Mathematisches Forschungsinstitut Oberwolfach

Report No. 49/2013

DOI: 10.4171/OWR/2013/49

Uniform Distribution Theory and Applications

Organised by
Michael Gnewuch, Kaiserslautern
Frances Y. Kuo, Sydney
Harald Niederreiter, Linz/Dhahran
Henryk Woźniakowski, New York/Warszawa

29 September – 5 October 2013

Abstract. The topics of the workshop were recent progress in the theory of uniform distribution theory (also known as discrepancy theory) and new developments in its applications in analysis, approximation theory, computer science, numerics, pseudo-randomness and stochastic simulation.

Mathematics Subject Classification (2000): 11K38, 65D30, 43A99, 41A65, 60G15.

Introduction by the Organisers

Uniform distribution or *discrepancy theory* studies distribution irregularities of point sets and sequences in Euclidean spaces and on manifolds. It relies on methods from number theory and analysis as well as on methods from geometry, probability theory and combinatorics. It has a wide range of applications as, e.g., in computational finance, computational geometry, numerical analysis, pseudo-random number generation, and stochastic simulation. An important application is the evaluation of multivariate or infinite-dimensional integrals. Integration methods based on the theory of uniform distribution are known as quasi-Monte Carlo (QMC) methods.

The workshop was organized by Michael Gnewuch (Kaiserslautern), Frances Y. Kuo (Sydney), Harald Niederreiter (Linz/Dhahran), and Henryk Woźniakowski (New York/Warszawa), and was attended by 51 participants from many different areas of mathematics and computer science, many of them PhD students or young postdocs. The program consisted of five survey lectures of 60 minutes and 30 talks of 30 minutes. The main topics of the workshop were

- *Star discrepancy and small ball conjecture.* The "great open problem of discrepancy theory", the asymptotic behavior of the star discrepancy (the L_∞-norm of the discrepancy function), is intimately related to the small ball conjecture and the Kolmogorov entropy of mixed derivative Sobolev spaces. Recent progress has been made in the understanding of the great open problem by improved lower bounds and by new results on the behavior of different norms of the discrepancy function. The survey lecture on this topic was delivered by Dmitriy Bilyk (Minneapolis).
- *Construction of low-discrepancy points and sequences.* Special constructions of QMC points are very useful not only in numerical integration, but, e.g., also for function approximation, for integral equations or for experimental design. Recently a lot of progress has been made in the (fast and efficient) construction of lattice rules, polynomial lattice rules, nets of higher order convergence, point sets on the sphere and other sets. The survey lecture on this topic was delivered by Josef Dick (Sydney).
- *Tractability of multivariate integration.* High-dimensional approximation problems are getting more and more important in mathematics and in applications. Tractability studies high-dimensional problems and their explicit dependence on the dimension d. In this field many new developments took place over the last few years and QMC methods were often used to establish new results, especially in multivariate integration. The survey lecture on this topic was delivered by Friedrich Pillichshammer (Linz).
- *Discrepancy, minimal energy points, and integration on the sphere.* Good sample points on the sphere are important in several applications. In the last two years important theoretical results were discovered (as, e.g., the solution of the conjecture of Korevaar and Meyers on spherical t-designs) and several point constructions have been found to perform well in numerical tests. The survey lecture on this topic was delivered by Peter Grabner (Graz).
- *Multilevel algorithms for infinite-dimensional integration and SDEs.* In the last few years multilevel algorithms turned out to be efficient tools for multivariate and infinite-dimensional integration, and for the simulation of stochastic differential equations. Nowadays there is a huge number of research articles that employ multilevel algorithms, e.g., in option pricing, engineering or information-based complexity. Recently it has been shown that in many problem settings the combination of multilevel methods and QMC point sets yields optimal algorithms. The survey lecture on this topic was delivered by Klaus Ritter (Kaiserslautern).

The atmosphere of the workshop was lively and active. The excellent facilities at Oberwolfach and the long afternoon breaks were extensively used by the participants to discuss and work together. The organizers would like to use the opportunity to thank the Mathematisches Forschungsinstitut Oberwolfach for the hospitality and to acknowledge the support of several young researchers through the Oberwolfach Leibniz Graduate Students Program.

Workshop: Uniform Distribution Theory and Applications

Table of Contents

Christoph Aistleitner
On measures of pseudorandomness for binary sequences 2843

Dmitriy Bilyk
Small ball inequalities and discrepancy 2844

Johann S. Brauchart (joint with Lou Fang, Josef Dick)
Stolarsky's invariance principle and discrepancy of point sets on the sphere 2845

Steffen Dereich (joint with Michael Scheutzow, Reik Schottstedt)
Approximation by empirical measures 2848

Josef Dick (joint with Fritz Pillichshammer)
Construction of low-discrepancy sequences 2848

Benjamin Doerr
A lower bound for the star discrepancy of a random point set 2850

Carola Doerr (joint with François-Michel De Rainville)
Constructing low star discrepancy point sets with genetic algorithms . . . 2851

Henri Faure
On the evolution of upper bounds for low-discrepancy sequences 2852

Takashi Goda (joint with Josef Dick)
Construction of interlaced scrambled polynomial lattice rules 2854

Peter Grabner
Discrepancy, minimal energy points, and integration on the sphere 2856

Mario Hefter (joint with Klaus Ritter)
Embeddings of weighted tensor product spaces 2859

Stefan Heinrich (joint with Thomas Daun)
Complexity of parametric integration in various smoothness classes 2862

Roswitha Hofer (joint with Harald Niederreiter)
New ways for constructing generating matrices of digital sequences 2865

Peter Kritzer (joint with Henri Faure)
Star discrepancy bounds for (t, m, s)-nets and (t, s)-sequences 2866

Thomas Kühn
Approximation numbers of Sobolev embeddings – Sharp constants and tractability 2868

Gerhard Larcher
Discrepancy estimates for sequences: New results and open problems ... 2870

Pierre L'Ecuyer (joint with Richard Simard)
The poor lattice structure of certain long-period multiple recursive generators ... 2871

Gunther Leobacher
On SDEs with discontinuous drift ... 2872

Sebastian Mayer (joint with Tino Ullrich, Jan Vybíral)
Tractability results for classes of ridge functions ... 2874

Dirk Nuyens (joint with Josef Dick, Frances Y. Kuo, Quoc T. Le Gia, Christoph Schwab)
Construction of interlaced polynomial lattice rules with SPOD weights .. 2876

Friedrich Pillichshammer
Tractability of multivariate integration ... 2878

Leszek Plaskota
Automatic integration using asymptotically optimal subdivision strategy . 2881

Daniel Potts
Multivariate approximation with trigonometric polynomials ... 2883

Klaus Ritter
Multi-level Monte Carlo algorithms ... 2885

Daniel Rudolf (joint with J. Dick)
Discrepancy bounds of Markov chain quasi-Monte Carlo ... 2885

Ian H. Sloan (joint with Peter Grabner)
Lower bounds and separation for spherical designs ... 2887

Stefan Steinerberger
Three problems for discrepancy and QMC ... 2890

Vladimir Temlyakov
On covering in finite dimensional Banach spaces ... 2892

Shu Tezuka
Recent results on $(t, \mathbf{e}, s)$-sequences ... 2895

Robert Tichy
Limit Theorems for Discrepancy Functions ... 2897

Maryna Viazovska (joint with Andriy Bondarenko, Danylo Radchenko)
Well-separated spherical designs ... 2899

Grzegorz W. Wasilkowski
Tractability of approximation of ∞-variate functions with bounded mixed partial derivatives ... 2901

Markus Weimar
On lower bounds for integration of multivariate permutation-invariant functions 2904

Arne Winterhof
A survey on recursive nonlinear pseudorandom number generators 2906

Robert S. Womersley
Well distributed points on the sphere – What do you want? 2909

Abstracts

On measures of pseudorandomness for binary sequences

Christoph Aistleitner

In 1997 Mauduit and Sárközy [7] started a series of papers in which they introduced three measures of pseudorandomness for finite binary sequences, which have attracted large interest since then. These three measures of pseudorandomness are the well-distribution measure $W(E_N)$, the normality measure $\mathcal{N}(E_N)$, and the correlation measure of order k, $C_k(E_N)$. Let $E_N = (e_1, \ldots, e_N) \in \{-1, 1\}^N$ denote a finite binary sequence. Then the definitions of the three measures of pseudorandomness are given as follows.

For $M \in \mathbb{N},\ a \in \mathbb{Z}$ and $b \in \mathbb{N}$ set

$$U(E_N, M, a, b) = \sum \{e_{a+jb} :\ 1 \le j \le M,\ 1 + a + jb \le N \text{ for all } j\}.$$

In words: $U(E_N, M, a, b)$ is the discrepancy of E_N with respect to a certain arithmetic progression in $\{1, \ldots, N\}$. Then the **well-distribution measure** $W(E_N)$ is given by

$$W_N(E_N) := \max\{|U(E_N, M, a, b)|,\ \text{where } 1 \le a + b \text{ and } a + Mb \le N\}.$$

For $k \in \mathbb{N},\ M \in \mathbb{N}$ and $X \in \{-1, 1\}^k$ let

$$T(E_N, M, X) = \#\{n :\ 0 \le n < M,\ n + k \le N,\ (e_{n+1}, \ldots, e_{n+k}) = X\}.$$

In words: $T(E_N, M, X)$ counts the number of occurrences of a pattern X among an initial part of E_N. Then the **normality measure** $\mathcal{N}(E_N)$ is given by

$$\mathcal{N}(E_N) := \max_k \max_X \max_M \left| T(E_N, M, X) - \frac{M}{2^k} \right|,$$

where the maxima are taken over $k \le \log_2 N,\ X \in \{-1, 1\}^k$ and $0 < M \le N+1-k$.

Furthermore, for $k \in \mathbb{N},\ M \in \mathbb{N}$ and $D = (d_1, \ldots, d_k) \in \mathbb{N}^k$ with $0 \le d_1 < \cdots < d_k < N$ let

$$V(E_N, M, D) = \sum \{e_{n+d_1} \ldots e_{n+d_k} :\ 1 \le n \le M,\ n + d_k \le N\}.$$

In words: $V(E_N, M, D)$ quantifies the correlation among k elements of E_N, which are relatively positioned according to D. Then the **correlation measure of order k,** $C_k(E_N)$, is given by

$$C_k(E_N) := \max\{|V(E_N, M, D)| :\ M, D \text{ satisfy } M + d_k \le N\}.$$

The minimal and typical values of these measures of pseudorandomness have been studied intensively by Alon, Kohayakawa, Mauduit, Moreira and Rödl [4, 5] in two papers. Here "minimal" means the smallest possible value of the measure of pseudorandomness for a given value of N, and "typical" means the behavior of the measure of pseudorandomness for an i.i.d. random binary sequence E_N.

They proved many results concerning these problems, but left several others open. The solutions of some of those are discussed in the present talk. In particular, the existence of a limit distribution of the normalized measures $W(E_N)/\sqrt{N}$ and $\mathcal{N}(E_N)/\sqrt{N}$ for random E_N and $N \to \infty$ is proved, and it is shown that a strongly improved upper bound for the minimal value of $\mathcal{N}(E_N)$ can be deduced from a construction of a normal number with small discrepancy due to Levin [6].

The talk is based on the speaker's papers [1, 2, 3].

References

[1] C. Aistleitner, *On the limit distribution of the well-distribution measure of random binary sequences*, J. Theor. Nombres Bordeaux, to appear.

[2] C. Aistleitner, *On the limit distribution of the normality measure of random binary sequences*, preprint, available at http://arxiv.org/abs/1301.6454.

[3] C. Aistleitner, *Normal numbers and the normality measure*, Combin. Probab. Comput. **22** (2013), 342–345.

[4] N. Alon, Y. Kohayakawa, C. Mauduit, C. G. Moreira, and V. Rödl, *Measures of pseudorandomness for finite sequences: minimal values*, Combin. Probab. Comput. **15** (2006), 1–29.

[5] N. Alon, Y. Kohayakawa, C. Mauduit, C. G. Moreira, and V. Rödl, *Measures of pseudorandomness for finite sequences: typical values*, Proc. Lond. Math. Soc. (3) **95** (2007), 778–812.

[6] M. B. Levin, *On the discrepancy estimate of normal numbers*, Acta Arith. 88 (1999), 99–111.

[7] C. Mauduit and A. Sárközy, *On finite pseudorandom binary sequences. I. Measure of pseudorandomness, the Legendre symbol*, Acta Arith. **82** (1997), 365–377.

Small ball inequalities and discrepancy

Dmitriy Bilyk

We discuss the connections between the so-called "small ball inequality" in harmonic analysis, uniform lower bounds for the star-discrepancy, and some problems in probability theory and approximation. The small ball conjecture deals with hyperbolic sums of multiparameter Haar functions; it speculates that

$$n^{\frac{d-2}{2}} \left\| \sum_{|R|=2^{-n}} \alpha_R h_R \right\| \gtrsim 2^{-n} \sum_{|R|=2^{-n}} |\alpha_R|,$$

where h_R are L^∞-normalized d-dimensional Haar functions adapted to dyadic rectangles $R \subset [0,1]^d$ and α_R are arbitrary real coefficients. While this problem is solved in dimension $d = 2$, it is wide open in higher dimensions (the first progress in this direction was achieved by the author with Lacey and Vagharshakyan). This conjecture (which is known to be sharp) arises naturally in problems of probability theory (small deviation asymptotics of the Brownian sheet) and approximation (entropy numbers for classes of multivariate functions with mixed derivative in L^2). The connections of the small ball conjecture to discrepancy theory is less direct. It comes from Roth's idea that the behavior of the discrepancy function is essentially defined by the part of its Haar expansion corresponding to rectangles

of volume $|R| \approx \frac{1}{N}$, where N is the number of points. The *signed* version of the small ball conjecture

$$\left\| \sum_{|R|=2^{-n}} \alpha_R h_R \right\| \gtrsim n^{\frac{d}{2}}, \quad \alpha_R = \pm 1$$

bears a very strong resemblance to the conjectured lower bound for the star-discrepancy

$$\|D_N\|_\infty \gtrsim (\log N)^{\frac{d}{2}}.$$

This inequality is also known in dimension $d = 2$ and most proof techniques work equally well in both settings. The sharpness of the small ball conjecture suggests that $d/2$ may perhaps be the correct exponent in the discrepancy bound.

Stolarsky's invariance principle and discrepancy of point sets on the sphere

Johann S. Brauchart
(joint work with Lou Fang, Josef Dick)

Let $\mathbb{S}^d$ be the unit sphere in the Euclidean space $\mathbb{R}^{d+1}$ endowed with the normalized surface area measure σ_d. *Stolarsky's invariance principle* (cf. [11]) states that the sum of all distances determined by N points on $\mathbb{S}^d$ plus their discrepancy does not depend on the choice of the points. In fact, this principle encodes much more: **(I)** In the form of

$$\left[D^C_{\mathbb{L}_2}(X_N)\right]^2 = C_d \left[\int_{\mathbb{S}^d} \int_{\mathbb{S}^d} |\mathbf{x}-\mathbf{y}| \, \mathrm{d}\,\sigma_d(\mathbf{x}) \, \mathrm{d}\,\sigma_d(\mathbf{y}) - \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N |\mathbf{x}_j - \mathbf{x}_k| \right] \tag{1}$$

it provides a convenient way of computing the spherical cap $\mathbb{L}_2$-discrepancy,

$$D^C_{\mathbb{L}_2}(X_N) := \left(\int_0^\pi \int_{\mathbb{S}^d} \left| \frac{|X_N \cap C(\mathbf{z};\theta)|}{N} - \sigma_d(C(\mathbf{z};\theta)) \right|^2 \mathrm{d}\,\sigma_d(\mathbf{z}) \, \sin\theta \, \mathrm{d}\,\theta \right)^{1/2},$$

of an N-point set $X_N = \{\mathbf{x}_1, \dots, \mathbf{x}_N\} \subset \mathbb{S}^d$. (Here, $C(\mathbf{z};\theta)$ is a spherical cap with center $\mathbf{z} \in \mathbb{S}^d$ that contains all points $\mathbf{x} \in \mathbb{S}^d$ with inner product $\mathbf{x}\cdot\mathbf{z} > \cos\theta$.) **(II)** Stolarsky's invariance principle is deeply rooted in the theory of error estimates for Quasi Monte Carlo numerical integration methods (QMC methods) that utilize reproducing kernel Hilbert space (RKHS) techniques; namely, the right-hand side of (1) can be interpreted as the *worst-case numerical integration error* of QMC methods for functions from the unit ball in a certain Sobolev space $\mathbb{H}^s(\mathbb{S}^d)$ with smoothness index $s = (d+1)/2$ (whereas $\mathbb{H}^0(\mathbb{S}^d) = \mathbb{L}_2(\mathbb{S}^d, \sigma_d)$) that becomes a RKHS when endowed with the distance kernel $1 - C_d\,|\mathbf{x}-\mathbf{y}|$; see [5] and below. **(III.a)** The study of the sum of distances is closely related to the potential-theoretical regime of the *discrete minimal Riesz-τ energy problem* (cf. [10]) and, in particular, to the investigation of the leading term(s) of the asymptotics of the Riesz-τ energy for $\tau > -2$ (cf. [6] and also [2]). Observe that the square-bracketed

expression in (1) singles out the second term in the asymptotic expansion of the maximum sum of distances when $N \to \infty$. **(III.b)** The double integral in (1) is a potential-theoretical quantity; that is, it is the maximum of the distance integral $\mathcal{I}[\mu] := \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} |\mathbf{x}-\mathbf{y}| \, \mathrm{d}\,\mu(\mathbf{x}) \, \mathrm{d}\,\mu(\mathbf{y})$, where μ is a Borel probability measure supported on $\mathbb{S}^d$. The surface area measure σ_d is its unique maximizer.

The power and elegance of the RKHS approach is demonstrated in [5], wherein it is shown that Stolarsky's invariance principle emerges in a natural way when studying the error of numerical integration on the sphere. Indeed, the symmetric and positive definite integral kernel

$$K(\mathbf{x},\mathbf{y}) := \int_0^\pi \int_{\mathbb{S}^d} \chi_{C(\mathbf{z};\theta)}(\mathbf{x}) \chi_{C(\mathbf{z};\theta)}(\mathbf{y}) \; \mathrm{d}\,\sigma_d(\mathbf{z}) \, \sin\theta \; \mathrm{d}\,\theta, \tag{2}$$

integrating a product of indicator functions for the spherical cap $C(\mathbf{z};\theta)$, uniquely defines a RKHS $\mathcal{H}$ consisting of functions $f : \mathbb{S}^d \to \mathbb{R}$ of the form

$$f(\mathbf{x}) = \int_0^\pi \int_{\mathbb{S}^d} g(\mathbf{z};\theta) \, \chi_{C(\mathbf{x};\theta)}(\mathbf{z}) \; \mathrm{d}\,\sigma_d(\mathbf{z}) \, \sin\theta \; \mathrm{d}\,\theta \tag{3}$$

for some potential function $g \in \mathbb{L}_2(\mathbb{S}^d \times [0,\pi], \sigma_d)$, provided with the inner product

$$(f_1, f_2)_K := \int_0^\pi \int_{\mathbb{S}^d} g_1(\mathbf{z};\theta) \, g_2(\mathbf{z};\theta) \; \mathrm{d}\,\sigma_d(\mathbf{z}) \, \sin\theta \; \mathrm{d}\,\theta. \tag{4}$$

The error of approximating the integral of a function $f \in \mathcal{H}$ using a QMC method with node set $X_N = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset \mathbb{S}^d$,

$$\frac{1}{N} \sum_{j=1}^N f(\mathbf{x}_j) - \int_{\mathbb{S}^d} f(\mathbf{x}) \; \mathrm{d}\,\sigma_d(\mathbf{x}) = (f, \mathcal{R}(\mathcal{H}, Q[X_N]; \mathbf{x})_K)_K,$$

is the inner product of f with the "representer"

$$\mathcal{R}(\mathcal{H}, Q[X_N]; \mathbf{x})_K := \int_0^\pi \int_{\mathbb{S}^d} \Delta_{X_N}(\mathbf{z};\theta) \, \chi_{C(\mathbf{x};\theta)}(\mathbf{z}) \; \mathrm{d}\,\sigma_d(\mathbf{z}) \, \sin\theta \; \mathrm{d}\,\theta,$$

which is of the form (3) and has as potential function the *local discrepancy function*

$$\Delta_{X_N}(\mathbf{z};\theta) := \frac{1}{N} \sum_{j=1}^N \chi_{C(\mathbf{x}_j;\theta)}(\mathbf{z}) - \sigma_d(C(\mathbf{z};\theta)).$$

A standard argument involving the application of the Cauchy Schwarz inequality yields that the worst-case error squared is the inner product of the representer with itself. It gives the left-hand side of (1) when using (4) and, when expanded, reduces to the worst-case error formula

$$\left[\mathrm{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d))\right]^2 = \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N K(\mathbf{x}_j, \mathbf{x}_k) - \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} K(\mathbf{x},\mathbf{y}) \, \mathrm{d}\,\sigma_d(\mathbf{x}) \; \mathrm{d}\,\sigma_d(\mathbf{y}).$$

Direct computation of the integral (2) yields a closed form representation (distance kernel) and analysis of its Laplace-Fourier expansion allows to identify the corresponding Sobolev space with appropriate norm. Substitution of the distance kernel into the worst-case error formula gives the right-hand side of (1).

