5th Workshop on

High-Dimensional Approximation

2013

PROGRAM and ABSTRACTS

February 11 – 15, 2013

Canberra, Australia

5th Workshop on

High-Dimensional Approximation

2013

PROGRAM and ABSTRACTS

February 11 – 15, 2013

Canberra, Australia

Organized by

Mathematical Sciences Institute, Australian National University

Scientific Program Committee

Jochen Garcke, Universität Bonn, Fraunhofer-Institute SCAI, Germany
Michael Griebel, Universität Bonn, Fraunhofer-Institute SCAI, Germany
Wolfgang Hackbusch, Max Planck Institute Leipzig, Germany
Markus Hegland, ANU, Australia, IAS Technische Universität München, Germany
Frances Kuo, University of New South Wales, Australia
Christoph Schwab, ETH Zürich, Switzerland
Ian Sloan, University of New South Wales, Australia
Greg Wasilkowski, University of Kentucky, USA
Henryk Woźniakowski, Columbia University, USA, Warsaw University, Poland
Harry Yserentant, Technische Universität Berlin, Germany

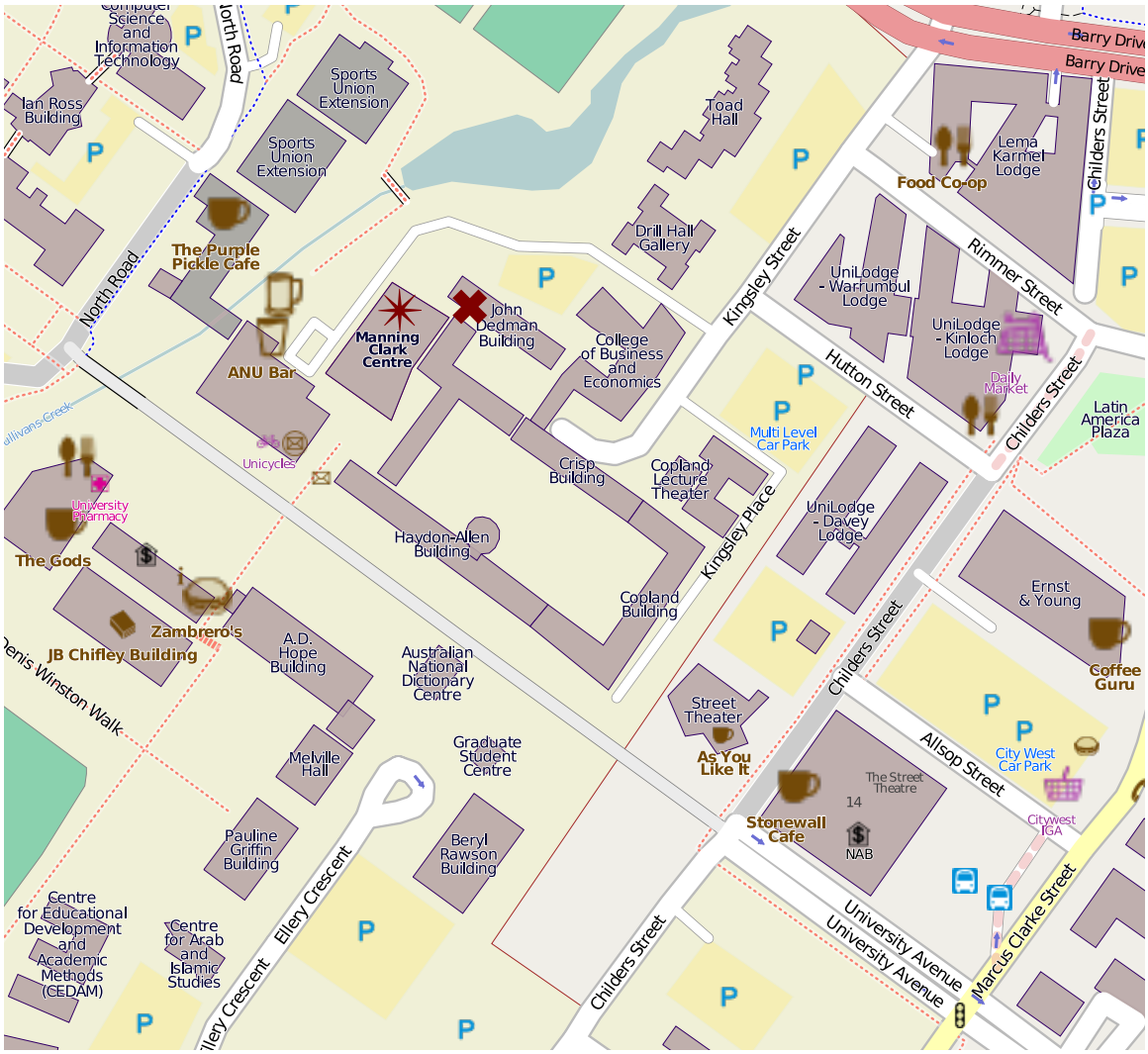
Organization

Jochen Garcke
Valeriy Khakhutskyy
Christoph Kowitz
Brittany Shoard

High-Dimensional Approximation, February 11 – 15, 2013

	Monday	Tuesday	Wednesday	Thursday	Friday
8:00 am	Registration				
8:30 am	Address of Welcome				
9:00 am	Wong, Harding	Gerstner, Kuo	Mayer, Zwicknagl	Webster, Kumar	Nichols, Migliorati
9:30 am					
10:00 am	Coffee Break	Coffee Break	Coffee Break	Coffee Break	Coffee Break
10:30 am	Kowitz, Hupp	Nuyens, Gnewuch	Dinh, Rieger	Schneider, Espig	Hegland, Schwab
11:00 am					
11:30 am	Lunch Break	Lunch Break	Lunch Break	Lunch Break	
12:00 am					
12:30 am					
1:00 pm	Aistleitner, Le Gia	Woźniakowski, Bohn	Oettershagen, Hinrichs	Hackbusch, Daum	
1:30 pm					
2:00 pm					
2:30 pm	Coffee Break	Coffee Break	Coffee Break	Coffee Break	
3:00 pm	Rudolf, Brauchart	Wissel, Pflüger	Yu, Zhang	Garcke, Khakhutskyy	
3:30 pm					
4:00 pm					
4:30 pm					
5:00 pm					
5:30 pm					
6:00 pm			Dinner @University House		
6:30 pm					
7:00 pm					
7:30 pm					
8:00 pm					
8:30 pm					
9:00 pm					
9:30 pm					
10:00 pm					

Directions and Orientation



Location

The workshop will be held at the Manning Clark Centre of The Australian National University.

Directions

There are many ways to get around Canberra. Below is some useful information about Bus and Taxi transport around the ANU, the Airport and surrounding areas.

Taxi

If you are catching a taxi to the ANU Mathematical Sciences Institute, ask to be taken to Kingsley Street, Drill Hall Gallery. From here, it is only a short walk to the Mathematical Sciences Building (John Dedman, Building 27). A Taxi will generally cost around $40 and will take roughly 15 minutes. Pricing and time may vary depending on traffic.

Airport Shuttle

Airport express buses to the city are cheaper but they will not take you directly to your destination. The buses stop at 3 locations marked in the map below. The most convenient stop for getting yourself to the ANU is West Row, Civic.

Shuttle prices:

- $10 one way
- $20 return

Action Buses

Canberra buses are a cheap and easy way of getting around town once you're here.

