Report from Dagstuhl Seminar 15391

# Algorithms and Complexity for Continuous Problems


**Edited by**

**Aicke Hinrichs[1], Joseph F. Traub*[2], Henryk Woźniakowski[3], and Larisa Yaroslavtseva[4]**

1 **Universität Linz, AT,** aicke.hinrichs@jku.at
2 **Columbia University – New York, US**
3 **Columbia University – New York, US**
4 **Universität Passau, DE,** larisa.yaroslavtseva@uni-passau.de



## Abstract

From 20.09.15 to 25.09.15, the Dagstuhl Seminar 15391 Algorithms and Complexity for Continuous Problems was held in the International Conference and Research Center (IBFI), Schloss Dagstuhl. During the seminar, participants presented their current research, and ongoing work and open problems were discussed. Abstracts or the presentations given during the seminar can be found in this report. The first section describes the seminar topics and goals in general. Links to extended abstracts or full papers are provided, if available.



**Seminar** September 20–25, 2015 – http://www.dagstuhl.de/15391

**1998 ACM Subject Classification** E.1 Data Structures, F.2 Analysis of Algorithms and Problem Complexity

**Keywords and phrases** High Dimensional Problems, Tractability, Random coefficients, Multi-level algorithms, computational stochastic processes, Compressed sensing, Learning theory

**Digital Object Identifier** 10.4230/DagRep.5.9.57

**Edited in cooperation with** Daniel Rudolf


## 1 Executive Summary


*Aicke Hinrichs*
*Henryk Woźniakowski*
*Larisa Yaroslavtseva*




This was already the 12th Dagstuhl Seminar on Algorithms and Complexity for Continuous Problems over a period of 24 years. It brought together researchers from different communities working on computational aspects of continuous problems, including computer scientists, numerical analysts, applied and pure mathematicians. Although the seminar title has remained the same, many of the topics and participants change with each seminar and each seminar in this series is of a very interdisciplinary nature.

Continuous computational problems arise in diverse areas of science and engineering. Examples include path and multivariate integration, approximation, optimization, as well as operator equations. Typically, only partial and/or noisy information is available, and the aim is to solve the problem within a given error tolerance using the minimal amount

* Joseph F. Traub (June 24, 1932 – August 24, 2015): http://www.cs.columbia.edu/~traub/.

of computational resources. For example, in high-dimensional integration one wants to compute an $\epsilon$-approximation to the integral with the minimal number of function evaluations. Here it is crucial to identify first the relevant variables of the function. Understanding the complexity of such problems and construction of efficient algorithms is both important and challenging. The current seminar attracted 35 participants from nine different countries all over the world. About 30 % of them were young researchers including PhD students. There were 25 presentations covering in particular the following topics:

- High-dimensional problems
- Tractability
- Computational stochastic processes
- Compressive sensing
- Random media
- Computational finance
- Noisy data
- Learning theory
- Biomedical learning problems
- Markov chains

There were three introductory talks to recent developments in PDE with random coefficients, learning theory and compressive sensing. A joint session with the Dagstuhl Seminar 15392 "Measuring the Complexity of Computational Content: Weihrauch Reducibility and Reverse Analysis" stimulated the transfer of ideas between the two different groups present in Dagstuhl.

The work of the attendants was supported by a variety of funding agencies. This includes the Deutsche Forschungsgemeinschaft, the Austrian Science Fund, the National Science Foundation (USA), and the Australian Research Council.

As always, the excellent working conditions and friendly atmosphere provided by the Dagstuhl team have led to a rich exchange of ideas as well as a number of new collaborations. Selected papers related to this seminar will be published in a special issue of the Journal of Complexity.

## 2 Table of Contents

# 3 Overview of Talks

## 3.1 Maximum Improvement Algorithm for Global Optimization of Brownian Motion

*James M. Calvin (NJIT – Newark, US)*



Two common approaches to optimizing an unkown random function are to choose the next point to maximize the conditional probability that the function value is less than some amount below the current record minimum, and to choose the next point to maximize the expected decrease below the current record minimum. We construct algorithms based on each approach, and describe error bounds.

## 3.2 On lattice rules, approximation and Lebesgue constants

*Ronald Cools (KU Leuven, BE)*


**Joint work of** Cools, Ronald; Nuyens, Dirk; Suryanarayana, Gowri

In this talk I start with introducing lattice rules for numerical integration with the trigonometric degree as quality criterion. I focus on Fibonacci lattice rules and lattice rules with minimal number of points for the trigonometric degree. Then I consider the points of these lattice rules for approximation and introduce the trigonometric approximation degree. This approximation degree is calculated for the lattices used throughout this talk.

To investigate the quality of point sets for approximation the Lebesgue constant is often used. The Lebesgue constant for trigonometric approximation for 1- and 2-dimensional lattices is investigated. We reveal some nice structures and make the link between the Dirichlet kernel and the reproducing kernel that was used to obtain minimal lattice rules in two dimensions.

## 3.3 Weak convergence for semi-linear SPDEs

*Sonja Cox (University of Amsterdam, NL)*


**Joint work of** Cox, Sonja; Jentzen, Arnulf; Kuniawan, Ryan
**Main reference** A. Jentzen, R. Kurniawan, "Weak convergence rates for Euler-type approximations of semilinear stochastic evolution equations with nonlinear diffusion coefficients," arXiv:1501.03539v1 [math.PR], 2015.
**URL** http://arxiv.org/abs/1501.03539v1

In recent work by Jentzen and Kurniawan weak convergence of both spatial and temporal discretizations for semi-linear SPDEs was proven. Their approach required the non-linear terms in the SPDE to be four times Fréchet differentiable as operators on a Hilbert space. In particular, their results can not be applied to non-linear terms arising from Nemytskii operators. In my talk I will explain how this problem can be overcome by working the more general Banach space setting.

