "Sequentially Semi-Separable Matrices." We will discuss these interconnections and matrix structures, and showcase some new matrix structures that have been derived from this work, pointing towards future research and more general interconnections.

28

LU factorization via QR like for matrices in quasiseparable form

Yuli Eidelmann
School of Mathematical Sciences, Tel Aviv University, Israel
joint work with Patrick Dewilde

We develop a new algorithm for matrices with quasiseparable representations. This algorithm is based on using of unitary transformations and does not contain the recursive iversion of matrices as in previous works by other authors. This fact allows to expect better numerical behaviour of our algorithm.

State Reconstruction and Robust Control for Large Scale Systems

Michel Verhaegen
Delft Center for Systems and Control, TU Delft, Netherlands

We will consider the class of discretizated PDE's or network connected systems that through (spatial) lifting give rise to large scale ODE's. In this lecture three matrix structures are considered that enable to solve filtering and control problems with complexity linear in the order of the ODE or better. The first class threated e.g. in [1] considers heterogenous PDE's in 1 spatial dimension where the system properties can change along the spatial dimension. The Sequential Semi-Separable (SSS) structure of the system matrices of the lifted state space model is exploited in the solution of Riccati equations for H_2 and H_∞ control problems. The second class of systems characterize a connection of identical subsystems in a given network topology with fixed interaction between the networks when their is a connection between subsystems. For this special network topology $O(1)$ computational solutions to Robust control problems will be presented [2]. Finally the class of ODE's with sparse system matrices is considered. For this class approximate distributed solution are derived to the moving horizon state reconstruction problem [3]. A trade-off analysis will be presented to trade the distributed nature of the solution w.r.t. the accuracy of the state reconstruction.

References:

[1] Justin K. Rice and Michel Verhaegen. Distributed Control: A Sequentially Semi-Separable Approach for Spatially Heterogeneous Linear Systems. IEEE-AC 54(6):1270-1283, 2009.

[2] Paolo Massioni and Michel Verhaegen. Distributed Control for Identical Dyna- mically Coupled Systems: A Decomposition Approach. IEEE-AC, 54(1):124-135, 2009.

[3] Aleksander Haber and Michel Verhaegen. Moving horizon estimation for large-scale interconnected systems. To appear in IEEE-AC, 2013.

30

Inner-Outer Factorization for Matrices

Patrick Dewilde
Institute for Advanced Study, Technische Universität München, Germany

Inner-outer factorization, originally covered by the Beurling invariant subspace theorem in Hardy space theory, has developed over the years as maybe the most central method in dynamical system theory, the key to solve estimation and control problems in an elegant way. It generalizes even further to matrix theory, leading to unexpected results, in particular a numerically stable method to compute the LU-factorization of a large class of doubly infinite (block) matrices (the class contains finite matrices). I will give a survey of the method, show how it solves LU-factorization, discuss potential (tensor) extensions and give some pertinent examples.

PhD Forum

Anisotropic Mesh Adaptation for the Minimization of the Ambrosio-Tortorelli Functional with Application to Quasi-static Brittle Fracture Propagation

Marco Artina
Mathematics, Technische Universität München, Germany

Francfort and Marigo presented in 1998 a model for quasi-static brittle fracture which requires the minimization of the Mumford-Shah functional, representing the energy of the system. The minimization of this functional represents a very challenging issue since it is non-smooth and non-convex. The numerical approximation of this problem can be issued via a Gamma-approximation on the energy functional proposed by Ambrosio and Tortorelli where a smooth indicator function identies the fracture. Then, we resort to an adaptive nite element approach based on piecewise linear elements. Nevertheless, similarly to early work by Chambolle et al. but dierently from recent approaches by Sli et al. where isotropic meshes are used, in this work we investigate how anisotropic meshes can lead to signi cant improvements in terms of the balance between accuracy and complexity. In fact, the employment of these grids allows us to shortly follow the propagation of the fracture by rening it only in a very thin neighborhood of the crack. Moreover, the main gain achievable is a relevant reduction of the number of element needed to obtain with good condence the expected behavior of the crack with respect to the techniques used in other works. In this talk, we first present the derivation of a novel anisotropic a posteriori error estimator driving the mesh adaptation for the approximation of the Ambrosio-Tortorelli model. Then, we provide several numerical results which corroborate the accuracy as well as the computational saving led by an anisotropic mesh adaptation procedure.