For more information about bus services visit: `https://www.action.act.gov.au/tourist_information.html`

Ticket prices:

- Adult single trip with 90 minute transfer: $4.20
- Adult daily: $8.00
- Concession and student single trip with 90 minute transfer: $2.10
- Concession and student daily: $4.00

Top Sights and Activities in Canberra

- Australian War Memorial
- National Gallery of Australia
- National Museum of Australia
- Cockngton Green Gardens
- National Zoo and Aquarium
- National Botanic Gardens

Conference Dinner, Wednesday, February 13th

The 5th High-Dimensional Approximation conference dinner will be held on Wednesday the 13th of February in the Drawing room of University House. The event will start at 6pm. Dinner will be served in the form of an Aussie BBQ!

PROGRAM

8.30–9.30 **Registration**
9.00–10.20 **Session**

1. **An error expansion for the generalized combination technique** – *Matthias Wong*
2. **Challenges solving high dimensional problems on supercomputers** – *Brendan Harding and Markus Hegland*

10.20–10.50 **Coffee Break**
10.50–12.10 **Session**

1. **An opticom method for computing eigenpairs in linear gyrokinetics** – *Christoph Kowitz and Markus Hegland*
2. **Efficient Communication in the Sparse Grid Combination Technique** – *Philipp Hupp, Riko Jacob and Markus Hegland*

12.10–13.30 **Lunch**
13.30–15.10 **Session**

1. **On the inverse of the star-discrepancy** – *Christoph Aistleitner*
2. **A multiscale method for the Stokes problem on bounded domains** – *Andrew Chernih and Thong Le Gia*

15.10–15.40 **Coffee Break**
15.40–17.00 **Session**

1. **Hybrid Slice Sampler** – *Daniel Rudolf*
2. **Testing and constructing sequences of point configurations on high-dimensional spheres and applications** – *Johann Brauchart*

9.00–10.20 **Session**

1. **Randomized and non-randomized multilevel quasi-Monte Carlo path simulation** – *Thomas Gerstner and Marco Noll*

2. **Multi-level quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients** – *Frances Kuo, Christoph Schwab and Ian Sloan*

10.20–10.50 **Coffee Break**

10.50–12.10 **Session**

1. **Integration of permutation-invariant functions** – *Dirk Nuyens, Gowri Suryanarayana and Markus Weimar*

2. **New Results on Infinite-Dimensional Integration in the Deterministic and Randomized Setting** – *Michael Gnewuch, Jan Baldeaux and Josef Dick*

12.10–13.30 **Lunch**

13.30–15.10 **Session**

1. **Exponential Convergence and Tractability** – *Henryk Woźniakowski*

2. **On dimensionality reduction for high-dimensional machine learning with sparse grids** – *Bastian Bohn and Michael Griebel*

15.10–15.40 **Coffee Break**

15.40–17.00 **Session**

1. **Identification of arbitrarily oriented low-dimensional subspaces in high-dimensional data** – *Daniel Wissel*

2. **Efficient High-Dimensional Regularization in the Sparse Grid Basis** – *Markus Hegland and Dirk Pflüger*

9.00–10.20 **Session**

1. **On Weighted Hilbert Spaces And Integration of Functions of Infinitely Many Variables** – *Sebastian Mayer*

2. **Expansion kernels and their approximation properties** – *Barbara Zwicknagl*

10.20–10.50 **Coffee Break**

10.50–12.10 **Session**

1. **Sampling recovery and cubature on sparse grids** – *Dung Dinh*

2. **Approximation error bounds and the concentration of measure effect** – *Christian Rieger*

12.10–13.30 **Lunch**

13.30–15.10 **Session**

1. **Integration in reproducing kernel Hilbert spaces** – *Jens Oettershagen and Michael Griebel*

2. **The Curse of Dimensionality for Numerical Integration of Smooth Functions** – *Aicke Hinrichs*

15.10–15.40 **Coffee Break**

15.40–17.00 **Session**

1. **Spectral Sparse Grid Methods for Elliptic Equations on Non-periodic Domains** – *Haijun Yu and Jie Shen*

2. **A hyper-spherical sparse grid approach for high-dimensional discontinuity detection** – *Guannan Zhang, Max Gunzburger and Clayton Webster*

9.00–10.20 **Session**

1. **Adaptive sparse grid stochastic collocation methods for high-dimensional random data identification problems** – *Clayton Webster, Max Gunzburger and Guannan Zhang*

2. **Construction of Tensor Chain using cross approximation techniques** – *Naraparaju Kishore Kumar, Schneider Jan and Espig Mike*

10.20–10.50 **Coffee Break**

10.50–12.10 **Session**

1. **Approximation of arbitrary tensor networks** – *Jan Schneider*

2. **Alternating Least Squares Optimisation in Tensor Format Representations** – *Mike Espig, Wolfgang Hackbusch and Aram Khachatryan*

12.10–13.30 **Lunch**

13.30–15.10 **Session**

1. **Maximum norm estimates for SVD truncated tensors** – *Wolfgang Hackbusch*

2. **Solution of the Monge-Kantorovich transport problem in high dimensions** – *Fred Daum*

15.10–15.40 **Coffee Break**

15.40–17.00 **Session**

1. **Using Hyperbolic Cross Approximation to measure and compensate Covariate Shift** – *Jochen Garcke and Thomas Vanck*

2. **Parallel Solution of Regression Problems Using Sparse Grids and Alternating Direction Method of Multipliers** – *Valeriy Khakhutskyy, Dirk Pflüger and Markus Hegland*

9.00–10.20 **Session**

1. **Weighted Hilbert spaces in the porous flow problem and associated numerical challenges** – *James Nichols*

2. **Adaptive polynomial approximation in high dimensional PDE with stochastic data** – *Giovanni Migliorati, Fabio Nobile and Raul Tempone*

10.20–10.50 **Coffee Break**

10.50–12.10 **Session**

1. **On tensor trains and signaling cascades** – *Jochen Garcke and Markus Hegland*

2. **Quadrature approach to Bayesian Inverse Problems** – *Christoph Schwab and Claudia Schillings*

ABSTRACTS

An error expansion for the generalized combination technique

Matthias Wong

Australian National University, Australia; matthias.wong@anu.edu.au

The sparse grid combination technique proposed in Griebel (1992) proved immensely useful in multidimensional applications. Traditional error analysis of the technique focused on an error splitting assumption. However, in applications the combination technique displays a wider range of phenomenon than the error splitting suggests. For example the technique may fail altogether. Moreover, the error splitting analysis do not apply to applications which uses component grids different from the traditional ones in Griebel (1992). In this paper, we start with a work by Hegland (2001) which generalized the sparse grid combination technique to include different component grids. We derive a new error expansion of any such combination technique and an corollary error bound. The new error expansion does not assume the traditional error splitting. When the splitting is assumed, the convergence rate $\mathcal{O}(h^2 \log(h^{-1}))$ is immediately derived.

Challenges solving high dimensional problems on supercomputers

Brendan Harding[1] *and Markus Hegland*[2]

Australian National University, Australia
[1]brendan.harding@anu.edu.au
[2]markus.hegland@anu.edu.au

Challenges for efficiently and effectively using petascale and exascale computers include managing many levels of parallelism and the ability to cope with run-time errors. Without robustness, applications developed for these machines will have little chance of completing successfully. The sparse grid combination technique approximates the solution to a given problem by taking the linear combination of its solution on multiple grids. It has been applied successfully to many high dimensional problems due to its ability to tackle the curse of dimensionality. In this presentation we discuss several approaches to fault-tolerance using the combination technique. The first of these is the observation that implementing within the MapReduce model means we can utilise the existing fault-tolerance in this model. In addition, this presents another level of parallelism which is easily exploited. Finally, we describe a novel approach in which the solution is computed on some extra grids which can be used for alternate combinations if other grids experience failure. We include some results based on the solution of the 2 and 3–dimensional scalar advection PDE.