### 3.4 General multilevel adaptations for stochastic approximation algorithms

*Steffen Dereich (Universität Münster, DE)*


Joint work of Dereich, Steffen; Müller-Gronbach, Thomas
Main reference S. Dereich, T. Müller-Gronbach, "General multilevel adaptations for stochastic approximation algorithms," arXiv:1506.05482v1 [math.PR], 2015.
URL http://arxiv.org/abs/1506.05482v1

We analyse multilevel adaptations of stochastic approximation algorithms. In contrast to the classical multilevel Monte Carlo algorithm of Mike Giles one now deals with a parameterised family of expectations and the aim is to compute parameters for which the expectation is zero. We propose a new algorithm and provide an upper bound for its error. Under similar assumptions as in [2] we recover the same order of convergence in the computation of zeroes as the ones originally derived in the computation of single expectations.

#### References

1 S. Dereich and T. Müller-Gronbach. *General multilevel adaptations for stochastic approximation algorithms*, arXiv:1506.05482.
2 M. B. Giles. *Multi-level Monte Carlo path simulation*. Operations Research, 56(3):607–617, 2008.

### 3.5 On exit times of diffusions from a domain

*Stefan Geiss (University of Jyväskylä, FI)*


Joint work of Bouchard, Bruno; Geiss, Stefan; Gobet, Emmanuel
Main reference B. Bouchard, S. Geiss, E. Gobet, "First time to exit of a continuous Itô process: general moment estimates and $\mathbf{L}_1$-convergence rate for discrete time approximations," arXiv:1307.4247v2 [math.PR], 2014.
URL http://arxiv.org/abs/1307.4247v2

We establish general moment estimates for the discrete and continuous exit times of a general Itô process in terms of the distance to the boundary. These estimates serve as intermediate steps to obtain strong convergence results for the approximation of a continuous exit time by a discrete counterpart, computed on a grid. In particular, we prove that the discrete exit time of the Euler scheme of a diffusion converges in the $\mathbf{L}_1$ norm with an order 1/2 with respect to the mesh size. This rate is optimal. The talk is based on [1].

#### References

1 B. Bouchard, S. Geiss, and E. Gobet: *First time to exit of a continuous Itô process: general moment estimates and* $\mathbf{L}_1$*-convergence rate for discrete time approximations*, In revision for *Bernoulli*.
2 B. Bouchard and S. Menozzi. *Strong approximations of BSDEs in a domain. Bernoulli*, 15(4):1117–1147, 2009.
3 D. J. Higham, X. Mao, M. Roj, Q. Song, and G. Yin. *Mean exit times and the multilevel Monte Carlo method. SIAM/ASA J. Uncertain. Quantif.* 1(1):2–18, 2013.

## 3.6 Universality of Weighted Anchored and ANOVA Spaces

*Michael Gnewuch (Universität Kiel, DE)*


**Joint work of** Gnewuch, Michael; Hefter, Mario; Hinrichs, Aicke; Ritter, Klaus

We present upper and lower error bounds for high- and infinite-dimensional integration. We study spaces of integrands with weighted norms and consider deterministic and randomized algorithms. Interesting examples of norms are norms induced by an anchored function space decomposition or the ANOVA decomposition.

In some settings (depending on the class of integrands we consider, the weighted norm, the class of algorithms we admit and the way we account for the computational cost) one can derive good or even optimal error bounds directly. If one changes the weighted norm, a correspondent direct error analysis can be much more involved and complicated. The focus of the talk is to discuss new results on function space embeddings of weighted spaces which allow for an easy transfer of error bounds.

## 3.7 Optimal Strong Approximation of the One-dimensional Squared Bessel Process

*Mario Hefter (TU Kaiserslautern, DE)*


**Joint work of** Hefter, Mario; Calvin, James M.; Herzwurm, André

We consider the one-dimensional squared Bessel processes given by the stochastic differential equation (SDE)

$$dX_t = 1\,dt + 2\sqrt{X_t}\,dW_t, \quad X_0 = x_0, \quad t \in [0,1], \tag{1}$$

and study strong (pathwise) approximation of the solution $X$ at the final time point $t = 1$. This SDE is a particular instance of a Cox-Ingersoll-Ross (CIR) process where the boundary point zero is accessible. We consider numerical methods that have access to values of the driving Brownian motion $W$ at a finite number of time points. We show that the polynomial convergence rate of the $n$-th minimal errors for the class of adaptive algorithms as well as for the class of algorithms that rely on equidistant grids are equal to infinity and $1/2$, respectively. As a consequence, we obtain that the parameters appearing in the CIR process affect the convergence rate of strong approximation.

A key step in the proofs consists of identifying the pathwise solution of (1) and link this problem to global optimization under the Wiener measure.

## 3.8 Complexity of parametric SDEs

*Stefan Heinrich (TU Kaiserslautern, DE)*



We consider the problem of strong solution of scalar stochastic differential equations depending on a parameter. We seek to find numerical approximations for all parameter values simultaneously.

