Report from Dagstuhl Seminar 19341

Algorithms and Complexity for Continuous Problems

Edited by

Dmitriy Bilyk[1], Aicke Hinrichs[2], Frances Y. Kuo[3], and Klaus Ritter[4]

1 **University of Minnesota – Minneapolis, US,** dbilyk@math.umn.edu
2 **Johannes Kepler Universität Linz, AT,** aicke.hinrichs@jku.at
3 **UNSW Sydney, AU,** f.kuo@unsw.edu.au
4 **TU Kaiserslautern, DE,** ritter@mathematik.uni-kl.de

Abstract

From 18.08. to 23.08.2019, the Dagstuhl Seminar 19341 Algorithms and Complexity for Continuous Problems was held in the International Conference and Research Center (LZI), Schloss Dagstuhl. During the seminar, participants presented their current research, and ongoing work and open problems were discussed. Abstracts of 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 August 18–23, 2019 – http://www.dagstuhl.de/19341
2012 ACM Subject Classification Theory of computation → Approximation algorithms analysis, Theory of computation → Lower bounds and information complexity, Mathematics of computing → Stochastic differential equations, Mathematics of computing → Approximation, Mathematics of computing → Quadrature
Keywords and phrases applied harmonic analysis, computational stochastics, computing and complexity in infinite dimensions, discrepancy theory, tractability analysis
Digital Object Identifier 10.4230/DagRep.9.8.26
Edited in cooperation with David Krieg, JKU Linz, AT, david.krieg@jku.at

1 Executive Summary

Dmitriy Bilyk
Aicke Hinrichs
Frances Y. Kuo
Klaus Ritter

This was already the 13th Dagstuhl Seminar on Algorithms and Complexity for Continuous Problems over a period of 28 years. It brought together researchers from different communities working on complexity of continuous problems. Such problems, which originate from numerous areas, including physics, chemistry, finance, and economics, can almost never be solved analytically, but rather only approximately to within some error threshold. The complexity analysis ideally includes the construction of (asymptotically) optimal algorithms. 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. The current seminar attracted 41 participants from nine different countries all over the world. About 30% of them were young researchers including PhD students. There were 34 presentations.

The following topics were covered:

Tractability analysis of high-dimensional problems: Tractability analysis is an area of applied mathematics and theoretical computer science that studies the minimal computational resources needed for the approximate solution of problems with a huge number of variables, and it can be seen as a unifying theme for the preceding seminars in this series. Many concrete problems from applications have been analyzed in this context, new algorithms were developed, approaches to break the curse of dimensionality were established, but there remain a number of important open problems. Tractability analysis will serve as a guideline and a tool for establishing complexity results and for constructing algorithms for infinite dimensional problems.

Computational stochastics: The focus was on weak and strong approximation as well as on the quadrature problem for stochastic ordinary or partial differential equations, i.e., on models with a random dynamics in a finite- or infinite-dimensional state space. A major topic was the complexity analysis for stochastic differential equations under non-standard assumptions.

Computing and complexity in infinite dimensions: Computational problems with infinitely many variables naturally arise in rather different application areas. Results and techniques from tractability analysis are available and thus permit one to study infinite dimensional problems as the limit of finite dimensional ones. Moreover, the availability of generic types of algorithms, like the multivariate decomposition method or the multi-level approach, will contribute to the complexity analysis and practical application in integration and approximation problems of infinitely many variables.

Discrepancy theory: Classical discrepancy theory is concerned with the question how uniformly finite point sets can be distributed. The geometric notion of discrepancy is intimately connected to the complexity of integration for functions from certain function classes. For problems in both fixed low dimension and high dimension, there are intriguing open questions whose solution would impact both fields of discrepancy theory and tractability studies.

Computational/applied harmonic analysis: Harmonic analysis plays an increasingly important role both in discrepancy theory and tractability analysis. One highlight is the proof of the currently best known lower bound for the star discrepancy in fixed dimension, which showed close connections between different areas, so similar techniques could be used to establish better bounds for the celebrated small ball problem for Gaussian processes. Equally important for the workshop is that many of the interesting spaces of functions occurring in numerical problems are well suited to the application of harmonic analysis.

As we understand better and better, these subjects are highly interrelated, and they are probably the most active and promising ones in the fields for the next decade. Bringing together a mix of junior and senior researchers from these diverse but interrelated subjects in a Dagstuhl seminar resulted in considerable progress both for the theory and the applications in these areas.

Seminars in applied mathematics and theoretical computer science typically consist of presentations, followed by short discussions in the plenum, and numerous informal discussions in smaller groups. In this seminar, we added another new feature. A moderator was assigned to three preselected talks (based on their particular relevance and on the experience of the speaker) in order to inspire a longer, in-depth discussion in the plenum. The three speakers were Jan Víbyral, Erich Novak, and Martin Hutzenthaler. The talks were scheduled as the

first talks on Tuesday, Wednesday and Thursday. It was indeed very inspiring to witness the long and deep discussions following these special talks. We feel that this format was successful and should be used also in other workshops and conferences of the community.

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 Bound on the expected number of function evaluations required to approximate the minimum of a smooth Gaussian process

James M. Calvin (NJIT – Newark, US)

We consider the problem of approximating the minimum of a function using sequentially chosen points at which to evaluate the function. Given a random function, we want an algorithm that approximates the minimum to a prescribed accuracy with few function evaluations on average.

In this talk we consider the function to be a centered stationary Gaussian process on the unit interval with three-times continuously differentiable paths. We assume that the covariance function of the process has positive second and fourth spectral moments.

We describe an algorithm that takes as input an error tolerance ϵ and confidence level γ, and stops when the probability that the error exceeds ϵ is at most γ. For our probability model and algorithm, the expected number of function evaluations required, in terms of the error tolerance ϵ, is of order $\log(1+1/\epsilon)\log\log(1+1/\epsilon)$.

