Report from Dagstuhl Seminar 23351

# Algorithms and Complexity for Continuous Problems

# Dmitriy Bilyk<sup>\*1</sup>, Michael Gnewuch<sup>\*2</sup>, Jan Vybíral<sup>\*3</sup>, Larisa Yaroslavtseva<sup>\*4</sup>, and Kumar Harsha<sup>†5</sup>

- $1 \quad University \ of \ Minnesota Minneapolis, \ US. \ dbilyk@math.umn.edu$
- 2 Universität Osnabrück, DE. michael.gnewuch@uni-osnabrueck.de
- 3 Czech Technical University Prague, CZ. jan.vybiral@fjfi.cvut.cz
- 4 Universität Graz, AT. larisa.yaroslavtseva@uni-passau.de
- 5 Universität Osnabrück, DE. kumar.harsha@uos.de

#### — Abstract -

The Dagstuhl Seminar 23351 was held at the Leibniz Center for Informatics, Schloss Dagstuhl, from August 27 to September 1, 2023. This event was the 14th in a series of Dagstuhl Seminars, starting in 1991. During the seminar, researchers presented overview talks, recent research results, work in progress and open problems. The first section of this report describes the goal of the seminar, the main seminar topics, and the general structure of the seminar. The third section contains the abstracts of the talks given during the seminar and the forth section the problems presented at the problem session.

Seminar August 27 - September 1, 2023 - https://www.dagstuhl.de/23351

**2012 ACM Subject Classification** Mathematics of computing  $\rightarrow$  Approximation; Mathematics of computing  $\rightarrow$  Numerical analysis; Mathematics of computing  $\rightarrow$  Probabilistic algorithms

Keywords and phrases computational stochastics, infinite-variate problems, quasi-Monte Carlo, sampling, tractability analysis

Digital Object Identifier 10.4230/DagRep.13.8.106

# 1 Executive Summary

Michael Gnewuch (Universität Osnabrück, DE) Dmitriy Bilyk (University of Minnesota – Minneapolis, US) Jan Vybíral (Czech Technical University – Prague, CZ) Larisa Yaroslavtseva (Universität Graz, AT)

The goal of the Dagstuhl Seminar was to bring together researchers that work in various fields united by a common theme of complexity of continuous problems. In this context, complexity refers to the computational effort required to solve the problem approximately, up to a given error. For these, often high- or even infinite-dimensional, problems it is desirable to have theoretical bounds on the complexity as well as explicit constructions of (nearly) optimal algorithms. Their applications range over natural sciences, economics, statistics, engineering, computer science (including computer graphics, global optimization, and machine learning) and other areas of interest. The focus of the seminar was on the following highly interrelated topics:

Except where otherwise noted, content of this report is licensed

under a Creative Commons BY 4.0 International license

Algorithms and Complexity for Continuous Problems, *Dagstuhl Reports*, Vol. 13, Issue 8, pp. 106–128 Editors: Dmitriy Bilyk, Michael Gnewuch, Jan Vybíral, and Larisa Yaroslavtseva



<sup>\*</sup> Editor / Organizer

<sup>&</sup>lt;sup>†</sup> Editorial Assistant / Collector

#### Dmitriy Bilyk, Michael Gnewuch, Jan Vybíral, and Larisa Yaroslavtseva

**Tractability analysis of high-dimensional problems.** Tractability analysis has its roots in information-based complexity. It studies continuous problems, where the instances are typically *d*-variate functions. The focus is on how the problem complexity depends on the error tolerance and on *d*. Topics discussed during the seminar were the study of recent notions of tractability and of new problem settings.

**Computational stochastics.** The focus was on strong approximation and quadrature of stochastic differential equations (SDEs) with non-globally Lipschitz continuous coefficients. Systematic investigations on the numerical approximation of such SDEs have started in the last decade and are still in their infancy.

**Infinite-variate problems: algorithms and complexity.** Many applications rely on models with countably many variables. The complexity analysis of the resulting continuous problems may be viewed as the limit of tractability analysis for d-variate functions, where d tends to infinity. For many problems sharp complexity bounds have been proved with the help of generic types of algorithms, as multilevel algorithms and multivariate decomposition methods, but mainly in the case where the spaces of input functions are weighted reproducing kernel Hilbert spaces (RKHSs) based on product weights or RKHSs of increasing smoothness. A major question was how to tackle important function spaces, which exhibit a different underlying structure.

Well-distributed point configurations: discrepancy, quasi-Monte Carlo methods, dispersion. Many problems in approximation theory and complexity rely on the existence and construction of good (random or deterministic) point distributions with respect to discrepancy, integration errors, strength of cubature formulas, dispersion, or discrete energy. Numerous questions in this area are still open: in various regimes sharp bounds for discrepancy and dispersion still remain unknown and explicit efficient constructions are still lacking. These problems were intensely discussed during the seminar.

Linear and standard information in approximation theory. A classical problem of approximation theory, learning theory and data analysis is the approximation of an unknown multivariate function from incomplete information. The information can be given by a limited number of function evaluations (standard information) or by evaluating a finite number of arbitrary linear functionals (linear information). Even in the simplest setting of approximation in a norm of a Hilbert space, the relation between using standard information and linear information was understood only recently and there is still a number of challenging open questions in this area. In several seminar talks interesting ideas to tackle these questions and new results were presented.

During the seminar we had five **overview talks** of 60 minutes (incl. discussion; one for each of the main research topics) given by leading experts in the respective area to enable researchers from different fields to follow the regular talks. The speakers and the titles of these lectures were (ordered according to the list of main research topics):

- Peter Kritzer (RICAM Linz), Tractability analysis.
- Thomas Müller-Gronbach (Unversität Passau), On the complexity of strong approximation of SDEs with a non-Lipschitz drift coefficient.
- Klaus Ritter (RPTU Kaiserslautern-Landau), Infinite-variate integration and L<sup>2</sup>-approximation.
- Stefan Steinerberger (University of Washington Seattle), Well-distributed point configurations.
- Mario Ullrich (JKU Linz), On optimal approximation based on random samples.

These talks were delivered at the first three days of the seminar. Additionally, we had 24 **regular talks** of 30 minutes (incl. discussion), many of them presented by young researchers. Due to the time format, there was plenty of time for discussions and collaborations in smaller groups.

Originally we planned an additional poster session for the participants who do not present a talk. Since finally the number of time slots for talks was sufficient to ensure that everyone who wanted to present a talk could actually do so, there was no need for a poster session anymore.

On Wednesday morning we had a **problem session** to share interesting open problems and conjectures and to initiate further discussions. Several researchers used the opportunity to present interesting and challenging open problems. One problem was even solved in a discussion of some participants directly after the problem session. The detailed problem formulations can be found at the end of this report.