The problem is approached within a general scheme of solving parameter-dependent numerical problems by multilevel methods, developed previously by T. Daun and the author in a series of papers. First we obtain suitable convergence results for the Banach space valued Euler-Maruyama scheme in spaces of martingale type 2. Then we develop a multilevel scheme involving two embedded Banach spaces, where discretization is balanced with approximation of the embedding. Finally, the parametric problem is cast into this embedded Banach space setup, from which a multilevel method for the strong solution of parametric stochastic differential equations results.

We obtain convergence rates for various smoothness classes of input functions. Furthermore, the optimality of these rates is established by proving matching lower bounds. Thus, the complexity of this problem is established in the sense of information-based complexity theory.

## 3.9 Quasi-Monte Carlo conquers the Rendering Industry

*Alexander Keller (NVIDIA GmbH – Berlin, DE)*



Quasi-Monte Carlo methods for image synthesis have been under investigation for over 20 years. Although the deterministic approach has been shown to be superior to corresponding Monte Carlo methods, adoption had been rare for a long time. However, coincident with the rendering industry changing to path tracing algorithms for light transport simulation, recently there is a huge interest and growing adoption of quasi-Monte Carlo methods.

We point out how Dagstuhl seminars brought about this change that even influenced standard textbooks, review the state of the art in the rendering industry, and discuss current algorithmic and mathematical questions in light transport simulation.

## 3.10 Efficient truncation for integration in weighted anchored and ANOVA spaces

*Peter Kritzer (Universität Linz, AT)*


**Joint work of** Kritzer, Peter; Pillichshammer, Friedrich, Wasilkowski, Greg W.
**Main reference** P. Kritzer, F. Pillichshammer, G. W. Wasilkowski, "Very low truncation dimension for high dimensional integration under modest error demand," arXiv:1506.02458v2 [math.NA], 2015.
**URL** http://arxiv.org/abs/1506.02458v2

We consider the problem of numerical integration for weighted anchored and ANOVA Sobolev spaces of $s$-variate functions, where $s$ is large. Under the assumption of sufficiently fast decaying weights, we show in a constructive way that such integrals can be approximated by quadratures for functions $f_k$ with only $k$ variables, where $k = k(\varepsilon)$ depends solely on the error demand $\varepsilon$. Moreover $k(\varepsilon)$ does not depend on the function being integrated, i.e., is the same for all functions from the unit ball of the space.

## 3.11 Approximation in multivariate periodic Gevrey spaces

*Thomas Kühn (Universität Leipzig, DE)*


**Joint work of** Kühn, Thomas; Petersen, Martin

The classical Gevrey classes, already introduced in 1918, play an important role in analysis, especially in the context of PDEs. They consist of $C^\infty$-functions on $\mathbb{R}^d$ whose derivatives satisfy certain growth conditions. All Gevrey classes contain non-analytic functions. For periodic functions $f$ , these growth conditions on the derivatives can be expressed equivalently by decay conditions on the Fourier coefficients of $f$. Using this approach, we define periodic Gevrey spaces $G^{s,c}(\mathbb{T}^d)$ on the $d$-dimensional torus $\mathbb{T}^d$, where $s \in (0, 1)$ is a smoothness parameter and $c > 0$ a fine parameter.

There is a rich literature on approximation of functions of *finite smoothness*, as well as for classes of *analytic functions*, but only quite few results are available for $C^\infty$-functions. The talk is devoted to this "intermediate" case, more precisely to estimates for approximation numbers $a_n$ of the embeddings $G^{s,c}(\mathbb{T}^d) \hookrightarrow L_2(\mathbb{T}^d)$. In particular, we determine the exact asymptotic rate of $a_n$ as $n \to \infty$. Not surprisingly, this rate is sub-exponential and faster than polynomial. Moreover, we give two-sided preasymptotic estimates, i.e. for small $n$, with special emphasis on the dependence of the hidden constants on the dimension $d$. These results allow an interpretation in the language of IBC, concerning different notions of tractability.

## 3.12 Minimax signal detection in statistical inverse problems

*Peter Mathè (Weierstraß Institut – Berlin, DE)*


**Joint work of** Mathé, Peter; Clèment Marteau
**Main reference** C. Marteau, P. Mathé, "General regularization schemes for signal detection in inverse problems," Mathematical Methods of Statistics, 23(3):176–200, 2014.
**URL** http://dx.doi.org/10.3103/S1066530714030028

We shall consider inverse problems in Hilbert space under Gaussian white noise. Usually, the problem of reconstructing the unknown signal, say $f$, from the noisy observation $Y = Tf + \sigma\xi$ is considered. Instead, we are interested in the nonparametric test problem, and we ask $f = f_0$ for some given function. We shall exhibit how optimality for such test problem can be defined. Lower bounds have been established, previously. We emphasize that many of the common regularization schemes, both with and/or without discretization can be used to yield order optimal tests. This is joint work with Clèment Marteau, Univ. Toulouse, [1].

### References

1 C. Marteau, and P. Mathé, *General regularization schemes for signal detection in inverse problems*, Mathematical Methods of Statistics, 23 (2014) pp. 176–200.

## 3.13 On tough quadrature problems for SDEs with bounded smooth coefficients

*Thomas Müller-Gronbach (Universität Passau, DE)*


**Joint work of** Müller-Gronbach, Thomas; Yaroslavtseva, Larisa

We study the problem of approximating the expected value $E(f(X(1)))$ of a function $f$ of the solution $X(1)$ of a SDE at time 1 based on a finite number of evaluations of $f$ and the coefficients of the SDE. We present classes of SDEs with bounded smooth coefficients such that this problem can not be solved with a polynomial error rate in the worst case sense.