In [4] this approach is extended by replacing the indicator function of an spherical cap with the truncated power function $\mathbf{x} \mapsto (\mathbf{x}\cdot\mathbf{z}-\cos\theta)_+^{\beta-1}$. The new kernel

$$\mathcal{K}_\beta(\mathbf{x},\mathbf{y}) := \int_0^\pi \int_{\mathbb{S}^d} (\mathbf{x}\cdot\mathbf{z}-\cos\theta)_+^{\beta-1}\,(\mathbf{y}\cdot\mathbf{z}-\cos\theta)_+^{\beta-1}\,\mathrm{d}\,\sigma_d(\mathbf{z})\sin\theta\,\mathrm{d}\,\theta \tag{5}$$

can be expressed as a sum of a constant multiple of a (signed) $(2\beta-1)$-power of the Euclidean distance and a special case of a Kampé de Fériet function. The latter reduces to a polynomial if β is an integer. The analysis of its Laplace-Fourier expansion yields that the RKHS uniquely defined by K_β coincides with $\mathbb{H}^s(\mathbb{S}^d)$ with $s=\beta+(d-1)/2$. A generalized Stolarsky's invariance principle allows one to compute the $\mathbb{L}_2$-discrepancy of the following local discrepancy function

$$\Delta_{\mathcal{P}_N,\beta}(\mathbf{z},t) := \frac{1}{N}\sum_{j=1}^N (\mathbf{x}_j\cdot\mathbf{z}-t)_+^{\beta-1} - \int_{\mathbb{S}^d}(\mathbf{y}\cdot\mathbf{z}-t)_+^{\beta-1}\,\mathrm{d}\sigma_d(\mathbf{y}), \quad \mathbf{z}\in\mathbb{S}^d, t\in[-1,1].$$

A *spatial extension* of Stolarsky's invariance principle into the Euclidean space $\mathbb{R}^{d+1}$ utilizes truncated spherical cones (anchored at infinity) as test sets. It arises in the study of numerical integration of functions defined on $\mathbb{R}^{d+1}$, $d\geq 2$; that is, $\int_{\mathbb{R}^{d+1}} f(\mathbf{x})\psi(\mathbf{x})\,\mathrm{d}\,\lambda_{d+1}(\mathbf{x})$, where ψ is a probability density function (typically a normal or related distribution) and λ_{d+1} is the Lebesgue measure on $\mathbb{R}^{d+1}$ ([8]).

A point set with small spherical cap $\mathbb{L}_2$-discrepancy will have large sum of distances and vice versa. It has been noted in [9] that despite extensive optimization the putative minimizers of the spherical cap $\mathbb{L}_2$-discrepancy are in very good agreement with the conjectured asymptotic behavior given in [6]. Numerical results for constructed point sets on the sphere (spherical digital nets and Fibonacci points) are conjectured to have optimal behavior ([1, 3]). It should be remarked that a sequence of low-discrepancy point sets on $\mathbb{S}^d$ is "almost" a QMC design sequence for $\mathbb{H}^s(\mathbb{S}^d)$ with $s\in(d/2,(d+1)/2)$; cf. [7].

References

[1] C. Aistleitner, J. S. Brauchart, and J. Dick. Point sets on the sphere $\mathbb{S}^2$ with small spherical cap discrepancy. *Discrete Comput. Geom.*, 48(4):990–1024, 2012.

[2] J. S. Brauchart. Optimal discrete Riesz energy and discrepancy. *Unif. Distrib. Theory*, 6(2):207–220, 2011.

[3] J. S. Brauchart and J. Dick. Quasi-Monte Carlo rules for numerical integration over the unit sphere $\mathbb{S}^2$. *Numer. Math.*, 121(3):473–502, 2012.

[4] J. S. Brauchart and J. Dick. A Characterization of Sobolev Spaces on the Sphere and an Extension of Stolarsky's Invariance Principle to Arbitrary Smoothness. *Constr. Approx.*, 3:1–49, 2013.

[5] J. S. Brauchart and J. Dick. A simple proof of Stolarsky's invariance principle. *Proc. Amer. Math. Soc.*, 141(6):2085–2096, 2013.

[6] J. S. Brauchart, D. P. Hardin, and E. B. Saff. The next-order term for optimal Riesz and logarithmic energy asymptotics on the sphere. In *Recent advances in orthogonal polynomials, special functions, and their applications*, volume 578 of *Contemp. Math.*, pages 31–61. Amer. Math. Soc., Providence, RI, 2012.

[7] J. S. Brauchart, E. B. Saff, I. H. Sloan, and R. S. Womersley. QMC designs: Optimal order Quasi Monte Carlo integration schemes on the sphere. arXiv:1208.3267v1 [math.NA] and in press: Math. Comp., 2013.

[8] L. Fang, J. S. Brauchart, and J. Dick. Spatial low-discrepancy sequences, spherical cone discrepancy, and applications in financial modelling. in preperation.
[9] M. Gräf. *Efficient algorithms for the computation of optimal quadrature points on Riemannian manifolds*. Dissertation, Chemnitz University of Technology, 2013.
[10] D. P. Hardin and E. B. Saff. Discretizing manifolds via minimum energy points. *Notices Amer. Math. Soc.*, 51(10):1186–1194, 2004.
[11] K. B. Stolarsky. Sums of distances between points on a sphere. II. *Proc. Amer. Math. Soc.*, 41:575–582, 1973.

Approximation by empirical measures

Steffen Dereich
(joint work with Michael Scheutzow, Reik Schottstedt)

We analyse approximation of probability measures by empirical measures. Two asymptotic estimates are provided for the Wasserstein L^p-distance: a Pierce-type estimate and a high-resolution formula.

References

[1] S. Dereich, M. Scheutzow, R. Schottstedt, *Constructive quantization: approximation by empirical measures*, arXiv:1108.5346, preprint 2011. (http://arxiv.org/abs/1108.5346).

Construction of low-discrepancy sequences

Josef Dick
(joint work with Fritz Pillichshammer)

We consider equidistribution properties of point sets and sequences in the s-dimensional unit-cube $[0,1)^s$ measured by their L_q discrepancy. For a finite set $P_{N,s} = \{\boldsymbol{x}_0, \dots, \boldsymbol{x}_{N-1}\}$ the local discrepancy function is defined as

$$\Delta_{P_{N,s}}(\boldsymbol{t}) = \frac{A_N([\boldsymbol{0}, \boldsymbol{t}), P_{N,s})}{N} - t_1 t_2 \cdots t_s,$$

where $\boldsymbol{t} = (t_1, t_2, \dots, t_s) \in [0,1]^s$ and $A_N([\boldsymbol{0}, \boldsymbol{t}), P_{N,s})$ denotes the number of indices n with $\boldsymbol{x}_n \in [0, t_1) \times \cdots \times [0, t_s) =: [\boldsymbol{0}, \boldsymbol{t})$. The local discrepancy function measures the difference between the portion of points in an axis parallel box containing the origin and the volume of this box. Hence it measures the irregularity of distribution of a point set in $[0,1)^s$.

Let $q \in [1, \infty]$. The L_q discrepancy of $P_{N,s}$ is defined as the L_q-norm of the local discrepancy function

$$L_{q,N}(P_{N,s}) = \|\Delta_{P_{N,s}}\|_{L_q} = \left(\int_{[0,1]^s} |\Delta_{P_{N,s}}(\boldsymbol{t})|^q \, \mathrm{d}\boldsymbol{t} \right)^{1/q} \tag{1}$$

with the obvious modifications for $q = \infty$. For an infinite sequence S_s in $[0,1)^s$ the L_q discrepancy $L_{q,N}(S_s)$ is the L_q discrepancy of the first N points of S_s.

A sequence is uniformly distributed modulo one if and only if its L_q discrepancy tends to zero for growing N. Furthermore, the L_q discrepancy can also be linked to the integration error of a quasi-Monte Carlo rule, see, e.g. [1, 5, 7].

One of the questions on irregularities of distribution is concerned with the precise order of convergence of the smallest possible values of the L_q discrepancy as N goes to infinity. While this problem is completely solved for $q \in (1, \infty)$, the case $q \in \{1, \infty\}$ appears to be much more difficult. In particular, for $q = \infty$ and $s \geq 3$ the exact asymptotic order of the smallest possible value of the L_∞ discrepancy is still unknown (for $s = 2$ it is known to be $(\log N)/N$). There are many people who conjecture that the sharp order of magnitude for the smallest possible value of the L_∞ discrepancy of N-element point sets in $[0, 1)^s$ is $(\log N)^{s-1}/N$. But there are also other opinions such as, for example, $(\log N)^{s/2}/N$, see [2]. Although there is some recent remarkable progress, which is surveyed in [2], the exact determination of the sharp order of magnitude for the smallest possible value of L_∞ discrepancy remains unknown.

We survey the development of the problem beginning with Roth's seminal lower bound on the L_2 discrepancy from 1954 to the recent constructions of point sets and sequences with optimal order of L_2 discrepancy, on which we put our main focus. In detail, we discuss Roth's lower bound for the L_2 discrepancy of finite point sets and its extensions to L_q discrepancy and to infinite sequences. We discuss existence results for point sets and sequences whose orders of magnitude of the L_2 discrepancy match the lower bounds. To do so, we introduce digital nets and sequences which provide the basic ideas for the explicit constructions. Walsh functions provide the main analytical tool to obtain discrepancy bounds. The ideas for the explicit constructions are motivated by the particular structure of the Walsh coefficients of the local discrepancy function. We introduce the explicit constructions of point sets with optimal L_2 discrepancy by Chen and Skriganov [3] and the explicit constructions of point sets and sequences with optimal L_2 discrepancy of [6]. We describe the extensions of the latter results to the L_q discrepancy by [8] and [4]. We briefly discuss recent extensions of the L_q discrepancy results to exponential Orlicz norms.

Finally, we conjecture that the constructions in [4] and [6] achieve the optimal order of the L_∞ discrepancy.

References

[1] J. Beck and W. W. L. Chen, *Irregularities of Distribution*, Cambridge University Press, Cambridge, 1987.

[2] D. Bilyk and M. Lacey, *The Supremum Norm of the Discrepancy Function: Recent Results and Connections*, In: *Monte Carlo and Quasi-Monte Carlo Methods 2012*, J. Dick, F. Y. Kuo, G. W. Peters and I. H. Sloan (eds.), to appear, Springer, Berlin Heidelberg New York, 2014.

[3] W. W. L. Chen and M. M. Skriganov, *Explicit constructions in the classical mean squares problem in irregularity of point distribution*, J. Reine Angew. Math. **545** (2002), 67–95.

[4] J. Dick, *Discrepancy bounds for infinite-dimensional order two digital sequences over* $\mathbb{F}_2$, submitted for publication.

[5] J. Dick and F. Pillichshammer, *Digital Nets and Sequences. Discrepancy Theory and Quasi-Monte Carlo Integration*, Cambridge University Press, Cambridge, 2010.
[6] J. Dick and F. Pillichshammer, *Optimal $\mathcal{L}_2$ discrepancy bounds for higher order digital sequences over the finite field $\mathbb{F}_2$*, to appear in Acta Arith., 2014.
[7] M. Drmota and R. F. Tichy, *Sequences, discrepancies and applications*, Lecture Notes in Mathematics 1651, Springer Verlag, Berlin, 1997.
[8] M. M. Skriganov, *Harmonic analysis on totally disconnected groups and irregularities of point distributions*, J. Reine Angew. Math. **600** (2006), 25–49.

A lower bound for the star discrepancy of a random point set

Benjamin Doerr

The main objective of this talk is to show that the star discrepancy of N random points in the s-dimensional unit cube is of order $\Omega(\sqrt{s/N})$, matching the upper bound given by Heinrich, Novak, Wasilkowski, and Woźniakowski [5].

Let $N, s \in \mathbb{N}$. Let $P \subseteq [0,1]^s$ with $|P| = N$. For $x \in [0,1]^s$, let us call the set $[0,x] := \prod_{i=1}^s [0,x_i]$ a *box*, and denote by $\mathcal{B} := \{[0,x] \mid x \in [0,1]^s\}$ the set of all these boxes. Denoting the Lebesgue measure of a measurable set B by $\lambda(B)$, the *star discrepancy* of P now is defined by

$$D^*(P) := \sup_{B \in \mathcal{B}} \left| \tfrac{1}{N} |P \cap B| - \lambda(B) \right|.$$

While the classic view on low-discrepancy point sets is to regard the asymptotics in N assuming the dimension s to be fixed (see, e.g., Niederreiter [6]), more recently it was observed that taking the dimension s as a constant (and thus also treating terms like 2^s as constant in asymptotic statements) can give misleading results in those practical applications where s is large. Heinrich, Novak, Wasilkowski, and Woźniakowski [5] started the quest for bounds and construction that have an explicit, and ideally polynomial, dependence on s. Among other results, they show that the minimal star discrepancy of an N-point set in the s-dimensional unit cube is $O(\sqrt{s/N})$. This bound is witnessed by a random point set already. See [1] for an elementary proof of these results that also gives good values for the implicit constants. See [3] for a recent survey on discrepancy results with explicit dependence on the dimension s.

No matching lower bounds for the minimal star discrepancy are known, the best one is $\Omega(s/N)$ by Hinrichs [4] (of course assuming $s = O(N)$). Closing this gap is one of the big open problems in this field.

Surprisingly, not even a lower bound for the discrepancy of a random point set is known. This is the objective of this talk, where we give a simple proof that the upper bounds given in [1, 5] are asymptotically tight.

Theorem 1. *There is an absolute constant K such that the following is true. Let $N, s \in \mathbb{N}$ such that $s \leq N$. Let P be a set of N points chosen independently and uniformly at random from $[0,1]^s$. Then the expected star discrepancy satisfies $E[D^*(P)] \geq K\sqrt{s/N}$. The probability that $D^*(P)$ is less than $K\sqrt{s/N}$, is at most $\exp(-\Theta(s))$.*

A proof of this result can be found in [2].

References

[1] C. Aistleitner and M. Hofer, *Probabilistic discrepancy bounds for Monte Carlo point sets*, Math. Comp., to appear. DOI 10.1090/S0025-5718-2013-02773-1

[2] B. Doerr, *A lower bound for the discrepancy of a random point set*, J. Complexity, to appear. DOI 10.1016/j.jco.2013.06.001.

[3] M. Gnewuch, *Entropy, randomization, derandomization, and discrepancy.* In L. Plaskota and H. Woźniakowski, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2010*, volume 23 of *Proceedings in Mathematics and Statistics* (2012), 43–78. Springer.

[4] A. Hinrichs, *Covering numbers, Vapnik-Červonenkis classes and bounds for the star-discrepancy*, J. Complexity **20** (2004), 477–483.

[5] S. Heinrich, E. Novak, G. W. Wasilkowski, and H. Woźniakowski, *The inverse of the star-discrepancy depends linearly on the dimension*, Acta Arith. **96** (2001), 279–302.

[6] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*, volume 63 of *CBMS-NSF Regional Conference Series in Applied Mathematics* (1992). Society for Industrial and Applied Mathematics (SIAM).

Constructing low star discrepancy point sets with genetic algorithms

Carola Doerr

(joint work with François-Michel De Rainville)

Point sets of small star discrepancy are relevant in numerous applications. While there has been substantial work on the development of low star discrepancy *point sequences* with good asymptotic behavior, not much is known about the efficient construction of low star discrepancy *point sets* for fixed parameter settings; i.e., we do not know currently how to construct for fixed dimension d and a fixed maximal number of points n a point set $X \subseteq [0,1)^d$ of size $|X| \leq n$ such that $\mathrm{disc}^*_\infty(X)$ is as small as possible. In this talk we present and discuss a new genetic algorithm that addresses this discrepancy optimization problem.

The algorithm itself is fairly simple. Still—even without any parameter tuning—it is able to compute point sets of much lower star discrepancy value than previous approaches found in the literature. In fact, in 61 out of 62 test cases our algorithm computes better point sets. They exhibit, on average, a star discrepancy value that is 31 percent smaller.

The genetic algorithm can easily be adapted to optimize *inverse star discrepancies*; i.e., for a given dimension d and a given discrepancy threshold ε it tries to find the smallest possible n such that there exists a point set X of size n and discrepancy at most ε.

An extended abstract describing this work can be found in [1]. Our algorithm is a recombination of the algorithms presented in [2] and [3]. The former was originally developed to optimize L_2 discrepancies. It serves in our genetic algorithm as a module for creating new solution candidates. The evaluation of the generated point sets is done either by the exact algorithm by Dobkin, Eppstein, and Mitchel [4] or (when the exact method is infeasible) by the threshold accepting algorithm from [3].

References

[1] C. Doerr, F.-M. De Rainville, *Constructing Low Star Discrepancy Point Sets with Genetic Algorithms*, Proc. of Genetic and Evolutionary Computation Conference (GECCO 2013), 789–796.

[2] F.-M. De Rainville, C. Gagné, O. Teytaud, and D. Laurendeau, *Optimizing low-discrepancy sequences with an evolutionary algorithm*, ACM Trans. Model. Comput. Simul. **22** (2012), 1–9:25.

[3] M. Gnewuch, M. Wahlström, and C. Winzen, *A new randomized algorithm to approximate the star discrepancy based on threshold accepting*, SIAM J. Numer. Anal. **50** (2012), 781–807.

[4] D. P. Dobkin, D. Eppstein, and D. P. Mitchell, *Computing the discrepancy with applications to supersampling patterns*, ACM Trans. Graph. **15** (1996), 354–376.

On the evolution of upper bounds for low-discrepancy sequences

Henri Faure

The two well-known families of multi-dimensional low-discrepancy sequences for which explicit upper bounds exist are the Halton sequences [7] and the so-called (t, s)-sequences introduced by Sobol' (in base 2) [14], Faure (in prime bases) [2] and generalized by Niederreiter in arbitrary bases [10].

Upper bounds for general (t, s)-sequences have been established by Niederreiter in 1987 [10] and remained unchanged for about twenty years, while in the meantime many constructions have been proposed to improve their asymptotic behavior involving the quality parameter t (see among others Niederreiter and Xing [11] and Niederreiter and Özbudak [12]). The next improvement was provided by Kritzer in 2006 [9], who was able to reduce by a factor of about two the leading constant c_s obtained in [10]. In the same paper, Kritzer also states a conjecture for (t, s)-sequences in even bases that would substantially improve his own bounds.

Recently, there has been a renewed interest in Halton sequences due to an important improvement discovered by Atanassov in 2004 [1]. His results provided a drastic change in our understanding of how bounds on the discrepancy vary with the dimension s for these sequences, obtaining a factor $1/s!$ in previous bounds by Faure in 1982 [2]. Moreover, using linear digit scramblings, Atanassov has been able to modify Halton sequences in such a way that their asymptotic behavior is the same as Niederreiter–Xing sequences. A careful comparison of these improved bounds through the leading constant c_s, for both Halton sequences and (t, s)-sequences, can be found in [4, Section 2.3].

Next, in 2010, Faure, Lemieux, and Wang [6] proposed extensions of the methods of Atanassov for Halton sequences to derive bounds for (t, s)-sequences by an approach completely different from the classical way used by Sobol', Faure, Niederreiter and Kritzer, even though this approach could not reach the currently best constants c_s obtained by Kritzer in 2006.

From that time onwards, results have been appearing at an increasing pace:

– First, using their extension of Atanassov's method for (t, s)-sequences, Faure and Lemieux [5] (2012) got very close to the conjecture of Kritzer for even bases, improving for instance by a factor $3/2^{s-1}$ the constant c_s in base 2.
– Then, going deeply in the classical method and thanks to two clever counting lemmas of Kritzer, Faure and Kritzer (2013) [3] proved his conjecture for the constant c_s in even bases. Together with the bound for odd bases in 2006 from [9], this result gives the currently best known constants c_s for general (t, s)-sequences.
– Almost at the same time, Tezuka (2013) [15] proposed an extension of (t, s)-sequences to a larger family he called $(t, \mathbf{e}, s)$-sequences where $\mathbf{e} = (e_1, \ldots, e_s)$ is a s-tuple of positive integers (in this framework, (t, s)-sequences are obtained with the s-tuple $\mathbf{e} = (1, \ldots, 1)$). It is quite remarkable that an adaptation of Atanassov's method also applies to $(t, \mathbf{e}, s)$-sequences and gives a new type of upper bounds quite different from usual ones, see [15, Theorem 2 and Corollary 1]. This extension is also the starting point of new constructions of (t, s)-sequences using global functions fields by Hofer and Niederreiter [8] and Niederreiter and Yeo [13]. This way, the constants from [3] can be further improved with special (t, s)-sequences, at least for sufficiently large s (see [8, 13, Section 5]).
– At present, using a new approach for the study of sequences in even bases, Faure and Lemieux are able to improve the previous explicit upper bounds in any base from [5], hence getting a better behavior in the non-asymptotic regime, still with the same leading constant c_s.
– And finally, thanks to an adaptation of this approach to $(t, \mathbf{e}, s)$-sequences, we are also able to improve Tezuka's bound in [15, Theorem 2] in the case of an even base. We will report on these latest developments in our proposed talk.