# 2 Table of Contents

<b>Executive Summary</b> Michael Gnewuch, Dmitriy Bilyk, Jan Vybíral, and Larisa Yaroslavtseva 106
Overview of Talks
Constructing Low-Discrepancy Point Sets: Subset Sampling and Optimal Sets François Clément
On the existence of optimal shallow networks Steffen Dereich
Data compression using lattices for machine learningKumar Harsha112
Results and open questions around the adaption problem Stefan Heinrich
Instant Neural Graphics Primitives <i>Alexander Keller</i>
Tractability in unweighted function classes David Krieg
Tractability analysis <i>Peter Kritzer</i>
Optimal confidence intervals for randomized quadrature Robert J. Kunsch
High-dimensional approximation: Transforming periodic Approximations vs. Rando Fourier Features Laura Lippert
On the minimal dispersion in the unit cube Alexander Litvak
Fixed-radius spherical cap discrepancy Michelle Mastrianni
Complexity of composite linear problems Peter Mathé
Comparison of Two Search Criteria for Lattice-based Kernel Approximation Weiwen Mo
On the complexity of strong approximation of SDEs with a non-Lipschitz drift coefficient Thomas Müller-Gronbach
Training of DNNs with lattice rules         Dirk Nuyens       117
Super-polynomial Accuracy of Median-of-means Zexin Pan
Sampling recovery in the uniform norm Kateryna Pozharska

23351

#### 110 $\label{eq:23351} \textbf{Algorithms and Complexity for Continuous Problems}$

Randon with Pe Pawel I	nized Euler algorithm for SDEs with drift in integral form and its connection rturbed SGD Przybylowicz
Infinite-	variate Integration and $L^2$ -Approximation
Klaus H	Ritter $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $119$
Integrat <i>Robin F</i>	tion and $L^2$ -Approximation on Gaussian Spaces Rüßmann
Large h <i>Mathias</i>	oles in large point sets Sonnleitner
Multi-le quantifi	evel quasi-Monte Carlo methods for kernel interpolation in uncertainty cation
Abiram	$i Srikumar \ldots 120$
Well-dis	stributed Point Configurations
Stefan S	Steinerberger
point co	ing diversity in StableDiffusion with genetic crossover and well distributed onfigurations Textsaud 120
On anti	$Teglauu  \dots  \dots  \dots  \dots  \dots  \dots  \dots  \dots  \dots $
Mario l	Inal approximation based on random samples <i>Ullrich</i> 121
Samplir Tino Ul	ng numbers of smoothness classes via $l_1$ minimization $llrich \ldots \ldots$
Some C Christia	Computational Approaches to the Pair Correlation Statistic an $Wei\beta$
The fixe Laurence	ed vector randomised lattice algorithm for high-dimensional integration <i>we Wilkes</i>
Randon Marcin	nized approximation of summable sequences: adaptive and non-adaptive $Wnuk$ 123
	1
Open pro	blems
Low dis	crepancy sets in 2D Bilak 123
Standar	rd information vorsus linear information for L approximation
David F	$Krieg \ldots \ldots$
Are line Mathias	ear algorithms non-adaptive?
4 rogula	wr graphs from the way der Corput sequence
Stefan S	Steinerberger
On a pr	coblem of classes of optimal information
Mario <sup>1</sup>	$\mathcal{I}$
Participa	$nts \dots \dots$

#### **3** Overview of Talks

# 3.1 Constructing Low-Discrepancy Point Sets: Subset Sampling and Optimal Sets

François Clément (Sorbonne University – Paris, FR)

License 
Creative Commons BY 4.0 International license
François Clément
Joint work of François Clément, Carola, Doerr, Kathrin Klamroth, Luís Paquete

Low-discrepancy sequences such as Sobol' or Halton have been designed to have small discrepancy values asymptotically. However, for practical applications, only a finite number of points will be possible, often too small to reach the asymptotic regime. In this talk, I present two different approaches to construct point sets with low discrepancy values.

The first is via the Star Discrepancy Subset Selection Problem, which consists in selecting from a fixed point set P of size n the best subset of size k, best being the subset with the smallest  $L_{\infty}$  star discrepancy. We tackle this problem both via exact methods in dimensions 2 and 3 [1] and via heuristics in higher dimensions [2]. In particular, we observe that we are able to obtain sets with 10-50% smaller discrepancy for all tested dimensions and number of points.

The second is the construction of optimal point sets for the  $L_{\infty}$  star discrepancy. This has been a rather overlooked problem, with the main previous work by White in 1977 solving for up to 6 points in dimension 2. We provide non-linear programming formulations to solve the problem for n smaller than 20 in dimension 2 and smaller than 8 in dimension 3 [3]. In particular, the structure of the local discrepancy values over  $[0, 1)^2$  for our optimal sets is very different to that of previously known sets such as Fibonacci or Sobol'.

#### References

- François Clément; Carola Doerr; Luís Paquete, Star discrepancy subset selection: Problem formulation and efficient approaches for low dimensions, Journal of Complexity, 70:101645, 2022
- 2 François Clément; Carola Doerr; Luís Paquete, Heuristic approaches to obtain lowdiscrepancy point sets via subset selection, 2023
- 3 François Clément; Carola Doerr; Kathrin Klamroth, Luís Paquete, Constructing Optimal L<sub>∞</sub> Star Discrepancy Sets, https://arxiv.org/abs/2311.17463, 2024

#### 3.2 On the existence of optimal shallow networks

Steffen Dereich (Universität Münster, DE)

License ☺ Creative Commons BY 4.0 International license © Steffen Dereich Joint work of Steffen Dereic, Arnulf Jentzen, Sebastian Kassing

In this talk, we discuss existence of global minima in optimisation problems over shallow neural networks. More explicitly, the function class over which we minimise is the family of all functions that can be expressed as artificial residual or feedforward neural networks with one hidden layer featuring a specified number of neurons with ReLU (or Leaky ReLU) activation. We give existence results. Moreover, we provide counterexamples that illustrate the relevance of the assumptions imposed in the theorems.

#### 3.3 Data compression using lattices for machine learning

Kumar Harsha (Universität Osnabrück, DE)

License ☺ Creative Commons BY 4.0 International license © Kumar Harsha Joint work of Kumar Harsha, Michael Gnewuch, Marcin Wnuk

The mean squared error is one of the standard loss functions in supervised machine learning. However, calculating this loss for enormous data sets can be computationally demanding. Modifying a earlier approach [1], we present algorithms to reduce extensive data sets to a smaller size using rank-1 lattice rules. With this compression strategy in the precomputation step, every lattice point gets a pair of weights depending on the original data and responses, representing its relative importance. As a result, the compressed data makes iterative loss calculations in optimization steps much faster. The proposed compression strategy is highly beneficial for regression problems without an analytical solution. Our derivation of the formulas for the weights assumes that input functions have a convergent Fourier series and that the relevant Fourier coefficients form a hyperbolic cross. Accordingly, we have analyzed our algorithms' error for functions whose Fourier coefficients decay sufficiently fast such that they lie in Wiener algebras.

#### References

1 Josef Dick; Michael Feischl, A quasi-Monte Carlo data compression algorithm for machine learning, Journal of Complexity, 67:101587, 12 2021

#### 3.4 Results and open questions around the adaption problem

Stefan Heinrich (RPTU – Kaiserslautern, DE)

**License** O Creative Commons BY 4.0 International license O Stefan Heinrich

In the framework of Information-Based Complexity we consider the adaption problem for linear solution operators in the randomized setting. Along with surveying recent advances we discuss a number of remaining/arising/related open questions.