An opticom method for computing eigenpairs in linear gyrokinetics

Christoph Kowitz[1] *and Markus Hegland*[2]

[1]Technische Universität München, Germany; kowitz@in.tum.de
[2]Australian National University, Australia; markus.hegland@anu.edu.au

One model for simulating hot plasmas occuring in experimental devices for magnetically confined fusion are the five-dimensional gyrokinetic equations, which can resolve the plasmas microturbulence. An analysis of the eigenvalues of the linearized gyrokinetic problem reveals the unstable modes driving the microturbulence already. The linearized problem is a four-dimensional one and it could thus heavily profit from using sparse grids. Since the highly efficient and parallelized simulation code GENE is already at hand, a combination technique approach is used.

Applying the classical combination technique to the gyrokinetic eigenvalue problem can be done for the eigenvalues, whereas the combination of the respective eigenvectors requires an additional scaling. The optimized combination technique can be used to do that appropriate scaling. For that the underlying minimization function has to be reformulated to suit the eigenvalue problem. A method for computing eigenvalues by solving an optimization problem is employed to then retrieve the eigenpair using the combination technique. A few test scenarios are investigated, showing the applicability of the method.

Efficient Communication in the Sparse Grid Combination Technique

Philipp Hupp[1]*, Riko Jacob*[2] *and Markus Hegland*[3]

[1]ETH Zürich, Switzerland; philipp.hupp@inf.ethz.ch
[2]ETH Zürich, Switzerland; rjacob@inf.ethz.ch
[3]Australian National University, Australia; Markus.Hegland@anu.edu.au

Sparse grids are designed to break the curse of dimensionality and al-low to work in high dimensions. In comparison to regular full grids, sparsegrids exclude highly mixed terms. Thereby the number of grid points issignificantly reduced while the accuracy of the solution deteriorates onlyslightly. Numerical problems on sparse grids can either be solved workingdirectly in the sparse grid space which results in more complicated dataaccess patterns.

Another approach is the combination technique which combines usual full grids,called combination grids, to form the sparse grid. To use iter-ative methods the solutions on the combination grids are aggregated toform the sparse grid solution. Afterwards the combined solution on thesparse grid is propagated back to the combination grids to perform thenext iteration.

In high performance computing parallelism and memory access arethe factors driving the development of efficient algorithms. Working onthe combination grids is inherently parallel as the combination grids areindependent of each other. As well, full grids are well understood andmemory efficient algorithms for problems on full grids are known and canbe exploited. However, aggregating the combination grid solutions to formthe sparse grid solution and distributing the sparse grid solution to thecombination grid has not yet been examined.This talk introduces an algorithmic idea to efficiently tackle this task:Working with the hierarchical surpluses, the combination grids are splitinto the hierarchical subspaces. Data is gathered for and distributed fromthe hierarchical subspaces. Communicating to and from each individualsubspace minimizes the total amount of data to be communicated. Also,several hierarchical subspaces can be merged together to reduce the

num-ber of messages passed. However, this increases the amount of data tobe communicated. In the extreme case that all hierarchical subspacesare merged and the combination grids communicate only with one singlenode, the sparse grid, the number of messages passed is minimal.

On the inverse of the star-discrepancy

Christoph Aistleitner

University of New South Wales, Australia; aistleitner@math.tugraz.at

The *inverse of the star-discrepancy* $N^*(d,\varepsilon)$ denotes the smallest possible cardinality of a point set in the d-dimensional unit cube $[0,1]^d$ having star-discrepancy at most ε. By a result of Heinrich, Novak, Wasilkowski and Woźniakowski, we have

$$N^*(d,\varepsilon) \leq c_{\text{abs}} d \varepsilon^{-2}.$$

On the other hand, a lower bound of Hinrichs states that

$$N^*(d,\varepsilon) \geq c_{\text{abs}}\ d \varepsilon^{-1}.$$

Thus the dependence of $N^*(d,\varepsilon)$ on the dimension d is precisely known, while the dependence on ε is an open problem.
Reporting on current work in progress, we present a new method which leads to a significant improvement of the known bounds on the inverse of the star-discrepancy. Furthermore, we show how the same method can be inappropriately applied to an other important open problem in discrepancy theory.

A multiscale method for the Stokes problem on bounded domains

Andrew Chernih[1] *and Thong Le Gia*[2]

University of New South Wales, Australia
[1]andrew@andrewch.com
[2]qlegia@unsw.edu.au

In this work, we investigate the application of radial basis functions (RBFs) for the approximation with collocation of the Stokes problem. The approximate solution is constructed in a multi-level fashion, each level using compactly supported radial basis functions with decreasing scaling factors. We use symmetric collocation and give sufficient conditions for convergence and stability analysis is also presented. Numerical experiments support the theoretical results.

Hybrid Slice Sampler

Daniel Rudolf

University of New South Wales, Australia; d.rudolf@unsw.edu.au

We study a hybrid slice sampler Markov chain for the approximation of a distribution. This distribution might be given by a possibly non-normalized density function.For the simple slice sampler it is assumed that one can sample the uniform distribution on the slices of the density. This assumption is very restrictive in high dimensions. We assume that on every slice a Markov chain with uniform limit distribution is given. Roughly spoken, the hybrid slice sampler works as follows: First, choose a slice and then do one step with the Markov chain on the slice.We show a lower bound of the absolute spectral gap of the hybrid slice sampler in terms of the absolute spectral gap of the simple slice sampler and properties of the Markov chains on the slices.

Testing and constructing sequences of point configurations on high-dimensional spheres and applications

Johann Brauchart

University of New South Wales, Australia; j.brauchart@unsw.edu.au

The explicit and efficient construction of a sequence of point configurations on the unit sphere $\mathbb{S}^d$ in the Euclidean space $\mathbb{R}^{d+1}$ with good distribution properties is a challenging task, more so in higher dimension d, as one has to cope with the severe restrictions emerging from the topology and the symmetry of the d-sphere. Our main motivation and application is the numerical integration of functions defined on $\mathbb{S}^d$ [and in $\mathbb{R}^{d+1}$ (by providing additional radial information)]. Consequently, as a quality criterion we use the worst-case numerical integration error (WCE) for equal-weight (QMC) rules for functions from the unit ball in a Sobolev space $\mathbb{H}^s(\mathbb{S}^d)$ over the unit sphere. For sufficient high smoothness s ($s > d/2$) this function space becomes a reproducing kernel Hilbert space. Provided with a particular type of reproducing kernel (generalized distance kernel), the corresponding WCE has a simple closed form (and can be given a geometric interpretation as a discrepancy generalizing the so-called spherical cap $\mathbb{L}_2$-discrepancy). Good point configurations have small WCE (for a given s). This leads to the concept of *QMC design sequences for* $\mathbb{H}^s(\mathbb{S}^d)$, where the node sets, by definition, have optimal order of WCE (arXiv:1208.3267v1 [math.NA]). The introduction of weights (weighted QMC rules) and their adjustment by minimizing the associated WCE improves the performance of the rules in certain cases when equal weights would give considerably worse results (say, e.g., random points). We report on our joint work in progress with Rob Womersley on these so-called *weighted QMC design sequences*.

By including radial information one can generate certain "well-distributed" point set sequences in $\mathbb{R}^{d+1}$. We introduce the concept of spherical cone discrepancy to measure the quality of such node sets and present work in progress (including applications in financial modeling) from joint work with Lou Fang and Josef Dick.

This talk is based on work in progress with Lou Fang, Josef Dick and Rob Womersley.