References

[1] E. I. Atanassov, On the discrepancy of the Halton sequences, *Math. Balkanica, New Series* **18.1-2** (2004), 15–32.
[2] H. Faure, Discrépance de suites associées à un système de numération (en dimension s), *Acta Arith.* **61** (1982), 337–351.
[3] H. Faure and P. Kritzer, New star discrepancy bounds for (t, m, s)-nets and (t, s)-sequences, *Monatsh. Math.*, to appear.
[4] H. Faure and C. Lemieux, Generalized Halton sequences in 2008: A comparative study, *ACM Trans. Model. Comp. Sim.*, **19**, 4 (2009), Article 15.
[5] H. Faure and C. Lemieux, Improvements on the star discrepancy of (t, s)-sequences, *Acta Arith.* **154.1** (2012), 61–78.
[6] H. Faure, C. Lemieux, X. Wang, Extensions of Atanassov's methods for Halton sequences, *Monte Carlo and Quasi-Monte Carlo Methods 2010*, L. Plaskota and H. Woźniakowski (Eds.), Springer, (2012), 345–362.
[7] J. H. Halton, On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals, *Numer. Math.* **2** (1960), 184–90.
[8] R. Hofer and H. Niederreiter, A construction of (t, s)-sequences with finite row generating matrices using global function fields, *Finite Fields Appl.*, **21** (2013), 97–110.
[9] P. Kritzer, Improved upper bounds on the star discrepancy of (t, m, s)-nets and (t, s)-sequences, *J. Complexity,* **22** (2006), 336–347.
[10] H. Niederreiter, Point sets and sequences with small discrepancy, *Monats. Math.*, **104** (1987), 273–337.

[11] H. Niederreiter and C. P. Xing, Quasirandom points and global functions fields. In: S. Cohen and H. Niederreiter (Eds) Finite Fields and Applications, London Math. Soc. Lectures Notes Series, **233** (1996), 269–296.
[12] H. Niederreiter and F. Özbudak, Low-discrepancy sequences using duality and global function fields, *Acta Arith.* **130** (2007), 79–97.
[13] H. Niederreiter and S. J. A. Yeo, Halton-Type Sequences From Global Function Fields, *Science China* **56-1**, to appear.
[14] I. M. Sobol', The distribution of points in a cube and the approximate evaluation of integrals, *U.S.S.R. Comput. Math. and Math. Phys.* **7** (1967), 86–112.
[15] S. Tezuka, On the discrepancy of generalized Niederreiter sequences, *J. Complexity*, in press.

Construction of interlaced scrambled polynomial lattice rules

Takashi Goda
(joint work with Josef Dick)

Higher order scrambled digital nets are randomized quasi-Monte Carlo (RQMC) rules which have been introduced in [1] to approximate the integral of smooth functions defined over the s-dimensional unite cube. It has been shown that higher order scrambled digital nets achieve the optimal rate of convergence of the root mean square error for numerical integration of smooth functions. In this talk, we attempt to replace the randomly scrambled digital nets by randomly scrambled polynomial lattice point sets, which allows us to obtain a better dependence on the dimension while still achieving the optimal rate of convergence. We call our RQMC rules *interlaced scrambled polynomial lattice rules* whose definition will be introduced in the following. For a prime b, we denote by $\mathbb{F}_b$ the finite field with b elements and by $\mathbb{F}_b[x]$ the set of polynomials over $\mathbb{F}_b$.

Let $p \in \mathbb{F}_b[x]$ be irreducible with $\deg(p) = m \in \mathbb{N}$ and let $\boldsymbol{q} = (q_1, \ldots, q_{ds}) \in (\mathbb{F}_b[x])^{ds}$. An interlaced scrambled polynomial lattice point set consisting of b^m points $\{\boldsymbol{x}_n : n \in \mathbb{F}_b[x], 0 \le \deg(n) < m\}$ is constructed by

$$\boldsymbol{x}_n = \mathcal{D}_d\left(\Pi_1\left(v_m\left(\frac{nq_1}{p}\right)\right), \ldots, \Pi_{ds}\left(v_m\left(\frac{nq_{ds}}{p}\right)\right)\right) \in [0,1)^s,$$

where v_m, $\Pi_1, \ldots, \Pi_{ds}$ and $\mathcal{D}_d$ are defined as follows:

- v_m denotes the mapping from the field of the formal Laurent series over $\mathbb{F}_b$ to the interval $[0,1)$ by

$$v_m\left(\sum_{l=t}^{\infty} \kappa_l x^{-l}\right) = \sum_{l=\max(1,t)}^{m} \kappa_l b^{-l},$$

  where t is an arbitraty integer and $\kappa_l \in \mathbb{F}_b$ for all l.
- For $1 \le j \le ds$, Π_j denotes a randomly chosen set of permutations for scrambling, that is,

$$\Pi_j = \{\pi_{j,\xi_{j,1},\ldots,\xi_{j,k-1}} : k \in \mathbb{N}, \xi_{j,1}, \ldots, \xi_{j,k-1} \in \mathbb{F}_b\},$$

where $\pi_{j,\xi_{j,1},\ldots,\xi_{j,k-1}}$ is a random permutation of $\{0,\ldots,b-1\}$. Let $x_j \in [0,1)$ with its b-adic expansion

$$x_j = \frac{\xi_{j,1}}{b} + \frac{\xi_{j,2}}{b^2} + \cdots .$$

Applying Π_j to x_j, we obtain $y_j \in [0,1)$ such that

$$y_j = \Pi_j(x_j) = \frac{\pi_j(\xi_{j,1})}{b} + \frac{\pi_{j,\xi_{j,1}}(\xi_{j,2})}{b^2} + \cdots .$$

- $\mathcal{D}_d$ denotes the mapping from $[0,1)^{ds}$ to $[0,1)^s$ by digitally interlacing every d components. Let $\boldsymbol{x} \in [0,1)^{ds}$ with $\boldsymbol{x} = (x_1,\ldots,x_{ds})$ and consider the b-adic expansion of each coordinate

$$x_j = \frac{\xi_{j,1}}{b} + \frac{\xi_{j,2}}{b^2} + \cdots .$$

Applying $\mathcal{D}_d$ to $\boldsymbol{x}$, we obtain $\boldsymbol{y} = (y_1,\ldots,y_s) \in [0,1)^s$, where for $1 \le j \le s$

$$y_j = \sum_{a=1}^{\infty} \sum_{r=1}^{d} \xi_{(j-1)d+r,a} b^{-r-(a-1)d}.$$

We consider weighted function spaces with general weights $(\gamma_u)_{u\subseteq\{1,\ldots,s\}}$, whose elements have square integrable partial mixed derivatives of order up to $\alpha \in \mathbb{N}$ in each variable, and derive an upper bound on the variance of the estimator for interlaced scrambled polynomial lattice rules, which is given as

$$\mathrm{Var}[\hat{I}(f)] \le V_{\alpha,\boldsymbol{\gamma}}^2(f) B_{\alpha,d,\boldsymbol{\gamma}}(\boldsymbol{q},p),$$

where $V_{\alpha,\boldsymbol{\gamma}}(f)$ is a weighted Hardy and Krause variation of f of order α (see [1, Subsection 3.2]), and $B_{\alpha,d,\boldsymbol{\gamma}}(\boldsymbol{q},p)$ is a function which depends only on the interlaced scrambled polynomial lattice point sets but does not depend on f. Moreover, we can show that there is a concise formula for $B_{\alpha,d,\boldsymbol{\gamma}}(\boldsymbol{q},p)$, so that $B_{\alpha,d,\boldsymbol{\gamma}}(\boldsymbol{q},p)$ can be used as a quality criterion for searching good sets of polynomials $\boldsymbol{q} = (q_1,\ldots,q_{ds})$.

Employing $B_{\alpha,d,\boldsymbol{\gamma}}(\boldsymbol{q},p)$ as a quality criterion, the component-by-component (CBC) construction can be used to obtain explicit constructions of finding good sets of polynomials. The CBC construction proceeds as follows:

(1) Choose an irreducible polynomial $p \in \mathbb{F}_b[x]$ with $\deg(p) = m$.
(2) Set $q_1 = 1$.
(3) For $\tau = 2,\ldots,ds$, find q_τ by minimizing $B_{\alpha,d,\boldsymbol{\gamma}}((q_1,\ldots,q_{\tau-1},\tilde{q}_\tau),p)$ as a function of $\tilde{q}_\tau \in \mathbb{F}_b[x]$ such that $\deg(\tilde{q}_\tau) < m$ and $\tilde{q}_\tau \ne 0$.

For p and $\boldsymbol{q}$ found by this algorithm, we obtain the bound on $B_{\alpha,d,\boldsymbol{\gamma}}(\boldsymbol{q},p)$ as

$$B_{\alpha,d,\boldsymbol{\gamma}}(\boldsymbol{q},p) \le \frac{1}{(b^m-1)^{1/\lambda}} \left[\sum_{\emptyset\ne u\subseteq\{1,\ldots,s\}} \gamma_u^\lambda C_{\alpha,d,\lambda}^{|u|} \right]^{\frac{1}{\lambda}},$$

for any $1/(2\min(\alpha,d)+1) < \lambda \le 1$, where $C_{\alpha,d,\lambda}$ is a constant independent of m and s. By choosing the interlacing factor $d \ge \alpha$, the bound implies a convergence rate of the variance $\mathrm{Var}[\hat{I}(f)]$ of order $N^{-2\alpha-1+\delta}$ for any $\delta > 0$.

We finally discuss a dependence of the variance $\mathrm{Var}[\hat{I}(f)]$ on the dimension s by adding one more notation

$$T_{\lambda,a} := \limsup_{s\to\infty} \left[\frac{1}{s^a} \sum_{\emptyset\neq u\subseteq\{1,\ldots,s\}} \gamma_u^\lambda C_{\alpha,d,\lambda}^{|u|} \right].$$

From the above bound on $B_{\alpha,d,\boldsymbol{\gamma}}(\boldsymbol{q},p)$, it is obvious to have the following facts:

(1) Assume that $T_{\lambda,0} < \infty$ for some $1/(2\min(\alpha,d)+1) < \lambda \leq 1$. Then $\mathrm{Var}[\hat{I}(f)]$ is bounded independently of the dimension.
(2) Assume that $T_{\lambda,a} < \infty$ for some $1/(2\min(\alpha,d)+1) < \lambda \leq 1$ and $a > 0$. Then the bound of $\mathrm{Var}[\hat{I}(f)]$ depends at most polynomially on the dimension with its degree a/λ.

For more information on this work, we refer to [2].

References

[1] J. Dick, *Higher order scrambled digital nets achieve the optimal rate of the root mean square error for smooth integrands*, Ann. Statist. **39** (2011), 1372–1398.
[2] T. Goda, J. Dick, *Construction of interlaced scrambled polynomial lattice rules of arbitrary high order*, preprint. Available at `http://arxiv.org/abs/1301.6441v2.`

Discrepancy, minimal energy points, and integration on the sphere

Peter Grabner

Different types of constructions have been used to find "good" configurations of N points $X_N = \{x_1,\ldots,x_N\}$ on a manifold M, especially the sphere $\mathbb{S}^d$. Of course the construction depends on what quantitative measure is used for the configuration. Several such measures have been used and shall be discussed in the sequel:

Discrepancy. Discrepancy given by

$$D(X_N) = \sup_{C\in\mathcal{C}} \left| \frac{1}{N}\sum_{n=1}^{N} \chi_C(x_n) - \sigma(C) \right|, \tag{1}$$

is an easy to understand concept; $D(X_N)$ just measures the maximal deviation of the discrete distribution from the limiting distribution σ using a set $\mathcal{C}$ of "test sets" (in statistics this is called the Kolmogorov-Smirnov statistics). On the other hand, the precise value of the discrepancy of a point set is rather difficult to compute. Thus discrepancy is usually estimated rather than computed directly. In the simplest one-dimensional case there are two classical estimates for discrepancy, namely the Erdős-Turán inequality and LeVeque's inequality (cf. [10]). Both inequalities have been generalised to the spherical case and used for estimating the discrepancy of point sets constructed by the various methods.

The $\mathbb{S}^d$ version of the Erdős-Turán inequality has been given indepently by the author [6] and Li and Vaaler [12] and reads as

$$D(X_N) \leq \frac{C_1(d)}{M} + \sum_{\ell=1}^{M} \frac{C_2(d)}{\ell} \sum_{m=1}^{Z(d,\ell)} \frac{1}{N} \left| \sum_{n=1}^{N} Y_{\ell,m}(x_n) \right| \tag{2}$$

valid for all positive integer values of M. Here $C_1(d)$ and $C_2(d)$ denote (explicitely known) constants, $Y_{\ell,m}$ $(m = 1, \ldots, Z(d,\ell))$ denote an orthonormal system of spherical harmonics of order ℓ, and $Z(d,\ell)$ denotes the dimension of the space of these spherical harmonics.

Only recently, a spherical version of the LeVeque inequality was found (cf. [13]):

$$D(X_N) \leq A(d) \left(\sum_{\ell=0}^{\infty} \ell^{-(d+1)} \sum_{m=1}^{Z(d,\ell)} \left(\frac{1}{N} \sum_{n=1}^{N} Y_{\ell,m}(x_n) \right)^2 \right)^{\frac{1}{d+2}} \tag{3}$$

with an explicit constant $A(d)$. Both inequalities (2) and (3) specialise to their classical versions for $d = 1$.

Minimal energy point sets. Using mutually repelling forces on N particles to distribute them on a surface M is a rather compelling idea. The motivation for this could be taken from physical experiments, where electric charges distribute themselves in a way that minimises the sum of the mutual energies

$$E_s(X_N) = \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \|x_i - x_j\|^{-s} \tag{4}$$

for $s = 1$ (cf. [5, 14]). The study of the precise distribution of the charges is the subject of classical potential theory (cf. [11]), which shows that the energy integral

$$I_s(\mu) = \iint_{M \times M} \|x - y\|^{-s} \, d\mu(x) \, d\mu(y). \tag{5}$$

has a unique minimiser amongst all Borel probability measures supported on M, the harmonic measure on M. This measure depends highly on the curvature of the surface, and thus differs from the surface measure, except for surfaces with high symmetry, like the sphere. For larger values of s there is no physical experiment, which can be used to describe the charge distribution, nevertheless, the intuition and the result remain the same – there exists a unique equilibrium measure depending on s on M – if $s < \dim(M)$.

The minimisation of such discrete energy expressions is a problem attributed to M. Fekete. Minimal energy point sets are thus called Fekete-points. The case $s < \dim(M)$ can be investigated by methods from classical potential theory (cf. [11]). In this case the unique minimiser $\mu_M^{(s)}$ of $I_s(\mu)$ is the weak limit of the measures ν_N (cf. [9]). For $s \geq \dim(M)$ the situation changes completely. The corresponding energy integral diverges for all probability measures. Techniques from geometric measure theory could be applied in [2, 8] to show that the limiting distribution

$\mu_M^{(s)}$ of the minimal energy distributions is the normalised $\dim(M)$-dimensional Hausdorff measure on M, if M is rectifiable.

Although the limiting distribution of minimal energy point sets X_N^* for $s \geq d$ on the sphere has been determined in [8], almost nothing is known about quantitative results. The only – and very weak – estimate for the discrepancy of minimal energy point sets in the singular case is due to S. Damelin and the author [4] and gives $D(X_N^*) = \mathcal{O}\left(\sqrt{\frac{\log\log N}{\log N}}\right)$ for $s = d$.

In [13] inequality (3) was applied to minimal energy point sets X_N^* for $-2 < s < 0$. This gives bounds for the discrepancy $D(X_N^*) \ll N^{-\frac{d-s}{d(d+2)}}$. For $s = 0$ in [3] the bound $D(X_N^*) \ll N^{-\frac{1}{d+2}}$ was obtained.

It should be mentioned that from the theory of irregularities of distribution [1] it is known that for all sets $X_N \subset \mathbb{S}^d$ the inequality

$$D(X_N) \gg N^{-\frac{1}{2}-\frac{1}{2d}} \tag{6}$$

holds for the spherical cap discrepancy.

It is also known that inequality (6) is best possible up to a factor $\sqrt{\log N}$. The existence of point sets X_N with $D(X_N) \ll N^{-\frac{1}{2}-\frac{1}{2d}}\sqrt{\log N}$ uses a probabilistic argument; up to now no explicit construction of such a point set is known. All the known estimates for the discrepancy of point sets differ from the lower bound (6) by a power of N.

Best packing. The problem of maximising the minimal distance

$$\delta(X_N) = \min_{1\leq i<j\leq N} \|x_i - x_j\| \tag{7}$$

of N distinct points on the sphere (cf. [7]) or more general manifolds occurs as the limiting case $s \to \infty$ in the energy minimisation. The problem of arranging points maximising the minimal mutual distance is usually attributed to the Dutch botanist P. M. L. Tammes, who used it to explain the distribution of pores on pollen grains (cf. [15]).

References

[1] J. Beck and W. W. L. Chen, *Irregularities of distribution*, Cambridge Tracts in Mathematics, vol. 89, Cambridge University Press, Cambridge, 1987.

[2] S. V. Borodachov, D. P. Hardin, and E. B. Saff, *Asymptotics for discrete weighted minimal Riesz energy problems on rectifiable sets*, Trans. Amer. Math. Soc. **360** (2008), no. 3, 1559–1580 (electronic).

[3] J. S. Brauchart, *Optimal logarithmic energy points on the unit sphere*, Math. Comp. **77** (2008), no. 263, 1599–1613.

[4] S. B. Damelin and P. J. Grabner, *Energy functionals, numerical integration and asymptotic equidistribution on the sphere*, J. Complexity **19** (2003), 231–246, Corrigendum, **20** (2004), 883–884.

[5] T. Erber and G. M. Hockney, *Equilibrium configurations of n equal charges on a sphere*, J. Phys. A: Math. Gen. **24** (1991), L1369–L1377.

[6] P. J. Grabner, *Erdős-Turán type discrepancy bounds*, Monatsh. Math. **111** (1991), 127–135.

[7] W. Habicht and B. L. van der Waerden, *Lagerung von Punkten auf der Kugel*, Math. Ann. **123** (1951), 223–234.

[8] D. P. Hardin and E. B. Saff, *Minimal Riesz energy point configurations for rectifiable d-dimensional manifolds*, Adv. Math. **193** (2005), no. 1, 174–204.
[9] A. B. J. Kuijlaars and E. B. Saff, *Asymptotics for minimal discrete energy on the sphere*, Trans. Amer. Math. Soc. **350** (1998), no. 2, 523–538.
[10] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley-Interscience, New York, 1974.
[11] N. S. Landkof, *Foundations of modern potential theory*, Springer-Verlag, New York, 1972, Translated from the Russian by A. P. Doohovskoy, Die Grundlehren der mathematischen Wissenschaften, Band 180.
[12] X.-J. Li and J. D. Vaaler, *Some trigonometric extremal functions and the Erdös-Turán type inequalities*, Indiana Univ. Math. J. **48** (1999), no. 1, 183–236.
[13] F. J. Narcowich, X. Sun, J. D. Ward, and Z. Wu, *LeVeque type inequalities and discrepancy estimates for minimal energy configurations on spheres*, J. Approx. Theory **162** (2010), no. 6, 1256–1278.
[14] E. A. Rakhmanov, E. B. Saff, and Y. M. Zhou, *Electrons on the sphere*, Computational methods and function theory 1994 (Penang), Ser. Approx. Decompos., vol. 5, World Sci. Publ., River Edge, NJ, 1995, pp. 293–309.
[15] P. M. L. Tammes, *On the origin of number and arrangement of the places of exit on the surface of pollen grains*, Recueil des travaux botaniques néerlandais **27** (1930), 1–84.

Embeddings of weighted tensor product spaces

Mario Hefter
(joint work with Klaus Ritter)

Embedding theorems deal with scales $(F_s^\alpha)_\alpha$ of function spaces on a common domain of dimension $s \in \mathbb{N}$, and one of the aims is to characterize those pairs of spaces F_s^α and F_s^β that permit a continuous embedding $i_s^{\alpha,\beta} : F_s^\alpha \hookrightarrow F_s^\beta$. A major application of embedding theorems in information-based complexity, approximation theory, and numerical mathematics is as follows: The existence of a continuous embedding $i_s^{\alpha,\beta}$ with norm $\|i_s^{\alpha,\beta}\|$ implies

$$e_n(F_s^\alpha) \le \|i_s^{\alpha,\beta}\| \cdot e_n(F_s^\beta) \tag{1}$$

for many quantities e_n of interest, like n-th minimal errors or n-widths.

In the classical approach one studies the asymptotic behavior of $e_n(F_s^\alpha)$ as n tends to infinity with α and s being fixed, and the mere existence of continuous embeddings can already be exploited, since (1) yields $e_n(F_s^\alpha) = O\left(e_n(F_s^\beta)\right)$. In particular, if $F_s^\alpha = F_s^\beta$ as vector spaces with equivalent norms, then the sequences $(e_n(F_s^\alpha))_n$ and $(e_n(F_s^\beta))_n$ are weakly equivalent.

In contrast, tractability analysis studies the explicit dependence of $e_n(F_s^\alpha)$ on n and on the dimension s, which is crucial to fully understand the impact of a high dimension on the computational or approximation problem at hand. We refer to [3, 4, 5] for a comprehensive study and further references. Moreover, tractability analysis enables the study of the limiting case $s = \infty$, i.e., of computational or approximation problems for functions with infinitely many variables. Exploiting the existence of continuous embeddings $i_s^{\alpha,\beta}$ or the equivalence of norms on $F_s^\alpha = F_s^\beta$ for all $s \in \mathbb{N}$ in tractability analysis requires a tight control of the dependence of the norms of the respective embeddings on the dimension s.