The results presented are contained in the following papers:

#### References

- Stefan Heinrich, Randomized complexity of parametric integration and the role of adaption I. Finite dimensional case, arXiv:2306.13471
- 2 Stefan Heinrich, Randomized Complexity of Parametric Integration and the Role of Adaption II. Sobolev spaces, arXiv:2306.13499
- 3 Stefan Heinrich, Randomized Complexity of Vector-Valued Approximation, to appear in: A. Hinrichs, P. Kritzer, F. Pillichshammer (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2022, Springer-Verlag, see also arxiv:2306.13697

#### 3.5 Instant Neural Graphics Primitives

Alexander Keller (NVIDIA – Berlin, DE)

License 

 Creative Commons BY 4.0 International license
 Alexander Keller

 Joint work of Alexander Keller, Thomas Müller, Alex Evans, Christoph Schied
 Main reference Thomas Müller, Alex Evans, Christoph Schied, Alexander Keller: "Instant neural graphics primitives with a multiresolution hash encoding", ACM Trans. Graph., Vol. 41(4), pp. 102:1–102:15, 2022.
 URL https://doi.org//10.1145/3528223.3530127

Neural graphics primitives, parameterized by fully connected neural networks, can be costly to train and evaluate. We reduce this cost with a versatile new input encoding that permits the use of a smaller network without sacrificing quality, thus significantly reducing the number of floating point and memory access operations: a small neural network is augmented by a multiresolution hash table of trainable feature vectors whose values are optimized through stochastic gradient descent. The multiresolution structure allows the network to disambiguate hash collisions, making for a simple architecture that is trivial to parallelize on modern GPUs.

# 3.6 Tractability in unweighted function classes

David Krieg (Johannes Kepler Universität Linz, AT)

License 

 © Creative Commons BY 4.0 International license
 © David Krieg

 Main reference David Krieg: "Tractability of sampling recovery on unweighted function classes", CoRR, Vol. abs/2304.14169, 2023.
 URL https://doi.org//10.48550/ARXIV.2304.14169

It is well-known that the integration problem as well as the problem of sampling recovery in the  $L_2$ -norm suffer from the curse of dimensionality on many classical function classes. This includes unweighted Korobov spaces (Sobolev spaces with mixed smoothness) as well as classical smoothness classes such as Hölder classes or the classes  $C^k([0, 1]^d)$ . We show that those problems become (polynomially) tractable, even for  $L_p$  approximation with  $1 \le p < \infty$ , if we further impose the condition that the sum of Fourier coefficients is bounded.

The classes where we obtain tractability are unweighted (invariant under a reordering of the variables) and bigger than the corresponding weighted Korobov spaces. Tractability is achieved by the use of non-linear algorithms, while linear algorithms cannot do the job.

In fact, the tractability result is a relatively simple implication of powerful results by Rauhut and Ward [Appl. Comput. Harmon. Anal. 40 (2016), pp. 321-351] on  $\ell_1$ minimization. The approach is not limited to the Fourier system, but it remains to be seen whether any "interesting" classes of nonperiodic functions can be tackled.

#### 3.7 Tractability analysis

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

License 🐵 Creative Commons BY 4.0 International license

© Peter Kritzer

Joint work of Peter Kritzer, Josef Dick, Adrian Ebert, Onyekachi Emenike, Fred J. Hickernell, Christian Irrgeher, Gunther Leobacher, Freidrich Pillichshammer, Henryk Wozniakowski

The information complexity of the problem of approximating an operator  $S_d : \mathcal{F}_d \to \mathcal{G}_d$ for normed spaces  $\mathcal{F}_d$  and  $\mathcal{G}_d$  is defined as the minimal number  $\operatorname{comp}(\varepsilon, d)$  of information evaluations needed to find an  $\varepsilon$ -approximation. A large literature specifies conditions under which the information complexity for a sequence of numerical problems defined for dimensions  $d \in \{1, 2, \ldots\}$  grows at a moderate rate when the error threshold  $\varepsilon$  decreases and/or the dimension increases, i.e., the sequence of problems is tractable.

In this talk, we give an overview of basic ideas in tractability analysis, highlight classical results for linear problems defined on Hilbert spaces, and outline some more recent results on this subject.

#### 3.8 Optimal confidence intervals for randomized quadrature

Robert J. Kunsch (RWTH Aachen, DE & TU Chemnitz, DE)

License O Creative Commons BY 4.0 International license

© Robert J. Kunsch Main reference Robert J. Kunsch: "Linear Monte Carlo quadrature with optimal confidence intervals", CoRR,

Vol. abs/2309.09059, 2023.

URL https://doi.org//10.48550/ARXIV.2309.09059

We study the numerical integration of functions from isotropic Sobolev spaces  $W_p^s([0,1]^d)$ using finitely many function evaluations within randomized algorithms, aiming for the smallest possible probabilistic error guarantee  $\varepsilon > 0$  at confidence level  $1 - \delta \in (0, 1)$ . This error criterion is more challenging than the classical root-mean-squared error which is usually taken for Monte Carlo algorithms within information-based complexity and where only nand  $\varepsilon$  are put in relation. For spaces consisting of continuous functions, non-linear Monte Carlo methods with optimal confidence properties have already been known, in few cases even linear methods that succeed in that regard. [1] In this talk we promote a method called stratified control variates (SCV), see [2], and by it show that already linear methods achieve optimal probabilistic error rates in the high smoothness regime without the need to adjust algorithmic parameters to the uncertainty  $\delta$ . We also analysed a version of SCV in the low smoothness regime where  $W_p^s([0,1]^d)$  may contain functions with singularities. Here, we observe a polynomial dependence of the error on  $\delta^{-1}$  which cannot be avoided for linear methods. This is worse than what is known to be possible using non-linear algorithms where only a logarithmic dependence on  $\delta^{-1}$  occurs if we tune in for a specific value of  $\delta$ . This new way of looking at randomized integration still leaves many questions open, for instance, precise lower bounds for linear methods in the low smoothness regime, as well as universal methods or optimality in mixed-smoothness spaces.

#### References

- 1 Robert J. Kunsch; Daniel Rudolf, Optimal confidence for Monte Carlo integration of smooth functions, Advances in Computational Mathematics, 45:3095–3122, 2019
- 2 Robert J. Kunsch, Linear Monte Carlo quadrature with optimal confidence intervals, arXiv:2309.09059 [math.NA], 2023

# **3.9 High-dimensional approximation: Transforming periodic Approximations vs. Rando Fourier Features**

Laura Lippert (TU Chemnitz, DE)

**License** O Creative Commons BY 4.0 International license O Laura Lippert

We study the problem of scattered-data approximation on  $\mathbb{R}^d$ , where we have given sample points and the corresponding function evaluations. We compare two approaches. In the first one we transform functions on  $\mathbb{R}^d$  to  $\mathbb{T}^d$ , apply an approximation based on hyperbolic wavelet regression and transform the resulting function back. In fact, we transform the sample points to points on  $\mathbb{T}^d$ , evaluate the periodic basis functions at these points, create a matrix and solve the matrix equation with an LSQR-algorithm to get an approximation.

The other approach involves random Fourier features, where we draw frequencies  $\omega_j \in \mathbb{R}^d$  at random and learn coefficients  $a_j$  from the given data to construct the approximant, i.e.

$$f(\cdot) \approx \sum_{j} a_{j} \exp\left(\langle \omega_{j}, \cdot \rangle\right).$$

We give error estimates for both cases, involving different function spaces. We truncate the ANOVA decomposition to approximate high-dimensional functions of low effective dimension.