3.2 Lattice Algorithms for Multivariate Approximation in Periodic Spaces with General Weight Parameters

Ronald Cools (KU Leuven, BE), Frances Y. Kuo (UNSW Sydney, AU), Dirk Nuyens (KU Leuven, BE), and Ian Sloan

Main reference Ronald Cools, Frances Y. Kuo, Dirk Nuyens, Ian H. Sloan: "Lattice algorithms for multivariate approximation in periodic spaces with general weight parameters", CoRR, Vol. abs/1910.06604, 2019.
URL http://arxiv.org/abs/1910.06604

This talk summarizes a recent manuscript by the authors on the theoretical foundation for the construction of lattice algorithms for multivariate L_2 approximation in the worst case setting, for functions in a periodic space with general weight parameters. Our construction leads to an error bound that achieves the optimal rate of convergence for lattice algorithms.

3.3 Convergence in Hölder and Sobolev norms for approximations of Gaussian fields

Sonja Cox (University of Amsterdam, NL)

Joint work of Sonja Cox, Kristin Kirchner
Main reference Sonja G. Cox, Kristin Kirchner: "Regularity and convergence analysis in Sobolev and Hölder spaces for generalized Whittle-Matérn fields", CoRR, Vol. abs/1904.06569, 2019.
URL https://arxiv.org/abs/1904.06569

In models involving a Gaussian field one frequently assumes the covariance operator to be given by a negative fractional power of a second-order elliptic differential operator of the form $L := -\nabla \cdot (A\nabla) + \kappa^2$. Whittle-Matérn fields form an well-known example of such a model. Such covariance operators allow for a reasonable amount of model flexibility (adjustable correlation length and the smoothness of the field) whilst being relatively easy to simulate. In our work we established optimal strong convergence rates in Hölder and Sobolov norms for Galerkin approximations of such Gaussian random fields. More specifically, we considered both spectral Galerkin methods and finite element methods. The latter, although significantly more tedious to analyse, are more suitable for non-stationary fields on non-standard domains.

3.4 CLTs for stochastic approximation schemes under non-standard assumptions

Steffen Dereich (Universität Münster, DE)

We establish new CLTs for Ruppert-Polyak averaged stochastic gradient descent schemes. Instead of isolated attractors we consider attracting manifolds. On the event of convergence we prove a stable limit theorem which is of the optimal order $n^{-1/2}$.

3.5 The spectral decomposition of discrepancy kernels on manifolds

Martin Ehler (Universität Wien, AT)

We study the spectral decomposition of discrepancy kernels when restricted to compact kernels of $\mathbb{R}^d$. For restrictions to the Euclidean ball in odd dimensions, to the rotation group SO(3), and to the Grassmannian manifold, we compute the kernel's Fourier coefficient and determine their asymptotics.

3.6 Nested multilevel Monte Carlo and use of approximate random variables

Michael Giles (University of Oxford, GB)

Joint work of Mike Giles, Oliver Sheridan-Methven
Main reference Michael B. Giles: "Multilevel Monte Carlo methods", Acta Numer., Vol. 24, pp. 259–328, 2015.
URL http://dx.doi.org/10.1017/S096249291500001X

The multilevel Monte Carlo (MLMC) method has been used for a wide variety of stochastic applications. In this talk we consider its use in situations in which input random variables can be replaced by similar approximate random variables which can be computed much more cheaply. A nested MLMC approach is adopted in which a two-level treatment of the approximated random variables is embedded within a standard MLMC application. We analyse the resulting nested MLMC variance in the specific context of an SDE discretisation in which Normal random variables can be replaced by approximately Normal random variables, and provide numerical results to support the analysis.

3.7 Mixed Randomized Sequences, Negative Dependence, and Probabilistic Discrepancy Bounds

Michael Gnewuch (Universität Osnabrück, DE)

Joint work of Michael Gnewuch, Benjamin Doerr, Nils Hebbinghaus, Marcin Wnuk

We consider sampling schemes in the d-dimensional unit cube $[0,1]^d$. A simple example would be Monte Carlo (MC) points $\mathbf{X} := (X_i)_{i=1}^n$, which are independent and uniformly distributed in $[0,1]^d$. It is known that MC points satisfy the probabilistic star discrepancy bound

$$\mathrm{disc}^*(\mathbf{X}) \leq c\sqrt{d/n} \tag{1}$$

with positive probability (Heinrich et al. 2001, Aistleitner 2011), where the smallest value for the constant so far, $c = 2.5287$, was achieved in [2]. This bound is a pre-asymptotic bound, since it gives useful information for a moderate number of points n (only depending linearly on d) and the dependence of all constants on the number of points n and the dimension d is made explicit. So far there is no sampling scheme known that satisfies a better pre-asymptotic bound for the star discrepancy.

Our goal is to identify those sampling schemes $\mathbf{X} := (X_i)_{i=1}^n$, whose points are "well spreaded" in $[0,1]^d$ in the sense that the probabilistic bound for the star discrepancy of $X_1, \ldots, X_n$ is (essentially) not worse than bound (1). One sufficient condition is that the sampling scheme satisfies certain negative dependence properties.

If $\mathbf{X}$ satisfies, e.g., a certain negative dependence property with respect to arbitrary axis-parellel boxes anchored in 0, then a discrepancy bound of the form

$$\mathrm{disc}^*(\mathbf{X}) \leq c\sqrt{d/n}\sqrt{\ln(1+n/d)},$$

c small, holds with positive probability, see [3].

If $\mathbf{X}$ satisfies even the corresponding negative dependence property with respect to arbitrary differences of axis-parellel boxes anchored in 0, then a discrepancy bound of the form (1) for small c holds with positive probability, see [2].

Examples of sampling schemes that satisfy these negative dependence properties include, apart from MC points, Latin hypercube sampling, see [2], generalized stratified sampling or certain mixed randomized sequences, see [3].

The notion of negative dependence used is a relaxation of the notion of (upper and lower) negative orthant dependence. The relaxation allows for a parameter $\gamma \in [1, \infty)$ that in the case of negative orthant dependence is fixed to be one. In a project started at the Dagstuhl seminar 19341, we recently have been able to show that the negative dependence property for Latin hypercube samples proved in [2] for parameters $\gamma = \gamma(d) = e^d$ can actually only be proved for a $\gamma = \gamma(d)$ that grows at least of the order $\Omega(\sqrt{d})$ as d tends to infinity, see [1]. That is, although Latin hypercube sampling definitely does not satisfy negative orthant dependence with respect to arbitrary differences of axis-parellel boxes anchored in 0, it satisfies the corresponding new relaxed negative dependence property with (dimension-dependent) parameter γ.