Randomized and non-randomized multilevel quasi-Monte Carlo path simulation

Thomas Gerstner[1] *and Marco Noll*[2]

University of Frankfurt, Germany
[1]gerstner@math.uni-frankfurt.de
[2]noll@math.uni-frankfurt.de

We present theoretical and numerical results with regard to convergence rates, for a combination of the multilevel Monte Carlo method with quasi-Monte Carlo number generation in finance applications. A complexity theorem, applicable for all path simulation methods that use the multilevel approach, shows that for many payoff functions in finance the convergence rate is improved significantly by using quasi-random numbers instead of random numbers. The theoretical results are confirmed by numerical examples with several types of options. We also discuss randomization approaches, which appear to play a crucial role in variance estimation.

Multi-level quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients

Frances Kuo[1]*, Christoph Schwab*[2] *and Ian Sloan*[3]

[1]School of Mathematics and Statistics, University of New South Wales, Australia; f.kuo@unsw.edu.au
[2]Seminar for Applied Mathematics, ETH Zurich, Switzerland; schwab@sam.math.ethz.ch
[3]School of Mathematics and Statistics, University of New South Wales, Australia; i.sloan@unsw.edu.au

Quasi-Monte Carlo (QMC) methods are applied to multi-level Finite Element (FE) discretizations of elliptic partial differential equations (PDEs) with a random coefficient, to estimate expected values of linear functionals of the solution. The expected value is considered as an infinite-dimensional integral in the parameter space corresponding to the randomness induced by the random coefficient. We use a multi-level algorithm, with the number of QMC points depending on the discretization level, and with a level-dependent dimension truncation strategy. In some scenarios, we show that the overall error is $\mathcal{O}(h^2)$, where h is the finest FE mesh width, or $\mathcal{O}(N^{-1+\delta})$ for arbitrary $\delta > 0$, where N is the maximal number of QMC sampling points. For these scenarios, the total work is essentially of the order of one single PDE solve at the finest FE discretization level. The analysis exploits regularity of the parametric solution with respect to both the physical variables (the variables in the physical domain) and the parametric variables (the parameters corresponding to randomness). Families of QMC rules with "POD weights" ("product and order dependent weights") which quantify the relative importance of subsets of the variables are found to be natural for proving convergence rates of QMC errors that are independent of the number of parametric variables. Conditions on the data of the problem to achieve a certain rate of convergence coincide with the sufficient conditions obtained in Cohen, De Vore, and Schwab (2010).

Integration of permutation-invariant functions

Dirk Nuyens[1], *Gowri Suryanarayana*[2] *and Markus Weimar*[3]

[1]KU Leuven, Belgium; dirk.nuyens@cs.kuleuven.be
[2]KU Leuven, Belgium; gowri.suryanarayana@cs.kuleuven.be
[3]University Jena, Germany; markus.weimar@uni-jena.de

We study integration of functions which are invariant under permutation (of a subset) of their arguments. We find conditions for (strong) polynomial tractability and study the application of shifted lattice rules.

New Results on Infinite-Dimensional Integration in the Deterministic and Randomized Setting

Michael Gnewuch[1], *Jan Baldeaux*[2] *and Josef Dick*[3]

[1]University of New South Wales, Sydney, Australia; mig@informatik.uni-kiel.de
[2]University of Technology Sydney, Australia; Jan.Baldeaux@uts.edu.au
[3]University of New South Wales, Sydney, Australia; josef.dick@unsw.edu.au

In this talk we present new upper and lower error bounds for the infinite-dimensional numerical integration problem on weighted Hilbert spaces with norms induced by an underlying function decomposition of ANOVA or anchored type. Here the weights model the relative importance of different groups of variables. We have results for randomized and deterministic algorithms, and our error bounds are in both settings sharp. The upper error bounds are based on randomized or deterministic multilevel algorithms and/or on deterministic changing dimension algorithms.

The focus of the talk will be on the randomized ANOVA setting, which is technically more demanding. Our analysis in paper [1] refines and extends the analysis provided in [F. J. Hickernell, T. Müller-Grobach, B. Niu, K. Ritter, J. Complexity 26 (2010), 229-254], and leads to matching upper and lower (i.e., optimal) error bounds. As an illustrative example, we discuss the unanchored Sobolev space and employ randomized quasi-Monte Carlo (QMC) multilevel algorithms based on scrambled polynomial lattice rules.

The talk is based on joint work with Jan Baldeaux (UTS, Sydney) and Josef Dick (UNSW, Sydney). Our findings are presented in full detail in the references below.

References

[1] J. Baldeaux, M. Gnewuch. Optimal randomized multilevel algorithms for infinite-dimensional integration on function spaces with ANOVA-type decomposition. arXiv:1209.0882v1 [math.NA], Preprint 2012.

[2] J. Dick, M. Gnewuch. Infinite-dimensional integration in weighted Hilbert spaces: anchored decompositions, optimal deterministic algorithms, and higher order convergence. arXiv:1210.4223 [math.NA], Preprint 2012.

[3] M. Gnewuch. Lower error bounds for randomized multilevel and changing dimension algorithms. arXiv:1209.1808 [math.NA],Preprint 2012. (To appear in: J. Dick, F. Y. Kuo, G. W. Peters, I. H. Sloan (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2012, Springer.

Exponential Convergence and Tractability

Henryk Woźniakowski

Columbia University and University of Warsaw , USA and Poland; onk62@hydra.mimuw.edu.pl

For analytic functions we can try to study tractability with respect to $\log 1/\epsilon$ and d. Here ϵ is an error bound and d is the number of variables of a multivariate problem.

I will cover recent results obtained together with J. Dick, P. Kritzer and F. Pillichshammer on multivariate integration and approximation defined over weighted Korobov spaces of analytic functions. We find necessary and sufficient conditions on weights of the Korobov space to get various notions of tractability such as strong polynomial, polynomial, quasi-polynomial and weak tractability.

On dimensionality reduction for high-dimensional machine learning with sparse grids

Bastian Bohn[1] *and Michael Griebel*[2]

University of Bonn, Germany
[1]bohn@ins.uni-bonn.de
[2]griebel@ins.uni-bonn.de

High-dimensional ($d > 3$) machine learning tasks are often tackled by support vector machines, radial basis function neural networks and other methods which employ data-centered discretizations spacesto explore given input data points and the manifold they reside on.For many applications a large amount of data ($n > 10^3$) is available and a straightforward application of the aforementioned techniques is not feasible. Algorithms with non-data-based tensor-product grid discretizations directly allow for linear runtime with respect to n. However the curse of dimensionality (i.e.the size $\mathcal{O}(M^d)$ of the discretization space – where M denotes the amount of basis functions in one space direction) prohibits an application to higher-dimensional problems. To this end, sparse grid methods have been successfully employed in this context to reduce the amount of basis functions considerably.Nevertheless, for dimensions $d \gg 10$ these algorithms are only viable if the dimension m of the embedded manifold on which the high-dimensional data points lie is small enough ($m \ll d$) and if the problem is solved in this intrinsic coordinate system. To this end, we will introduce a dimension-adaptive variant which explores the structure of the ANOVA (analysis of variance) decomposition of the solution.Furthermore, the dimension of the discretization space can be significantly reduced by exploiting the non-uniqueness of the parametrization of the embedded manifold. Finally, we present results for the application to a real-world car-crash simulation data-set.

Identification of arbitrarily oriented low-dimensional subspaces in high-dimensional data

Daniel Wissel

Institute for Numerical Simulation, University of Bonn, Germany; wissel@ins.uni-bonn.de

We present an efficient method to detect low-dimensional subspaces in high-dimensional data. The subspaces can be arbitrarily oriented and intersecting each other. The method relys on local Euclidean distance information and elementary geometric properties. The complexity of the most expensive subroutine does not depend on the dimension of the data space, but only on the dimension of the search space.