#### References

- 1 Laura Lippert; Daniel Potts, Variable Transformations in combination with Wavelets and ANOVA for high-dimensional approximation, arXiv:2207.12826, 2022
- 2 Laura Lippert; Daniel Potts; Tino Ullrich, Fast Hyperbolic Wavelet Regression meets ANOVA, Numerische Mathematik 154, 155–207, 2023

#### 3.10 On the minimal dispersion in the unit cube

Alexander Litvak (University of Alberta – Edmonton, CA)

License  $\textcircled{\mbox{\scriptsize \ensuremath{\varpi}}}$ Creative Commons BY 4.0 International license  $\textcircled{\mbox{$\odot$}}$  Alexander Litvak

We discuss ideas which lead to recent improvements of upper bounds for the minimal dispersion of a point set in the unit cube and its inverse. We also discuss sharpness of bounds. The talk is partially based on a joint work with G. Livshyts.

### 3.11 Fixed-radius spherical cap discrepancy

Michelle Mastrianni (University of Minnesota – Minneapolis, US)

License 

 © Creative Commons BY 4.0 International license
 © Michelle Mastrianni

 Joint work of Dmitriy Bilyk, Stefan Steinerberger
 Main reference Dmitriy Bilyk, Michelle Mastrianni, Stefan Steinerberger: "Single radius spherical cap discrepancy via gegenbadly approximable numbers", CoRR, Vol abs/2308.00694, 2023
 URL https://doi.org/10.48550/arXiv.2308.00694

A seminal result of Beck shows that for any set of N points on the d-dimensional sphere, there always exists a spherical cap such that the number of points in the cap deviates from the expected value by at least  $N^{(1/2-1/2d)}$ . We refine the result by removing a layer of

averaging: we show that, when d is not 1 mod 4, there exists a set of real numbers such that for each r > 0 in the set one is always guaranteed to find a spherical cap with radius r for which Beck's result holds. The main ingredient is a generalization of the notion of badly approximable numbers to the setting of Gegenbauer polynomials, which we call Gegenbadly approximable numbers. These are fixed numbers x such that the sequence of Gegenbauer polynomials evaluated at x avoids being close to 0 in a precise quantitative sense.

#### 3.12 Complexity of composite linear problems

Peter Mathé (Weierstraß Institut – Berlin, DE)

```
    License 

            © Creative Commons BY 4.0 International license
            © Peter Mathé

    Joint work of Peter Mathé, Bernd Hofmann

            Main reference
            Bernd Hofmann, Peter Mathé: "The degree of ill-posedness of composite linear ill-posed problems with focus on the impact of the non-compact Hausdorff moment operator", CoRR, Vol abs/2111.01036, 2021.
            URL https://doi.org/10.48550/arXiv.2111.01036
```

We consider compact composite linear operators in Hilbert space, where the composition is given by some compact operator followed by some non-compact one possessing a non-closed range. Focus is on the impact of the non-compact factor on the overall behavior of the decay rates of the singular values of the composition.

# 3.13 Comparison of Two Search Criteria for Lattice-based Kernel Approximation

Weiwen Mo (KU Leuven, BE)

License 
 © Creative Commons BY 4.0 International license
 © Weiwen Mo

 Joint work of Weiwen Mo, Frances Y. Kuo, Dirk Nuyens, Abirami Srikumar, Ian H. Sloan
 Main reference Frances Y. Kuo, Weiwen Mo, Dirk Nuyens, Ian H. Sloan, Abirami Srikumar: "Comparison of Two

Search Criteria for Lattice-based Kernel Approximation", CoRR, Vol. abs/2304.01685, 2023. URL https://doi.org//10.48550/ARXIV.2304.01685

The kernel interpolant in a reproducing kernel Hilbert space is optimal in the worst-case sense among all approximations of a function using the same set of function values. In this talk, we compare two search criteria to construct lattice point sets for use in lattice-based kernel approximation. The first candidate,  $P_n^*$ , is based on the power function that appears in machine learning literature. The second candidate,  $S_n^*$ , is a search criterion used for generating lattices for approximation using truncated Fourier series. We find that the empirical difference in error between the lattices constructed using the two search criteria is marginal. The criterion  $S_n^*$  is preferred as it is computationally more efficient and has a bound with a superior convergence rate.

# 3.14 On the complexity of strong approximation of SDEs with a non-Lipschitz drift coefficient

Thomas Müller-Gronbach (Universität Passau, DE)

We study the complexity of pathwise approximation in p-th mean of the solution of a stochastic differential equation at the final time based on finitely many evaluations of the driving Brownian motion. First, we briefly review the case of equations with globally Lipschitz continuous coefficients, for which an error rate of at least 1/2 in terms of the number of evaluations of the driving Brownian motion is always guaranteed by using the equidistant Euler-Maruyama scheme. Then we illustrate that giving up global Lipschitz continuity may lead to non-polynomial error rates for the Euler-Maruyama scheme or even for any method based on finitely many evaluations of the driving Brownian motion. Finally, we turn to recent complexity results in the case of equations with a drift coefficient that is not globally Lipschitz continuous. Here we focus on scalar equations with a Lipschitz continuous diffusion coefficient and a drift coefficient that satisfies piecewise smoothness assumptions or has fractional Sobolev regularity.

# 3.15 Training of DNNs with lattice rules

Dirk Nuyens (KU Leuven, BE)

License 
Creative Commons BY 4.0 International license
Dirk Nuyens
Joint work of Dirk Nuyens, Alex Keller, Frances Kuo, Ian Sloan

Following Mishra, Rusch (2021) and Longo, Mishra, Rusch, Schwab (2021) we consider the so-called "generalization error" which describes how well a trained DNN can approximate a function on the whole domain. In fact this is just the L2 approximation error if measured in the L2-norm. We start from the assumption that the underlying function is periodic and in any Korobov space with smoothness alpha. This is motivated by the periodic model of the random diffusion field when doing uncertainty quantification.

# 3.16 Super-polynomial Accuracy of Median-of-means

Zexin Pan (Stanford University, US)

```
    License 
        G Creative Commons BY 4.0 International license

            Zexin Pan
            Joint work of Zexin Pan, Art Owen

    Main reference Zexin Pan, Art B. Owen: "Super-polynomial accuracy of multidimensional randomized nets using the median-of-means", CoRR, Vol. abs/2208.05078, 2022.
    URL https://doi.org//10.48550/ARXIV.2208.05078
```

Digital net is an important class of Quasi-Monte Carlo methods used for multidimensional integration. In the talk, I am going to show digital net randomized by linear scrambling and digital shift exhibits surprising concentration behavior. Taking the median of several digital net averages can exclude outliers and boost the convergence rate from nearly cubic to super-polynomial when the integrand is smooth. I will end with some discussions on the difficulties in building confidence intervals.

#### 3.17 Sampling recovery in the uniform norm

Kateryna Pozharska (National Academy of Sciences of Ukraine – Kyiv, UA & TU Chemnitz, DE)

License 
 © Creative Commons BY 4.0 International license
 © Kateryna Pozharska

 Joint work of David Krieg, Kateryna Pozharska, Mario Ullrich, Tino Ullrich
 Main reference David Krieg, Kateryna Pozharska, Mario Ullrich, Tino Ullrich: "Sampling recovery in the uniform norm", CoRR, Vol. abs/2305.07539, 2023.
 URL https://doi.org//10.48550/ARXIV.2305.07539

In the talk, I will present the results of our joint work with David Krieg, Mario Ullrich and Tino Ullrich [1] on the recovery of functions based on function evaluations.