In this talk we consider scales of weighted tensor product Hilbert spaces, which are most often studied in tractability analysis. The starting point for the construction of these spaces is a reproducing kernel k on a domain $D \times D$ and a sequence $(\gamma_j)_{j\in\mathbb{N}}$ of positive weights. By assumption, the Hilbert space $H(1+k)$ with reproducing kernel $1+k$ is the orthogonal sum of the space $H(1)$ of constant functions and the space $H(k)$. The corresponding norm of $f \in H(1+\gamma_j k)$ is therefore given by

$$\|f\|^2_{1+\gamma_j k} = P(f)^2 + \frac{1}{\gamma_j} \cdot \|f - P(f)\|^2_k,$$

where P denotes the orthogonal projection onto $H(1)$. The second scale is derived from an equivalent norm

$$\|f\|^2_{1+l_{\gamma_j}} = \langle f, f\rangle + \frac{1}{\gamma_j} \cdot \|f - P(f)\|^2_k$$

on the same vector space $H = H(1+k) = H(1+\gamma_j k)$. By assumption, $\langle\cdot,\cdot\rangle$ is a properly normalized symmetric bilinear form on H that is continuous on $H(k)$, and actually we get a new reproducing kernel Hilbert space $H(1+l_{\gamma_j})$ in this way. As it turns out,

$$\forall\, f \in H : P(f) = \langle f, 1\rangle \tag{2}$$

forms a particular instance, since (2) is equivalent to $H(1)$ and $H(k)$ being orthogonal in the spaces $H(1+l_{\gamma_j})$, too.

The first result deals with embeddings $i_s^{\boldsymbol{\eta},\boldsymbol{\gamma}} : H(K_s^{\boldsymbol{\eta}}) \hookrightarrow H(L_s^{\boldsymbol{\gamma}})$ between the tensor product spaces $H(K_s^{\boldsymbol{\eta}}) = \bigotimes_{j=1}^s H(1+\eta_j k)$ and $H(L_s^{\boldsymbol{\gamma}}) = \bigotimes_{j=1}^s H(1+l_{\gamma_j})$ of functions on D^s, where $\boldsymbol{\eta} = (\eta_j)_{j\in\mathbb{N}}$ and $\boldsymbol{\gamma} = (\gamma_j)_{j\in\mathbb{N}}$ are arbitrary sequences of positive weights. Hence the reproducing kernels $K_s^{\boldsymbol{\eta}}$ and $L_s^{\boldsymbol{\gamma}}$ are given by

$$K_s^{\boldsymbol{\eta}}(\mathbf{x},\mathbf{y}) = \prod_{j=1}^s (1+\eta_j k(x_j,y_j)), \qquad \mathbf{x},\mathbf{y} \in D^s,$$

and

$$L_s^{\boldsymbol{\gamma}}(\mathbf{x},\mathbf{y}) = \prod_{j=1}^s (1+l_{\gamma_j}(x_j,y_j)), \qquad \mathbf{x},\mathbf{y} \in D^s.$$

Here we present a particular result for summable weights, i.e.,

$$\sum_{j\in\mathbb{N}} \gamma_j < \infty. \tag{3}$$

This condition often arises in the context of tractability analysis and was first encountered in [6]. If (2) is satisfied, then (3) is equivalent to

$$\sup_{s\in\mathbb{N}} \max\left(\|\imath_s^{\boldsymbol{\gamma},\boldsymbol{\gamma}}\|, \|(\imath_s^{\boldsymbol{\gamma},\boldsymbol{\gamma}})^{-1}\|\right) < \infty,$$

i.e., we have a uniform equivalence of the norms on the spaces $H(K_s^{\boldsymbol{\gamma}})$ and $H(L_s^{\boldsymbol{\gamma}})$ for $s \in \mathbb{N}$. If (2) is not satisfied, then (3) is equivalent to the existence of $0 < c' < 1 < \tilde{c}$ such that $\sup_{s\in\mathbb{N}} \max\left(\|\imath_s^{c'\boldsymbol{\gamma},\boldsymbol{\gamma}}\|, \|(\imath_s^{\tilde{c}\boldsymbol{\gamma},\boldsymbol{\gamma}})^{-1}\|, \|\imath_s^{\boldsymbol{\gamma},\tilde{c}\boldsymbol{\gamma}}\|, \|(\imath_s^{\boldsymbol{\gamma},c'\boldsymbol{\gamma}})^{-1}\|\right) < \infty$, i.e.,

multiplying the weights γ_j by suitable constant factors leads to uniformly bounded norms of the embeddings corresponding to $H(K_s^{c'\boldsymbol{\gamma}}) \subseteq H(L_s^{\boldsymbol{\gamma}}) \subseteq H(K_s^{\tilde{c}\boldsymbol{\gamma}})$ and $H(L_s^{c'\boldsymbol{\gamma}}) \subseteq H(K_s^{\boldsymbol{\gamma}}) \subseteq H(L_s^{\tilde{c}\boldsymbol{\gamma}})$.

For the second result we consider the limit $s \to \infty$ of the reproducing kernels $K_s^{\boldsymbol{\eta}}$ and $L_s^{\boldsymbol{\gamma}}$, namely,

$$K^{\boldsymbol{\eta}}(\mathbf{x},\mathbf{y}) = \prod_{j=1}^{\infty}(1+\eta_j k(x_j,y_j)), \qquad \mathbf{x},\mathbf{y} \in \mathfrak{X}^{\boldsymbol{\eta}},$$

where $\mathfrak{X}^{\boldsymbol{\eta}} = \{\mathbf{x} \in D^{\mathbb{N}} : \prod_{j=1}^{\infty}(1+\eta_j k(x_j,x_j)) < \infty\}$, and

$$L^{\boldsymbol{\gamma}}(\mathbf{x},\mathbf{y}) = \prod_{j=1}^{\infty}(1+l_{\gamma_j}(x_j,y_j)), \qquad \mathbf{x},\mathbf{y} \in \mathfrak{Y}^{\boldsymbol{\gamma}},$$

where $\mathfrak{Y}^{\boldsymbol{\gamma}} = \{\mathbf{y} \in D^{\mathbb{N}} : \prod_{j=1}^{\infty}(1+l_{\gamma_j}(y_j,y_j)) < \infty\}$. Here we present a particular result, which again deals with summable weights. If (2) is satisfied, then (3) is equivalent to $H(K^{\boldsymbol{\gamma}}) = H(L^{\boldsymbol{\gamma}})$. If (2) is not satisfied, then (3) is equivalent to the existence of $0 < c' < 1 < \tilde{c}$ such that $H(K^{c'\boldsymbol{\gamma}}) \subseteq H(L^{\boldsymbol{\gamma}}) \subseteq H(K^{\tilde{c}\boldsymbol{\gamma}})$ and $H(L^{c'\boldsymbol{\gamma}}) \subseteq H(K^{\boldsymbol{\gamma}}) \subseteq H(L^{\tilde{c}\boldsymbol{\gamma}})$. Due to the closed graph theorem the respective embeddings are continuous, and in the case (2) we have equivalence of the norms on the spaces $H(K^{\boldsymbol{\gamma}})$ and $H(L^{\boldsymbol{\gamma}})$.

We refer to [2] for analytic properties of the spaces $H(K^{\boldsymbol{\eta}})$ and $H(L^{\boldsymbol{\gamma}})$ as well as for more general weighted superpositions of tensor products of reproducing kernel Hilbert spaces. Implications for tractability analysis and for computational or approximation problems for functions with infinitely many variables will be studied in the forthcoming paper [1].

Two prominent examples from tractability analysis are given by $D = [0,1]$ and $k^{(1)}(x,y) = \min(x,y)$ as well as $k^{(2)}(x,y) = 1/2 + (x^2+y^2)/2 - \max(x,y)$ for $x,y \in D$. We have $H(1+\gamma_j k^{(i)}) = W_2^1([0,1])$ for $i = 1,2$ and

$$\|f\|^2_{1+\gamma_j k^{(1)}} = f^2(0) + \frac{1}{\gamma_j} \cdot \int_0^1 (f')^2(x)\,dx$$

as well as

$$\|f\|^2_{1+\gamma_j k^{(2)}} = \left(\int_0^1 f(x)\,dx\right)^2 + \frac{1}{\gamma_j} \cdot \int_0^1 (f')^2(x)\,dx$$

for $f \in W_2^1([0,1])$, and these settings are called the anchored decomposition and the ANOVA decomposition of the space $W_2^1([0,1])$. An obvious choice of $\langle\cdot,\cdot\rangle$ leads to

$$\|f\|^2_{1+l_{\gamma_j}} = \int_0^1 f^2(x)\,dx + \frac{1}{\gamma_j} \cdot \int_0^1 (f')^2(x)\,dx,$$

which yields another decomposition of the space $W_2^1([0,1])$. See [3, Sec. A.2].

Our results applied to this example reveals that uniform equivalence of the norms on $\bigotimes_{j=1}^s H(1+\gamma_j k^{(1)})$ and $\bigotimes_{j=1}^s H(1+\gamma_j k^{(2)})$ holds iff $\sum_{j\in\mathbb{N}} \sqrt{\gamma_j} < \infty$, and the latter is also equivalent to uniform equivalence of the norms on $\bigotimes_{j=1}^s H(1+$

$\gamma_j k^{(1)})$ and $\bigotimes_{j=1}^{s} H(1 + l_{\gamma_j})$. However, uniform equivalence of the norms on $\bigotimes_{j=1}^{s} H(1 + \gamma_j k^{(2)})$ and $\bigotimes_{j=1}^{s} H(1 + l_{\gamma_j})$ is equivalent to $\sum_{j\in\mathbb{N}} \gamma_j < \infty$.

References

[1] J. Dick, M. Gnewuch, M. Hefter, A. Hinrichs, and K. Ritter, in preparation, 2013.

[2] M. Gnewuch, S. Mayer, and K. Ritter, *On weighted Hilbert spaces and integration of functions of infinitely many variables*, J. Complexity (2013), to appear.

[3] E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems. Vol. 1: Linear Information*, EMS Tracts in Mathematics, European Mathematical Society (EMS), Zürich, 2008.

[4] ———, *Tractability of Multivariate Problems. Vol. 2: Standard Information for Fundtionals*, EMS Tracts in Mathematics, European Mathematical Society (EMS), Zürich, 2010.

[5] ———, *Tractability of Multivariate Problems. Vol. 3: Standard Information for Operators*, EMS Tracts in Mathematics, European Mathematical Society (EMS), Zürich, 2012.

[6] I. H. Sloan and H. Woźniakowski, *When are quasi-Monte Carlo algorithms efficient for high dimensional integrals?*, J. Complexity 14 (1998), 1–33.

Complexity of parametric integration in various smoothness classes

Stefan Heinrich

(joint work with Thomas Daun)

The complexity of definite parametric integration was studied in [6], [3], and [9], while in [1] the complexity of both definite and indefinite parametric integration was considered. Parametric definite integration is a problem intermediate between integration and approximation. Parametric indefinite integration can be viewed as a model for the solution of parametric initial value problems in the sense that it is a partial, but typical case, and some of the methods developed here will be used in the study of parametric initial value problems, see [2].

These results are a continuation of [1] and we study both definite and indefinite integration. So far definite parametric integration was considered only for isotropic classes and, in [3], for a specific anisotropic class (Sobolev case with no smoothness in the integration variable). Indefinite parametric integration was only studied for C^r. In [1] we gave a general (multilevel) scheme for Banach space valued integration of functions belonging to

$$C^r(X) \cap C^{r_1}(Y), \tag{1}$$

where X and Y are Banach spaces such that Y is continuously embedded into X, from which the upper bounds for parametric integration in the C^r-case were derived.

Now we further explore the range given in (1) by considering classes of functions with dominating mixed derivatives and other types of non-isotropic smoothness. In contrast to the C^r case, these classes allow to treat different smoothnesses for the parameter dependence and for the basic (nonparametric) integration problem. We want to understand the typical behaviour of the complexity in these classes and the relation between the deterministic and randomized setting, this way clarifying in

which cases and to which extend randomized methods are superior to deterministic ones.

We recall some notation from information-based complexity theory [8, 7], see also [4] for the precise notions used here. Let F be a nonempty set, G a normed linear space, $S : F \to G$ an arbitrary mapping, let K be a nonempty set, and let Λ be a set of mappings from F to K. We interpret F as the set of inputs, S as the solution operator, that is, the mapping that sends the input $f \in F$ to the exact solution Sf, and Λ is understood as the class of admissible information functionals. Thus, the tuple

$$P = (F, G, S, K, \Lambda) \tag{2}$$

describes the abstract numerical problem under consideration. Moreover, the deterministic n-th minimal error is denoted by $e_n^{\mathrm{det}}(S, F)$ and the randomized n-th minimal error is denoted by $e_n^{\mathrm{ran}}(S, F)$. So $e_n^{\mathrm{det}}(S, F)$, respectively $e_n^{\mathrm{ran}}(S, F)$, is the minimal possible error among all deterministic, respectively randomized algorithms that use at most n information functionals.

For $d_0 \in \mathbb{N}$, $Q_0 = [0, 1]^{d_0}$ we study definite and indefinite integration of functions depending on a parameter $s \in Q_0$. Let $r_0, r \in \mathbb{N}_0$ and let $C^{r_0,r}(Q_0, Q)$ be the space of continuous functions $f : Q_0 \times Q \to \mathbb{K}$ having for $\alpha = (\alpha_0, \alpha_1)$, $\alpha_0 \in \mathbb{N}_0^{d_0}$, $\alpha_1 \in \mathbb{N}_0^{d}$ with $|\alpha_0| \le r_0$, $|\alpha_1| \le r$ continuous partial derivatives $\frac{\partial^{|\alpha|} f(s,t)}{\partial s^{\alpha_0} \partial t^{\alpha_1}}$, endowed with the norm

$$\|f\|_{C^{r_0,r}(Q_0,Q)} = \max_{|\alpha_0| \le r_0, |\alpha_1| \le r} \sup_{s \in Q_0, t \in Q} \left| \frac{\partial^{|\alpha|} f(s,t)}{\partial s^{\alpha_0} \partial t^{\alpha_1}} \right|.$$

Let furthermore $r_1 \in \mathbb{N}_0$ and put

$$F = B_{C^{0,r}(Q_0,Q)} \cap B_{C^{r_0,r_1}(Q_0,Q)}.$$

The definite parametric integration operator $S_0 : C(Q_0 \times Q) \to C(Q_0)$ is given by

$$(S_0 f)(s) = \int_Q f(s,t)dt \quad (s \in Q_0),$$

and the indefinite parametric integration operator $S_1 : C(Q_0 \times Q) \to C(Q_0 \times Q)$ by

$$(S_1 f)(s,t) = \int_{[0,t]} f(s,u)du \quad (s \in Q_0, t \in Q).$$

We consider standard information consisting of values of f, so the class of information functionals is $\Lambda = \{\delta_{s,t} : \ s \in Q_0, t \in Q\}$, where $\delta_{s,t}(f) = f(s,t)$. In the terminology of (2), the definite parametric integration problem is described by the tupel

$$\Pi_0 = (B_{C^{0,r}(Q_0,Q)} \cap B_{C^{r_0,r_1}(Q_0,Q)}, C(Q_0), S_0, \mathbb{K}, \Lambda)$$

and the indefinite parametric integration problem by

$$\Pi_1 = (B_{C^{0,r}(Q_0,Q)} \cap B_{C^{r_0,r_1}(Q_0,Q)}, C(Q_0 \times Q), S_1, \mathbb{K}, \Lambda).$$

The following theorem gives the complexity of definite and indefinite parametric integration. Below $\wedge$ and $\vee$ mean logical conjunction and disjunction, respectively.

Theorem. *Let $r_0, r, r_1 \in \mathbb{N}_0$, $r \geq r_1$, $d, d_0 \in \mathbb{N}$, $\iota \in \{0,1\}$. Then the deterministic minimal errors satisfy*

$$\begin{aligned}
&e_n^{\det}(S_\iota, F) \asymp n^{-\upsilon_1} && \text{if} \quad \tfrac{r_0}{d_0} > \tfrac{r_1}{d} \\
&n^{-\frac{r_0}{d_0}} \preceq e_n^{\det}(S_\iota, F) \preceq n^{-\frac{r_0}{d_0}} (\log n)^{\frac{r_0}{d_0}+1} && \text{if} \quad \tfrac{r_0}{d_0} = \tfrac{r_1}{d} > 0 \\
&e_n^{\det}(S_\iota, F) \asymp n^{-\frac{r_0}{d_0}} && \text{if} \quad \tfrac{r_0}{d_0} = \tfrac{r_1}{d} = 0 \vee \tfrac{r_0}{d_0} < \tfrac{r_1}{d},
\end{aligned}$$

where

$$\upsilon_1 \quad = \quad \frac{\frac{r_0}{d_0}}{\frac{r_0}{d_0} + \frac{r}{d} - \frac{r_1}{d}} \, \frac{r}{d} \, .$$

Moreover, the randomized minimal errors fulfill

$$\begin{aligned}
&e_n^{\mathrm{ran}}(S_\iota, F) \asymp n^{-\frac{r}{d}-\frac{1}{2}} && \text{if} \quad \tfrac{r_0}{d_0} > \tfrac{r_1}{d} + \tfrac{1}{2} \wedge r = r_1 \\
&e_n^{\mathrm{ran}}(S_\iota, F) \asymp n^{-\upsilon_2} (\log n)^{\frac{1}{2}} && \text{if} \quad \tfrac{r_0}{d_0} > \tfrac{r_1}{d} + \tfrac{1}{2} \wedge r > r_1 \\
&n^{-\frac{r_0}{d_0}} (\log n)^{\frac{1}{2}} \preceq e_n^{\mathrm{ran}}(S_\iota, F) \preceq n^{-\frac{r_0}{d_0}} (\log n)^{\frac{r_0}{d_0}+\frac{3}{2}} && \text{if} \quad \tfrac{r_0}{d_0} = \tfrac{r_1}{d} + \tfrac{1}{2} \\
&e_n^{\mathrm{ran}}(S_\iota, F) \asymp n^{-\frac{r_0}{d_0}} (\log n)^{\frac{r_0}{d_0}-\frac{r_1}{d}} && \text{if} \quad \tfrac{r_1}{d} < \tfrac{r_0}{d_0} < \tfrac{r_1}{d} + \tfrac{1}{2} \\
&n^{-\frac{r_0}{d_0}} \preceq e_n^{\mathrm{ran}}(S_\iota, F) \preceq n^{-\frac{r_0}{d_0}} (\log \log n)^{\frac{r_0}{d_0}+1} && \text{if} \quad \tfrac{r_0}{d_0} = \tfrac{r_1}{d} > 0 \\
&e_n^{\mathrm{ran}}(S_\iota, F) \asymp n^{-\frac{r_0}{d_0}} && \text{if} \quad \tfrac{r_0}{d_0} = \tfrac{r_1}{d} = 0 \vee \tfrac{r_0}{d_0} < \tfrac{r_1}{d},
\end{aligned}$$

with

$$\upsilon_2 \quad = \quad \frac{\frac{r_0}{d_0}}{\frac{r_0}{d_0} + \frac{r}{d} - \frac{r_1}{d}} \left(\frac{r}{d} + \frac{1}{2} \right).$$

Next we give an example of an particular class. If $r_1 = r$, then $F = B_{C^{r_0,r}(Q_0,Q)}$, which is a class of dominating mixed smoothness, more precisely, the smoothness with respect to the parameter variables s and the smoothness with respect to the variables t are combined in such a way.

Corollary. *Let $r_0, r \in \mathbb{N}_0$, $r_1 = r$, $d, d_0 \in \mathbb{N}$, $\iota \in \{0,1\}$. Then*

$$\begin{aligned}
e_n^{\det}(S_\iota, F) \quad &\asymp_{\log} \quad n^{-\min\left(\frac{r}{d}, \frac{r_0}{d_0}\right)} \\
e_n^{\mathrm{ran}}(S_\iota, F) \quad &\asymp_{\log} \quad n^{-\min\left(\frac{r}{d}+\frac{1}{2}, \frac{r_0}{d_0}\right)},
\end{aligned}$$

where $\asymp_{\log}$ denotes the asymptotic notation neglecting logarithmic factors.

References

[1] Th. Daun, S. Heinrich, Complexity of Banach space valued and parametric integration, to appear in the Proceedings of Monte Carlo and Quasi-Monte Carlo Methods 2012.

[2] Th. Daun, S. Heinrich, Complexity of parametric initial value problems for systems of ODEs, submitted.
[3] S. Heinrich, The multilevel method of dependent tests, Advances in Stochastic Simulation Methods (N. Balakrishnan, V. B. Melas, S. M. Ermakov, editors), pp. 47–62, Birkhäuser, Boston, Basel, Berlin 2000.
[4] S. Heinrich, Monte Carlo approximation of weakly singular integral operators, J. Complexity 22 (2006), 192–219.
[5] S. Heinrich, B. Milla, The randomized complexity of indefinite integration, J. Complexity 27 (2011), 352–382.
[6] S. Heinrich, E. Sindambiwe, Monte Carlo complexity of parametric integration, J. Complexity 15 (1999), 317-341.
[7] E. Novak, Deterministic and Stochastic Error Bounds in Numerical Analysis, Lecture Notes in Mathematics 1349, Springer, 1988.
[8] J. F. Traub, G. W. Wasilkowski, and H. Woźniakowski, Information-Based Complexity, Academic Press, New York, 1988.
[9] C. Wiegand, Optimal Monte Carlo and Quantum Algorithms for Parametric Integration, PhD Thesis 2005, University of Kaiserslautern; Shaker Verlag Aachen, 2006.