References

1 B. Doerr, M. Gnewuch. *On negative dependence properties of Latin hypercube samples and scrambled nets*. In preparation.

2 M. Gnewuch, N. Hebbinghaus. *Discrepancy bounds for a class of negatively dependent random points including Latin hypercube samples*. Preprint 2018 (submitted).

3 M. Wnuk, M. Gnewuch, N. Hebbinghaus. *On negatively dependent sampling schemes, variance reduction, and probabilistic upper discrepancy bounds*. Preprint, arXiv:1904.10796. (To appear in: D. Bylik, J. Dick, F. Pillichshammer (Eds.), Proceedings of the RICAM Special Semester on Multivariate Algorithms and their Foundations in Number Theory, Linz 2018, DeGruyter.)

3.8 Multilevel Monte Carlo methods for estimating the expected value of sample information

Takashi Goda (University of Tokyo, JP)

Joint work of Michael B. Giles, Takashi Goda, Tomohiko Hironaka, Howard Thom

Main reference Michael B. Giles, Takashi Goda: "Decision-making under uncertainty: using MLMC for efficient estimation of EVPPI", Statistics and Computing, Vol. 29(4), pp. 739–751, 2019.

URL http://dx.doi.org/10.1007/s11222-018-9835-1

Motivated by applications to medical decision making, we study Monte Carlo estimation of the expected value of partial perfect information (EVPPI) and the expected value of sample information (EVSI). Both EVPPI and EVSI are defined as nested expectations, for which the standard (nested) Monte Carlo methods requires $O(\varepsilon^{-3})$ or $O(\varepsilon^{-4})$ computational costs to achieve the root-mean-square accuracy ε. To reduce these costs to $O(\varepsilon^{-2})$, we introduce antithetic multilevel Monte Carlo (MLMC) estimators for these quantities in this study. Under some assumptions on decision models, the antithetic property of the MLMC estimator enables to prove such a computational complexity for estimating EVPPI (Giles and Goda, 2019). The result can be extended to EVSI, by directly using the Bayes' formula and showing auxiliary results on the MLMC estimation of nested ratio expectations (Hironaka, Giles, Goda and Thom, in preparation). Numerical experiments support our theoretical analysis.

3.9 Adaptive Quantile Computation for Brownian Bridge in Change-Point Analysis

Mario Hefter (TU Kaiserslautern, DE)

Joint work of Mario Hefter, Jürgen Franke, André Herzwurm, Klaus Ritter, Stefanie Schwaar

In change-point analysis, weighted partial sum processes are used to detect changes. A well-known test statistic for change-points is their maximum. Asymptotically, its distribution is specified by the supremum of a weighted Brownian bridge, for which the distribution function is not known in general such that critical values have to be calculated numerically by simulation. We construct an adaptive Monte Carlo algorithm for generating weighted Brownian bridges with the goal of approximating the distribution of their suprema. We compare the new method with the classical algorithm based on evaluating the stochastic process on an equidistant grid. For prescribed approximation quality, the new algorithm provides a much faster calculation of, e.g., critical values.

3.10 Overcoming the curse of dimensionality for parabolic PDEs

Martin Hutzenthaler (Universität Duisburg-Essen, DE)

Joint work of Weinan E, Martin Hutzenthaler, Arnulf Jentzen, Thomas Kruse, Tuan Anh Nguyen, Philippe von Wurstemberger
Main reference Martin Hutzenthaler, Arnulf Jentzen, Thomas Kruse, Tuan Anh Nguyen, Philippe von Wurstemberger: "Overcoming the curse of dimensionality in the numerical approximation of semilinear parabolic partial differential equations", CoRR, Vol abs/1807.01212, 2018.
URL https://arxiv.org/abs/1807.01212

For a long time it is well-known that high-dimensional linear parabolic partial differential equations (PDEs) can be approximated by Monte Carlo methods with a computational effort which grows polynomially both in the dimension and in the reciprocal of the prescribed accuracy. In other words, linear PDEs do not suffer from the curse of dimensionality. For general semilinear PDEs with Lipschitz coefficients, however, it remained an open question whether these suffer from the curse of dimensionality. This talk explains a new numerical approximation algorithm introduced in [1] and [2] which overcomes the curse of dimensionality in the numerical approximation of general semilinear heat equations with gradient-independent nonlinearities.

References

1 E, W., Hutzenthaler, M., Jentzen, A., and Kruse, T. Linear scaling algorithms for solving high-dimensional nonlinear parabolic differential equations. *arXiv:1605.00856* (2016).
2 Hutzenthaler, M., Jentzen, A., Kruse, T., Nguyen, T. A., and von Wurstemberger, P. Overcoming the curse of dimensionality in the numerical approximation of semilinear parabolic partial differential equations. *arXiv preprint arXiv:1807.01212* (2018).

3.11 Quasi-Monte Carlo Methods and Artificial Neural Networks

Alexander Keller (NVIDIA, DE)

Joint work of Gonçalo Mordido, Matthijs Van Keirsbilck, Alexander Keller
URL https://developer.nvidia.com/gtc/2019/video/S9389

The average human brain has about 10^{11} nerve cells, where each of them may be connected to up to 10^4 others. We therefore investigate the question whether there are algorithms for artificial neural networks that are linear in the number of neurons, while the number of connections incident to a neuron is bounded by a constant.

Representing artificial neural networks by paths, we offer two approaches to answer this question: First, we derive an algorithm that quantizes a trained artificial neural network such that the resulting complexity is linear [1]. Second, we demonstrate that training networks, whose connections are determined by uniform sampling can achieve a similar precision as using fully connected layers. Due to sparsity upfront, these networks can be trained much faster. Finally, we explain how generating the paths using quasi-Monte Carlo methods, especially the Sobol' low discrepancy sequence, leads to a new parallel hardware architecture for artificial neural networks.