The main emphasis is made on the uniform recovery, however we provide a method to transfer results for  $L_2$ -approximation to rather general seminorms (including  $L_p$ ,  $1 \le p \le \infty$ ). The underlying (least squares) algorithm is based on a random construction and subsampling based on the solution of the Kadison-Singer problem.

Besides an explicit bound for the corresponding sampling widths, we also obtain some interesting inequalities between the sampling, Kolmogorov and Gelfand widths. Namely, we show that there is an absolute constant c > 0, such that for a compact topological space Dand a compact subset F of C(D) the following holds

$$g_{2n}^{\rm lin}(F, L_{\infty}) \leq c \sqrt{n} \, d_n(F, L_{\infty}). \tag{1}$$

The bound in (1) is optimal up to constants, also if we consider only convex and symmetric F and replace the Kolmogorov width  $d_n$  by the Gelfand width  $c_n$  on the right hand side. This means, that there are linear algorithms using 2n samples that are as good as all algorithms using arbitrary linear information up to a factor of at most  $\sqrt{n}$ . A result that cannot be true without oversampling [2]. Moreover, our results imply that linear sampling algorithms are optimal up to a constant factor for many reproducing kernel Hilbert spaces.

**Acknowledgements.** KP would like to acknowledge support by the Philipp Schwartz Fellowship of the Alexander von Humboldt Foundation.

#### References

- 1 David Krieg; Kateryna Pozharska; Mario Ullrich; Tino Ullrich, Sampling recovery in the uniform norm, arXiv: 2305.07539v2, 2023
- 2 Erich Novak, Deterministic and stochastic error bounds in numerical analysis, Lecture Notes in Mathematics 1349, Springer-Verlag, 1988

# 3.18 Randomized Euler algorithm for SDEs with drift in integral form and its connection with Perturbed SGD

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

License 

Creative Commons BY 4.0 International license

Pawel Przybylowicz

We will present a version of randomized Euler algorithm for approximation of solutions of stochastic differential equations when the drift is in integral form. We will show upper estimates for the error and some lower bounds in the worst-case model. We will show some relationship of the randomized Euler algorithm with perturbed SGD used in machine learning.

# 3.19 Infinite-variate Integration and L<sup>2</sup>-Approximation

Klaus Ritter (RPTU – Kaiserslautern, DE)

**License** O Creative Commons BY 4.0 International license O Klaus Ritter

In this survey we consider two basic computational problems, integration and  $L^2$ -approximation, for functions of countably infinite real variables. For motivation we briefly refer to stochastic models based on iid-sequences of random variables, e.g., series expansions of stochastic processes and random fields.

Thereafter we discuss appropriate cost models (information cost) for infinite-variate problems as well as the general structure of the functions spaces (RKHS of tensor product form) that have been studied so far. Finally, we present results (decay of the minimal errors) and key ingredients of the proofs for tensor products of weighted spaces and of spaces of increasing smoothness.

# 3.20 Integration and L<sup>2</sup>-Approximation on Gaussian Spaces

Robin Rüßmann (RPTU – Kaiserslautern, DE)

License ☺ Creative Commons BY 4.0 International license © Robin Rüßmann Joint work of Robin Rüßmann, Michael Gnewuch, Klaus Ritter

We study integration and approximation on reproducing kernel Hilbert spaces with Gaussian kernels of tensor product form. We find new results in the infinite-variate case, and we improve some known results in the finite-variate case.

Our most important tool is a close relation between Gaussian spaces and Hermite spaces of infinite smoothness and tensor product form.

This relation allows us to transform any algorithm for integration or approximation on the Gaussian space into an algorithm for the same problem on the Hermite space and vice versa, preserving error and cost, which means that known upper and lower error bounds for one function space setting also apply to the other setting.

### 3.21 Large holes in large point sets

Mathias Sonnleitner (Universität Passau, DE)

 $\begin{array}{c} \mbox{License} \ensuremath{\,\textcircled{\textcircled{}}}\ensuremath{\,\textcircled{}}\ensuremath{\,ensuremath{\,ensuremath{}}\ensuremath{\,ensuremath{\,ensuremath{}}\ensuremath{\,ensuremath{\,ensuremath{\,ensuremath{}}\ensuremath{\,ensuremath{\,ensuremath{\,ensuremath{}}\ensuremath{\,ens$ 

Dispersion is a measure of equidistribution of point sets. We give a brief introduction and survey known results with focus on a large number of points compared to the dimension. From the literature on coverings of Euclidean space and convex body approximation, we deduce implications for the spherical case and point to open problems.

# 3.22 Multi-level quasi-Monte Carlo methods for kernel interpolation in uncertainty quantification

Abirami Srikumar (UNSW Sydney, AU)

License ☺ Creative Commons BY 4.0 International license ◎ Abirami Srikumar Joint work of Abirami Srikumar, Alexander D. Gilbert, Michael Giles, Frances Y. Kuo, Ian H. Sloan

As high-dimensional problems become increasingly prevalent in many applications, the effective evaluation of these problems within the limits of our current technology poses a great hurdle due to the exponential increase in computational cost as dimensionality increases. One class of strategies for evaluating such problems efficiently are quasi-Monte Carlo (QMC) methods.

In this talk, we explore the effectiveness of multi-level quasi-Monte Carlo methods for approximating solutions to elliptic partial differential equations with stochastic coefficients that depend periodically on the stochastic parameters. In particular, we are interested in fast approximation using kernel-based lattice point interpolation. We present some regularity results, theory on the convergence properties of errors of such approximations and the results of numerical experiments.

# 3.23 Well-distributed Point Configurations

Stefan Steinerberger (University of Washington – Seattle, US)

License  $\textcircled{\mbox{\scriptsize \ensuremath{\varpi}}}$  Creative Commons BY 4.0 International license  $\textcircled{\mbox{$\mathbb{O}$}}$  Stefan Steinerberger

This is a survey talk to describe several different problems in the broad framework of welldistributed point configurations. Problems discussed include: low-discrepancy sets of points, minimizers of energy functionals, greedy sequences, problems on the sphere  $S^2$  and minimizer of the logarithmic energy as well as the connection to the crystallization conjecture.

# 3.24 Improving diversity in StableDiffusion with genetic crossover and well distributed point configurations

Olivier Teytaud (Meta AI Research – Tournon-sur-Rhone, FR)

License 
Creative Commons BY 4.0 International license
Olivier Teytaud
Joint work of Olivier Teytaud, Mariia Zameshina, Laurent Najman, Mathurin Videau

Latent diffusion models excel at producing high-quality images from text. Yet, concerns appear about the lack of diversity in the generated imagery. To tackle this, we introduce Diverse Diffusion, a method for boosting image diversity beyond gender and ethnicity, spanning into richer realms, including color diversity. Diverse Diffusion is a general unsupervised technique that can be applied to existing text-to-image models. Our approach focuses on finding vectors in the Stable Diffusion latent space that are distant from each other. We generate multiple vectors in the latent space until we find a set of vectors that meet the desired distance requirements and the required batch size. To evaluate the effectiveness of our diversity methods, we conduct experiments examining various characteristics, including color

#### Dmitriy Bilyk, Michael Gnewuch, Jan Vybíral, and Larisa Yaroslavtseva

diversity, LPIPS metric, and ethnicity/gender representation in images featuring humans. The results of our experiments emphasize the significance of diversity in generating realistic and varied images, offering valuable insights for improving text-to-image models. Through the enhancement of image diversity, our approach contributes to the creation of more inclusive and representative AI-generated art.