Efficient High-Dimensional Regularization in the Sparse Grid Basis

Markus Hegland[1] *and Dirk Pflüger*[2]

[1]Australian National University, Australia; markus.hegland@anu.edu.au
[2]Universität Stuttgart, Germany; Dirk.Pflueger@ipvs.uni-stuttgart.de

Penalized least squares methods have a long history in machine learning. Penalties or regularisation are used to control overfitting (fitting of the errors) by suppressing small scales. Such a filtering works when the underlying signal is dominated by large scales. This is essentially a smoothness condition.

We consider here learning methods which use finite dimensional function spaces, and, in particular, piecewise multilinear functions. Conventional approaches employ ansatz functions which are associated to the data to learn from. Therefore, they typically scale at least quadratically or even worse in the size of the training data set. They are thus unsuited to deal with arbitrarily large data sets. Mesh- or grid-based approaches, in contrast, result in algorithms that scale only linearly in the size of the data set to train from. But as they suffer the curse of dimensionality, they are seldom considered in learning. To mitigate the curse, we employ sparse grids in their spatially adaptive variant.

The classical penalization with the squared norm of the gradient leads to sparse matrices in the corresponding linear system in the nodal point basis. It it thus well suitable for the combination technique. There, diagonal penalties do not work. In the hierarchical basis, however, the classical approach leads to dense stiffness matrices and thus high effort. But here, diagonal penalties are well-suited and sufficient for regularization. They are even more flexible and more stable with respect to the choice of grids. We discuss the rationale for diagonal penalties and the reason why this works.

On Weighted Hilbert Spaces And Integration of Functions of Infinitely Many Variables

Sebastian Mayer

Universität Bonn, Institut für numerische Simulation, Germany; mayer@ins.uni-bonn.de

For functions of infitinely many variables $x_1, x_2, \dots$ taking values in a non-empty set D, the study of quadrature problems and their complexity is often based on weighted reproducing kerneld Hilbert spaces $H(K)$ as models for the function spaces where integrands are taken from.

These kernel spaces are constructed from a univariate kernel k on D by building the kernel $K(\mathbf{x}, \mathbf{y}) = \sum_{u \subset \mathbb{N}, |u| < \infty} \gamma_u \prod_{j \in u} k(x_j, y_j)$ with a family of weights $(\gamma_u)_u$.

This talk, based on a joint work with Micheal Gnewuch and Klaus Ritter, is devoted to some aspects of the analytical foundations of the weighted kernel approach. First we address conditions on the kernel k and the weights γ_u which allow for the construction sketched above. Then we turn to the question how integration on $H(K)$ can be defined based on a probability measure ρ on D such that $H(k) \subseteq L_1(\rho)$.

Expansion kernels and their approximation properties

Barbara Zwicknagl

Institute for Applied Mathematics, Bonn university, Germany; zwicknagl@iam.uni-bonn.de

We discuss approximation properties of kernels with multiscale structure. The latter are positive definite kernel functions that possess series expansions in terms of simple basis functions. Typical examples of such basis functions include monomials, wavelet-type function systems, or eigenfunctions of Laplace-type operators. The approximation properties of trial spaces built by translates of such series kernels are discussed. We characterize the interplay between the multiscale structure of the kernel and the choice of data points, and present applications to the effective numerical solution of variational problems on manifolds.

This is partly based on joint works with M. Griebel, C. Rieger, and R. Schaback.

Sampling recovery and cubature on sparse grids

Dung Dinh

Vietnam National University, Hanoi, Viet Nam; dinhzung@gmail.com

Let $X_n = \{x^j\}_{j=1}^n$ be a set of n points in the d-cube $[0,1]^d$, and $\Phi_n = \{\varphi_j\}_{j=1}^n$ a family of n functions in the space $L_q([0,1]^d)$, $0 < q \le \infty$. We consider the approximate recovery in $L_q([0,1]^d)$ of functions f on $[0,1]^d$ from the sampled values $f(x^1), ..., f(x^n)$, by the linear sampling algorithm

$$L_n(X_n, \Phi_n, f) \ := \ \sum_{j=1}^{n} f(x^j)\varphi_j.$$

Functions f to be recovered are from the unit ball of Besov type spaces $B_{p,\theta}^{\alpha,\beta}$ of a "hybrid" of mixed smoothness α and isotropic smoothness β. We constructed optimal linear sampling algorithms $L_n(X_n^*, \Phi_n^*, \cdot)$ on special sparse grids X_n^* and a family Φ_n^* of linear combinations of integer or half

integer translated dilations of tensor products of B-splines. We computed the asymptotic of the error of the optimal recovery. As consequences we obtained the asymptotic of optimal cubature formulas for integration of functions from from the unit ball of $B^{\alpha,\beta}_{p,\theta}$.

Approximation error bounds and the concentration of measure effect

Christian Rieger

Institute for Numerical Simulation, University of Bonn, Germany; rieger@ins.uni-bonn.de

We present special a priori error estimates for function reconstructions that take into account the concentration of measure effect. The aim is to understand why in some cases the curse of dimensionality can be alleviated or even circumvented.

As observed by many authors, the latter does often not affect the asymptotic convergence rates but rather the constants which blow up with dimension. Therefore to achieve small errors, one would need many more degrees of freedom than actual computers allow for.

We show how, in our situation, the concentration of measure effect allows to explicitly characterize the constants in the error bounds in terms of geometrical information such as metrics. As observed in literature, this allows not only to understand why certain experiments work in practice but also to systematically identify settings in which good behavior of the constants is guaranteed. This is joint work with Michael Griebel (Uni Bonn) and Petr Onyshchuk.

Integration in reproducing kernel Hilbert spaces

Jens Oettershagen[1] *and Michael Griebel*[2]

[1]Institute for Numerical Simulation, Germany; oettersh@ins.uni-bonn.de
[2]University of Bonn, Germany; griebel@ins.uni-bonn.de

The approximation of multivariate integrals is among the most important tasks in numerical analysis with numerous applications in computational finance, physics and econometrics.

In this talk we will deal with quadrature rules that are based on the so called 'spline algorithm', which is known to be optimal within a reproducing kernel Hilbert space for a fixed set of linear information.

Moreover, if the information is not fixed but given in terms of point-evaluation functionals at free nodes, the question of a good or even optimal choice of their design arises. We will compute univariate quadrature rules that fulfill certain optimality conditions and use sparse tensor products (i.e. sparse grids) to tackle high-dimensional integration problems.

The Curse of Dimensionality for Numerical Integration of Smooth Functions

Aicke Hinrichs

Friedrich-Schiller University Jena, Germany; a.hinrichs@uni-jena.de

We prove the curse of dimensionality for multivariate integration for a number of classes of smooth functions. In particular, for the class of r times continuously differentiable d-variate functions whose values are at most one the curse holds iff the bound on all derivatives up to order r does not go to zero faster than $d^{-1/2}$. We also consider the case of infinitely many differentiable functions and prove the curse if the bounds on the successive derivatives are appropriately large. The proof technique is based on a volume estimate of a neighborhood of the convex hull of n points which decays exponentially quickly if n is small relative to d.

This is joint work with Erich Novak, Henryk Woźniakowski and Mario Ullrich.

Spectral Sparse Grid Methods for Elliptic Equations on Non-periodic Domains

Haijun Yu[1] *and Jie Shen*[2]

[1]Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China; hyu@lsec.cc.ac.cn
[2]Purdue University, United States; shen7@purdue.edu

In this talk, we introduce Chebyshev sparse grid methods. Using the fast Chebyshevtransform on sparse grid and almost orthogonal polynomial bases in Galerkin form, we are able to build fast Chebyshev Galerkin and Chebyshev-Legendre Galerkin solvers for elliptic equations in higher-dimensional non-periodic domains. The Chebyshev Galerkin method is extended further to unbounded domains using delicated mapping functions. As an example, the eigenvalue problem of the one-dimensional n-particale Schrödinger equation is solved using this new built sparse grid method.