Damping Noise-Folding and Enhanced Support Recovery in Compressed Sensing

Steffen Peter
Mathematics, Technische Universität München, Germany

The practice of compressive sensing suffers very importantly in terms of the efficiency/accuracy trade-off when acquiring noisy signals prior to measurement. It is rather common to find results treating the noise affecting the measurements, avoiding in this way to face the so-called noise-folding phenomenon, related to the noise in the signal, eventually amplified by the measurement procedure. After a short introduction into the field of compressive sensing, in this talk, we present a new decoding procedure, combining ℓ_1-minimization followed by a selective least p-powers, which not only is able to reduce this component of the original noise, but also has enhanced properties in terms of support identification with respect to the sole ℓ_1-minimization. We prove such features, providing relatively simple and precise theoretical guarantees. We additionally confirm and support the theoretical estimates by extensive numerical simulations, which give a statistics of the robustness of the new decoding procedure with respect to more classical ℓ_1-minimization.

Computing Semiclassical Quantum Dynamics with the Herman-Kluk Propagator

David Sattlegger
Mathematics, Technische Universität München, Germany

The evolution of quantum systems such as the dynamics of molecules is governed by the time dependent Schrödinger equation. Finding solutions to it poses a variety of numerical challenges. Even for small molecules the physical system has numerous degrees of freedom, e.g. 75 for carbon dioxide. The resulting high dimensional partial differential equations is practically impossible to solve by common approximation schemes. What is more, we have multiple scales in our system, since the mass of an electron is very small compared to the mass of a proton. In addition, the resulting solutions are highly oscillatory and there are geometric structures such as symplecticity that one would like to preserve in numerical schemes.

To nevertheless obtain satisfactory results, one relies on a combination of model reduction and numerical techniques in a physical framework somewhere between quantum and classical molecular dynamics. A majority of computations in molecular dynamics rely on the Born-Oppenheimer approximation. It describes the motion of the nuclei driven by the potential energy surfaces that are created by the electrons, thus separating the different scales and reducing the dimensionality of the system.

In this talk we shall consider the the Herman-Kluk propagator which is an asymptotic solution to the resulting semi-classical Schrödinger equation and proved to be a successful method in chemical physics. It is a Fourier integral operator that relies on the computation of classical trajectories in order to approximate - in an asymptotic sense - the time propagation operator of the Schrödinger equation. We will discuss its approximation properties, develop a numerical method for actual computations, and prove its stability. The results will be illustrated by numerical experiments.

How to Steer High-Dimensional Cucker-Smale Systems to Consensus Using Low-Dimensional Information Only

Benjamin Scharf
Mathematics, Technische Universität München, Germany

Dynamical systems of Cucker-Smale type can be used to describe the consensus formation of interacting agents. There are two situations: If the difference between the velocities of the agents is not two large in comparison to the distances of the agents, the system tends to consensus. Otherwise, when there is no self-organization, it was shown recently that one can steer the system to consensus using a sparse control acting only on the agent farthest away from the mean velocity. However, in real-life complex situations the dimension of the agents might be very large (thousands and more) and numerical simulations might be extremely expensive. In this talk we will present an idea to use Johnson-Lindenstrauss embeddings to reduce the system to a low-dimensional Cucker-Smale system. The main question is: Can we choose the agent on which we want to infer control only using the information from the low-dimensional system and thereby steer the high-dimensional system to consensus?

Quasi-linear Compressed Sensing and Algorithmic Strategies

Juliane Sigl
Mathematics, Technische Universität München, Germany

We consider quasi-linear compressed sensing and reconstruct rapidly decaying signals through greedy strategies. The findings are applied to analyze simulated asteroseismic and phase retrieval problems, but may also have wider applications. For sparse signals that do not necessarily satisfy rapid decay assumptions, but can be well approximated by sparse vectors, we discuss iterative hard-thresholding and also develop an iterative soft-thresholding algorithm.