We also present genetic crossovers for combining latent variables into a better latent variable.

After the seminar, we start a cool collaboration with Carola, Stefan, Dmitry, Laurent, for producing fantastic point configurations in high-dimensional point spaces.

Joint work ESIEE and Meta.

#### 3.25 On optimal approximation based on random samples

Mario Ullrich (Johannes Kepler Universität Linz, AT)

License 
Creative Commons BY 4.0 International license
Mario Ullrich

- Joint work of Mario Ullrich, Matthieu Dolbeault, David Krieg, Katerina Pozharska, Mathias Sonnleitner, Tino Ullrich
  - $\label{eq:URL https://arxiv.org/search/?query=Ullrich\%2C+Mario\&searchtype=author\&abstracts=hide&order=-announced_date_first&size=50$

In this talk we give an overview of some recent and not so recent developments in the area of approximation of functions based on function evaluations. The emphasis is on information-based complexity and the worst-case setting, i.e., we ask for the minimal number of information (aka measurements) needed by any algorithm to achieve a prescribed error for all inputs, basically ignoring implementation issues.

However, it turned out that in many cases, certain (unregularized) least squares methods based on "random" information, like function evaluations, can catch up with arbitrary algorithms based on arbitrary linear information, i.e., the best we can do theoretically.

#### 3.26 Sampling numbers of smoothness classes via $l_1$ minimization

Tino Ullrich (TU Chemnitz, DE)

License 

 © Creative Commons BY 4.0 International license
 © Tino Ullrich

 Joint work of Thomas Jahn, Tino Ullrich, Felix Voigtlaender
 Main reference Thomas Jahn, Tino Ullrich, Felix Voigtlaender: "Sampling numbers of smoothness classes via l<sub>1</sub>-minimization", Journal of Complexity, Vol. 79, p. 101786, 2023.
 URL https://doi.org//10.1016/j.jco.2023.101786

Using techniques developed recently in the field of compressed sensing we show new upper bounds for general (nonlinear) sampling numbers of (quasi-)Banach smoothness spaces in  $L_2$ . In particular, we show that in relevant cases such as mixed and isotropic weighted Wiener classes or Sobolev spaces with mixed smoothness, sampling numbers in  $L_2$  can be upper bounded by best n-term trigonometric widths in  $L_{\infty}$ . We describe a recovery procedure from m function values based on  $l_1$ -minimization (basis pursuit denoising). With this method, a significant gain in the rate of convergence compared to recently developed linear recovery methods is achieved. In this deterministic worst-case setting we see an additional speed-up of  $m^{-1/2}$  (up to log factors) compared to linear methods in case of weighted Wiener spaces. For

their quasi-Banach counterparts even arbitrary polynomial speed-up is possible. Surprisingly, our approach allows to recover mixed smoothness Sobolev functions belonging on the *d*-torus with a logarithmically better rate of convergence than any linear method can achieve when 1 and*d*is large. This effect is not present for isotropic Sobolev spaces.

#### 3.27 Some Computational Approaches to the Pair Correlation Statistic

Christian Weiß (Hochschule Ruhr-West - Mülheim, DE)

License  $\textcircled{\textbf{ commons BY 4.0}}$  Creative Commons BY 4.0 International license  $\textcircled{ \ \ }$  Christian Weiß

The pair correlation statistic measures the behavior of gaps between the first elements of a sequence on a local scale. Although a generic independent, identically distributed random sequence drawn from uniform distribution has so-called Poissonian pair correlations, there are only few explicitly known such examples. The main reason why they are difficult to find is that it is in general hard to calculate the pair correlation statistic. In this talk, we therefore concentrate on computational aspects of the pair correlation and present some possible approaches. For instance, geometric properties like the gap structure, an underlying lattice structure of the sequence or the discrepancy turn out to be useful tools. We also present some recent combinatorial results.

# 3.28 The fixed vector randomised lattice algorithm for high-dimensional integration

Laurence Wilkes (KU Leuven, BE)

License 

 © Creative Commons BY 4.0 International license
 © Laurence Wilkes

 Joint work of Frances Y. Kuo, Dirk Nuyens, Laurence Wilkes
 Main reference Frances Y. Kuo, Dirk Nuyens, Laurence Wilkes: "Random-prime-fixed-vector randomised lattice-based algorithm for high-dimensional integration", J. Complex., Vol. 79, p. 101785, 2023.
 URL https://doi.org//10.1016/J.JCO.2023.101785

The fixed vector algorithm offers a very simple solution to the problem of producing a lattice-based randomised algorithm for numerical integration with the optimal randomised error. We shift the construction of the generating vector of the lattice to a precomputation so that the only randomised element is to choose the number of function evaluations from a suitable range. In the talk, we will look at the existence result for such a fixed generating vector, a method to construct the vector and finally look at how the algorithm can be generalised to work in the half-period cosine space and with a lower smoothness parameter.

# 3.29 Randomized approximation of summable sequences: adaptive and non-adaptive

Marcin Wnuk (Universität Osnabrück, DE)

License 
Creative Commons BY 4.0 International license
Marcin Wnuk
Joint work of Marcin Wnuk, Erich Novak, Robert Kunsch

We are considering approximation of the identical embedding of finite-dimensional spaces  $S_M : \ell_1^M \to \ell_\infty^M$  by randomized algorithms. We establish a lower bound on the *n*-th minimal error which is, roughly speaking, of the form: there exists an  $\varepsilon > 0$  such that if for all  $M \in \mathbb{N}$ 

 $e^{\operatorname{ran-non}}(n(M), S_M) < \varepsilon$ 

then

 $n(M) \ge C\sqrt{\log(M)}.$ 

Here  $e^{\operatorname{ran-non}}(n, S)$  denotes the *n*-th minimal error which can be obtained when using randomized non-adaptive algorithms. This has at least two interesting consequences:

- 1. We are able to give an example of a sequence of linear problems for which the *n*-th minimal error for randomized adaptive algorithms decreases much faster than the minimal error for randomized non-adaptive algorithms-a phenomenon which was only recently observed by S.Heinrich in the context of vector-valued mean computation.
- 2. It enables us to show that the only operators which can be arbitrarily well approximated by randomized non-adaptive algorithms are compact operators- i.e. exactly the same operators which can be arbitrarily well approximated by deterministic algorithms.

#### 4 Open problems

#### 4.1 Low discrepancy sets in 2D

Dmitriy Bilyk (University of Minnesota – Minneapolis, US)

License  $\textcircled{\mbox{\scriptsize \ensuremath{\textcircled{} \mbox{\scriptsize \ensuremath{\\} \mbox{\scriptsize \ensuremath{\textcircled{} \mbox{\scriptsize \ensuremath{\\} \mbox{\scriptsize \ensuremath{\\} \mbox{\scriptsize \ensuremath{\textcircled{} \mbox{\scriptsize \ensuremath{\\} \mbox{ \ensuremath{\\} \mbox{\scriptsize \ensuremath{\\} \mbox{\scriptsize \ensuremath{\\} \mbox{\scriptsize \ensuremath{} \mbox{\scriptsize \ensuremath{} \mbox{\scriptsize \ensuremath{\\} \mbox{ \ensuremath{\\} \mbox{\scriptsize \ensuremath{\\} \mbox{\scriptsize \ensuremath{\\} \mbox{\scriptsize \ensuremath{\\} \mbox{\& \ensuremath{\\} \mbox{\& \ensuremath{\\} \mbox{\& \ensuremath{\\} \mbox{\& \ensuremath{\\} \mbox{\& \ensuremath{\\} \mbox{\& \ensuremath{\} \mbox{\& \ensuremath{\\} \mbox{\& \ensuremath{\} \mbox{\& \ensuremath{\} \mbox{\& \ensuremath{\}} \mbox$ 

#### **Problem 1: Permutation sets**

This problem is concerned with new constructions of two-dimensional low discrepancy sets. Two standard examples of sets achieving the lowest possible growth of the discrepancy  $\mathcal{O}(\log N)$  are the van der Corput set and the Fibonacci lattice.