A hyper-spherical sparse grid approach for high-dimensional discontinuity detection

Guannan Zhang[1]*, Max Gunzburger*[2] *and Clayton Webster*[3]

[1]Oak Ridge National Lab, United States; zhangg@ornl.gov
[2]Florida State University, United States; mgunzburger@fsu.edu
[3]Oak Ridge National Lab, United States; webstercg@ornl.gov

High-dimensional discontinuity detection is significantly important to several areas of science and engineering. For example, in uncertainty quantification (UQ), it is directly related to risk assessment and predicting rare events. More generally, in an N-dimensional Euclidean space, the location of the discontinuity is generally an N-1-dimensional manifold, and therefore, the most difficult challenge is to accurately and efficiently represent such manifold. Conventional adaptive sparse-grid hierarchical interpolation has been employed to characterize such manifolds, however, the mesh refinement densely places grid points around the discontinuity, and fails to "sparsely" represent the

N-1 dimensional manifold.We propose a novel method for identifying jump discontinuities in high dimensional spaces by incorporating a hyper-spherical coordinate system (HSCS) into the sparse-grid approximation framework. The basic idea is to transform the Cartesian coordinate system to an N-dimensional HSCS and treat the manifold as the N-1 dimensional function in the subspace constituted by the N-1 angle coordinates. Then a sparse-grid approximation can be constructed in the subspace where the function value at each grid point is estimated by solving a nonlinear equation with use of the bisection or Newton method. This novel technique identifies the discontinuity with a reduced number of sparse grid points compared to existing methods. Several numerical examples will be presented that show the increased efficiency and accuracy of our approach for detecting the discontinuities.

Adaptive sparse grid stochastic collocation methods for high-dimensional random data identification problems

Clayton Webster[1], *Max Gunzburger*[2] *and Guannan Zhang*[3]

[1]Oak Ridge National Laboratory, United States; webstercg@ornl.gov
[2]Florida State University, United States; gunzburger@fsu.edu
[3]Oak Ridge National Laboratory, United States; zhangg@ornl.gov

Characterizing stochastic model inputs for physical and engineering problems relies on approximations in high-dimensional spaces, particularly in the case when experimental data or targets are affected by large amounts of uncertainty. To identify these high-dimensional random parameters we integrate several adaptive sparse grid stochastic collocation methods (aSGSC), constructed from both global and local multi-resolution interpolating polynomials, with 1) a SPDE-constrained least squares adjoint-based approach and 2), a Bayesian inference approach that uses Markov Chain Monte Carlo (MCMC) sampling. The advantage of our first technique is that it allows for optimal identification of statistical moments or even the whole posterior probability density function (PPDF) of the input random data, given the probability distribution of some response of a system of PDEs. Whereas, our second technique improves the computational efficiency of exploring the PPDF, required by Bayesian inference. Rigorously derived error estimates will be used to show the efficiency of these novel approaches at predicting the behavior of the stochastic parameters.

Construction of Tensor Chain using cross approximation techniques

Naraparaju Kishore Kumar[1], *Schneider Jan*[2] *and Espig Mike*[3]

[1]IIT Gandhinagar, India; thimmaki@gmail.com
[2]Max-Planck institute for Mathematics in the Sciences, Leipzig, Germany; jschneid@mis.mpg.de
[3]Max-Planck institute for Mathematics in the Sciences, Leipzig, Germany; Mike.Espig@mis.mpg.de

This talk deals with the approximation of d-dimensional tensors, as discrete representations of arbitrary functions $f(x_1, \ldots, x_d)$ on $[0,1]^d$, in the so-called Tensor Chain format using Skeleton/Cross Approximation methods. For example, the tensor chain approximation of three dimensional array looks as

$$A = (a_{ijk}) \approx \sum_{m_1=1}^{r_1} \sum_{m_2=1}^{r_2} \sum_{m_3=1}^{r_3} u_i^{(m_1,m_2)} v_j^{(m_2,m_3)} w_k^{(m_3,m_1)}.$$

where u_i, v_j, w_k depends only on one variable. The summation indices m_1, m_2, m_3 are referred to as auxiliary indices. The tensor chain in dimension 3 is shown in Fig. 1.

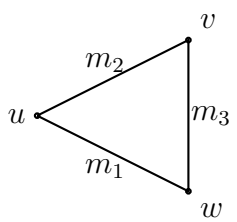


Figure 1: Tensor Chain graph for $d = 3$

Here the nodes symbolize the spatial variables and the edges show the summation indices.

The question arised whether skeleton/Cross Approximation methods would also be able to construct the Tensor Chain format with the ranks r_1, r_2, r_3 which are different from 1. But there was for some time the believe in the community that this is not possible. The main aim of this talk is to show that such an approximation is possible i.e it is possible to approximate a multidimensional array in Tensor Chain format with the ranks $r_1, r_2, ..., r_d$ which are different from 1 using matrix Cross Approximation methods. The complete algorithm is described and the complexity of the algorithm is shown to be linear in d. Some numerical examples are given to validate the theoritical results.

Approximation of arbitrary tensor networks

Jan Schneider

MPI for Mathematics in Sciences, Leipzig., Germany; jschneid@mis.mpg.de

We present ongoing work concerning the implementation of an approximation procedure for arbitrary tensor network structures. The method is a generalization of earlier Tensor Train - or Tensor Chain - versions based on successive pseudo skeleton decompositions. To validate the promising concept we show some numerical experiments in d=4.

Alternating Least Squares Optimisation in Tensor Format Representations

Mike Espig[1], *Wolfgang Hackbusch*[2] *and Aram Khachatryan*[3]

Max Planck Institute for Mathematics in the Sciences, Germany
[1]mike.espig@alopax.de
[2]Wolfgang.Hackbusch@mis.mpg.de
[3]Aram.Khachatryan@mis.mpg.de

The problem considered is formally a quadratic minimisation problem on the tensor product space $\mathcal{V} = \bigotimes_{\nu=1}^{d} \mathbb{R}^{m_\nu}$. The optimisation task is introduced by the quadratic function $f : \mathcal{V} \to \mathbb{R}$

$$f(v) := \frac{1}{\|b\|^2} \left[\frac{1}{2} \langle Av, v \rangle_{\mathcal{V}} - \langle b, v \rangle_{\mathcal{V}} \right], \tag{1}$$

where $A \in \mathbb{R}^{m_1 \cdots m_d \times m_1 \cdots m_d}$ is a positive definite matrix ($A > 0$, $A^t = A$) and $b \in \mathcal{V}$. In our analysis, a tensor u is represented in a tensor format U, i.e. $u = U(p_1, \ldots, p_L)$. Where a tensor format $U : P := \bigtimes_{\mu=1}^{L} P_\mu \to \mathcal{V}$ is a multilinear map from the cartesian product of parameter spaces P_μ ($d \leq L$) into the tensor space $\mathcal{V}$ and $(p_1, \ldots, p_L)^t \in P$ is called representation system of u respect to the tensor format U. The solution $v^* = \mathrm{argmin}_{v \in \mathcal{V}} f(v)$ is approximated respect to the tensor format representation U, i.e. we are looking for a representation system $(p_1^*, \ldots, p_L^*)^t \in P$ such that for

$$\begin{aligned} F &:= f \circ U : P \to \mathcal{V} \to \mathbb{R} \\ F(p_1, \ldots, p_L) &= \frac{1}{\|b\|^2} \left[\frac{1}{2} \langle AU(p_1, \ldots, p_L), U(p_1, \ldots, p_L) \rangle_{\mathcal{V}} - \langle b, U(p_1, \ldots, p_L) \rangle_{\mathcal{V}} \right] \end{aligned} \tag{2}$$

we have

$$F(p_1^*, \ldots, p_L^*) = \inf_{(p_1, \ldots, p_L)^t \in P} F(p_1, \ldots, p_L).$$