New ways for constructing generating matrices of digital sequences

Roswitha Hofer
(joint work with Harald Niederreiter)

Most of the well-known s-dimensional low-discrepancy sequences can be constructed via the digital method, which was introduced by Niederreiter [4] and generalized earlier forms by Sobol [6] and Faure [1]. The digital methods constructs a sequence $(\boldsymbol{x}_n)_{n\geq 0}$ in $[0,1]^s$ as follows.:
Choose a finite field $\mathbb{F}_q$ with cardinality q, and put $Z_q = \{0,1,\ldots,q-1\} \subset \mathbb{Z}$. Choose

(i) bijections $\psi_r : Z_q \to \mathbb{F}_q$ for all integers $r \geq 0$, satisfying $\psi_r(0) = 0$ for all sufficiently large r;
(ii) *generating matrices* $C^{(i)} := (c_{j,r}^{(i)})_{j\geq 1, r\geq 0} \in \mathbb{F}_q^{\mathbb{N}\times\mathbb{N}_0}$ for $1 \leq i \leq s$;
(iii) bijections $\lambda_{i,j} : \mathbb{F}_q \to Z_q$ for $1 \leq i \leq s$ and $j \geq 1$.

The ith coordinate $x_n^{(i)}$ of the nth point $\boldsymbol{x}_n$ of the sequence is computed as follows. Given an integer $n \geq 0$, let $n = \sum_{r=0}^{\infty} z_r(n) q^r$ be the digit expansion of n in base q, with all $z_r(n) \in Z_q$ and $z_r(n) = 0$ for all sufficiently large r. Carry out the matrix-vector product

$$C^{(i)} \cdot \begin{pmatrix} \psi_0(z_0(n)) \\ \psi_1(z_1(n)) \\ \vdots \end{pmatrix} =: \begin{pmatrix} y_{n,1}^{(i)} \\ y_{n,2}^{(i)} \\ \vdots \end{pmatrix} \quad \text{and put} \quad x_n^{(i)} = \sum_{j=1}^{\infty} \lambda_{i,j}(y_{n,j}^{(i)}) q^{-j} \in [0,1].$$

Note that the distribution of the generated sequence mainly depends on the choice of the generating matrices.

This talk considers different methods for constructing generating matrices that are qualified to generate low-discrepancy sequences. More exactly it compares the

famous classical Niederreiter construction [5] and the Xing-Niederreiter construction [7] — both build up the generating matrices row-by-row — with the recent columnwise concepts in [2] and [3].

References

[1] H. Faure, *Discrépance de suites associées à un système de numération (en dimension s)*, Acta Arith. **41** (1982), 337–351.

[2] R. Hofer, *A construction of low-discrepancy sequences involving finite-row digital (t, s)-sequences*, Monatsh. Math. **171** (2013), 77–89.

[3] R. Hofer and H. Niederreiter, *A construction of (t, s)-sequences with finite-row generating matrices using global function fields*, Finite Fields Appl. **21** (2013), 97–110.

[4] H. Niederreiter, *Point sets and sequences with small discrepancy*, Monatsh. Math. **104** (1987), 273–337.

[5] H. Niederreiter, *Low-discrepancy and low-dispersion sequences*, J. Number Theory **30** (1988), 51–70.

[6] I.M. Sobol', *Distribution of points in a cube and approximate evaluation of integrals (Russian)*, Ž. Vyčisl. Mat. i Mat. Fiz. **7** (1967), 784–802.

[7] C. Xing and H. Niederreiter, *A construction of low-discrepancy sequences using global function fields*, Acta Arith. **73** (1995), 87–102.

Star discrepancy bounds for (t, m, s)-nets and (t, s)-sequences

Peter Kritzer

(joint work with Henri Faure)

The star discrepancy, which is one of most important measures of uniformity of distribution, of a point set $\mathcal{P}$ with N points $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-1}$ in the s-dimensional unit cube $[0,1)^s$ is defined as

$$D_N^*(\mathcal{P}) := \sup \left| |\{n : 0 \le n < N, \boldsymbol{x}_n \in [\boldsymbol{0}, \boldsymbol{a})\}| N^{-1} - \lambda_s([\boldsymbol{0}, \boldsymbol{a})) \right|,$$

where the supremum is extended over all $\boldsymbol{a} \in [0,1]^s$. For an infinite sequence $\mathcal{S}$, $D_N^*(\mathcal{S})$ denotes the star discrepancy of the first N points of $\mathcal{S}$.

We consider the star discrepancy of two classes of point sets that are heavily used in quasi-Monte Carlo methods for numerical integration, namely (t, m, s)-nets and (t, s)-sequences, as introduced by Sobol', Faure, and Niederreiter. The quality of distribution of a given point set is linked to its quality as the node set in a quasi-Monte Carlo integration rule via the well-known Koksma-Hlawka inequality.

The following definition of a (t, m, s)-net is due to Niederreiter [4].

Definition 1. *Let $b \ge 2$, $s \ge 1$, and $0 \le t \le m$ be integers. A point set $\mathcal{P}$ with $N = b^m$ points in $[0,1)^s$ is a (t, m, s)-net in base b, if every subinterval $J = \prod_{j=1}^{s} \left[\frac{a_j}{b^{d_j}}, \frac{a_j+1}{b^{d_j}} \right)$ of $[0,1)^s$, with integers $d_j \ge 0$ and $0 \le a_j < b^{d_j}$ for $1 \le j \le s$ and of volume b^{t-m}, contains exactly b^t points of $\mathcal{P}$.*

Infinite analogues of (t, m, s)-nets are provided by (t, s)-sequences.

Definition 2. *Let $b \geq 2$, $s \geq 1$ and $t \geq 0$ be integers. An infinite sequence $\mathcal{S}$ is a (t,s)-sequence in base b, if for all $m > t$ and all $k \geq 0$ the points*

$$\boldsymbol{x}_{kb^m}, \boldsymbol{x}_{kb^m+1}, \ldots, \boldsymbol{x}_{(k+1)b^m-1}$$

form a (t,m,s)-net in base b.

An obvious question is how small or large the star discrepancy of (t,m,s)-nets and (t,s)-sequences is. While there are only few results on lower bounds for the discrepancy of these point sets that make use of their structure, and one usually has to resort to general lower bounds, there have been many results on upper star discrepancy bounds for (t,m,s)-nets and (t,s)-sequences exploiting the particular distribution properties of these.

An upper discrepancy bound for the star discrepancy of arbitrary (t,m,s)-nets was first shown by Niederreiter in [4]. This bound can be formulated as follows.

Theorem 1 (Niederreiter)**.** *For the star discrepancy of any (t,m,s)-net $\mathcal{P}$ in base b with $m > 0$ we have*

$$ND_N^*(\mathcal{P}) \leq B(s,b)b^t(\log N)^{s-1} + \mathcal{O}\left(b^t(\log N)^{s-2}\right),$$

where the constant in the $\mathcal{O}$-notation does not depend on N, and where $B(s,b) = \left(\frac{b-1}{2\log b}\right)^{s-1}$ if either $s = 2$ or $b = 2, s = 3, 4$; otherwise $B(s,b) = \frac{1}{(s-1)!}\left(\frac{\lfloor b/2 \rfloor}{\log b}\right)^{s-1}$.

A similar bound for the discrepancy of the first N terms of a (t,s)-sequence, using slightly different terms $C(s,b)$ (dependent on s and b) instead of $B(s,b)$, and involving an additional $\log N$-factor have also been shown in [4]. Here, we present the latest of a series of results that have been aimed at reducing the leading terms $B(s,b)$ and $C(s,b)$ in Niederreiter's discrepancy bounds. Indeed, in the paper [1], we show discrepancy bounds of a similar flavor as the previous ones, where we can replace the terms $B(s,b)$ and $C(s,b)$ by smaller expressions. Regarding (t,m,s)-nets we have the following theorem.

Theorem 2 (Faure, Kritzer)**.** *Let $\mathcal{P}$ be a (t,m,s)-net in base b with $N = b^m$ points, $m > 0$. Then it is true that*

$$ND_N^*(\mathcal{P}) \leq E(s,b)b^t(\log N)^{s-1} + \mathcal{O}(b^t(\log N)^{s-2}) \text{ with}$$

$$E(s,b) = \begin{cases} \frac{1}{(s-1)!}\frac{b^2}{2(b^2-1)}\left(\frac{b-1}{2\log b}\right)^{s-1} & \text{if } b \text{ is even,} \\ \\ \frac{1}{(s-1)!}\frac{1}{2}\left(\frac{b-1}{2\log b}\right)^{s-1} & \text{if } b \text{ is odd.} \end{cases}$$

Again, the implied constant in the $\mathcal{O}$-notation does not depend on N.

The proof of our result is different from earlier proofs of similar results in [4] and later [3]. To be more precise, the proofs in [3, 4] are based on a double induction method, whereas the proof of Theorem 2 is based on a single induction argument, also using an adapted method from a different proof in [4].

A corresponding bound can be shown for the discrepancy of the first N elements of a (t,s)-sequence in base b, provided that $N \geq \max\{b, b^t\}$ (see [1] for details).

We remark that our new bounds are the currently best known discrepancy bounds for (t, m, s)-nets and (t, s)-sequences with respect to the leading coefficients.

From a numerical point of view, we compare, for exemplary instances, our new leading coefficients to earlier findings by Faure and Lemieux [2]. These results show that we could improve the leading factors in the discrepancy bound by a factor of roughly 2/3. Furthermore, we consider another setting for numerical comparisons, where we do not only concentrate on the leading term, but on the global discrepancy bounds, including all terms of lower order. In this case, we see that our new result is better than previous results for low dimensions, whereas the discrepancy bounds of Faure and Lemieux are favorable in higher dimensions.

References

[1] H. Faure and P. Kritzer *New star discrepancy bounds for (t, m, s)-nets and (t, s)-sequences*, Monatsh. Math. **172** (2013), 55–75.

[2] H. Faure and C. Lemieux *Improvements on the star discrepancy of (t, s)-sequences*, Acta Arith. **154** (2012), 61–78.

[3] P. Kritzer, *Improved upper bounds on the star discrepancy of (t, m, s)-nets and (t, s)-sequences*, J. Complexity **22** (2006), 336–347.

[4] H. Niederreiter, *Point sets and sequences with small discrepancy*, Monatsh. Math. **104** (1987), 273–337.

Approximation numbers of Sobolev embeddings – Sharp constants and tractability

Thomas Kühn

In the first part of this talk, optimal linear approximations (approximation numbers) in the context of periodic Sobolev spaces $H^s(\mathbb{T}^d)$ of fractional smoothness $s > 0$ on the d-dimensional torus are studied. The error is always measured in $L_2(\mathbb{T}^d)$. For several equivalent norms, including the classical one, we get the optimal decay rate of the approximation numbers (which is of course well known) and also the exact order of the constants in dependence on s and d. The approximation numbers of a bounded linear operator $T : X \to Y$ between Banach spaces are defined as

$$a_n(T) := \inf\{ \|T - A\| : rank(A) < n\} \quad , \quad n \in \mathbb{N} .$$

For compact operators between Hilberts spaces, these are just the singular numbers.

The following two-sided estimate can be found in many monographs, e.g. in [3], Chapter 2, Theorems 4.1 and 4.2,

$$c_s(d) n^{-s/d} \le a_n(I_d : H^s(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) \le C_s(d) n^{-s/d} ,$$

where the constants $c_s(d)$ and $C_s(d)$, only depending on s and d, were not explicitly determined. We show that, for fixed $s > 0$, these constants decay polynomially in

d as $d \to \infty$. The most complete results are obtained for the norm

$$\|f \mid H^s(\mathbb{T}^d)\|^{\#} = \left[\sum_{k \in \mathbb{Z}^d} |c_k(f)|^2 \left(1 + \sum_{j=1}^{d} |k_j|\right)^{2s} \right]^{1/2} ,$$

where

$$c_k(f) = (2\pi)^{-d/2} \int_{\mathbb{T}^d} f(x) e^{-ikx} \, dx$$

are the Fourier coefficients of $f \in L_2(\mathbb{T}^d)$. For this norm we show, for all $s > 0$ and every $d \in \mathbb{N}$,

$$\lim_{n \to \infty} n^{s/d} a_n(I_d : H^s(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) = \left(\frac{2}{\sqrt[d]{d!}}\right)^s \asymp \left(\frac{2e}{d}\right)^s .$$

Moreover we give estimates with specific constants for moderately large n, for instance for $n \geq 6^d$.

For small n in the so-called pre-asymptotic range $2 \leq n \leq 2^d$, we have

$$\left(\frac{1}{2 + \log_2 n}\right)^s \leq a_n(I_d : H^s(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) \leq \left(\frac{\log_2(2d+1)}{\log_2 n}\right)^s .$$

In the second part of the talk, we interpret these results in the context of information-based complexity. More precisely, we show that the approximation problem

$$I_d : H^s(\mathbb{T}^d) \to L_2(\mathbb{T}^d) \quad , \quad d \in \mathbb{N},$$

with respect to the above norm is weakly tractable, if $s > 1$, and intractable, if $0 < s < 1$. Here, as usual, weak tractabiliy means that the information complexity

$$n(\varepsilon, d) := \min\{n \in \mathbb{N} : a_n(I_d : H^s(\mathbb{T}^d) \to L_2(\mathbb{T}^d) \leq \varepsilon\}$$

satisfies

$$\lim_{1/\varepsilon + d \to \infty} \frac{\log n(\varepsilon, d)}{1/\varepsilon + d} = 0 .$$

In other words, $n(\varepsilon, d)$ does not increase exponentially, neither in $1/\varepsilon$ as $\varepsilon \to 0$, nor in d as $d \to \infty$.

In a forthcoming paper [2] we investigate similar questions for Sobolev spaces with dominating mixed smoothness.

References

[1] T. Kühn, W. Sickel and T. Ullrich, *Approximation numbers of Sobolev embeddings – Sharp constants and tractability*, J. Complexity (2013), 22 pages, Online First

[2] T. Kühn, W. Sickel and T. Ullrich, *Approximation numbers of embeddings of mixed Sobolev spaces*, Preprint 2013

[3] V. N. Temlyakov, *Approximation of Periodic Functions*, Nova Science, New York, 1993.

Discrepancy estimates for sequences: New results and open problems

Gerhard Larcher

In my talk I presented two of my recent results on discrepancy theory and two open problems connected with these results. In the first part we reported about questions concerning the discrepancy of Halton-Kronecker sequences.

A Halton-Kronecker sequence is a sequence $z_n = (x_n, y_n) \in [0,1)^{s+t}$ where $(x_n)_{n\geq 0}$ is a Halton sequence in $[0,1)^s$ and $(y_n)_{n\geq 0} = ((\{n\alpha_1\}, \ldots, \{n\alpha_t\}))_{n\geq 0}$ is a Kronecker sequence in $[0,1)^t$.

It was shown by Niederreiter [5] that this sequence is uniformly distributed in $[0,1)^{s+t}$ if and only if the Kronecker part $(y_n)_{n\geq 0}$ is uniformly distributed in $[0,1)^t$. Niederreiter also gave discrepancy estimates for the sequence $(z_n)_{n\geq 0}$ in dependence on simultaneous approximation properties of $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_t)$.

In [3] the author of this report could show the following metric result for the discrepancy of Halton-Kronecker sequences:

Theorem 1. *For almost all $\boldsymbol{\alpha} \in [0,1)^t$ the discrepancy D_N of a Halton-Kronecker sequence satisfies*

$$D_N = O\left(\frac{(\log N)^{s+t+\epsilon}}{N}\right)$$

for all $\epsilon > 0$.

The method to prove this theorem is an adaption of techniques which were developed by Beck [1] in probabilistic diophantine approximation.

Until now there is no **explicite** Halton-Kronecker sequence with $s, t \geq 1$ known which satisfies this discrepancy bound which is valid for almost all $\boldsymbol{\alpha}$.

So we rise the following open problem:

Open Problem 1. *Does the discrepancy of the following simple 2-dimensional Halton-Kronecker sequence*

$$z_n = (\phi_2(n), \{n \cdot \sqrt{2}\})_{n\geq 0}$$

where $(\phi_2(n))_{n\geq 2}$ is the van der Corput sequence in base 2 satisfy

$$D_N = O\left(\frac{(\log N)^{2+\epsilon}}{N}\right)$$

or not?

In the second part of the talk we reported on new developments concerning the best uniform distribution constant for sequences in the unit interval.

It is known from Schmidt [7] that for every sequence $(x_n)_{n\geq 0}$ in $[0,1)$ we have

$$D_N \geq c \cdot \frac{\log N}{N} \tag{1}$$

for infinitely many N. The order $\frac{\log N}{N}$ ist best possible. We ask now for c^*, the supremum over all constants c, such that (1) holds. It is known from [2] and from [6] that

$$0.06015\ldots \leq c^* \leq 0.222\ldots$$

holds.

Using and refining a method of Liardet (see [4]) we give an easier and more illustrative proof for this lower bound.

By further refining the method we can show that even $c^* \geq 0.06458\ldots$ holds.

We give some hints, how a further analysis of the method can lead to further improvements of the lower bound for c^*. Hence we state

Open Problem 2. *Improve the above estimate for c^* for example by the method indicated in this talk.*

References

[1] J. Beck. Probabilistic diophantine approximation, I. Kronecker-sequences. *Annales of Mathematics* **(140)** (1994), pp. 451–502.

[2] R. Bejian. Minoration de la discrepance d'une suite quelconque sur T . *Acta Arithmetica* **(41)** (1982), pp. 185–202.

[3] G. Larcher. Probabilistic diophantine approximation and the distribution of Halton-Kronecker sequences. J. Complexity, **(29)** (2013), pp. 397-423.

[4] P. Liardet. Discrepance sur le cercle. Primaths I, Univ. Marseille (1979), pp. 7–11.

[5] H. Niederreiter. On the discrepancy of some hybrid sequences. *Acta Arithmetica* **(138)** (2009), pp. 373–398.

[6] V. Ostromoukhov. Recent progress in improvement of extreme discrepancy and star discrepancy of one-dimensional sequences. *In: Monte-Carlo and Quasi-Monte Carlo Methods 2008*, P. L'Ecuyer and A. B. Owen (eds.), Springer, New York, (2009) pp. 561–572.

[7] W.M. Schmidt. Irregularities of distribution VII. *Acta Arith.* **(66)** (1972), 21:45–50.

The poor lattice structure of certain long-period multiple recursive generators

Pierre L'Ecuyer

(joint work with Richard Simard)

Random number generators based on linear recurrences modulo a prime number m are widely used for simulation on computers, mainly because they are very fast and their behavior is easy to characterize mathematically. In particular, they have very long periods under simple conditions, the uniformity of the set of points they produce from all possible seeds can be measured precisely in a mathematical sense, and a good uniformity can be obtained via a proper choice of parameters.

In their quest for longer periods together with high speed, some authors have proposed high-order recurrences with very special structure. In particular, in a series of nearly a dozen articles, L.-Y. Deng and his co-authors proposed various special cases of linear multiple recursive generators (MRGs) of large order k for m near 2^{31}, with extremely long periods, where the coefficients a_j satisfy certain

conditions that can make the implementation faster; see [2, 1, 3] and other references therein. In particular, to reduce the number of multiplications modulo m in the implementation, either most coefficients of the recurrence are set to zero (just a few are nonzero) or all nonzero coefficients take the same value, or at most two values.

In this talk, we study the lattice structure of the point sets produced by these special types of MRGs over their entire period. We show that this lattice structure is always bad, regardless of the choice of parameters. More specifically, for each proposed type of MRG, there is a small set of coordinates $i_1, \ldots, i_s$, where $0 = i_1 < \cdots < i_s \leq k+1$ and $3 \leq s \leq 5$, for which the set of all points of the form $(u_{n+i_1}, \ldots, u_{n+i_s})$ taken from the sequence $\{u_n,\, n \geq 0\}$ of successive output values produced by the MRG from all possible initial states, always has a lattice structure with large spacings between the hyperplanes that contain all the points, in s dimensions. We show that this can be easily detected by empirical statistical tests. We also show that due to the particular structure of these MRGs, a poor initialization of the state can have a very long-lasting impact, because of the limited diffusion capacity of the recurrence. The details will be published in [3]. This type of behavior strongly suggests that these MRGs should be avoided.

References

[1] L.-Y. Deng, J.-J. H. Shiau, and H. H.-S. Lu. Large-order multiple recursive generators with modulus $2^{31} - 1$. *INFORMS Journal on Computing*, 24:636–647, 2012.

[2] L.-Y. Deng and H. Xu. A system of high-dimensional, efficient, long-cycle and portable uniform random number generators. *ACM Transactions on Modeling and Computer Simulation*, 13(4):299–309, 2003.

[3] P. L'Ecuyer and R. Simard. On the lattice structure of a special class of multiple recursive random number generators. *INFORMS Journal on Computing*, 2014. to appear.

On SDEs with discontinuous drift

Gunther Leobacher

We consider the following autonomous stochastic differential equation (SDE)

$$dX = b(X)dt + \sigma(X)dW \tag{1}$$

It is well-known that if b and σ are both Lipschitz, then there exists a unique strong solution of the SDE.