References

1 Gonçalo Mordido, Matthijs Van keirsbilck, Alexander Keller. *Instant Quantization of Neural Networks using Monte Carlo Methods*. https://arxiv.org/abs/1905.12253, 2019

3.12 Fast simulation of non-stationary Gaussian random fields

Kristin Kirchner (ETH Zürich, CH)

Joint work of Lukas Herrmann, Kristin Kirchner, Christoph Schwab
Main reference Lukas Herrmann, Kristin Kirchner, Christoph Schwab: "Multilevel Approximation of Gaussian Random Fields: Fast Simulation", Mathematical Models and Methods in Applied Sciences(ja), 2019.
URL http://dx.doi.org/10.1142/S0218202520500050

We propose and analyze multilevel algorithms for the fast simulation of possibly non-stationary Gaussian random fields (GRFs for short) indexed, e.g., by a bounded domain $\mathcal{D} \subset \mathbb{R}^d$ or by a compact d-manifold $\mathcal{M}$. A *colored* GRF $\mathcal{Z}$, admissible for our algorithms, solves the stochastic fractional-order equation $\mathcal{A}^\beta \mathcal{Z} = \mathcal{W}$ for some $\beta > d/4$, where $\mathcal{A}$ is a linear, local, second-order elliptic differential operator in divergence form and $\mathcal{W}$ is white noise. We thus consider GRFs with covariance operators of the form $\mathcal{C} = \mathcal{A}^{-2\beta}$.

The proposed algorithms numerically approximate samples of $\mathcal{Z}$ on nested sequences $\{\mathcal{T}_\ell\}_{\ell \geq 0}$ of regular, simplicial partitions $\mathcal{T}_\ell$ of $\mathcal{D}$ and $\mathcal{M}$, respectively. Work and memory to compute one approximate realization of the GRF $\mathcal{Z}$ on the triangulation $\mathcal{T}_\ell$ with consistency $\mathcal{O}(N_\ell^{-\rho})$, for some consistency order $\rho > 0$, scale essentially linear in $N_\ell = \#(\mathcal{T}_\ell)$, independent of the possibly low regularity of the GRF. The algorithms are based on a sinc quadrature for an integral representation of (the application of) the negative fractional-order elliptic operator $\mathcal{A}^{-\beta}$. For the proposed numerical approximation, we prove bounds of the computational cost and the consistency error.

3.13 The power of random information

David Krieg (Johannes Kepler Universität Linz, AT), Aicke Hinrichs (Johannes Kepler Universität Linz, AT), Erich Novak (Universität Jena, DE), Joscha Prochno (Universität Graz, AT), and Mario Ullrich (Johannes Kepler Universität Linz, AT)

Main reference Aicke Hinrichs, David Krieg, Erich Novak, Joscha Prochno, Mario Ullrich: "On the power of random information", CoRR, Vol. abs/1903/006081, 2019.
URL https://arxiv.org/abs/1903.00681

We study problems like recovering a function from a finite number of function values. Usually, it is assumed that these function values can be computed at arbitrary points. In this talk, we assume that we do not get to choose the points. We compare the quality of random sampling points with the quality of optimal sampling points. How much do we loose?

References

1 Aicke Hinrichs, David Krieg, Erich Novak, Joscha Prochno, Mario Ullrich. *Random sections of ellipsoids and the power of random information*. arXiv:1901.06639 [math.FA]
2 Aicke Hinrichs, David Krieg, Erich Novak, Joscha Prochno, Mario Ullrich. *On the power of random information*. arXiv:1903.00681 [math.NA]
3 David Krieg, Mario Ullrich. *Function values are enough for L_2-approximation*. arXiv:1905.02516 [math.NA]

3.14 Exponential tractability of linear tensor product problems

Peter Kritzer (Österreichische Akadamie der Wissenschaften – Linz, AT)

Joint work of Fred J. Hickernell, Peter Kritzer, Henryk Wozniakowski
Main reference Fred J. Hickernell, Peter Kritzer, Henryk Wozniakowski: "Exponential tractability of linear tensor product problems", CoRR, Vol. abs/1811.05856, 2018.
Main reference https://arxiv.org/abs/1811.05856

We consider the approximation of compact linear operators defined over tensor product Hilbert spaces. Necessary and sufficient conditions on the singular values of the problem under which we can or cannot achieve different notions of exponential tractability were given by Papageorgiou, Petras, and Wozniakowski in 2017. Here we present an alternative proof method based on a more recent result to obtain these conditions. As opposed to the algebraic setting, several tractability notions cannot be achieved for non-trivial cases in the exponential setting.

3.15 Optimal confidence for Monte Carlo integration of smooth functions

Robert J. Kunsch (RWTH Aachen, DE) and Daniel Rudolf (Universität Göttingen, DE)

Main reference Robert J. Kunsch, Daniel Rudolf: "Optimal confidence for Monte Carlo integration of smooth functions", CoRR, Vol. abs/1809/09890, 2018.
URL https://arxiv.org/abs/1809.09890

We study the information-based complexity of approximating integrals of smooth functions at absolute precision $\varepsilon > 0$ with confidence level $1 - -\delta \in (0, 1)$ using function evaluations within randomized algorithms. The probabilistic error criterion is new in the context of integrating smooth functions. In previous research, Monte Carlo integration was studied in terms of the expected error (or the root mean squared error), for which linear methods achieve optimal rates of the error $e(n)$ in terms of the number n of function evaluations. In our context, usually methods that provide optimal confidence properties exhibit non-linear features. The optimal probabilistic error rate $e(n, \delta)$ for multivariate functions from classical isotropic Sobolev spaces $W_p^r(G)$ with sufficient smoothness on bounded Lipschitz domains $G \subset \mathbb{R}^d$ is determined. It turns out that the integrability index p has an effect on the influence of the uncertainty δ in the complexity. In the limiting case $p = 1$ we see that deterministic methods cannot be improved by randomization. In general, higher smoothness reduces the additional effort for diminishing the uncertainty. Finally, we add a discussion about this problem for function spaces with mixed smoothness.