The alternating least squares (ALS) algorithm is recursively defined. Suppose that the k-th iterate $\underline{p}^k = (p_1^k, \ldots, p_L^k)^t$ and the first $\mu-1$ components $p_1^{k+1}, \ldots, p_{\mu-1}^{k+1}$ of the $(k+1)$-th iterate $\underline{p}^{k+1}$ have been determined. The basic step of the ALS algorithm is to compute the minimum norm solution

$$p_\mu^{k+1} := \operatorname*{argmin}_{q_\mu \in P_\mu} F(p_1^{k+1}, \ldots, p_{\mu-1}^{k+1}, q_\mu, p_{\mu+1}^k, \ldots, p_L^k).$$

Thus, in order to obtain $\underline{p}^{k+1}$ from $\underline{p}^k$, we have to solve successively L ordinary least squares problems. In our talk, the convergence of the ALS method is analysed.

References

[1] M. Espig and W. Hackbusch and A. Khachatryan. On the Convergence of Alternating Least Squares Optimisation in Tensor Format Representations, in preparation.

Maximum norm estimates for SVD truncated tensors

Wolfgang Hackbusch

Max-Planck Institute for Mathematics in the Sciences, Germany; wh@mis.mpg.de

The standard HOSVD truncation of tensors yields a controlled error with respect to the corresponding Hilbert norm, usually L2-norm. If the tensor represents a multivariate function, the aim may be to evaluate the function at certain points. Therefore, we need maximum norm estimates instead. Usually, the maximum norm cannot be estimated by the L2-norm. We show that, nevertheless, this is possible for the HOSVD truncated tensors.

Solution of the Monge-Kantorovich transport problem in high dimensions

Fred Daum

Raytheon, United States; frederick_e_daum@raytheon.com

We solve the Monge-Kantorovich transport problem in high dimensions by computing the flow of particles induced by Bayes rule with a linear constant coefficient first order highly underdetermined PDE, similar to the Gauss divergence law. We solve this PDE using the Fourier transform, but we avoid computing the inverse Fourier transform by use of Monte Carlo integration and the generalized matrix inverse (roughly analogous to a pseudodifferential operator). We show numerical results for high dimensional nonlinear filter problems (d = 1 to 30) with a wide variety of parameters (stable and unstable plants, initial uncertainty of the state vector, measurement accuracy, and dimension of the plant). This algorithm works for arbitrary smooth nowhere vanishing probability densities, including highly non-Gaussian multimodal densities, and hence the theory can be applied to essentially any estimation or Bayesian decision problem. We solve the fundamental and well known problem of particle degeneracy by using a log-homotopy of the conditional probability density, but we do not use resampling of particles or a proposal density, and we do not use Metropolis-Hastings or MALA or Hamiltonian Monte Carlo or any other Markov chain Monte Carlo method (MCMC), and our method is not anything like variational Bayes algorithms. It turns out that it is crucial to select an optimal set of points in k-space at which to evaluate the Fourier transform, and somewhat surprisingly we must enforce neutral charge density of particles to very high accuracy along the

flow in order to guarantee the existence of a solution to our linear underdetermined PDE. Our algorithm is roughly ten orders of magnitude faster than standard particle filters for the same estimation accuracy.

Using Hyperbolic Cross Approximation to measure and compensate Covariate Shift

Jochen Garcke[1] *and Thomas Vanck*[2]

[1]University of Bonn, Germany; garcke@ins.uni-bonn.de
[2]Technische Universität Berlin, Germany; vanck@math.tu-berlin.de

The concept of covariate shift in data analysis describes a displacement of the training and test distribution while the conditional remains the same. One can address this problem by using individual weights for each training datapoint, which emphasizes the training points close to the test data set so that these get a higher significance.

We propose a new method for calculating such weights by minimizing a Fourier series approximation of distance measures like Pearson chi-squared or Kullback-Leibler. To be able to use the propsed approach for higher dimensional data, we employ the so-called hyperbolic cross approximation.

Results show that the new approach can compete with the latest methods and that on real life data an improved performance can be obtained.

Parallel Solution of Regression Problems Using Sparse Grids and Alternating Direction Method of Multipliers

Valeriy Khakhutskyy[1]*, Dirk Pflüger*[2] *and Markus Hegland*[3]

[1]Technische Universität München, Germany; khakhutv@in.tum.de
[2]Universität Stuttgart, Germany; Dirk.Pflueger@ipvs.uni-stuttgart.de
[3]Australian National University, Australiavim te ; markus.hegland@anu.edu.au

We consider the application of sparse grids for solving regression problems in data mining. Regression has a lot of applications in data mining in a large variety of different areas, e.g. prediction of real estate prices, healthcare risks, or photometric redshift of galaxies. Compared to established machine learning techniques, like artificial neural networks or support vector machines, sparse grids produce an approximant that is easier to analyze and to interpret. High-dimensional regression problems present a challenge for sparse grids and are usually tackled by either dimensionality reduction (e.g. using PCA), or by approximation of the high-dimensional solution by a combination of subsolutions in lower dimensional space (e.g. using ANOVA decomposition), or by parallelizing the solution algorithms.

In this work we focus on the last kind of techniques. We investigate the parallelization opportunities for solving high-dimensional machine learning problems with adaptive sparse grids using the alternating direction method of multipliers (ADMM). ADMM allows us to split the initially large problem into smaller ones. They can then be solved in parallel while their reduced problem sizes can even be small enough for an explicitly assembly of the system matrices. We present some computational experiments which use ADMM for hierarchical basis sparse grids approximation and in particular discuss the convergence rates.

Weighted Hilbert spaces in the porous flow problem and associated numerical challenges

James Nichols

University of New South Wales, Australia; james.ashton.nichols@gmail.com

We present recent results in the theory of applying QMC rules to integrating PDEs with random coefficients, as for example arises in the Darcy flow problem in porous media. Proving good convergence of rank-1 lattice rules requires new results for generalised spaces. We present those results, and discuss decisions motivated by the challenging numerics of this problem.

Adaptive polynomial approximation in high dimensional PDE with stochastic data

Giovanni Migliorati[1], *Fabio Nobile*[2] *and Raul Tempone*[3]

[1]Ecole Polytechnique Federale de Lausanne, Switzerland; giovanni.migliorati@gmail.com
[2]Ecole Polytechnique Federale de Lausanne, Switzerland; fabio.nobile@epfl.ch
[3]King Abdullah University of Science and Technology, Saudi Arabia; raul.tempone@kaust.edu.sa

In many PDE models the parameters are not known with enough accuracy, or they naturally feature randomness and can be treated therefore as random variables. The challenge is then to efficiently compute the statistics of the solution of the PDE or some quantities of interest (outputs), given the probability distribution of the random input parameters. This can be seen as an high dimensional approximation / integration problem.

If the parameter to solution map is smooth, one can look for a multivariate polynomial approximation of it (polynomial chaos expansion) [3, 5]. An approach that has been advocated recently ([9, 2, 1]) consists in evaluating the solution on randomly chosen parameters and doing a discrete L^2 projection on the polynomial space. This problem can be analyzed in a regression framework [4] with random design, where the regression function minimizes the L^2 risk. However, we consider here the case of noise-free observations on random points.