## 3.14 Optimal approximation of SDEs driven by fractional Brownian motion – An overview

*Andreas Neuenkirch (Universität Mannheim, DE)*



In this talk I give an overview on recent results concerning the optimal approximation of stochastic differential equations (SDEs) driven by fractional Brownian motion (fBm) with Hurst parameter $H > 1/4$. More precisely, I will consider the approximation of the solution at a fixed time point with respect to the root mean square error given an equidistant discretisation of the driving fBm.

While the scalar case has been analysed in detail [1] for $H > 1/2$ in 2008, in recent years several error bounds have been established in the multi-dimensional case. The picture is now as follows: Up to sub-polynomial terms the optimal convergence order is at least $\min\{2H - 1/2, 1\}$, due to results of Bayer et al. [2] and Hu et al. [3]. In the case of the fractional Lévy area, which corresponds to a particular two-dimensional SDE, the optimal convergence order is $2H - 1/2$ for $H > 1/2$, see [4].

I strongly suppose that as long the diffusion coefficients do not commute, the optimal convergence order is $2H - 1/2$.

### References

1 A. Neuenkirch (2008). *Optimal pointwise approximation of stochastic differential equations driven by fractional Brownian motion.* Stochastic Processes and their Applications 118 (12), 2294–2333
2 C. Bayer, P. K. Friz, S. Riedel, J. Schoenmakers (2013+). *From rough path estimates to multilevel Monte Carlo.* Working Paper
3 Y. Hu, Y. Lui, D. Nualart (2015+). *Rate of convergence and asymptotic error distribution of Euler approximation schemes for fractional diffusions.* Annals of Applied Probability. To appear
4 A. Neuenkirch, T. Shalaiko (2015+). *The maximum rate of convergence for the approximation of the fractional Lévy area at a single point.* Journal of Complexity. To appear

## 3.15 Multivariate integration over the Euclidean space for analytic functions and $r$-smooth functions

*Dong Nguyen (KU Leuven, BE)*


**Joint work of** Nuyens, Dirk

In this talk we study multivariate integration over $\mathbb{R}^s$ for weighted analytic functions, whose Fourier transform decays exponentially fast. We prove that the exponential convergence rate can be achieved by using a classical quasi-Monte Carlo method. More specific, we prove two convergence rates of $\mathcal{O}(\exp(-N^{\frac{1}{D(s)+B(s)}}))$ and $\mathcal{O}(\exp(-N^{\frac{1}{B(s)}}(\ln N)^{-\frac{D(s)}{B(s)}}))$, where $D(s)$ and $B(s)$ are respectively defined by the exponential decay of the Fourier transform and of the integrand, for two different function classes. We discuss work in progress to obtain a stronger convergence rate with less dependence on the dimension. Some numerical results demonstrate the theory.

## 3.16 A Universal Algorithm for Multivariate Integration

*Erich Novak (Universität Jena, DE)*


**Main reference** D. Krieg, E. Novak, "A Universal Algorithm for Multivariate Integration," arXiv:1507.06853v1 [math.NA], 2015.
**URL** http://arxiv.org/abs/1507.06853v1

We present an algorithm for multivariate integration over cubes that is unbiased and has optimal order of convergence (in the randomized sense as well as in the worst case setting) for all Sobolev spaces $H^{r,\mathrm{mix}}([0,1]^d)$ and $H^s([0,1]^d)$ for $s > d/2$.

## 3.17 Approximation with lattice points

*Dirk Nuyens (KU Leuven, BE)*


**Joint work of** Nuyens, Dirk; Gowri, Suryanarayana; Ronald, Cools; Frances Y., Kuo

We analyse the asymptotic worst case error of using tent transformed lattice points for approximation of functions in the half period cosine space. This space is the unanchored Sobolev space of smoothness one for a certain choice of parameters. This is a continuation of [1].

### References

1 G. Suryanarayana, D. Nuyens, and R. Cools. *Reconstruction and collocation of a class of non-periodic functions by sampling along tent-transformed rank-1 lattices.* Journal of Fourier Analysis and Applications, pp. 1–28, 2015.

## 3.18 Tensor product approximation of analytic functions

*Jens Oettershagen (Universität Bonn, DE)*


**Joint work of** Oettershagen, Jens; Griebel, Michael
**Main reference** M. Griebel, J. Oettershagen, "On tensor product approximation of analytic functions," INS Preprint, No. 1512, 2015.
**URL** http://wissrech.ins.uni-bonn.de/research/pub/oettershagen/INSPreprint1512.pdf

We prove sharp, two-sided bounds on sums of the form $\sum_{\mathbf{k}\in\mathbb{N}_0^d\setminus\mathcal{D}_{\mathbf{a}}(T)} \exp(-\sum_{j=1}^d a_j k_j)$, where $\mathcal{D}_{\mathbf{a}}(T) := \{\mathbf{k} \in \mathbb{N}_0^d : \sum_{j=1}^d a_j k_j \leq T\}$ and $\mathbf{a} \in \mathbb{R}_+^d$. These sums appear in the error analysis of tensor product approximation, interpolation and integration of $d$-variate analytic functions. Examples are tensor products of univariate Fourier-Legendre expansions or interpolation and integration rules at Leja points . Moreover, we discuss the limit $d \to \infty$, where we prove both, algebraic and sub-exponential upper bounds. As an application we consider tensor products of Hardy spaces, where we study convergence rates of a certain truncated Taylor series, as well as of interpolation and integration using Leja points.