The digit-reversing van der Corput set is a set of  $N = 2^n$  points of the following form

 $(0.x_1x_2\ldots x_n, 0.x_nx_{n-1}\ldots x_1),$ 

where the coordinates are written as binary fractions of length n, and the binary digits  $x_i$  take values zero or one. Many other variations are known (digit shifts and digit scrambling, cyclic shifts, constructions in bases other than two etc).

The Fibonacci lattice consists of  $N = F_n$ , where  $F_n$  is the  $n^{th}$  Fibonacci number, and has the form

$$\left(\frac{k}{F_n}, \left\{\frac{kF_{n-1}}{F_n}\right\}\right), \ k = 0, 1, \dots, F_n - 1.$$

This can be viewed as an approximation of the Kronecker lattice, i.e. points  $(k/N, \{k\alpha\})$ , where  $\alpha$  is the golden ratio. This generalizes to similar constructions using approximants to other badly approximable numbers.

One can easily observe that both of the examples presented above are of the form

$$\left(\frac{k}{N}, \frac{\sigma(k)}{N}\right), \ k = 0, 1, \dots, N-1,$$

where  $\sigma$  is a permutation of an N-elements set  $\{0, 1, \dots, N-1\}$ . Therefore, natural questions arise:

- Which other permutations  $\sigma$  produce low discrepancy sets?
- Which structural and combinatorial properties of the permutation  $\sigma$  are responsible for low discrepancy of the arising point sets?
- Do minimizers of the discrepancy always have a structure of such "permutation sets"? (at least, if restricted to the grid (i/N, j/N)?)
- Are  $L^2$  (or more generally,  $L^p$  discrepancies) minimized by "permutation sets"?
- Closely related question: For a set of points  $z_1, \ldots, z_N \in \mathbb{T}^2$ , where  $z_i = (x_i, y_i)$ , consider the tensor-product interaction energy of the form

$$E(z_1,\ldots,z_N) = \sum_{i,j} f(x_i - x_j) f(y_i - y_j).$$

Under which conditions on the function f are the minimizers of this energy "permutation sets"? And which permutations produce minimizers? (Through Warnock-type formulas, the periodic  $L^2$  discrepancy can be rewritten exactly as such an energy, and other versions of the  $L^2$  discrepancies are closely related to similar energies.)

Are there generalizations of permutation constructions to higher dimensions?

#### Problem 2: Projections of the vertices of the unit cube

The aforementioned van der Corput set can be realized as an image of the  $2^n$  vertices of the *n*-dimensional unit cube  $\{0,1\}^n$  under a linear mapping  $A : \mathbb{R}^n \to \mathbb{R}^2$  given by the matrix

$$A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2^2} & \cdots & \frac{1}{2^n} \\ \frac{1}{2^n} & \frac{1}{2^{n-1}} & \cdots & \frac{1}{2} \end{pmatrix}.$$

- Can other mappings be constructed which similarly produce low discrepancy sets?
- Can probabilistic results be proved for the discrepancy of random projections?
- Can this approach be extended to construct well distributed sets in higher dimensions?

### 4.2 Standard information versus linear information for L<sub>p</sub>-approximation

David Krieg (Johannes Kepler Universität Linz, AT)

License  $\textcircled{\mbox{\footnotesize \mbox{\footnotesize e}}}$  Creative Commons BY 4.0 International license  $\textcircled{\mbox{$ \mbox{$ \odot $} $}}$  David Krieg

Let D be a compact domain of unit volume and let C(D) be the space of continuous complexvalued functions on D. Consider a compact, convex and symmetric subset F of C(D). For  $1 \le p \le \infty$ , we want to compare the rate of convergence of the numbers

$$g_n(F, L_p) = \inf_{\substack{x_1, \dots, x_n \in D\\ \phi \colon \mathbb{C}^n \to L_p(D)}} \sup_{f \in F} \left\| f - \phi(f(x_1), \dots, f(x_n)) \right\|_p$$

and

$$c_n(F, L_p) = \inf_{\substack{L_1, \dots, L_n \in C(D)' \\ \phi : \mathbb{C}^n \to L_p(D)}} \sup_{f \in F} \left\| f - \phi \big( L_1(f), \dots, L_n(f) \big) \right\|_p$$

for  $n \to \infty$ . These numbers describe the minimal worst case error for  $L_p$ -approximation on F if n function values, respectively n arbitrary continuous linear measurements are allowed. The rate of convergence of a non-negative and non-increasing sequence  $(x_n)$ , denoted by  $\deg(x_n)$ , is the supremum of all t > 0 such that  $(x_n) \in \mathcal{O}(n^{-t})$ .

How much can we loose in the rate when restricting from arbitrary linear information to function evaluation? That is, the task is to study the loss function

$$\log_p(\alpha) = \sup_{\substack{D \& F \text{ as above:} \\ \deg(c_n(F,L_p)) = \alpha}} \left[ \deg(c_n(F,L_p)) - \deg(g_n(F,L_p)) \right].$$

How does the loss function behave as a function of  $\alpha > 0$  and  $1 \le p \le \infty$ ?

Results until 2012 are collected in [4, Section 29.1]. The univariate Sobolev spaces  $W_1^s([0,1])$  of integrability one show that  $loss_p(\alpha) \ge 1/2$  for all  $p \ge 2$ , at least if  $\alpha > 1$ . Recent upper bounds imply that in fact  $loss_p(\alpha) = 1/2$  if either p = 2 (see [2] together with [1]) or  $p = \infty$  (see [3]). Is it true that  $loss_p(\alpha) = 1/2$  holds for all  $p \ge 2$  and  $\alpha > 1$ ?

#### References

- J. Creutzig, P. Wojtaszczyk. Linear vs. nonlinear algorithms for linear problems. Journal of Complexity 20 (2004), 807–820.
- 2 M. Dolbeault, D. Krieg, M. Ullrich. A sharp upper bound for sampling numbers in L<sub>2</sub>. Applied and Computational Harmonic Analysis **63** (2023), 113–134.
- 3 D. Krieg, K. Pozharska, M. Ullrich, T. Ullrich. Sampling recovery in the uniform and other norms, preprint, https://arxiv.org/abs/2305.07539.
- 4 E. Novak, H. Woźniakowski. Tractability of multivariate problems. Volume III: Standard information for operators. Volume 18 of *EMS Tracts in Mathematics*. EMS, Zürich, 2012.