For any finite positive integer d, denote by Y a d-dimensional random variable distributed according to a known density ρ. We consider target functions $\phi = \phi(Y)$, and study the approximation properties of the random L^2 projection Π_Λ^M over the polynomial space $\mathbb{P}_\Lambda$ of random variables. We analyze stability and approximation properties of Π_Λ^M with respect to the number of sampling points M, the maximum polynomial degree w, and the smoothness of the function to approximate. In the sequel N denotes the number of terms retained in the polynomial space, *i.e.* $N = \#\Lambda = \dim\mathbb{P}_\Lambda$.

Using the random points $y_1, \ldots, y_M$ which are i.i.d. realizations of Y, we define the random L^2 projection of ϕ on the polynomial space $\mathbb{P}_\Lambda$ as

$$\Pi_\Lambda^M \phi = \underset{v \in \mathbb{P}_\Lambda}{\operatorname{argmin}} \frac{1}{M} \sum_{i=1}^{M} (\phi(y_i) - v(y_i))^2. \tag{3}$$

We denote by $\|\phi - \Pi_\Lambda^M \phi\|_{L^2_\rho}$ the approximation error in the L^2_ρ norm obtained when projecting the function ϕ on the polynomial space $\mathbb{P}_\Lambda$. This error is also a random variable, due to the randomness in the sampling points.

In [6] we proved optimality estimates (up to a logarithmic factor) in the monovariate case when the random points are sampled from bounded random variables with strictly positive probability

density functions. Our analysis of the random projection proves that the optimal convergence rate is achieved when the number of sampling points scales as the square of the dimension of the polynomial space. Moreover, it gives an insight on the role of smoothness and the conditioning of the random projection operator in the accuracy and stability of the discrete L^2 projection.

In the one-dimensional case ($N = 1$) with the uniform density $\rho \sim \mathbb{U}(-1, 1)$, the optimality results is the following:

Theorem 1 ([6]). *For any $\alpha \in (0, 1)$ and under the condition*

$$\frac{M}{3\log((M+1)/\alpha)} \geq 4\sqrt{3}\,\mathrm{w}^2 \tag{4}$$

it holds

$$\mathbb{P}\left(\|\phi - \Pi_\Lambda^M \phi\|_{L^2_\rho} \leq \left(1 + \sqrt{3\log\frac{M+1}{\alpha}}\right) \inf_{v\in\mathbb{P}_\mathrm{w}} \|\phi - v\|_{L^\infty}\right) \geq 1 - \alpha. \tag{5}$$

This theorem states that, with confidence level $1 - \alpha$, the discrete L^2 projection with random points is (nearly) optimal up to a logarithmic factor in M, provided M is large enough and satisfies the condition (4). In one dimension the space $\mathbb{P}_\Lambda$ consists of all the polynomials up to degree w, hence $N = \mathrm{w} + 1$. Therefore, condition (4) requires the number of sampling points to scale as the square of the dimension of the polynomial space. Under the same condition, the approximation problem is also stable independently of the maximum polynomial degree w.

Analogous one-dimensional estimates have been derived in [7], but in expectation instead of probability. The condition to have a stable and optimally convergent approximation is the same as (4), *i.e.* $M \sim N^2$.

We will present several numerical tests that confirm our theoretical results, and point out some significant differences between the one-dimensional and the multidimensional cases.

Then we will present the application of this methodology to compute Quantities of Interest associated to the solution of stochastic PDEs. We will deal with stochastic coefficients and with random domains, *i.e.* domains whose shape is described by random variables. Numerical examples in low and high dimensions will be shown as well.

Finally we present some recent results on adaptive approximation on polynomial spaces, when the dimension d of the parameter set Γ is finite and moderately large. We consider a class of isotropic problems (that include the elliptic model and the linear elasticity model, both with stochastic coefficients), for which all random variables have equal influence on the solution. In [8] (work in collaboration with A.Cohen and A.Chkifa), we proved sub-exponential convergence w.r.t. M, when the number of sampling points M scales as $M \sim N^2$ and the polynomial space is suitably adapted. In addition, this result establishes an equivalence with classical Stochastic Galerkin methods. Numerical evidence shows that in the multivariate case also a linear scaling $M \sim N$ yields a stable and optimally convergent adaptive projection, making this methodology well promising for high dimensional approximations.

The extension to the anisotropic framework is currently under investigation.

References

[1] G.Blatman and B.Sudret. *Adaptive sparse polynomial chaos expansion based on least angle regression.* Journal of Computational Physics, 230(6):2345 – 2367, 2011.

[2] M.S.Eldred. Design under uncertainty employing stochastic expansion methods. *International Journal for Uncertainty Quantification.* 1:119–146, 2011.

[3] G.R.Ghanem, P.D.Spanos: *Stochastic Finite Elements: a spectral approach.* Dover Publications, 2003.

[4] L.Gyorfi, M.Kohler, A.Krzyzak, and H.Walk. *A distribution-free theory of nonparametric regression.* Springer Series in Statistics. Springer-Verlag, New York, 2002.

[5] O.P.Le Maitre, O.M.Knio: *Spectral Methods for Uncertainty Quantification.* Springer, 2010.

[6] G.Migliorati, F.Nobile, E.von Schwerin, R.Tempone. *Analysis of the discrete L^2 projection on polynomial spaces with random evaluations.* submitted. Also available as MOX-report 46-2011.

[7] Cohen, A. and Davenport, M. A. and Leviatan, D. : *On the stability and accuracy of Least Squares approximations*, arXiv:1111.4422v1, 2011.

[8] A.Chkifa, A.Cohen, G.Migliorati, F.Nobile, R.Tempone. *Adaptive polynomial approximation based on discrete least squares with random evaluation and application to parametric and stochastic PDEs*. in preparation.

[9] S. Hosder, R. W. Walters, and M. Balch. *Point-collocation nonintrusive polynomial chaos method for stochastic computational fluid dynamics.* AIAA Journal, 48:2721–2730, 2010.

On tensor trains and signaling cascades

Jochen Garcke[1] *and Markus Hegland*[2]

[1]University of Bonn, Germany; garcke@ins.uni-bonn.de
[2]Australian National University, Australia; markus.hegland@anu.edu.au

We study the chemical master equations for chemical signaling cascades, one of the main components of biological switches. They have been shown to simultaneously per- form thresholding and signal amplification. This despite the intrinsic noise which is introduced by the chemical reactions which constitute the cascade stages.

Using the chemical master equation one obtains a tool for the determination of the marginal probability distributions over the domain of copy numbers of the species involved. This leads to a natural representation of intrinsic noise.

The solution of the master equations and even the representation of the resulting multidimensional probability distribution suffers under the curse of dimensionality. We propose the use of the recently introduced tensor-train framework for the representation of the arising high-dimensional objects. Theoretical considerations show a relation of the structure of a cascade to the structure of a tensor train. Numerical studies give evidence that the chemical master equation for a signaling cascade and its stationary solution have a structure in the tensor train format allowing its computation in higher dimensions.

Quadrature approach to Bayesian Inverse Problems

Christoph Schwab[1] *and Claudia Schillings*[2]

ETH Zurich, Switzerland
[1]schwab@math.ethz.ch
[2]claudia.schillings@sam.math.ethz.ch

We propose a quadrature approach for Bayesian Inverse Problems. The task is to compute statistical estimates for predicted responses of uncertain systems under given, noisy data. Problem classes considered are PDEs and large ODE systems with uncertain coefficients and parameters.

The Bayesian framework for response prediction is reduced to a parametric, deterministic infinite-dimensional quadrature problem. The integrand functions (posterior densities conditional on given data) are shown to belong to classes of functions with sparse polynomial expansions. Adaptive sparse tensor quadrature algorithms are proposed and analyzed; they are shown to exhibit dimension-independent rates of convergence which are governed only by the sparsity models of the prior distributions in the Bayesian estimate. Numerical experiments confirm the theoretical results.

Index