3.16 Uniform Recovery Guarantees for Least Squares Approximtion

Lutz Kämmerer (TU Chemnitz, DE)

Joint work of Lutz Kämmerer, Tino Ullrich, Toni Volkmer
Main reference Lutz Kämmerer: "Multiple Lattice Rules for Multivariate L_∞ Approximation in the Worst-Case Setting", CoRR, Vol. abs/1909.02290, 2019.
URL http://arxiv.org/abs/1909.02290

We recapitulate recent results for least squares approximation using random point sets. In particular, for the $L_2^(\mathbb{T}^d)$ approximation of functions from periodic Sobolev spaces $H^s_{\mathrm{mix}}(\mathbb{T}^d)$ of dominating mixed smoothness s, the uniform recovery guarantees $\sup_{\|f\|_{H^s_{\mathrm{mix}}}} \|f - \tilde{f}\|_{L_2} \lesssim n^{-s} \log^{ds} n$ hold, which is an improvement compared to best known so far sparse grid algorithms for small smoothness. Furthermore, the $L_\infty(\mathbb{T}^d)$ approximation using a set of rank-1 lattices as sampling nodes provides an efficient approximation algorithm that uses several least squares solutions in order to build up an approximation. This approach achieves the best possible main rate $s - 1/2$ in $1/n$ of the sampling error.

3.17 In the search for all zeros of smooth functions

Leszek Plaskota (University of Warsaw, PL)

We report results obtained in an on going research on the problem of finding the set of all zeros of functions $f \in C^r([0,1])$, $r \in \{0,1,2,\ldots\}$, such that $f^{(r)}$ is Hölder continuous with exponent $\varrho \in (0,1]$. We also allow $r = +\infty$, in which case f is infinitely many times continuously differentiable. Possible algorithms use information about values of f and/or its derivatives at n points. The error between the true solution $Z(f)$ and approximate solution $Z_n(f)$ is measured via the Hausdorff distance $d_H(Z(f), Z_n(f))$ between sets. We construct a nonadaptive algorithm using function evaluations at equally spaced points whose error converges to zero as $n \to +\infty$, for all functions f from our class. On the other hand, the convergence is arbitrarily slow. Specifically, for any sequence $\{Z_n\}_{n\geq 1}$ of approximations and for any positive sequence $\{\tau_n\}_{n\geq 1}$ converging to zero there are functions f^* having exactly one zero for which the errors $d_H(Z(f^*), Z_n(f^*))$ do not converge to zero or converge slower than τ_n.

We also note that the same results hold for finding zeros of functions from the corresponding class of multivariate functions, and for other problems, such as finding all fixed points or finding all global minima.

These results confirm a common belief that smoothness itself is not enough to have faster convergence of algorithms for those problems.

3.18 Convergence order of the Euler-Maruyama scheme in dependence of the Sobolev regularity of the drift

Michaela Szölgyenyi (Alpen-Adria-Universität Klagenfurt, AT)

Joint work of Michaela Szölgyenyi, Andreas Neuenkirch

We study the strong convergence rate of the Euler-Maruyama scheme for scalar SDEs with additive noise and irregular drift. We provide a framework for the error analysis by reducing it to a weighted quadrature problem for irregular functions of Brownian motion. By analysing the quadrature problem we obtain for abitrarily small $\epsilon > 0$ a strong convergence order of $(1+\kappa)/2 - \epsilon$ for a non-equidistant Euler-Maruyama scheme, if the drift has Sobolev-Slobodeckij-type regularity of order $\kappa \in (0,1)$.

3.19 On strong approximation of SDEs with a discontinuous drift coefficient

Thomas Müller-Gronbach (Universität Passau, DE) and Larisa Yaroslavtseva (Universität Passau, DE)

Recently a lot of effort has been invested in the literature to analyze the L_p-error of the Euler-Maruyama scheme in the case of stochastic differential equations (SDEs) with a drift coefficient that may have discontinuities in space. For scalar SDEs with a piecewise Lipschitz drift coefficient and a Lipschitz diffusion coefficient that is non-zero at the discontinuity points of the drift coefficient so far only an L_p-error rate of at least $1/(2p)-$ has been proven in the literature. In this talk we show that under the latter assumptions on the coefficients of the SDE the Euler-Maruyama scheme in fact achieves an L_p-error rate of at least $1/2$ for all $p \in [1, \infty)$ as in the case of SDEs with Lipschitz coefficients. We furthermore present a numerical method, which achieves an L_p-error rate of at least $3/4$ for all $p \in [1, \infty)$ if, additionally to the assumptions stated above, both the drift and the diffusion coefficients are piecewise differentiable with Lipschitz derivatives.

3.20 Algorithms and Complexity for Functions on General Domains

Erich Novak (Universität Jena, DE)

Error bounds and complexity bounds in numerical analysis and information-based complexity are often proved for functions that are defined on very simple domains, such as a cube, a torus, or a sphere. We study optimal error bounds for the approximation and integration and only assume that the domain is a bounded Lipschitz domain in $\mathbb{R}^d$. It is known that for many problems the order of convergence does not depend on the domain. We present examples for which the following is true:

1) Also the asymptotic constant does not depend on the shape of the domain, only of its volume.
2) There are explicit and uniform lower (or upper, respectively) bounds for the error that are only slightly smaller (or larger, respectively) than the asymptotic error bound.

3.21 Lattice algorithms for approximation: new constructions

Dirk Nuyens (KU Leuven, BE), Ronald Cools (KU Leuven, BE), Ian H. Sloan, and Frances Y. Kuo (UNSW Sydney, AU)

We derive a new CBC algorithm for the construction of good generating vectors for rank-1 lattice point sets which can be used for approximation in the Korobov space with general weights. The good news is that this construction is independent of the index set on which we represent our approximated function which makes for nice and fast construction algorithms in the case of product, POD and SPOD weights.

3.22 Randomized Euler scheme for strong approximation of SDEs under Sobolev-Slobodeckij smoothness

Paweł Przybyłowicz (AGH Univ. of Science & Technology-Krakow, PL)

Joint work of Paweł Przybyłowicz, Raphael Kruse

We investigate the problem of strong approximation of solution of the following scalar SDE

$$\begin{cases} dX(t) = a(t, X(t))dt + b(t)dW(t), \quad t \in [0, T], \\ X(0) = \eta, \end{cases} \tag{2}$$

driven by a standard one-dimensional Wiener process $W = (W(t))_{t\in[0,T]}$. We assume that $a = a(t, y)$ and $b = b(t)$ are only measurable with respect to the time variable t, and a is globally Lipschitz with respect to the space variable y.