We are interested in the case where b is not Lipschitz. For this case there exists an existence and uniqueness theorem as well:

Theorem 1 (Zvonkin [4],Veretennikov [3])**.** *Let $b : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ be bounded and measurable, let $\sigma : \mathbb{R}^d \longrightarrow \mathbb{R}^{d\times d}$ be Lipschitz and uniformly elliptic, i.e.,*

$$\exists \lambda > 0 : \forall \xi, x \in \mathbb{R}^d : \|\sigma(x)\xi\| \geq \lambda \|\xi\| \,.$$

Then there exists a unique strong solution to

$$dX = b(X)dt + \sigma(X)dW \,.$$

From some control/filtering problems arise SDEs of the general form (1) with discontinuous drift b and and with a diffusion coefficient σ that violates the uniform ellipticity condition. For example, the discontinuity could arise from a threshold strategy for paying dividends. See, for example, [1].

The same authors provide an existence and uniqueness theorem under conditions that conform to the control problem situation.

Theorem 2 (Leobacher, Szölgyenyi, Thonhauser). *Let $\sigma : \mathbb{R}^d \longrightarrow \mathbb{R}^{d\times d}$ be Lipschitz and satisfy*

$$\exists \lambda > 0 : \forall x \in \mathbb{R}^d : \|\sigma_{11}(x)^2 + \ldots + \sigma_{1d}(x)^2\| \geq \lambda \,.$$

Let $b : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ be continuous on $\mathbb{R}^d \backslash \{x \in \mathbb{R}^d : x_1 = 0\}$ and such that there exists Λ with $\|\frac{\partial^3}{\partial x_j^3} b(x)\|_\infty \leq \Lambda$ for all $x \in \mathbb{R}$ and all $j \neq 1$.

Then there exists a unique strong solution to

$$dX = b(X)dt + \sigma(X)dW \,.$$

The basic idea underlying both theorems is that the drift can be removed through a transform $Z = g(X)$, that is, g is invertible and Z is described by a SDE without drift and with Lipschitz diffusion coefficient.

From the perspective of numerical approximation this transform is of special interest since it theoretically allows the direct use of standard schemes, like Euler-Maruyama or Milstein method.

In particular, when using quasi-Monte Carlo simulation, it is necessary to remove the discontinuities to have the regularity required by these methods. However, the transform g is characterized by a differential equation, which is a d-dimensional elliptic PDE in the setup of Theorem 1 and an ODE in the setup of Theorem 2. In the latter case, the ODE can be solved by computing a double integral. But of course it is still expensive to compute the transform and even more expensive to compute its inverse.

Suppose that the left and right limits

$$\lim_{h\to 0\pm} \frac{b(h, x_2, \ldots, x_d)}{\sigma_{11}(h, x_2, \ldots, x_d)^2 + \ldots + \sigma_{1d}(h, x_2, \ldots, x_d)^2} =: c_\pm(x_2, \ldots, x_d)$$

exists and is C^3.

Define

$$g_1(t) := \begin{cases} \int_0^t \exp(c_+\tau)\, d\tau & x \geq 0 \\ \int_0^t \exp(-c_-\tau)\, d\tau & x \leq 0 \end{cases}$$

and consider the transform $g(x_1, \ldots, x_d) = (g_1(x_1), x_2, \ldots, x_d)$. Formally applying Itô's formula shows that $Z = g(X)$ is described by a SDE with Lipschitz drift. Thus there is a unique solution to this transformed SDE (which also can be approximated numerically). g is locally invertible, and applying Itô's formula to compute

In the latter case, the use of Itô's formula can be justified exactly as in [2].

Of course, g and g^{-1} can be computed explicitly, thus providing us with a practical method for solving the SDE also numerically. As an added value, the use of quasi-Monte Carlo is facilitated in that the formerly discontinuous drift is continuous in the transformed SDE.

References

[1] G. Leobacher, M. Szölgyenyi, and S. Thonhauser. Bayesian dividend optimization and finite time ruin probabilities. 2013. Submitted.

[2] G. Leobacher, M. Szölgyenyi, and S. Thonhauser. On the existence of solutions of a class of SDEs with discontinuous drift and singular diffusion. 2013. Preprint.

[3] A.JU. Veretennikov. On the strong solutions of stochastic differential equations. *Theory of Probability and its Applications*, 24(2), 1979.

[4] A. Zvonkin. A transformation of the phase space of a diffusion process that removes the drift. *Mathematics of the USSR, Sbornik*, 22(129), 1974.

Tractability results for classes of ridge functions

Sebastian Mayer

(joint work with Tino Ullrich, Jan Vybíral)

Approximating functions which depend on a large number of variables can be a hard to solve problem. In many settings, it is known to suffer from the curse of dimensionality even for very smooth functions. So, when smoothness is not enough, what additional inner structure of functions could help to overcome the curse. In this regard, we investigate *ridge functions* $f(x) = g(a \cdot x)$ in high dimensions d in our current work [3]. A ridge function is characterized by two components: a d-dimensional vector a, called the *ridge direction*, and a univariate function g, which is called the *profile*. Ridge functions provide a simple, coordinate-independent model to describe inherently one-dimensional structures hidden in a high-dimensional ambient space.

The setting. The problem we study is that of approximating ridge functions in the uniform norm in the worst-case by means of deterministic, adaptive algorithms. As information, the algorithms have only a limited amount of function values available. The function classes $\mathcal{F}$, where the ridge functions may be taken from, are introduced below. For the remainder of this text, let us call this setup the *sampling problem*. The key quantity, for which we find bounds from above and below, is the *n-th (adaptive) sampling number*

$$g_{n,d}^{\mathrm{ada}}(\mathcal{F}, L_\infty) := \inf\Big\{ \sup_{f\in\mathcal{F}} \|f - S(f)\|_\infty : \ S \text{ adaptive algorithm using } n \text{ function values}\Big\}.$$

So, let us finally introduce the ridge function classes. As domain Ω, we fix the closed Euclidean unit ball $\Omega = \bar{B}_2^d$. Then, we impose two constraints: firstly, we restrict the norm of the feasible ridge directions, namely we require $\|a\|_p \leq 1$ for some $0 < p \leq 2$; secondly, we define the feasible profiles to be those which have a certain order of *Lipschitz*-smoothness $\alpha > 0$. We denote the so obtained classes

by $\mathcal{R}_d^{\alpha,p}$. In a sense, the case $p = 2$ can be understood as all directions being possible. Letting p get smaller and smaller, then imposes a more and more strict sparsity constraint on the feasible directions.

Bounds on sampling numbers. What could an algorithm that exploits the ridge structure look like? At least for $\alpha > 1$, there is an intuitive idea: find a point where the first derivative of the profile g' is sufficiently large, do first-order Taylor around that point to recover the ridge direction a, and finally sample along a to approximate g. Unfortunately, this cannot work for ridge functions on the Euclidean unit ball as given by the class $\mathcal{R}_d^{\alpha,p}$. Let us just mention the crucial point here: whatever algorithm we take, sampling the whole relevant domain of the profile is only then possible if the algorithm samples exactly along the ridge direction. Otherwise, there is always a range the algorithm cannot reach; there is no guarantee to find a point where the first derivative g' is sufficiently large. It turns out that for functions from $\mathcal{R}_d^{\alpha,p}$, particularly in case $p = 2$, we cannot really do better than for general multi-variate Lipschitz functions. In terms of algorithms, this means that the best one can try is to spread the sampling points, in a certain sense, as uniformly as possible over the domain.

Let us formulate in more detail what bounds we get by the above reasoning. Both the lower and the upper bound on $g_{n,d}^{\mathrm{ada}}(\mathcal{R}_d^{\alpha,p}, L_\infty)$ are determined to a large extent by *(dyadic) entropy numbers* e_k. For the lower bound, it is the entropy numbers of the p-sphere in ℓ_2^d; for the upper bound, it is the entropy numbers of the domain $\bar{B}_2^d$ in $\ell_{p'}^d$, where p' is the dual index of p given by $1/\max\{1,p\} + 1/p' = 1$. In formulas, if we let k be the smallest integer such that $n \le 2^{k-1}$, then we have bounds

$$e_k(\mathbb{S}_p^{d-1}, \ell_2^d)^{2\alpha} \lesssim g_{n,d}^{\mathrm{ada}}(\mathcal{R}_d^{\alpha,p}, L_\infty) \lesssim g_{n,d}^{\mathrm{lin}}(\mathcal{R}_d^{\alpha,p}, L_\infty) \lesssim e_{k-\Delta}(\bar{B}_2^d, \ell_{p'}^d)^\alpha, \tag{1}$$

where Δ is a positive integer depending logarithmically on the dimension d and the smoothness parameter α. The behaviour of these entropy numbers is completely understood; for details, see [3, Subsection 2.3] and references therein, and [4]. If we just plug in these existing results on entropy numbers, then we obtain that sampling of ridge functions from $\mathcal{R}_d^{\alpha,2}$ is essentially as difficult as sampling of general multi-variate Lipschitz functions. For $p < 2$, the situation gets better, but still the bounds are far worse than those one would get for univariate Lipschitz functions. (Actually, we prove a slightly better lower bound than that shown in (1); this better bound is less easy to grasp, however.)

Tractability results. Translating the bounds on $g_{n,d}^{\mathrm{ada}}(\mathcal{R}_d^{\alpha,p}, L_\infty)$ into bounds on the *information complexity*

$$n(\varepsilon, d) := \inf\{n \in \mathbb{N} : g_{n,d}^{\mathrm{ada}}(\mathcal{F}, L_\infty) \le \varepsilon\},$$

we see a surprisingly diverse picture in terms of degrees of tractability. Namely, we see almost the entire spectrum of degrees of tractability as introduced in the recent monographs by Novak and Woźniakowski. To be precise, the sampling problem

(1) suffers from the curse of dimensionality if $p = 2$ and $\alpha < \infty$,
(2) never suffers from the curse of dimensionality if $p < 2$,

(3) is intractable if $p < 2$ and $\alpha \leq \frac{1}{1/p-1/2}$,
(4) is weakly tractable if $p < 2$ and $\alpha > \frac{1}{1/\max\{1,p\}-1/2}$,
(5) is quasi-polynomially tractable if $\alpha = \infty$.

Let us emphasize the case 1. Not only do we understand here that the ridge structure alone does not help to overcome the curse of dimensionality. It even tells us that neither adaptivity nor non-linearity of algorithms leads to improvements. In the remaining cases, it is less obvious what to learn from the results. Especially, for $p \leq 1$, the choice of the domain $\Omega = \bar{B}_2^d$ becomes somewhat artificial. The natural choice then would be the cube $\Omega = [-1, 1]^d$.

Actually, in a recent work [1] it has been shown that, when $\Omega = [0, 1]^d$, $\alpha > 1$, and $p \leq 1$, then the sampling problem is *polynomially tractable*, provided that the feasible ridge directions are component-wise larger than zero. The results are based on an adaptive algorithm which is basically constructed like the intuitive step-wise scheme we have sketched above. This time, it works efficiently because on $\Omega = [0, 1]^d$ it is possible to reach the whole relevant range of the profile's domain without knowing the ridge direction a priori; one just has to sample along the direction $(1, \ldots, 1)$.

We have already argued that a comparable algorithm cannot work whenever the domain is the Euclidean unit ball. But, of course, there is one remedy: make it explicit knowledge about the problem that the profiles' first derivatives are uniformly bounded away from zero in some point, say the origin. In other words, for $\alpha > 1$ and some $0 < \kappa < 1$, we add the constraint $|g'(0)| \geq \kappa$ for all feasible profiles g to the class $\mathcal{R}_d^{\alpha,p}$. In this we follow [2], where the approach has been worked out in broader generality. Now, the sampling problem becomes polynomially tractable, no matter what precise values the parameters $\alpha > 1$ and p take.

References

[1] A. Cohen, I. Daubechies, R. A. DeVore, G. Kerkyacharian, and D. Picard, *Capturing ridge functions in high dimensions from point queries*, Constr. Approx. **35**(2) (2012), 225–243.
[2] M. Fornasier, K. Schnass and J. Vybí ral, *Learning functions of few arbitrary linear parameters in high dimensions*, Found. Comput. Math. **12**(2) (2012), 229–262.
[3] S. Mayer, T. Ullrich, and J. Vybíral, *Entropy and sampling numbers of classes of ridge functions*, preprint.
[4] A. Hinrichs, and S. Mayer, *Entropy numbers of spheres in Banach and quasi-Banach spaces*, work in progress.

Construction of interlaced polynomial lattice rules with SPOD weights

Dirk Nuyens

(joint work with Josef Dick, Frances Y. Kuo, Quoc T. Le Gia, Christoph Schwab)

In this talk I start from a parametrised PDE to define an infinite dimensional integral which we want to approximate by a QMC rule. The integral is the expectation over the parameter space of a linear functional of the solution of the PDE. It turns out that the regularity of the solution in the parameter domain can be transferred

to the integration problem in the form of SPOD weights. We construct interlaced polynomial lattice rules for this setting using a fast CBC algorithm. This is a report on joint work with J. Dick, F. Y. Kuo, Q. T. Le Gia and Ch. Schwab [1].

More specifically we consider a parametrized PDE

$$A(y(\omega))\, u(x; y(\omega)) = f,$$

where A is a bounded linear operator parametrized by $y(\omega) \in \mathbb{R}^{\mathbb{N}}$. We are interested in

$$\mathbb{E}[G(u)] = \int_{\Omega} G(u(\cdot; y(\omega)))\, \mathrm{d}\mathbb{P}(\omega) = \int_{[0,1]^{\mathbb{N}}} F(y)\, \mathrm{d}y,$$

using a uniform distribution for y and assume G is a bounded linear operator. We further assume an affine parameter dependence

$$A(y) = A_0 + \sum_{j \geq 1} y_j\, A_j,$$

for which there is a $p \in (0, 1]$ such that

$$\sum_{j \geq 1} \|A_j\|^p < \infty.$$

We solve for u using a QMC Galerkin method

$$u(x; y) \approx u^h(x; y) \approx u^h_s(x; y),$$

with finite number of dimensions s, and approximate the expectation by a QMC rule

$$\mathbb{E}[G(u)] \approx \mathbb{E}[G(u^h)] \approx \mathbb{E}[G(u^h_s)] \approx \frac{1}{N} \sum_{k=1}^{N} G(u^h_s(y_k)).$$

In [1] we show how to construct "interlaced polynomial lattice rules" of order $\alpha = \lfloor 1/p \rfloor + 1$ with "SPOD weights", using a fast component-by-component algorithm, in $O(\alpha s N \log N + \alpha^2 s^2 N)$ operations, which achieve a convergence rate of $O(N^{-1/p})$ with the implied constant independent of s.

The SPOD weights, which stands for "smoothness-driven product and order dependent weights", take the form, for $\mathfrak{u} \subseteq \{1, \ldots, s\}$,

$$\gamma_{\mathfrak{u}} := \sum_{\nu_{\mathfrak{u}} \in \{1:\alpha\}^{|\mathfrak{u}|}} |\nu_{\mathfrak{u}}|! \prod_{j \in \mathfrak{u}} \left(2^{\delta(\nu_j, \alpha)} \beta_j^{\nu_j}\right),$$

making use of a bound on the regularity of F in terms of the parameter y

$$|(\partial_y^{\nu} F)(y)| \leq c\, |\nu|!\, \beta^{\nu} \quad \text{for all } \nu \in \mathbb{N}_0^{\mathbb{N}} \text{ with } |\nu| < \infty,$$

with $|\nu| = \sum_{j \geq 1} \nu_j$.

References

[1] J. Dick, F. Y. Kuo, Q. T. Le Gia, D. Nuyens, and Ch. Schwab, *Higher order QMC Galerkin discretization for parametric operator equations*, submitted, 2013.

Tractability of multivariate integration

Friedrich Pillichshammer

We study multivariate integration $I_s(f) = \int_{[0,1]^s} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ of functions f from some normed space $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$ of functions defined on $[0,1]^s$ by means of quasi-Monte Carlo (QMC) rules $Q_{N,s}(f) = (1/N)\sum_{n=1}^{N} f(\boldsymbol{t}_n)$ for $f \in \mathcal{H}$ and sample nodes $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N \in [0,1)^s$. We consider the worst-case error of a QMC rule.

Definition 1. *The* **worst-case (QMC) error** *of $Q_{N,s}$ based on the node set $\mathcal{P}$ is*

$$e(\mathcal{H}, \mathcal{P}) = \sup_{\|f\|_{\mathcal{H}} \le 1} |I_s(f) - Q_{N,s}(f)|.$$

The **initial error** *(used as reference value) is $e(\mathcal{H}, \emptyset) = \sup_{\|f\|_{\mathcal{H}} \le 1} |I_s(f)| = \|I_s\|$.*

Definition 2.

- *For $N, s \in \mathbb{N}$ the N***th minimal error** *is $e_{\mathcal{H}}(N,s) = \inf_{\substack{\mathcal{P} \subseteq [0,1]^s \\ \#\mathcal{P} = N}} e(\mathcal{H}, \mathcal{P})$.*
- *For $s \in \mathbb{N}$ and $\varepsilon \in (0,1)$ the* **(QMC) information complexity** *is*

$$N_{\mathcal{H}}(s, \varepsilon) = \min\{N \in \mathbb{N} \ : \ e_{\mathcal{H}}(N,s) \le \varepsilon\, e(\mathcal{H}, \emptyset)\}.$$

In many cases one has excellent asymptotic bounds on the Nth minimal error. For example, consider

$$\mathcal{F}^*_{s,q} = \{f \in \mathcal{W}^{\mathbf{1}}_q : f(\boldsymbol{x}) = 0 \ \text{ if } x_j = 1 \ \text{ for some } j \ \text{ and } \|f\|^*_{s,q} < \infty\},$$

where $\mathcal{W}^{\mathbf{1}}_q$ is the Sobolev space of functions defined on $[0,1]^s$ that are once differentiable in each variable and whose derivatives have finite L_q norm, with norm

$$\|f\|^*_{s,q} = \left\| \frac{\partial^s}{\partial \boldsymbol{x}} f \right\|_{L_q}, \ \text{where} \quad \frac{\partial^s}{\partial \boldsymbol{x}} f = \frac{\partial^s}{\partial x_1 \partial x_2 \ldots \partial x_s} f.$$

Then it is well-known that

$$e(\mathcal{F}^*_{s,q}, \mathcal{P}) = L_p(\mathcal{P}) \quad \text{for } \ 1/p + 1/q = 1,$$

where $L_p(\mathcal{P})$ is the L_p-discrepancy of the sample point set $\mathcal{P}$. From asymptotic results on the L_p discrepancy it follows that

$$e_{\mathcal{F}^*_{s,q}}(N,s) \asymp_{s,p} \frac{(\log N)^{(s-1)/2}}{N} \quad \text{for } q > 1$$

and

$$\frac{(\log N)^{(s-1)/2+\eta_s}}{N} \ll_s e_{\mathcal{F}^*_{s,1}}(N,s) \ll_s \frac{(\log N)^{s-1}}{N} \ \text{ for some } \ \eta_s \in (0, 1/2].$$

These results are excellent in an asymptotic sense for $N \to \infty$. However, often it is not clear how long one has to wait to see the good asymptotic behavior of the Nth minimal error, especially if the dimension s is large. This is the subject of tractability which means, roughly speaking, that the information complexity lacks a certain disadvantageous dependence on ε^{-1} and s, the curse of dimensionality.

Definition 3. *We have the* **curse of dimensionality** *if $\exists c, \tau > 0$ and ε_0 such that $N_{\mathcal{H}}(\varepsilon, s) \ge c\,(1+\tau)^s$ for all $\varepsilon \le \varepsilon_0$ and infinitely many s.*

It is known that many multivariate problems defined over standard spaces of functions suffer from the curse of dimensionality. Examples are integration of Lipschitz functions, of smooth one-periodic functions, of functions from $\mathcal{F}^*_{s,2}$, of monotone or convex functions .... The reason for this may be in the fact that for standard spaces all variables and groups of variables are equally important. As a way out Sloan and Woźniakowski suggested to consider weighted spaces, in which the importance of successive variables and groups of variables is monitored by corresponding weights, to vanquish the curse of dimensionality.

Definition 4. *We say that we have:*

- **Weak (QMC) Tractability (WT)** *if* $\lim_{s+\varepsilon^{-1}\to\infty} \frac{\log\ N_{\mathcal{H}}(\varepsilon,s)}{s+\varepsilon^{-1}} = 0$.
- **Polynomial (QMC) Tractability (PT)** *if* $\exists c, \tau_1, \tau_2 > 0$ *such that* $N_{\mathcal{H}}(\varepsilon, s) \le c\, s^{\tau_1}\, (\varepsilon^{-1})^{\tau_2}$ *for all* $s \in \mathbb{N}$, $\varepsilon \in (0,1)$.
- **Strong Polynomial (QMC) Tractability (SPT)** *if* $\exists c, \tau > 0$ *such that* $N_{\mathcal{H}}(\varepsilon, s) \le c\, (\varepsilon^{-1})^{\tau}$ *for all* $s \in \mathbb{N}$, $\varepsilon \in (0,1)$. *The exponent* τ^* *of SPT is the infimum of* τ *for which SPT holds.*

For example, consider weighted Korobov spaces $\mathcal{H} = \mathcal{H}(K_s)$ of periodic and smooth functions, defined via a reproducing kernel of the form

$$K_s(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\boldsymbol{h}\in\mathbb{Z}^s} \rho_{\boldsymbol{h}} \exp(2\pi \mathtt{i} \boldsymbol{h}\cdot(\boldsymbol{x}-\boldsymbol{y})) \quad \text{for all} \quad \boldsymbol{x}, \boldsymbol{y} \in [0,1]^s,$$

where $\rho_{\boldsymbol{h}} > 0$ and $\sum_{\boldsymbol{h}\in\mathbb{Z}^s} \rho_{\boldsymbol{h}} < \infty$ in order that the kernel is well defined. The smoothness of $f \in \mathcal{H}(K_s)$ is determined by the decay of the $\rho_{\boldsymbol{h}}$'s.