## 3.19 A linear functional strategy for regularized ranking

*Sergei Pereverzyev (RICAM – Linz, AT)*


**Main reference** G. Kriukova, O. Panasiuk, S. V. Pereverzyev, P. Tkachenko, "A Linear Functional Strategy for Regularized Ranking," RICAM Report 2015-13, 2015.
**URL** http://www.ricam.oeaw.ac.at/publications/reports/15/rep15-13.pdf

Regularization schemes are frequently used for performing ranking tasks. This topic has been intensively studied in recent years. However, to be effective a regularization scheme should be equipped with a suitable strategy for choosing a regularization parameter. In the present study we discuss an approach, which is based on the idea of a linear combination of regularized rankers corresponding to different values of the regularization parameter. The coefficients of the linear combination are estimated by means of the so-called linear functional strategy. We provide a theoretical justification of the proposed approach and illustrate them by numerical experiments. Some of them are related with ranking the risk of nocturnal hypoglycemia of diabetes patients.

## 3.20 Linear versus nonlinear approximation in the average case setting

*Leszek Plaskota (University of Warsaw, PL)*



We compare the average errors of linear and nonlinear approximations assuming that the coefficients in an orthogonal expansion are scaled i.i.d. random variables. We show that generally the $n$-term nonlinear approximation can be much better than linear approximation. On the other hand, if the scaling parameters decrease no faster than polynomially then the average error of nonlinear approximations does not converge to zero faster than that of linear approximations, as $n$ goes to infinity.

**References**

1 A. Cohen and J.-P. D'Ales, *Nonlinear approximation of random functions*. SIAM J. Appl. Math. 57 (1997) pp. 518–540.
2 J. Creutzig, T. Müller-Gronbach, K. Ritter, *Freer-knot spline approximation of stochastic processes*. J. Complexity 23 (1997) pp. 867–889.
3 R. A. DeVore, *Nonlinear approximation*. Acta Numerica 8 (1998) pp. 51–150.
4 R. A. DeVore and B. Jawerth, *Optimal nonlinear approximation*, Manuscripta Math. 63 (1992) pp. 469–478.
5 M. A. Kon and L. Plaskota, *Information-based nonlinear approximation: an average case setting*. J. Complexity 21 (2005), pp. 211–229.
6 J. Vybiral, *Average best m-term approximation*, Constr. Approx. 36 (2012), pp. 83–115.


## 3.21 Optimal adaptive solution of piecewise smooth systems of IVPs with unknown switching hypersurface

*Pawel Przybylowicz (AGH Univ. of Science & Technology-Krakow, PL)*


**Joint work of** Kacewicz, Boleslaw; Przybylowicz, Pawel
**Main reference** B. Kacewicz, Boleslaw, P. Przybylowicz, "Complexity of the derivative-free solution of systems of IVPs with unknown singularity hypersurface," Journal of Complexity, 31(1):75–91, 2015.
**URL** http://dx.doi.org/10.1016/j.jco.2014.07.002

We present results concerning optimal approximation of solutions of piecewise regular systems of IVPs. We assume that a right-hand side function is smooth everywhere except for an unknown smooth hypersurface, defined by zeros of an 'event' function $h$. We do not assume the knowledge of $h$, even in the weak sense of computing certain discrete information on $h$. We restrict ourselves to information defined only by values of the right-hand side function (computation of partial derivatives is not allowed).We show how to construct optimal algorithm that is rigorous (it is not based on heuristic arguments), it does not use information on the event function, and preserves the optimal error known for regular systems. The complexity of piecewise regular problems is consequently asymptotically the same as that for globally regular problems.

## 3.22 Compressive sensing and function reconstruction in high dimensions

*Holger Rauhut (RWTH Aachen, DE)*


**Main reference** S. Foucart, H. Rauhut, "A Mathematical Introduction to Compressive Sensing," ISBN 978-0-8176-4947-0, Applied and Numerical Harmonic Analysis, Birkhäuser/Springer, 2013.
**URL** http://www.springer.com/de/book/9780817649470

Compressive sensing is a recent field originating from mathematical signal processing which predicts that sparse or compressible vectors can be reconstructed from a few linear and non-adaptive measurements via efficient algorithms such as l1-minimization. It is a remarkable fact that uptodate all provably optimal measurement matrices are based on randomness. An important special case is the reconstruction of sparse signals from randomly selected Fourier coefficients. Extensions of this principle can be applied to the reconstruction of functions of many variables. Under standard smoothness assumptions this problem faces the curse of dimensionality. Introducing some non-standard smoothness spaces allowing for efficient sparse approximations, one may avoid the curse of dimension by using compressive sensing techniques for the reconstruction. This principle has applications for the numerical solution of high-dimensional parametric operator equations. The talk gives an overview on these topics.