Algorithms for Finite Projected Entangled Pair States

Michael Lubasch, Juan Ignacio Cirac, and Mari-Carmen Bañuls
Max-Planck Institut für Quantenoptik, Garching, Germany

Projected Entangled Pair States (PEPS) represent a variational ansatz for the wave function of a quantum many-body system, that naturally generalizes Matrix Product States (MPS), also known as Tensor Trains (TT), to arbitrary lattice dimensions and geometries. While MPS have proven extremely successful in the numerical simulation of 1-dimensional physical systems, their application to higher dimensional problems requires exponentially increasing resources. This could be understood lately from results emerging at the interface between quantum information and condensed matter theory. These results also motivated the construction of PEPS as the correct and promising generalization of MPS. However, the computational cost of the originally proposed PEPS algorithm is significantly higher than that of its MPS counterpart.

In recent years, several methods have been proposed for PEPS with the goal of improving the original algorithm. We analyze and compare the most promising of them within the context of finite PEPS. In conclusion, we introduce a new scheme that naturally interpolates between the cheapest and most imprecise and the most costly and most precise algorithms.

State of the Art of the Sparse Grid Combination Technique

Matthias Wong
Mathematics, Australien National University, Canberra, Australia

Consider an infinite dimensional vector space V with subspaces $V_{i,j}$ having dimension 2^{i+j}. We seek to approximate some $u \in V$ using $u_{i,j} \in V_{i,j}$, especially in the case when $i = j = n$ for some large values of n.

It is known that the linear combination

$$u^c_{n,n} = \sum_{i+j=n+1} u_{i,j} - \sum_{i+j=n} u_{i,j}$$

is a good approximation to $u_{n,n}$ under the so-called error splitting assumption. This special linear combination is the sparse grid combination technique introduced in Griebel et al (1992).

In this talk, we discuss the most up-to-date research concerning the combination technique. This includes theoretical results and practical variations including dimensional adaptivity and the Opticom method of Hegland et al (2006).

The Sparse Grid Combination Technique in Linear Gyrokinetics

Christoph Kowitz
Computer Science, Technische Universität München, Garching, Germany

The coming large fusion experiment ITER will heavily benefit from numerical simulations. One model for simulating the hot plasmas occurring in such a device are the gyrokinetic equations, which can resolve the microturbulence in the plasma. Due to their moderately high dimensionality they could profit from using sparse grids. Since the highly efficient and parallelized simulation code GENE is already at hand, the sparse grid combination technique can be used to create sparse grid solutions. In this way, the full parallelism of GENE can be used and on top of it, another layer of parallelism is introduced. This approach can on the one hand be one step toward exascale computations, since it the parallelism acts on top of the current application. On the other hand it can also be used to reduce the effect of hardware faults, which will probably occur more often on exascale architectures. We will present some results of the combination technique for the special case of linear computations in GENE. For that the optimized combination technique is used, which allows to adapt the coefficients for combination to the underlying problem, which is in our case an eigenvalue problem. It computes its combination coefficients out of an optimization problem, which is also including the search for single eigenvalues of the system. In the end, the method will have retrieved an approximation of the eigenvalue and a representation of the eigenvector in the basis of the partial solutions used for combination. This method can also be generalized to other basis functions than the partial solutions and might thus be also employed for more general problems than the gyrokinetic eigenvalue problem.

Unsolved Problems

At the end of the workshop and just before the conference dinner, we propose to have a group discussion on ”Outstanding Problems” to which we would like all our participants to participate with short (we think of around 5 mins) presentations of so far unsolved (and important) problems in the areas of the workshop. Each presentation would be followed by a short discussion on ideas where a solution might be found or how to deal with the proposed problem. To prepare for the session, we would like to know if you are willing to participate with one (or more?) such presentation. Although we want to be adaptive, and use the session also as a conclusion of the workshop, it would be nice if you could already indicate if you feel like giving such a presentation, maybe even with a title. We will adaptively (even during the workshop) collect the proposals and put them on our webpage. As you know, there have been some famous ”collections of new problems” in the past (one in the book of Wohlers, and MTNS did the exercise once), it is a good thing to continue the effort that may inspire future generations.