A very well studied case are Korobov spaces of **finite** smoothness $\alpha > 1$ and weights $\boldsymbol{\gamma} = \{\gamma_j\}_{j\ge 1}$. Here for $\boldsymbol{h} = (h_1, \ldots, h_s) \in \mathbb{Z}^s$ one puts $\rho_{\boldsymbol{h}} = \prod_{j=1}^s r_{\alpha,\gamma_j}(h_j)$, where $r_{\alpha,\gamma}(h) = \gamma |h|^{-\alpha}$ if $h \ne 0$ and $r_{\alpha,\gamma}(0) = 1$. The $\rho_{\boldsymbol{h}}$'s decay polynomially in the components of $\boldsymbol{h}$. We write $K_s = K_{s,\alpha,\boldsymbol{\gamma}}$. The parameter α guarantees the existence of some partial derivatives and the weights $\boldsymbol{\gamma}$ model the influence of the different components on the variation of $f \in \mathcal{H}(K_{s,\alpha,\boldsymbol{\gamma}})$.

Theorem 1 (Sloan & Woźniakowsk)**.** *For QMC integration in* $\mathcal{H}(K_{s,\alpha,\boldsymbol{\gamma}})$ *we have:*

- *WT, iff* $\lim_{s\to\infty} \frac{\sum_{j=1}^s \gamma_j}{s} = 0$.
- *PT, iff* $\limsup_{s\to\infty} \frac{\sum_{j=1}^s \gamma_j}{\log s} < \infty$.
- *SPT, iff* $\sum_{j=1}^\infty \gamma_j < \infty$. *If so, then the exponent of SPT satisfies*

$$\tau^* \le 2 \inf \left\{ \lambda \in (1/\alpha, 1] \ : \ \sum_{j=1}^\infty \gamma_j^\lambda < \infty \right\}.$$

In 2011 Dick, Kritzer, Larcher, Woźniakowski and the author started a series of papers where we study integration and approximation of functions from Korobov spaces of **infinite** smoothness. Fix $\omega \in (0,1)$ and put $\rho_{\boldsymbol{h}} = \omega^{\sum_{j=1}^s a_j |h_j|^{b_j}}$ for $\boldsymbol{h} = (h_1, h_2, \ldots, h_s) \in \mathbb{Z}^s$, with weight sequences $\boldsymbol{a} = \{a_j\}_{j\ge 1}$ and $\boldsymbol{b} = \{b_j\}_{j\ge 1}$ where $0 < a_1 \le a_2 \le \ldots$ and $\inf b_j > 0$. Now the $\rho_{\boldsymbol{h}}$'s decay exponentially fast

in the components of $\boldsymbol{h}$. We write $K_s = K_{s,\boldsymbol{a},\boldsymbol{b}}$. Functions from $\mathcal{H}(K_{s,\boldsymbol{a},\boldsymbol{b}})$ are analytic.

For infinite smoothness it is natural to demand more of the Nth minimal errors and of the information complexity than for finite smoothness. We are interested in obtaining uniform exponential convergence of $e_{\mathcal{H}}(N,s)$.

Definition 5. Uniform exponential convergence (UEXP) *means that* $\exists q \in (0,1)$, $p>0$ *and a function* $C:\mathbb{N}\to(0,\infty)$ *such that*

$$e_{\mathcal{H}}(N,s) \le C(s)\, q^{N^p} \quad \text{for all} \quad s,N\in\mathbb{N}.$$

UEXP implies $N_{\mathcal{H}}(\varepsilon,s) \le \left\lceil \left(\frac{\log C(s)+\log \varepsilon^{-1}}{\log q^{-1}}\right)^{1/p}\right\rceil$ for all $s\in\mathbb{N}$, $\varepsilon\in(0,1)$. Hence we need $O([\log \varepsilon^{-1}]^{1/p})$ function values to compute an ε-approximation to multivariate integrals. This is excellent, but we do not know how long we have to wait to see this asymptotic behavior, especially for large s. In this context we study the following notions of tractability:

Definition 6. *We say that we have:*

- **EC weak tractability (EC-WT)** *if* $\lim_{\log\varepsilon^{-1}+s\to\infty} \frac{\log N_{\mathcal{H}}(\varepsilon,s)}{\log\varepsilon^{-1}+s} = 0$.
- **EC polynomial tractability (EC-PT)** *if* $\exists c,\tau_1,\tau_2>0$ *such that* $N_{\mathcal{H}}(\varepsilon,s)\le c\, s^{\tau_1}\left(1+\log\varepsilon^{-1}\right)^{\tau_2}$ *for all* $s\in\mathbb{N}$, $\varepsilon\in(0,1)$.
- **EC strong polynomial tractability (EC-SPT)** *if* $\exists c,\tau>0$ *such that* $N_{\mathcal{H}}(\varepsilon,s)\le c\,(1+\log\varepsilon^{-1})^{\tau}$ *for all* $s\in\mathbb{N}$, $\varepsilon\in(0,1)$. *The exponent* τ^* *of EC-SPT is the infimum of* τ *for which EC-SPT holds.*

Theorem 2 (Kritzer, Pillichshammer & Woźniakowski)**.** *For QMC integration in* $\mathcal{H}(K_{s,\boldsymbol{a},\boldsymbol{b}})$ *we have:*

- *UEXP, iff* $B:=\sum_{j=1}^{\infty}\frac{1}{b_j}<\infty$*; if so then* $p=1/B$.
- *EC-PT* $\Rightarrow$ *UEXP*
- *EC-WT* $\Rightarrow$ $\lim_{N\to\infty} N^{\alpha} e_{\mathcal{H}(K_{s,\boldsymbol{a},\boldsymbol{b}})}(N,s)=0$ *for all* $\alpha>0$.
- *We have EC-WT* $\Leftrightarrow$ $\lim_{j\to\infty} a_j=\infty$ *and EC-WT+UEXP* $\Leftrightarrow$ $B<\infty$ *and* $\lim_{j\to\infty} a_j=\infty$.
- *EC-PT, EC-PT+UEXP, EC-SPT and EC-SPT+UEXP are equivalent.*
- *EC-SPT+UEXP holds iff* $B:=\sum_{j=1}^{\infty}\frac{1}{b_j}<\infty$ *and* $\alpha^*:=\liminf_{j\to\infty}\frac{\log a_j}{j}>0$. *The exponent of EC-SPT satisfies* $\tau^*\in[B,B+\min(B,(\log 3)/\alpha^*)]$.

The current state of the art of tractability theory is summarized in [1, 2, 3].

References

[1] Novak, E. and Woźniakowski, H.: *Tractability of Multivariate Problems, Volume I: Linear Information*. EMS, Zurich, 2008.

[2] Novak, E. and Woźniakowski, H.: *Tractability of Multivariate Problems, Volume II: Standard Information for Functionals*. EMS, Zurich, 2010.

[3] Novak, E. and Woźniakowski, H.: *Tractability of Multivariate Problems, Volume III: Standard Information for Operators*. EMS, Zurich, 2012.

Automatic integration using asymptotically optimal subdivision strategy

Leszek Plaskota

A common approach to the numerical approximation of the integral

$$I(f) = \int_a^b f(x)\,\mathrm{d}x$$

of a function $f : [a, b] \to \mathbb{R}$ relies on using composite quadrature rules. Ideally, we would like to have an automatic routine that, for a given function f and error tolerance ε, produces an approximation $Q(f)$ to $I(f)$ such that it uses as little function evaluations as possible, and its error

$$|I(f) - Q(f)| \le \varepsilon.$$

This is usually realized with the help of *adaption*. For a given interval, two simple quadrature rules are applied, one more accurate than the other. If the difference between them is sufficiently small, the integral in this interval is approximated by the more accurate quadrature. Otherwise, the interval is divided into two smaller subintervals and the above rule is recursively applied for each of the subintervals. This often results in a more efficient final subdivision of $[a, b]$ than the uniform subdivision of nonadaptive quadratures. See, e.g., [1] for an account on adaptive numerical integration.

We have to admit that there are results saying that adaptive quadratures are not better than nonadaptive quadratures. This holds, for instance, in the *worst case setting* over convex and symmetric classes of functions. There are also corresponding *adaption-does-not-help* results in other settings, see, e.g., [6, 7, 10, 11]. On the other hand, if the class is not convex and/or a different from the worst case error criterion is used to compare algorithms, then adaption can significantly help. A characteristic example is integration of piecewise smooth functions [9]. For those reasons, adaptive algorithms are frequent guests in computational practice. Adaptive *Simpson* quadratures or those based on four-point *Gauss-Lobatto* rule and its seven-point *Kronrod* extension [2, 8] are now standard elements of numerical packages such as QUADPACK or MATLAB.

Adaptive Simpson quadrature, first published in algorithm form in [5], is the oldest and probably most known example of automatic integration. Some modifications of the standard algorithm, like those in [4], were proposed to make it applicable for as many functions as possible. However, to the author's knowledge, there is no comprehensive analysis explaining the behavior of adaptive quadratures in a quantitative way. A common knowledge is that "adaptive quadratures do very well for rapidly varying functions".

In this talk, we analyze the adaptive Simpson quadratures from the point of view of computational complexity. Allowing *all* subdivision strategies, the goal is to find *optimal* strategy, for which the corresponding algorithm returns an ε-approximation to the integral using minimal number of integrand evaluations or, equivalently, the minimal number of subintervals. The main analysis is asymptotic

and done under the assumptions that f is four times continuously differentiable, $f^{(4)} \in ([a,b])$, and that its 4th derivative is positive.

To reach our goal, we first derive formulas for the asymptotic error of adaptive Simpson quadratures which were first presented in the unpublished master thesis [3]. We find that the optimal strategy produces the partition $a = x_0^* < \cdots < x_m^* = b$ such that

$$\int_a^{x_i^*} \left(f^{(4)}(x)\right)^{1/5} \mathrm{d}x \;=\; \frac{i}{m} \int_a^b \left(f^{(4)}(x)\right)^{1/5} \mathrm{d}x, \qquad i = 0, 1, \ldots, m.$$

This, up to a constant, is practically realized by the adaptive subdivision strategy that keeps the error on successive subintervals on the same level. The optimal error corresponding to the subdivision into m subintervals is then proportional to $L^{\mathrm{opt}}(f)\, m^{-4}$ where

$$L^{\mathrm{opt}}(f) = \left(\int_a^b \left(f^{(4)}(x)\right)^{1/5} \mathrm{d}x \right)^5.$$

For comparison, the errors for the standard adaptive and for nonadaptive (using uniform subdivision) quadratures are respectively proportional to $L^{\mathrm{std}}(f)\, m^{-4}$ and $L^{\mathrm{non}}(f)\, m^{-4}$ where

$$L^{\mathrm{std}}(f) = (b-a) \left(\int_a^b \left(f^{(4)}(x)\right)^{1/4} \mathrm{d}x \right)^4, \; L^{\mathrm{non}}(f) = (b-a)^4 \left(\int_a^b f^{(4)}(x)\, \mathrm{d}x \right).$$

Obviously, $L^{\mathrm{opt}}(f) \le L^{\mathrm{std}}(f) \le L^{\mathrm{non}}(f)$. Hence the optimal Simpson quadrature is especially effective when $L^{\mathrm{opt}}(f) \ll L^{\mathrm{std}}(f)$. An example is provided by the integral $\int_\delta^1 x^{-1/2}\, \mathrm{d}x$ with 'small' δ. E.g., if $\delta = 10^{-8}$ then $L^{\mathrm{opt}}(f)$, $L^{\mathrm{std}}(f)$, and $L^{\mathrm{non}}(f)$ are correspondingly of order 10^5, 10^8, and 10^{28}.

We show that the optimal strategy can be harnessed to automatic integration. The only serious problem of how to choose the acceptable error ε_1 for subintervals to obtain the final error ε with minimal cost is resolved by splitting the recursive subdivision process into two phases. In the first phase, the process is run with the acceptable error set to a 'test' level $\varepsilon_2 = \varepsilon$. Then the acceptable error is updated to

$$\varepsilon_1 = \varepsilon\, m_2^{-5/4}$$

where m_2 is the number of subintervals obtained from the first phase. In the second phase, the recursive subdivision is continued with the 'target' level ε_1.

As we noticed earlier, the main analysis is provided under the assumption that $f \in C^4([a,b])$ and $f^{(4)} > 0$. It turns out that, using additional arguments, the obtained results can be extended to functions with $f^{(4)} \ge 0$ and/or possible endpoint singularities, i.e., when $f^{(4)}(x)$ goes to $+\infty$ as x approaches a or b. The latter extension is especially important as endpoint singularities regularly appear when one approximates a weighted integral over an infinite interval by switching to a finite interval. For instance, for the integral $\int_0^\infty f(x) e^{-x}\, \mathrm{d}x$ one can apply the change of variables $u = e^{-x}$ to transform it to $\int_0^1 F(u)\, \mathrm{d}u$ with $F(u) = -f(-\ln u)$, which

usually results in a singularity at $u = 0$. For such integrals, the optimal strategy works perfectly well while the other two quadratures may lose the convergence rate m^{-4}.

The assumption $f^{(4)} \geq 0$ is of course restrictive. However, it allows us to exploit the full power of adaptive Simpson quadratures without introducing any additional safety measures to avoid possible premature algorithm terminations, which may happen as a result of inaccurate estimations of local errors. Actually, the problem of how to effectively deal with functions for which $f^{(4)}$ changes sign is known to be non-trivial and should be studied separately.

We stress that our technique for deriving the optimal automatic integration is by no means restricted to Simpson quadratures and can be also applied when other quadratures are used as basic components of composite rules.

References

[1] P.J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, 2nd ed., Academic Press, Orlando, 1984.

[2] W. Gander and W. Gautschi, *Adaptive quadrature - revisited*, BIT **40** (2000), 84–101.

[3] A. Kruk, *Is the adaptive Simpson quadrature optimal?* Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Master Thesis, 2012 (in Polish).

[4] J.N. Lyness, *Notes on the adaptive Simpson quadrature routine*, J. Assoc. Comput. Mach. **16** (1969), 483–495.

[5] W.M. McKeeman, *Algorithm 145: Adaptive numerical integration by Simpson's rule*, Commun. ACM **5** (1962), 604.

[6] E. Novak, *On the power of adaption*, J. Complexity **12** (1996), 199–238.

[7] L. Plaskota, *Noisy Information and Computational Complexity*, Cambridge University Press, Cambridge, 1996.

[8] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed., Cambridge University Press, New York, 2007.

[9] L. Plaskota and G.W. Wasilkowski, *Adaption allows efficient integration of functions with unknown singularities*, Numerische Mathematik **102** (2005), 123–144.

[10] J.F. Traub, G.W. Wasilkowski and H. Woźniakowski, *Information-Based Complexity*, Academic Press, New York, 1988.

[11] G.W. Wasilkowski, *Information of varying cardinality*, J. Complexity **1** (1986), 107–117.

Multivariate approximation with trigonometric polynomials

Daniel Potts

A straightforward discretisation of problems in d spatial dimensions with $N = 2^n$ grid points in each coordinate leads to an exponential growth N^d in the number of degrees of freedom. Even an efficient algorithm like the d-dimensional fast Fourier transform (FFT) uses $C\,d\,N^d \log N$ floating point operations. This is labelled as the curse of dimensions and the use of sparsity has become a very popular tool in such situations. For moderately high dimensional problems the use of sparse grids and the approximation on hyperbolic crosses has led to problems of total size $C_d N \log^{d-1} N$. Moreover, the approximation rate hardly deteriorates for functions in an appropriate scale of spaces of dominating mixed smoothness [12].

The FFT has been adapted to this thin discretisation as hyperbolic cross fast Fourier transform (HCFFT), which uses $C_d N \log^d N$ floating point operations, see [1, 4, 5]. However, these classical sparse grid discretisations are numerically unstable as shown in [8]. On the other hand, lattice rules are well known for the integration of functions of many variables [12, 3]. We discuss rank-1 lattice rules as spatial discretisation for the hyperbolic cross FFT. For given $M \in \mathbb{N}$ and $\mathbf{z} \in \mathbb{Z}^d$, we consider the *rank-1 lattice*

$$\mathcal{X}(\mathbf{z}, M) := \{\mathbf{x}_j = j\mathbf{z}/M \bmod 1,\ j = 0, \dots, M-1\} \subset \mathbb{T}^d.$$

The evaluation of trigonometric polynomials supported on a hyperbolic cross, i.e., the mapping from the hyperbolic cross in frequency domain to the rank-1 lattice in spatial domain reduces to a single one dimensional FFT and thus can be computed very efficiently and stable. For the inverse transform, mapping the samples of a trigonometric polynomial to its Fourier coefficients on the hyperbolic cross, we discuss the recently presented necessary and sufficient conditions on rank-1 lattices allowing a stable reconstruction of trigonometric polynomials supported on hyperbolic crosses and the generalisation to arbitrary index sets in the frequency domain. We suggest approaches for determining suitable rank-1 lattices using a component-by-component algorithm [7, 2]. In conjunction with numerical found lattices, we show that this new method outperforms the classical hyperbolic cross FFT for realistic problem sizes [9].

The use of a generalisation of rank-1 lattices as spatial discretisations offers an additional suitable possibility for sampling sparse trigonometric polynomials. To this end, we define for given $M \in \mathbb{N}$ and $\mathbf{r} \in \mathbb{R}^d$ the *generated set*

$$\Lambda(\mathbf{r}, M) := \{\mathbf{x}_j = j\mathbf{r} \bmod 1, j = 0, \dots, M-1\} \subset \mathbb{T}^d.$$

The fast computation of trigonometric polynomials on generated sets, can be realised by using the nonequispaced fast Fourier transform (NFFT), cf. [11]. A simple sufficient condition on a generated set $\Lambda(\mathbf{r}, M)$ allows the fast, unique and stable reconstruction of the frequencies of a d-dimensional trigonometric polynomial from its samples along $\Lambda(\mathbf{r}, M)$. In contrast to searching for suitable rank-1 lattices, we can use continuous optimisation methods in order to determine generated sets that are suitable for reconstruction, see [6].

Finally we discuss the approximation with lattice rules of functions in periodic Sobolev spaces of dominating mixed smoothness [10], similar as recently presented in [5], where sampling on generalised sparse grids were used.

This contribution is joint work with Lutz Kämmerer and Toni Volkmer (Universität Chemnitz, Germany).

References

[1] G. Baszenski and F.-J. Delvos. A discrete Fourier transform scheme for Boolean sums of trigonometric operators. In C. K. Chui, W. Schempp, and K. Zeller, editors, *Multivariate Approximation Theory IV*, ISNM 90, pages 15 – 24. Birkhäuser, Basel, 1989.

[2] R. Cools, F. Y. Kuo, and D. Nuyens. Constructing lattice rules based on weighted degree of exactness and worst case error. *Computing*, 87:63 – 89, 2010.

[3] J. Dick, F. Y. Kuo, and I. H. Sloan. High-dimensional integration: The quasi-monte carlo way. *Acta Numer.*, 22:133 – 288, 2013.
[4] M. Döhler, L. Kämmerer, S. Kunis, and D. Potts. NHCFFT, Matlab toolbox for the nonequispaced hyperbolic cross FFT. `http://www.tu-chemnitz.de/~lkae/nhcfft`.
[5] M. Griebel and J. Hamaekers. Fast discrete Fourier transform on generalized sparse grids. Univ. Bonn, INS Preprint No. 1305, 2013.
[6] L. Kämmerer. Reconstructing multivariate trigonometric polynomials by sampling along generated sets. *Monte Carlo and Quasi-Monte Carlo Methods 2012*, 2012.
[7] L. Kämmerer. Reconstructing hyperbolic cross trigonometric polynomials by sampling along rank-1 lattices. *SIAM J. Numer. Anal.*, 2013. in print.
[8] L. Kämmerer and S. Kunis. On the stability of the hyperbolic cross discrete Fourier transform. *Numer. Math.*, 117:581 – 600, 2011.
[9] L. Kämmerer, S. Kunis, and D. Potts. Interpolation lattices for hyperbolic cross trigonometric polynomials. *J. Complexity*, 28:76 – 92, 2012.
[10] L. Kämmerer, D. Potts, and T. Volkmer. Approximation of multivariate functions by trigonometric polynomials based on rank-1 lattice sampling. *Preprint 145, DFG Priority Program 1324*, 2013.
[11] J. Keiner, S. Kunis, and D. Potts. Using NFFT3 - a software library for various nonequispaced fast Fourier transforms. *ACM Trans. Math. Software*, 36:Article 19, 1 – 30, 2009.
[12] V. N. Temlyakov. *Approximation of Periodic Functions*. Computational Mathematics and Analysis Series. Nova Science Publishers Inc., Commack, NY, 1993.

Multi-level Monte Carlo algorithms

Klaus Ritter

In this talk we present a survey on multi-level Monte Carlo algorithms. In the first part we outline the basic idea of the stochastic multi-level approach, which may be considered as a generalization of control variates, a well-known variance reduction technique. We also provide a comparision with the classical, single-level approach, as well as pointers to the original work by Heinrich (1998), Kebaier (2005), and Giles (2008).

The second, main part of the talk is devoted to quadrature problems on the sequence space $D^{\mathbb{N}}$ with respect to a product probability measure. The corresponding function classes are unit balls in reproducing kernel Hilbert spaces, with kernel being weighted superpositions of tensor products of kernels on $D \times D$. We discuss upper and lower bounds for minimal errors as well as the role of (randomized) quasi-Monte Carlo rules as building blocks for the multi-level construction.