## 3.23 Perturbation theory of Markov chains

*Daniel Rudolf (Universität Jena, DE)*


**Joint work of** Rudolf, Daniel; Schweizer, Nikolaus
**Main reference** D. Rudolf, N. Schweizer, "Perturbation theory for Markov chains via Wasserstein distance," arXiv:1503.04123v2 [stat.CO], 2015.
**URL** http://arxiv.org/abs/1503.04123v2

Perturbation theory for Markov chains addresses the question how small differences in the transitions of Markov chains are reflected in differences between their distributions. We show bounds on the distance of the nth step distributions of two Markov chains when one of them satisfies a Wasserstein ergodicity condition. Our work is motivated by the recent interest in approximate Markov chain Monte Carlo (MCMC) methods in the analysis of big data sets. We illustrate our theory by showing quantitative estimates for an autoregressive model and an approximate version of the Metropolis-Hastings algorithm.

## 3.24 Generalized solution operators and topology

*Pawel Siedlecki (University of Warsaw, PL)*



It is known that a solution operator $S : F \times [0, \infty) \to \mathcal{P}(G)$ induces a certain type of a topological structure on a set $G$, i.e., a family of pseudometrics such that the sets of $\varepsilon$-approximations of solutions (i.e., $S(f, \varepsilon)$) are, almost, closed balls in these pseudometrics. We generalize this result to the case of solution operators $S : F \times P \to \mathcal{P}(G)$, where $P$ is a partially ordered set with some additional structure. We investigate what types of metric-like and topological structures are induced on a set $G$ in such a case, and how $S(f, \varepsilon)$ may be interpreted.

## 3.25 PDE with random coefficients – a survey

*Ian H. Sloan (University of New South Wales – Sydney, AU)*



This invited survey described recent algorithmic developments in partial differential equations with random coefficients treated as a high-dimensional problem. The prototype of such problems is the underground flow of water or oil through a porous medium, with the permeability of the material treated as a random field. (The stochastic dimension of the problem is high if the random field needs a large number of random variables for its effective description.). The talk introduced the problem, then explained different approaches to the problem, ranging from the polynomial chaos method initiated by Norbert Wiener to the Monte Carlo and Quasi-Monte Carlo methods. In recent years there have been significant progress in the development and analysis of algorithms in these areas. The talk aimed to encourage further interest and activity, especially from younger researchers.

## 3.26 Multi-Level Monte Carlo for Parametric Integration of a Discontinuous Function

*Jeremy Staum (Northwestern University – Evanston, US)*


**Joint work of** Rosenbaum, Imry; Staum, Jeremy

We consider parametric integration of a discontinuous function, focusing on the stochastic setting of random fields. We explore sets of assumptions under which the Multi-Level Monte Carlo method can be shown to improve computational complexity of parametric integration.

## 3.27 Analysis of Kernel-Based Learning Methods

*Ingo Steinwart (Universität Stuttgart, DE)*



The last decade has witnessed an explosion of data collected from various sources. Since in many cases these sources do not obey the assumptions of classical statistical approaches, new automated methods for interpreting such data have been developed in the machine learning community. Statistical learning theory tries to understand the statistical principles and mechanisms these methods are based on.

This talk begins by introducing some central questions considered in statistical learning. Then various theoretical aspects of a popular class of learning algorithms, which include support vector machines, are discussed. In particular, I will describe how classical concepts from approximation theory such as interpolation spaces and entropy numbers are used in the analysis of these methods. The last part of the talk considers more practical aspects including the choice of the involved loss function and some implementation strategies. In addition, I will present a data splitting strategy that enjoys the same theoretical guarantees as the standard approach but reduces the training time significantly.

## References


1. Caponnetto, A. and De Vito, E. (2007). *Optimal rates for regularized least squares algorithm. Found. Comput. Math.*, 7:331–368.
2. Eberts, M. and Steinwart, I. (2013). *Optimal regression rates for SVMs using Gaussian kernels. Electron. J. Stat.*, 7:1–42.
3. Eberts, M. and Steinwart, I. (2014). *Optimal learning rates for localized SVMs.* Technical Report 2014-002, Fakultät für Mathematik und Physik, Universität Stuttgart.
4. Mammen, E. and Tsybakov, A. (1999). *Smooth discrimination analysis. Ann. Statist.*, 27:1808–1829.
5. Mendelson, S. and Neeman, J. (2010). *Regularization in kernel learning. Ann. Statist.*, 38:526–565.
6. Platt, J. (1999). *Fast training of support vector machines using sequential minimal optimization.* In *Advances in Kernel Methods–Support Vector Learning*, pages 185–208. MIT Press, Cambridge, MA.
7. Rahimi, A. and Recht, B. (2008). *Random features for large-scale kernel machines.* In Platt, J., Koller, D., Singer, Y., and Roweis, S., editors, *Advances in Neural Information Processing Systems 20*, pages 1177–1184.
8. Smale, S. and Zhou, D.-X. (2003). *Estimating the approximation error in learning theory. Anal. Appl.*, 1:17–41.
9. Smale, S. and Zhou, D.-X. (2007). *Learning theory estimates via integral operators and their approximations. Constr. Approx.*, 26:153–172.
10. Steinwart, I. (2009). *Oracle inequalities for SVMs that are based on random entropy numbers. J. Complexity*, 25:437–454.
11. Steinwart, I., Hush, D., and Scovel, C. (2009). *Optimal rates for regularized least squares regression.* In Dasgupta, S. and Klivans, A., editors, *Proceedings of the 22nd Annual Conference on Learning Theory*, pages 79–93.
12. Steinwart, I., Hush, D., and Scovel, C. (2011). *Training SVMs without offset. J. Mach. Learn. Res.*, 12:141–202.
13. Steinwart, I. and Scovel, C. (2007). *Fast rates for support vector machines using Gaussian kernels. Ann. Statist.*, 35:575–607.
14. Steinwart, I. and Scovel, C. (2012). *Mercer's theorem on general domains: on the interaction between measures, kernels, and RKHSs. Constr. Approx.*, 35:363–417.
15. Urner, R., Wulff, S., and Ben-David, S. (2013). *PLAL: cluster-based active learning.* In Shalev-Shwartz, S. and Steinwart, I., editors, *COLT 2013 – The 26th Annual Conference on Learning Theory*, pages 376–397.
16. Williams, C. K. I. and Seeger, M. (2001). *Using the Nyström method to speed up kernel machines.* In Leen, T., Dietterich, T., and Tresp, V., editors, *Advances in Neural Information Processing Systems 13*, pages 682–688. MIT Press.
17. Yao, Y., Rosasco, L., and Caponnetto, A. (2007). *On early stopping in gradient descent learning. Constr. Approx.*, 26:289–315.