#### 4.3 Are linear algorithms non-adaptive?

Mathias Sonnleitner (Universität Passau, DE)

Let F, G be normed spaces (over  $\mathbb{R}$ ) and  $S: F \to G$  be linear and bounded. Suppose that for each  $f \in F$ , we have n pieces of adaptive information

$$N(f) = (L_1(f), L_2(f, L_1(f)), \dots, L_n(f, L_2(f, L_1(f)), \dots)), \quad f \in F$$

based on linear and continuous functionals on F. Let  $A = \varphi \circ N : F \to G$  be a linear algorithm with  $\varphi : \mathbb{R}^n \to G$ .

Can we find  $N^{\text{non}}(f) = (L_1(f), \dots, L_n(f)), f \in F$ , where  $L_1, \dots, L_n : F \to \mathbb{R}$  are linear and (linear)  $\tilde{\varphi} : \mathbb{R}^n \to G$  such that  $A = \tilde{\varphi} \circ N^{\text{non}}$ ?

This basic problem was solved completely using elementary means in a discussion with David Krieg, Robert Kunsch and Marcin Wnuk during the seminar.

#### 4.4 4-regular graphs from the van der Corput sequence

Stefan Steinerberger (University of Washington – Seattle, US)

Consider the following test for pseudorandomness: to any distinct, real numbers  $x_1, \ldots, x_n$ , we associate a 4-regular graph G as follows: using  $\pi$  to denote the permutation ordering the elements,  $x_{\pi(1)} < x_{\pi(2)} < \cdots < x_{\pi(n)}$ , we build a graph on  $\{1, \ldots, n\}$  by connecting i and i + 1 (cyclically) and  $\pi(i)$  and  $\pi(i + 1)$  (cyclically).



**Figure 1** The Graph for the first few digits of  $\sqrt{2}$ : 1, 41, 42, 13, 56, 23, 73. If the digits of  $\sqrt{2}$  behave like truly random numbers, these graphs have optimal expansion properties.

One reason why these graphs are interesting is the following result for the second eigenvalue  $\lambda_2$  of the adjacency matrix of such a graph.

▶ **Theorem 1** (Friedman, Theorem 1.2. in [1]). If  $x_1, \ldots, x_n$  are *i.i.d.* random variables chosen from an absolutely continuous distribution and G is the 4-regular graph constructed from them as above, then for any  $\varepsilon > 0$  there exists c > 0 such that with likelihood at least 1 - c/n, we have

 $|\lambda| \le 2\sqrt{3} + \varepsilon.$ 

As it turns out, this graph construction can indeed be effectively used for understanding the strength of a random number generator: if the numbers are truly random, the arising graphs are random in a "near-"optimal way [3].

One could now wonder what happens to these graphs when the underlying real numbers are far from "random" and two canonical examples are the Kronecker sequence and the van der Corput sequence. Fig. 2 shows the type of graph obtained from the van der Corput sequence in base 2

 $\frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{5}{8}, \frac{3}{8}, \frac{7}{8}, \frac{1}{16}, \dots$ 

We observe the emergence of some strange topological features that are currently unexplained.



**Figure 2** The first 4096 elements of the van der Corput sequence in base 2 (left) and the first  $2187 = 3^7$  elements of the van der Corput sequence in base 3 (right).

**Problem.** What is this emerging manifold? How does it depend on the base b of the van der Corput sequence?

Something similar happens to the Kronecker sequence but there it is perhaps slightly less mysterious (the continued fraction expansion can help to explain it). It has since become clear that this graph structure tends to exhibit all sorts of strange patterns for structured sequences and many examples are given in [2]. Very few of these cases are understood.

#### References

- J. Friedman, A proof of Alon's second eigenvalue conjecture and related problems, Mem. Amer. Math. Soc. 195 (2008), no. 910
- 2 D. Korssjoen, B. Li, S. Steinerberger, R. Tripathi, R. Zhang, Finding Structure in Sequences of Real Numbers via Graph Theory: a Problem List, Involve 15, 251-270 (2022).
- 3 S. Steinerberger, Using Expander Graphs to test whether samples are i.i.d, arXiv:2008.01153

# 4.5 On a problem of classes of optimal information

Mario Ullrich (Johannes Kepler Universität Linz, AT)

License 
 Creative Commons BY 4.0 International license
 Image: Mario Ullrich

 Joint work of Mario Ullrich, Mathias Sonnleitner
 Main reference Mathias Sonnleitner, Mario Ullrich: "On the power of iid information for linear approximation", CoRR, Vol. abs/2310.12740, 2023.

URL https://doi.org//10.48550/ARXIV.2310.12740

New open problem arise after recent results on the power of standard information (function values), and more general classes  $\Lambda$ , indicate that already under mild assumptions on  $\Lambda$  we might hope for optimal approximations, i.e., that the corresponding minimal errors are of the same order as minimal errors based on arbitrary information. Several of related open problems can be found in the survey paper [1], which was nearly finished during the Dagstuhl Seminar 23351. A (simple, and too optimistic) version of an open problem was presented at the seminar, and already disproved in a subsequent discussion with David Krieg and Erich Novak.

#### References

1 Mathias Sonnleitner, Mario Ullrich. On the power of iid information for linear approximation, arXiv:2310.12740, 2023.



 Dmitriy Bilyk University of Minnesota -Minneapolis, US François Clément Sorbonne University – Paris, FR Ronald Cools KU Leuven, BE Steffen Dereich Universität Münster, DE Benjamin Doerr Ecole Polytechnique -Palaiseau, FR Carola Doerr CNRS, Sorbonne University -Paris, FR Michael Giles University of Oxford, GB Michael Gnewuch Universität Osnabrück, DE Kumar Harsha Universität Osnabrück, DE Stefan Heinrich RPTU - Kaiserslautern, DE Alexander Keller NVIDIA, DE David Krieg Johannes Kepler Universität Linz, AT Peter Kritzer Österreichische Akademie der Wissenschaften - Linz, AT Thomas Kühn Universität Leipzig, DE

Robert J. Kunsch RWTH Aachen, DE & TU Chemnitz, DE Frances Y. Kuo UNSW Sydney, AU Laura Lippert TU Chemnitz, DE Alexander Litvak University of Alberta -Edmonton, CA Michelle Mastrianni University of Minnesota -Minneapolis, US Peter Mathé Weierstraß Institut - Berlin, DE Weiwen Mo KU Leuven, BE Thomas Müller-Gronbach Universität Passau, DE Erich Novak Friedrich-Schiller-Universität Jena, DE Dirk Nuyens KU Leuven, BE Zexin Pan Stanford University, US Kateryna Pozharska National Academy of Sciences of Ukraine – Kyiv, UA & TU Chemnitz, DE Pawel Przybylowicz AGH Univ. of Science & Technology-Krakow, PL

Klaus Ritter
 RPTU – Kaiserslautern, DE
 Daniel Rudolf
 Universität Passau, DE
 Robin Rüßmann

RPTU – Kaiserslautern, DE Mathias Sonnleitner

Universität Passau, DE

Abirami Srikumar UNSW Sydney, AU

Stefan Steinerberger
 University of Washington –
 Seattle, US

Olivier Teytaud
 Meta AI Research –
 Tournon-sur-Rhone, FR

Mario Ullrich Johannes Kepler Universität Linz, AT

Tino Ullrich TU Chemnitz, DE

Jan Vybíral
 Czech Technical University –
 Prague, CZ

Christian Weiß
 Hochschule Ruhr-West –
 Mülheim, DE

Laurence Wilkes KU Leuven, BE

Marcin Wnuk
 Universität Osnabrück, DE