We investigate behavior of the randomized Euler scheme X_n^{RE}, which evaluates a and b at randomly chosen points. By using Information-Based Complexity framework we show that randomized Euler scheme converges to the solution X of the underlying SDE but the convergence of X_n^{RE} to X may be arbitrarily slow ([5]). In order to get positive results we assume that b belongs to the Sobolev-Slobodeckij space $W^{\sigma,p}$, $\sigma \in (0, 1)$, $p > 2$. In this case we show that the $L^2(\Omega)$-error of the algorithm X_n^{RE} is $O(n^{-\min\{\frac{1}{2}-\frac{1}{p},\sigma\}})$. Moreover, we investigate corresponding lower bounds ([3]). In particular, this extends the results from [1], [2], [4], and [6], obtained for the randomized Euler scheme.

References

1 S. Heinrich, and B. Milla, The randomized complexity of initial value problems, *J. Complexity*, **24**: 77–88 (2008).
2 A. Jentzen, and A. Neuenkirch, A random Euler scheme for Carathéodory differential equations, *J. Comput. Appl. Math.*, **224**: 346–359 (2009).
3 R. Kruse, and P. Przybyłowicz, Approximation of solutions of SDEs with fractional Sobolev regularity, in preparation
4 R. Kruse, and Y. Wu, Error analysis of randomized Runge–Kutta methods for differential equations with time-irregular coefficients, *Comput. Methods Appl. Math.*, **17**: 479–498 (2017).
5 P. Przybyłowicz, On arbitrary slow rate of convergence for randomized Euler scheme, in preparation
6 P. Przybyłowicz, and P. Morkisz, Strong approximation of solutions of stochastic differential equations with time-irregular coefficients via randomized Euler algorithm, *Appl. Numer. Math.*, **78**: 80–94 (2014).

3.23 Tractability properties of discrepancy

Friedrich Pillichshammer (Johannes Kepler Universität Linz, AT)

Joint work of Josef Dick, Aicke Hinrichs, Friedrich Pillichshammer

Discrepancies are quantitative measures for the irregularity of distribution of point sets in $[0, 1]^d$ which are closely related to the error of quasi-Monte Carlo (QMC) integration rules. Classical results consider discrepancy with respect to its asymptotic dependence when the

size N of a point set tends to infinity. In this sense optimal results are known, but often these results give no information on the pre-asymptotic scale, especially when the dimension d is large.

In 2001 Heinrich, Novak, Wasilkowski and Woźniakowski [1] initiated the study of the dependence of discrepancy on the dimension d with a remarkable result for the star discrepancy. They showed that for every N and d there exists a N-point set in $[0,1]^d$ with classical star discrepancy of at most $C\sqrt{d/N}$, where C is a positive constant independent of N and d. Since then a lot of papers on this topic with exciting results have appeared. Nevertheless, a lot of problems are still open.

In this talk we give a review of this topic and present some new results concerning the periodic L_2 discrepancy and the discrepancy with respect to the exponential Orlicz norm.

References

1 S. Heinrich, E. Novak, G.W. Wasilkowski, and H. Woźniakowski: The inverse of the star-discrepancy depends linearly on the dimension. Acta Arith. 96: 279-302, 2001.

3.24 Large deviations in geometric functional analysis

Joscha Prochno (Universität Graz, AT)

Joint work of Joscha Prochno, Zakhar Kabluchko, Christoph Thäle

Large deviations are a classical topic in probability theory, but have only recently entered the scene of asymptotic geometric analysis. After giving a short introduction to the theory of large deviations, we present a large deviations principle for the q-norm length of a random vector chosen uniformly at random from the unit ball of ℓ_p^n. More precisely, we show that for $1 \leq p < \infty$ and $q > p$, the sequence $(n^{1/p-1/q}\|X\|_q)_{n\in\mathbb{N}}$ satisfies a large deviations principle with speed $n^{p/q}$ and rate

$$\mathbb{I}(z) = \begin{cases} \frac{1}{p}(z^q - -M_p(q))^{p/q} & \text{for } z^q \geq M_p(q), \\ \infty & \text{else.} \end{cases}$$

We shall also mention large deviations results that can be proved in the noncommutative setting of Schatten classes. In this case, the rate function is essentially the logarithmic energy plus some perturbation by a constant strongly connected to the famous Ullman distribution. As a consequence of the Sanov-type large deviations, one obtains a strong law of large numbers showing that the empirical spectral measure converges weakly almost surely to the Ullman distribution.

3.25 Wasserstein contraction and spectral gap of simple slice sampling

Daniel Rudolf (Universität Göttingen, DE)

Joint work of Daniel Rudolf, Viacheslav Natarovskii, Björn Sprungk
Main reference Viacheslav Natarovskii, Daniel Rudolf, Björn Sprungk: "Quantitative spectral gap estimate and Wasserstein contraction of simple slice sampling", CoRR, Vol abs/1903/03824, 2019.
URL https://arxiv.org/abs/1903.03824

We provide results on Wasserstein contraction of simple slice sampling for approximate sampling w.r.t. distributions with log-concave and rotational invariant Lebesgue densities. This leads to an explicit quantitative lower bound of the spectral gap of simple slice sampling. In addition to that this lower bound carries over to more general target distributions depending only on the volume of the (super-)level sets of their unnormalized density.

3.26 Complexity of stochastic integration

Stefan Heinrich (TU Kaiserslautern, DE)

We study the complexity of stochastic integration with respect to the Wiener sheet measure $\int_{[0,1]^d} f(t)dW_t$ of stochastic functions $f = f(t,\omega)$ with Besov $B^r_{pp}([0,1]^d)$ and Bessel potential $H^r_p([0,1]^d)$ regularity in t. We determine the complexity in the deterministic and randomized setting, which includes finding and analyzing algorithms of optimal order and proving matching lower bounds.

3.27 Potential Theory, inverse Laplacians and new Low(?)-Discrepancy Sequences

Stefan Steinerberger (Yale University – New Haven, US)

Main reference Stefan Steinerberger: "A Nonlocal Functional Promoting Low-Discrepancy Point Sets", CoRR, Vol. abs/1902.00441, 2019.
URL https://arxiv.org/abs/1902.00441
Main reference Stefan Steinerberger: "Dynamically Defined Sequences with Small Discrepancy", CoRR, Vol. abs/1902.03269, 2019.
URL https://arxiv.org/abs/1902.03269
Main reference Florian Pausinger: "Greedy energy minimization can count in binary: point charges and the van der Corput sequence", CoRR, Vol. abs/1905.09641, 2019.
URL https://arxiv.org/abs/1905.09641

We discuss a new way to construct sequences in the unit interval with very favorable distribution properties. Our construction is based on a greedy algorithm that uses the Green function of the fractional Laplacian as a kernel; we can prove that the discrepancy of this set is at least $N^{-1/2}\log(N)$ but presumably much stronger results hold true (and this is also backed up by numerical investigations). This seems to open several different lines of research.