## 3.28 On numerical integration of functions with mixed smoothness

*Mario Ullrich (Universität Linz, AT)*


Joint work of Ullrich, Mario; Ullrich, Tino
Main reference M. Ullrich, T. Ullrich, "The role of Frolov's cubature formula for functions with bounded mixed derivative," arXiv:1503.08846v1 [math.NA], 2015.
URL http://arxiv.org/abs/1503.08846v1

We prove upper bounds on the order of convergence of Frolov's cubature formula for numerical integration in function spaces of dominating mixed smoothness on the unit cube with homogeneous boundary condition. More precisely, we study worst-case integration errors for Besov $\mathbf{B}^s_{p,\theta}$ and Triebel-Lizorkin spaces $\mathbf{F}^s_{p,\theta}$ and our results treat the whole range of admissible parameters ($s \geq 1/p$). In particular, we treat the case of small smoothness which is given for Triebel-Lizorkin spaces $\mathbf{F}^s_{p,\theta}$ in case $1 < \theta < p < \infty$ with $1/p < s \leq 1/\theta$. The presented upper bounds on the worst-case error show a completely different behavior compared to "large" smoothness $s > 1/\theta$. In the latter case the presented upper bounds are optimal, i.e., they can not be improved by any other cubature formula. The optimality for "small" smoothness is open. Moreover, we present a modification of the algorithm which leads to the same bounds also for the larger spaces of periodic functions, and we discuss a randomized version of the algorithm. All results come with supporting numerical results.

## 3.29 Preasymptotic error bounds for multivariate approximation problems

*Tino Ullrich (Universität Bonn, DE)*


Joint work of Ullrich, Tino; Kühn, Thomas; Mayer, Sebastian
Main reference T. Kühn, S. Mayer, T. Ullrich, "Counting via entropy: new preasymptotics for the approximation numbers of Sobolev embeddings," arXiv:1505.00631v1 [math.NA], 2015.
URL http://arxiv.org/abs/1505.00631v1

We study the classical problem of finding the rate of convergence of the approximation numbers of isotropic and dominating mixed multivariate periodic Sobolev embeddings. Our particular focus is on so-called preasymptotic estimates, i.e., error estimates for rather small $n$. By pointing out an interesting relation to entropy numbers in finite dimensional spaces, we can precisely determine the preasymptotic rate of convergence for a family of isotropic norms defined through an additional "compressibility" parameter $p$ which enters the (sharp) preasymptotic error estimate as well as the asymptotic constant which is also determined exactly.

## 3.30 $(s,t)$-Weak tractability

*Markus Weimar (Universität Marburg, DE)*


**Joint work of** Siedlecki, Pawel; Weimar, Markus
**Main reference** P. Siedlecki, M. Weimarm "Notes on (s,t)-weak tractability: A refined classification of problems with (sub)exponential information complexity," Journal of Approximation Theory, Vol. 200, pp. 227–258, 2015.
**URL** http://dx.doi.org/10.1016/j.jat.2015.07.007

In the last 20 years a whole hierarchy of notions of tractability was proposed and analyzed by several authors. These notions are used to describe the computational hardness of continuous numerical problems in terms of the behavior of their information complexity $n(\epsilon, d)$ as a function of the accuracy $\epsilon$ and the dimension $d$; see [3]. In this talk we present the new notion of $(s,t)$-*weak tractability* defined by

$$\lim_{\epsilon^{-1}+d\to\infty} \frac{\ln n(\epsilon,d)}{\epsilon^{-s}+d^{t}} = 0 \quad \text{for fixed} \quad s,t \geq 0$$

which allows a refined classification of problems with (sub-/super-)exponentially growing information complexity. For compact linear Hilbert space problems $S = (S_d \colon H_d \to G_d)_{d\in\mathbb{N}}$ we provide characterizations of $(s,t)$-weak tractability in terms of the asymptotic decay of the sequence of singular values $(\lambda_{d,j})_{j\in\mathbb{N}}$ of $S_d$. In addition, the advantages of our new notion is illustrated by the example of embedding problems of periodic Sobolev spaces with hybrid smoothness $H^{a,b}(p,\mathbb{T}^d)$ which collect all $f \in L_2(\mathbb{T}^d)$ for which the norm

$$\left[\sum_{\mathbf{k}\in\mathbb{Z}^d} |c_{\mathbf{k}}|^2 \big(1+\sum_{j=1}^{d} |k_j|^p\big)^{2a/p} \prod_{j=1}^{d} (1+|k_j|^2)^b\right]^{1/2}$$

is finite; see [1]. In detail, we complement some conclusions drawn in [2] by showing the following complete tractability characterization:

▶ **Theorem 1.** *Let* $\gamma, \beta \in \mathbb{R}$, $p \in (0,\infty]$, $\alpha > 0$, *and* $s,t \in [0,\infty)$. *Consider* $\mathrm{id} = (\mathrm{id}_d)_{d\in\mathbb{N}}$ *given by*

$$\mathrm{id}_d \colon H^{\gamma+\alpha,\beta}(p,\mathbb{T}^d) \to H^{\gamma,\beta}(p,\mathbb{T}^d), \quad f \mapsto \mathrm{id}_d(f) = f,$$

*w.r.t. the worst case setting and* $\Lambda^{\mathrm{all}}$. *Then we have* $(s,t)$*-weak tractability if and only if*

$$s > \frac{p}{\alpha} \;\; \text{and} \;\; t > 0 \qquad \text{or} \qquad s > 0 \;\; \text{and} \;\; t > 1.$$

*In particular,*
- *UWT, QPT, PT, or SPT never holds,*
- *classical weak tractability holds if and only if* $\alpha > p$.

*Finally, we have the curse of dimensionality if and only if* $p = \infty$.

### References

1 M. Griebel and S. Knapek. *Optimized tensor-product approximation spaces. Constr. Approx.*, 16(4):525–540, 2000.
2 T. Kühn, S. Mayer, and T. Ullrich. *Counting via entropy: new preasymptotics for the approximation numbers of Sobolev embeddings.* arXiv:1505.00631, 2015.
3 E. Novak and H. Woźniakowski. *Tractability of Multivariate Problems. Vol. I–III.* EMS, Zürich, 2008–2012.

## 3.31 On SDEs with arbitrary slow convergence rate at the final time

*Larisa Yaroslavtseva (Universität Passau, DE)*


**Joint work of** Yaroslavtseva, Larisa; Jentzen, Arnulf; Müller-Gronbach, Thomas
**Main reference** A. Jentzen, T. Müller-Gronbach, L. Yaroslavtseva, "On stochastic differential equations with arbitrary slow convergence rates for strong approximation," arXiv:1506.02828v1 [math.NA], 2015.
**URL** http://arxiv.org/abs/1506.02828v1

In the recent article [Hairer, M., Hutzenthaler, M., & Jentzen, A., Loss of regularity for Kolmogorov equations, To appear in Ann. Probab. (2015)] it has been shown that there exist stochastic differential equations (SDEs) with infinitely often differentiable and bounded coefficients such that the Euler scheme converges to the solution in the strong sense but with no polynomial rate. Hairer et al.'s result naturally leads to the question whether this slow convergence phenomenon can be overcome by using a more sophisticated approximation method than the simple Euler scheme. In this talk we answer this question to the negative. We prove that there exist SDEs with infinitely often differentiable and bounded coefficients such that no approximation method based on finitely many observations of the driving Brownian motion converges in absolute mean to the solution with a polynomial rate. Even worse, we prove that for every arbitrarily slow convergence speed there exist SDEs with infinitely often differentiable and bounded coefficients such that no approximation method based on finitely many observations of the driving Brownian motion can converge in absolute mean to the solution faster than the given speed of convergence.

## Participants

- James M. Calvin
NJIT – Newark, US
- Ronald Cools
KU Leuven, BE
- Sonja Cox
University of Amsterdam, NL
- Steffen Dereich
Universität Münster, DE
- Stefan Geiss
University of Jyväskylä, FI
- Michael Gnewuch
Universität Kiel, DE
- Mario Hefter
TU Kaiserslautern, DE
- Stefan Heinrich
TU Kaiserslautern, DE
- Aicke Hinrichs
Universität Linz, AT
- Alexander Keller
NVIDIA GmbH – Berlin, DE
- Peter Kritzer
Universität Linz, AT
- Thomas Kühn
Universität Leipzig, DE
- Peter Mathé
Weierstraß Institut – Berlin, DE
- Thomas Müller-Gronbach
Universität Passau, DE
- Andreas Neuenkirch
Universität Mannheim, DE
- Dong Nguyen
KU Leuven, BE
- Erich Novak
Universität Jena, DE
- Dirk Nuyens
KU Leuven, BE
- Jens Oettershagen
Universität Bonn, DE
- Sergei Pereverzyev
RICAM – Linz, AT
- Leszek Plaskota
University of Warsaw, PL
- Pawel Przybylowicz
AGH Univ. of Science & Technology-Krakow, PL
- Holger Rauhut
RWTH Aachen, DE
- Klaus Ritter
TU Kaiserslautern, DE
- Daniel Rudolf
Universität Jena, DE
- Winfried Sickel
Universität Jena, DE
- Pawel Siedlecki
University of Warsaw, PL
- Ian H. Sloan
University of New South Wales – Sydney, AU
- Jeremy Staum
Northwestern University – Evanston, US
- Ingo Steinwart
Universität Stuttgart, DE
- Mario Ullrich
Universität Linz, AT
- Tino Ullrich
Universität Bonn, DE
- Markus Weimar
Universität Marburg, DE
- Larisa Yaroslavtseva
Universität Passau, DE
- Marguerite Zani
Université d'Orléans, FR