3.28 Dimension-independent convergence of Gaussian process regression

Aretha Teckentrup (University of Edinburgh, GB)

We consider the problem of interpolating a function $f : [0,1]^s \to \mathbb{R}$, where the input dimension s is potentially large. In particular, we study kernel based meshless methods such as kernel based interpolants and Gaussian process emulators. Using results from high-dimensional quadrature, we prove error estimates that are independent of s. The errors are measured in the L^2-norm or the supremum norm.

3.29 Sampling discretization error of integral norms for function classes

Vladimir N. Temlyakov (University of South Carolina – Columbia, US)

The new ingredient of this paper is that we consider infinitely dimensional classes of functions and instead of the relative error setting, which was used in previous papers on norm discretization, we consider the absolute error setting. We demonstrate how known results from two areas of research – supervised learning theory and numerical integration – can be used in sampling discretization of the square norm on different function classes.

3.30 Discrepancy, Dispersion and Fixed Volume Discrepancy

Mario Ullrich (Johannes Kepler Universität Linz, AT)

Joint work of Vladimir N. Temlyakov, Mario Ullrich
Main reference Vladimir N. Temlyakov, Mario Ullrich: "On the fixed volume discrepancy of the Fibonacci sets in the integral norms", CoRR, Vol. abs/1908.04658, 2019.
URL http://arxiv.org/abs/1908.04658

We present a bunch of recent results on the discrepancy and dispersion, especially in high dimensions, and give an introduction to a new geometric quantity – the fixed volume discrepancy. One of the implications that can be obtained from this new quantity can be stated like this: "Bad boxes" for the discrepancy cannot be "too small".

3.31 Approximation of shallow neural networks

Jan Vybíral (Czech Technical University – Prague, CZ)

Joint work of Jan Vybíral, Massimo Fornasier, Ingrid Daubechies, Karin Schnass, Tino Ullrich, Sebastian Mayer
Main reference Massimo Fornasier, Jan Vybíral, Ingrid Daubechies: "Identification of Shallow Neural Networks by Fewest Samples", CoRR, Vol. abs/1804.01592, 2018.
URL http://arxiv.org/abs/1804.01592

We address the structure identification and the uniform approximation of sums of ridge functions $f(x) = \sum_{i=1}^{m} g_i(a_i \cdot x)$ on $\mathbb{R}^d$, representing a general form of a shallow feed-forward neural network, from a small number of query samples. Higher order differentiation, as used in our constructive approximations, of sums of ridge functions or of their compositions, as in deeper neural network, yields a natural connection between neural network weight identification and tensor product decomposition identification. We prove that in the case of the shallowest feed-forward neural network, second order differentiation and tensors of order two (i.e., matrices) suffice. Based on multiple gathered approximated first and second order differentials, our general approximation strategy is developed as a sequence of algorithms to perform individual sub-tasks. We first perform an active subspace search by approximating the span of the weight vectors $a_1, \dots, a_m$. Then we use a straightforward substitution, which reduces the dimensionality of the problem from d to m. The core of the construction is then the stable and efficient approximation of weights expressed in terms of rank-1 matrices $a_i \otimes a_i$, realized by formulating their individual identification as a suitable nonlinear program. We prove the successful identification by this program of weight vectors being close to orthonormal and we also show how we can constructively reduce to this case by a whitening procedure, without loss of any generality.

3.32 Randomized Smolyak Algorithm: Explicit Cost Bounds and an Application to Infinite-Dimensional Integration

Marcin Wnuk (Universität Kiel, DE)

Joint work of Michael Gnewuch, Marcin Wnuk
Main reference Michael Gnewuch, Marcin Wnuk: "Explicit error bounds for randomized Smolyak algorithms and an application to infinite-dimensional integration", CoRR, Vol. abs/1903/02276, 2019.
URL https://arxiv.org/abs/1903.02276

The Smolyak method is a generic tool to tackle tensor product problems. Let $d \in \mathbb{N}$. Generally, for $n = 1, \dots, d$, separable Hilbert spaces of functions $F^{(n)}$, separable Hilbert spaces $G^{(n)}$, and bounded linear operators $S^{(n)} : F^{(n)} \to G^{(n)}$ are given. The tensor product problem is defined by the solution operator $S_d = \bigotimes_{n=1}^{d} S^{(n)}$, so for $F_d = \bigotimes_{n=1}^{d} F^{(n)}$ and $G_d = \bigotimes_{n=1}^{d} G^{(n)}$ we have

$$S_d : F_d \to G_d.$$

Suppose that for every $n = 1, \dots, d$, one has a sequence of algorithms $(U_j^{(n)})_{j \in \mathbb{N}}$ meant to approximate $S^{(n)}$. Usually, with growing j the algorithms $U_j^{(n)}$ give better approximation, but are at the same time more expensive. The algorithms $(U_j^{(n)})$ are often referred to as building

blocks. Define the algorithm differences $\Delta_j^{(n)} = U_j^{(n)} - -U_{j-1}^{(n)}, j \geq 2$ and $\Delta_1^{(n)} = U_1^{(n)}$. The d-variate Smolyak method of level $L \geq d$ is now given by

$$A(L,d) = \sum_{j \in Q(L,d)} \bigotimes_{n=1}^{d} \Delta_{j_n}^{(n)},$$

where $Q(L,d) = \{j = (j_1, \ldots, j_d) \in \mathbb{N}^d \mid \sum_n j_n \leq L\}$. One speaks of a randomized Smolyak method if the $(U_j^{(n)})_{n,j}$ are randomized algorithms.

In this talk, as an error criterion we consider the randomized error given by

$$e^{ran}(S_d, A(L,d)) = \left[\sup_{f \in F_d, \|f\|_{F_d}=1} \mathbb{E}\|(S_d - -A(L,d))f\|_{G_d}^2 \right]^{\frac{1}{2}},$$

and show under some regularity conditions on the building blocks that if for every $n = 1, \ldots, d$, the convergence rate of $(U_j^{(n)})_j$ is of the order $\mathcal{O}(\frac{1}{N^\alpha})$ then one also has for some positive constants C_0, C_1, not depending on N nor d

$$e^{ran}(A(L,d), S_d) \leq C_0 C_1^d \left(1 + \frac{\log(N)}{d-1}\right)^{(d-1)(\alpha+1)} N^{-\alpha}, \ d \geq 2, . \tag{3}$$

Here N denotes the cardinality of information used by the respective algorithms.

Our interest in the Smolyak method is twofold. Firstly, the upper bound (3) shows that the Smolyak method is quite an efficient generic tool to tackle tensor product problems in moderate dimension d. Secondly, even if d is very large, Smolyak method may be used as a building block of more complicated algorithms.

We illustrate the second statement with an example of inifinite-dimensional integration on weighted function spaces. There one considers input from some Hilbert space

$$H = \bigotimes_{n=1}^{\infty} H_n,$$

where for each $n \in \mathbb{N}$, H_n is a reproducing kernel Hilbert space. Moreover, with growing n the spaces H_n are assigned decreasing weights, meaning basically that even though one is considering as input functions of infinitely many variables, the impact of variables from higher coordinates gets smaller and smaller. Under some technical assumptions in this setting one may define in a sensible way the integral of functions from H. The problem is now to approximate the integral with the help of randomized algorithms. It turns out that our bound (3) in combination with the results of Plaskota and Wasilkowski on multivariate decomposition methods (MDMs) [3] and the embedding results of Gnewuch, Hefter, Hinrichs and Ritter [1] allows us to show that multivariate decomposition methods using our randomized Smolyak algorithms as building blocks achieve optimal convergence rate.

Details and further results may be found in [2].

References

1 Gnewuch Michael, Hefter Mario, Hinrichs Aicke, Ritter Klaus, 'Embeddings of weighted Hilbert spaces and applications to multivariate and infinite-dimensional integration', Journal of Approximation Theory, Volume 222, 2017

2 Gnewuch Michael, Wnuk Marcin 'Explicit error bounds for randomized Smolyak algorithms and application to infinite-dimensional integration', Preprint, ArXiv 1903.02276, 2019

3 Plaskota Leszek, Wasilkowski Grzegorz, 'Tractability of infinite-dimensional integration in the worst case and randomized settings', Journal of Complexity, Volume 27, 2011

3.33 Tractability for Volterra problems with convolution kernels

Henryk Wozniakowski (Columbia University – New York, US)

We show that the information complexity of the Volterra problems considered in this talk is the same (essentially) as the information complexity of multivariate approximation. Therefore the Volterra problems enjoy the same notions of tractability.

We also analyze the combinatory cost of Picard's algorithm for the Volterra problems. The bounds we obtain are not necessarily optimal.

3.34 Approximation complexity for additive random fields

Marguerite Zani (Université d'Orléans, FR)

Joint work of Marguerite Zani, Alexey Khartov, Mikhail A. Lifshits

We study the approximation complexity of additive random fields. For example for $Y_d(t) = \sum_{j=1}^{d} X_j(t)$ where the X_j are uncorrelated, square-integrable, centered random processes of dimension 1.

The complexity $n^{Y_d}(\varepsilon)$ in the average case setting is considered here. We give asymptotics for $\log n^{Y_d}(\varepsilon)$ when ε is fixed and d goes to infinity. We give results in case constant 1 is an eigenfunction of the covariance function associated to Y_d, and if not we can boil down to this situation.

Participants

- Dmitriy Bilyk
University of Minnesota – Minneapolis, US
- James M. Calvin
NJIT – Newark, US
- Ronald Cools
KU Leuven, BE
- Sonja Cox
University of Amsterdam, NL
- Steffen Dereich
Universität Münster, DE
- Benjamin Doerr
Ecole Polytechnique – Palaiseau, FR
- Martin Ehler
Universität Wien, AT
- Michael Giles
University of Oxford, GB
- Michael Gnewuch
Universität Osnabrück, DE
- Takashi Goda
University of Tokyo, JP
- Mario Hefter
TU Kaiserslautern, DE
- Stefan Heinrich
TU Kaiserslautern, DE
- Aicke Hinrichs
Johannes Kepler Universität Linz, AT
- Martin Hutzenthaler
Universität Duisburg-Essen, DE
- Lutz Kämmerer
TU Chemnitz, DE
- Alexander Keller
NVIDIA, DE
- Kristin Kirchner
ETH Zürich, CH
- David Krieg
Johannes Kepler Universität Linz, AT
- Peter Kritzer
österreichische Akadamie der Wissenschaften – Linz, AT
- Robert J. Kunsch
RWTH Aachen, DE
- Frances Y. Kuo
UNSW Sydney, AU
- Gunther Leobacher
Universität Graz, AT
- Thomas Müller-Gronbach
Universität Passau, DE
- Erich Novak
Universität Jena, DE
- Dirk Nuyens
KU Leuven, BE
- Friedrich Pillichshammer
Johannes Kepler Universität Linz, AT
- Leszek Plaskota
University of Warsaw, PL
- Joscha Prochno
Universität Graz, AT
- Pawel Przybylowicz
AGH Univ. of Science & Technology-Krakow, PL
- Klaus Ritter
TU Kaiserslautern, DE
- Daniel Rudolf
Universität Göttingen, DE
- Stefan Steinerberger
Yale University – New Haven, US
- Michaela Szölgyenyi
Alpen-Adria-Universität Klagenfurt, AT
- Aretha Teckentrup
University of Edinburgh, GB
- Vladimir N. Temlyakov
University of South Carolina – Columbia, US
- Mario Ullrich
Johannes Kepler Universität Linz, AT
- Jan Vybíral
Czech Technical University – Prague, CZ
- Marcin Wnuk
Universität Kiel, DE
- Henryk Wozniakowski
Columbia University – New York, US
- Larisa Yaroslavtseva
Universität Passau, DE
- Marguerite Zani
Université d'Orléans, FR