Discrepancy bounds of Markov chain quasi-Monte Carlo

Daniel Rudolf
(joint work with J. Dick)

Markov chain Monte Carlo sampling can be represented via $X_{i+1} = \varphi(X_i, U_i)$ for $i \geq 2$, with $X_1 = \psi(U_1)$ and the $U_i \sim U(0,1)^s$ are i.i.d. The state X_i is an element in $G \subseteq \mathbb{R}^d$, the function $\varphi\colon G \times [0,1]^s \to G$ is called update function and $\psi\colon [0,1]^s \to G$ is called generator function. The update function gives a transition kernel, say K. For $f\colon G \to \mathbb{R}$ let $\mathbb{E}_\pi(f) = \int_G f(x)\pi(\mathrm{d}x)$ be the desired mean and

$Pf(x) = \int_G f(y)K(x, \mathrm{d}y)$ be the Markov operator induced by the transition kernel K. We assume that the transition kernel is reversible with respect to a distribution π and that it is variance bounding, see [3]. Roughly, a Markov chain is variance bounding if the asymptotic variances for functionals with unit stationary variance are uniformly bounded. Equivalent to this is the assumption that $\Lambda < 1$ with

$$\Lambda = \sup\{\lambda \in \sigma(P - \mathbb{E}_\pi \mid L_2)\} \tag{1}$$

where $\sigma(P - \mathbb{E}_\pi \mid L_2)$ denotes the spectrum of $P - \mathbb{E}_\pi$ on L_2. For example let us consider the two state Markov chain which always jumps from one state to the other one. It is periodic and satisfies $\Lambda = -1$, thus it is variance bounding. With this toy example in mind let us point out that the Markov chain does not need to be uniformly or geometrically ergodic, it might even be periodic, and the distribution of X_i, for i arbitrarily large, is not necessarily close to π.

By a deterministic sequence $(u_i)_{i\geq 0}$ we generate the deterministic Markov chain $(x_i)_{i\geq 1}$ with $x_1 = \psi(u_0)$ and $x_i = \varphi(x_{i-1}, u_{i-1})$ where $i \geq 2$. The efficiency of this procedure is measured by the star-discrepancy, a generalized Kolmogorov-Smirnov test, between the stationary measure π and the empirical distribution $\widehat{\pi}_n(A) = \frac{1}{n}\sum_{i=1}^n 1_{x_i\in A}$, where $1_{x_i\in A}$ is the indicator function of a set $A \subseteq G$. We define the star-discrepancy $D^*_{\mathcal{A},\pi}$ of $S_n = \{x_1, \dots, x_n\}$ as the supremum of $|\pi(A) - \widehat{\pi}(A)|$ over all $A \in \mathcal{A}$, i.e.

$$D^*_{\mathcal{A},\pi}(S_n) = \sup_{A\in\mathcal{A}} |\widehat{\pi}(A) - \pi(A)|\,,$$

where $\mathcal{A}$ denotes a set of subsets of G. By inverting the iterates of the update function we also define a push-back discrepancy of the driver sequence (the test sets are pushed back). We show that for large $n \in \mathbb{N}$ both discrepancies are close to each other.

The main result is an estimate of $D^*_{\mathcal{A},\pi}(S_n)$ under the assumption that we have an approximation of $\mathcal{A}$, for $\delta > 0$ given by a so-called δ-cover Γ_δ of $\mathcal{A}$ with respect to π. This is based on a Hoeffding inequality for Markov chains. A sufficiently good δ-cover exists if π is absolutely continuous with respect to the Lebesgue measure and the set of test sets is the set of open axis-parallel boxes restricted to G anchored at $-\infty$, i.e. we consider

$$\mathcal{B} = \{(-\infty, x) \cap G \colon x \in \mathbb{R}^d\},$$

with $(-\infty, x) = \Pi_{i=1}^d(-\infty, x_i)$. By a Koksma-Hlawka inequality we have

$$\left|\mathbb{E}_\pi(f) - \frac{1}{n}\sum_{i=1}^n f(x_i)\right| \leq \|f\|_{H_1} D^*_{\mathcal{B},\pi}(S_n).$$

Thus a bound on the discrepancy leads to an error bound for the approximation of $\mathbb{E}_\pi(f)$. In particular, in [1] we show for all $n \geq 16$ that there exists a driver sequence $u_0, \dots, u_{n-1} \in [0, 1]^s$ such that $S_n = \{x_1, \dots, x_n\}$ given by

$$\begin{aligned} x_1 &= \psi(u_0) \\ x_{i+1} &= \varphi(x_i; u_i), \quad i = 1, \dots, n-1, \end{aligned}$$

satisfies

$$D^*_{\mathcal{B},\pi}(S_n) \le \sqrt{\frac{1+\Lambda_0}{1-\Lambda_0}} \cdot \frac{\sqrt{2}\,(\log \|\frac{d\nu}{d\pi}\|_2 + d\log n + 3d^2\log(5d))^{1/2}}{\sqrt{n}} + \frac{8}{n^{3/4}},$$

with $\nu = \mathbb{P}_\psi$, the probability measure induced by ψ, and $\Lambda_0 = \max\{\Lambda, 0\}$ where Λ is defined in (1). This improves the result of [2] to Markov chains which satisfy a significantly weaker convergence condition than uniform ergodicity. Further by the Koksma-Hlawka inequality we have that the sample average converges to the mean with $\mathcal{O}(n^{-1/2}(\log n)^{1/2})$.

References

[1] J. Dick and D. Rudolf, *Discrepancy estimates for variance bounding Markov chain quasi-Monte carlo*, submitted (2013).

[2] J. Dick, D. Rudolf, and H. Zhu, *Discrepancy bounds for uniformly ergodic Markov chain quasi-Monte carlo*, Preprint, Available at http://arxiv.org/abs/1303.2423 (2013).

[3] G. Roberts and J. Rosenthal, *Variance bounding Markov chains*, Ann. Appl. Probab. **18** (2008), no. 3, 1201–1214.

Lower bounds and separation for spherical designs

Ian H. Sloan

(joint work with Peter Grabner)

Spherical t-designs, with $t \in \mathbb{N}$, are point sets $X_N := \{x_1, x_2, \cdots, x_N\} \subset \mathbb{S}^d$ such that

$$\frac{1}{N}\sum_{j=1}^{N} p(x_j) = \frac{1}{\Omega_d}\int_{\mathbb{S}^d} p(x)\mathrm{d}\omega_d(x) \;\; \forall \;\; p \in \mathbb{P}_t(\mathbb{S}^d).$$

Here $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ is the unit d-dimensional sphere, ω_d is the measure on $\mathbb{S}^d$ and Ω_d the total measure of $\mathbb{S}^d$, while $\mathbb{P}_t(\mathbb{S}^d)$ is the set of polynomials on $\mathbb{S}^d$ of degree $\le t$, or equivalently, the set of all polynomials in $\mathbb{R}^{d+1}$ of total degree $\le t$ restricted to $\mathbb{S}^d$. That is, X_N is a spherical t-design if the cubature rule with these points and with equal weights

$$w_j = \frac{1}{N} \qquad \forall j = 1, \ldots, N$$

integrates exactly all polynomials of degree up to t. Spherical designs were first introduced by [1] in 1977.

We first rederive the lower bounds on N proved by [1], namely

$$N \ge \binom{d+t/2}{d} + \binom{d+t/2-1}{d} \quad \text{if } t \text{ is even}, \quad N \ge 2\binom{d+\lfloor t/2 \rfloor}{d} \quad \text{if } t \text{ is odd}.$$

The proof in [1] proceeds by applying the spherical design property to a well chosen (zonal) polynomial $q(x\cdot y)$ with fixed $y \in \mathbb{S}$, and with $q = q_t$ a non-negative

polynomial of degree t on $[-1,1]$. This gives

$$\frac{1}{N}\sum_{j=1}^{N} q(x_j \cdot y) = \frac{1}{\Omega_d}\int_{\mathbb{S}^d} q(x\cdot y)\mathrm{d}\omega(x)$$
$$= \frac{\Omega_{d-1}}{\Omega_d}\int_{-1}^{1} q(z)(1-z^2)^{d/2-1}\mathrm{d}z := \overline{q},$$

thus $\overline{q}$ is the spherical average of $q(x\cdot y)$, which is independent of y. Now choose $y = x_1$, and rearrange to obtain

$$N = \frac{1}{\overline{q}}\left(q(1) + \sum_{j=2}^{N} q(x_j \cdot x_1)\right) \geq \frac{q(1)}{\overline{q}}.$$

The polynomial that maximises the RHS (we prove the optimality) is

$$q_t^{\mathrm{D}}(z) = \begin{cases} \left(P_{t/2}^{(d/2,d/2-1)}(z)\right)^2 & \text{for } t \text{ even},\\ (1+z)\left(P_{t/2-1}^{(d/2-1,d/2-1)}(z)\right)^2 & \text{for } t \text{ odd}, \end{cases}$$

where $P_\tau^{\alpha,\beta}$ is the Jacobi polynomial with indices α,β of degree τ.

Denote the corresponding lower bound by N_t^{D}. We extend the argument of [1] to obtain the following separation result:

Theorem: *Suppose that $X_N = \{x_1,\cdots,x_N\} \subset \mathbb{S}^d$ is a spherical t-design with $N \asymp t^d$ and $N_t^{\mathrm{D}} \leq N \leq \rho 2N_t^{\mathrm{D}}$ for some fixed $\rho < 1$. Then $\exists\, c = c(\rho)$ such that*

$$\cos^{-1}(x_i \cdot x_j) \geq \frac{c(\rho)}{N^{1/d}} \quad \forall\, 1 \leq i \neq j \leq N.$$

The condition $N < 2N_t^{\mathrm{D}}$ is comfortably satisfied by the best spherical designs for $d = 2$ and small t in the Hardin and Sloane collection [2], but restricts the applicability of this separation result because the lower bounds of [1] were significantly improved by Yudin [3]. For $d = 2$ the relative improvement ranges from nothing for small t, to $\approx 8\%$ for large t. But for large d the Yudin bound has an exponentially larger constant. Thus the separation result can be useful for at most small d.

Yudin applied the spherical design property not to a polynomial, but to a specially designed zonal function $V(x\cdot y)$ with $V = V_t \in C([-1,1])$ and $V(z) \geq 0\ \forall\ z \in [-1,1]$, and one more essential property, still to come. Yudin's strategy (actually, the strategy is already indicated in [1]) makes use of the Legendre expansion of V,

$$V(z) = \sum_{n=0}^{\infty} \widehat{V}(n) Z(d,n) P_n(z), \quad \widehat{V}(n) := \frac{\Omega_{d-1}}{\Omega_d}\int_{-1}^{1} V(z)P_n(z)(1-z^2)^{d/2-1}\mathrm{d}z.$$

where $P_n(z) := P_n^{(d/2-1,d/2-1)}(z)/P_n^{(d/2-1,d/2-1)}(1)$ is the Legendre polynomial for dimension d and degree n, and $Z(d,n)$ is the dimension of the space of homogeneous harmonic polynomials of degree n. The extra condition on V is

$$\widehat{V}(n) \le 0 \text{ for all } n > t.$$

Applying the spherical t-design property to $V(x \cdot y)$, we obtain

$$\frac{1}{N}\sum_{i=1}^{N} V(x_i \cdot y) = \overline{V} + \sum_{n=t+1}^{\infty} \widehat{V}(n) Z(d,n) \frac{1}{N}\sum_{i=1}^{N} P_n(x_i \cdot y),$$

where $\overline{V}$ is the spherical average of $V(x \cdot y)$. Now form a double sum:

$$\frac{1}{N^2}\sum_{i,j=1}^{N} V(x_i \cdot x_j) = \overline{V} + \sum_{n=t+1}^{\infty} \widehat{V}(n) Z(d,n) \frac{1}{N^2}\sum_{i,j=1}^{N} P_n(x_i \cdot x_j) \le \overline{V},$$

where the inequality holds because $\widehat{V}(n) \le 0$ for $n > t$, and because (using the addition theorem for spherical harmonics) $\sum_{i,j=1}^{N} P_n(x_i \cdot x_j) \ge 0$. Equivalently

$$N \ge \frac{1}{N\overline{V}}\sum_{i,j=1}^{N} V(x_i \cdot x_j) = \frac{1}{N\overline{V}}[NV(1) + \sum_{i,j=1, i\ne j}^{N} V(x_i \cdot x_j)] \ge \frac{V(1)}{\overline{V}}.$$

By choosing V so as to maximise the RHS we then obtain the Yudin bound:

$$N \ge \frac{V_t^{\mathrm{Y}}(1)}{\overline{V_t^{\mathrm{Y}}}} =: N_t^{\mathrm{Y}}.$$

So how did Yudin construct his function V_t^{Y}? He took V_t^{Y} to be the convolution $v * w$ of two positive functions v and w , i.e.

$$V(x \cdot y) = \int_{\mathbb{S}^d} v(x \cdot u) w(u \cdot y) \mathrm{d}\omega_d(u),$$

in this way immediately ensuring the positivity of V_t^{Y}. And since

$$\widehat{V}(n) = c\widehat{v}(n)\widehat{w}(n),$$

the property $\widehat{V}(n) \le 0$ for $n > t$ follows if we can ensure that $\widehat{w}(n)$ has the opposite sign to $\widehat{v}(n)$ for $n > t$. For further details of Yudin's construction see [3].

The two ingredients in the Yudin convolution are

$$v_t^{\mathrm{Y}}(z) = \begin{cases} 0 \text{ if } z < \alpha_t \\ P_{t+1}(z) - P_{t+1}(\alpha_t) \text{ if } z \ge \alpha_t \end{cases}$$

$$w_t^{\mathrm{Y}}(z) = \begin{cases} 0 \text{ if } z < \alpha_t \\ c \text{ if } z \ge \alpha_t \end{cases}$$

where $\alpha_t = \cos\theta_t$ is the largest zero of P'_{t+1}. The support of V_t^{Y} is easily seen to be $[2\alpha_t^2 - 1, 1]$.

We have shown that Yudin's result is optimal among all constructions that follow the above strategy, and have simplified the proof.

We also extended the argument above to obtain the following result on the average separation of points of spherical designs: suppose that we have a sequence of spherical designs X_N with $N \asymp t^d$, and that $\beta_t \in \mathrm{supp}(V_t^{\mathrm{Y}})$ with $1 - \beta_t$ a constant fraction τ of $|\mathrm{supp}(V_t^{\mathrm{Y}})|$. Then we show

$$\#\{(i,j) : i \neq j,\ x_i \cdot x_j \geq \beta_t\} \leq cN.$$

Thus on the average, a spherical cap of radius $\cos^{-1}(\beta_t)$ centered at a point of the design contains at most a constant number of other points. The constant is small if τ is close to 1 and N is close to N_t^{Y}. (However, we don't yet know whether N close to N_t^{Y} is achievable).

Acknowledgements The support of the Australian Research Council is gratefully acknowledged.

References

[1] P. Delsarte, J.M. Goethals and J.J. Seidel *Spherical codes and designs*, Geom. Dedicata **6** (1977), 363–388.
[2] R.H. Hardin and N.J.A. Sloane. Spherical designs, http://neilsloane.com/sphdesigns
[3] V. Yudin, *Lower bounds for spherical designs*, Izvestiya Mathematics **61:3** (1997), 673–683.

Three problems for discrepancy and QMC

Stefan Steinerberger

Our setting is $[0,1]^2$ or $[0,1]^3$ and we discuss problems on discrepancy with different codimensions.

Points: QMC for harmonic functions. Let $u : \mathbb{R}^2 \to \mathbb{R}$ be harmonic, i.e.

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) u(x,y) = 0.$$

Any harmonic function satisfies a mean-value property, i.e. the value at any point coincides with the average function value on any circle centered at that point

$$u(x) = \frac{1}{|\partial B_r(x)|} \int_{\partial B_r(x)} u(z) d\sigma(z).$$

We wish to find a set of points (x_n) in $[0,1]^2$ such that

$$\int_{[0,1]^2} u\ dx\ dy\ dz \approx \frac{1}{N} \sum_{i=1}^{N} u(x_i).$$

Is there any sequence of points exploiting the mean-value property efficiently such that the integral can be quickly computed?

Lines: just beyond Schmidt's theorem? Let $\ell_1, \ell_2, \ldots, \ell_N$ be a set of N lines in $\mathbb{R}^3$. Let $\mathcal{H}^k$ denote the $k-$dimensional Hausdorff measure. We consider the set

$$A = \bigcup_{i=1}^{N} \ell_i \cap [0,1]^3$$

and a measure of discrepancy

$$D_N(A) = \sup_{Q \subset [0,1]^3} \left| \frac{\mathcal{H}^1(A \cap Q)}{\mathcal{H}^1(A)} - \mathcal{H}^3(Q) \right|,$$

where the supremum is taken over all cubes Q. Is there an inequality of the form

$$D_N(A) \geq c \frac{\log \mathcal{H}^1(A)}{\mathcal{H}^1(A)}?$$

Note that by taking all lines to be parallel to one of the axes, this would imply Schmidt's bound

$$D_N \geq c \frac{\log N}{N} \qquad \text{for sets in } [0,1]^2.$$

Conversely, any set of points with small discrepancy in $[0,1]^2$ can be trivially extended to a set of lines in $\mathbb{R}^3$ such that our discrepancy notion does not increase.

As such, the problem seems to be strictly more complicated than Schmidt's result but maybe not quite as complicated as the same question for sets of point in $[0,1]^3$.

Sets: a geometric uncertainty principle. It is trivially impossible to decompose $[0,1]^2$ into finitely many discs of equal radius. A quantitative version was studied in [1]. In measuring how much a set deviates from a ball, Fraenkel asymmetry has recently become an increasingly central notion: given a domain $\Omega \subset \mathbb{R}^n$, its Fraenkel asymmetry is defined via

$$\mathcal{A}(\Omega) := \inf_{B} \frac{|\Omega \triangle B|}{|\Omega|},$$

where the infimum ranges over all disks $B \subset \mathbb{R}^n$ with $|B| = |\Omega|$ and $\triangle$ is the symmetric difference

$$\Omega \triangle B = (\Omega \setminus B) \cup (B \setminus \Omega).$$

Fraenkel asymmetry is scale-invariant

$$0 \leq \mathcal{A}(\Omega) \leq 2.$$

As for deviation in size, we define the deviation from the smallest element in the partition via

$$D(\Omega_i) := \frac{|\Omega_i| - \min_{1 \leq j \leq N} |\Omega_j|}{|\Omega_i|},$$

which is scale invariant as well and satisfies

$$0 \leq D(\Omega_i) \leq 1.$$

Theorem. *Suppose we are given a partition*

$$[0,1]^2 = \bigcup_{i=1}^{N} \Omega_i.$$

For all N sufficiently large,

$$\left(\sum_{i=1}^{N} \frac{|\Omega_i|}{|\Omega|} \mathcal{A}(\Omega_i)\right) + \left(\sum_{i=1}^{N} \frac{|\Omega_i|}{|\Omega|} D(\Omega_i)\right) \geq \frac{1}{60000}.$$

This statement has applications in nodal domain estimates and spectral partition problems. The constant is clearly not sharp and any serious improvement seems to require drastically new insights. Is the extremal configuration given by a decomposition into hexagons (with obvious modifications at the boundary)? This would correspond to an optimal constant of $c \sim 0.07\ldots$

References

[1] S. Steinerberger, *A geometric uncertainty principle with an application to Pleijel's estimate*, arXiv:1306.3103

On covering in finite dimensional Banach spaces

Vladimir Temlyakov

We discuss construction of coverings of the unit ball of a finite dimensional Banach space. The well known technique of comparing volumes gives upper and lower bounds on covering numbers. This technique does not provide a construction of good coverings. Here we apply incoherent dictionaries for construction of good coverings. We use the following strategy. First, we build a good covering by balls with a radius close to one. Second, we iterate this construction to obtain a good covering for any radius. We mostly concentrate on the first step of this strategy.

Let X be a Banach space $\mathbb{R}^d$ with a norm $\|\cdot\|$ and let $B := B_X$ denote the corresponding closed unit ball:

$$B := B_X := \{x \in \mathbb{R}^d : \|x\| \leq 1\}. \tag{1}$$

The open unit ball will be denoted by $B^o := B_X^o$:

$$B^o := B_X^o := \{x \in \mathbb{R}^d : \|x\| < 1\}. \tag{2}$$

Notation $B(x,r) := B_X(x,r)$ and $B^o(x,r) := B_X^o(x,r)$ will be used respectively for closed and open balls with the center x and radius r. In case $r = 1$ we drop it from the notation: $B^o(x) := B^o(x,1)$. For a compact set A and a positive number ϵ we define the covering number $N_\epsilon(A)$ as follows

$$N_\epsilon(A) := N_\epsilon(A,X) := \min\{n \,|\, \exists x^1,\ldots,x^n : A \subseteq \cup_{j=1}^n B_X(x^j,\epsilon)\}.$$

The following proposition is well known.

Proposition 1. *For any d-dimensional Banach space X we have*

$$\epsilon^{-d} \leq N_\epsilon(B_X,X) \leq (1+2/\epsilon)^d.$$

This proposition describes the behavior of $N_\epsilon(B_X, X)$ when $\epsilon \to 0$. In the paper [8] we concentrate on the case when ϵ is close to 1. In particular, we discuss the following problem: How many balls $B^o(x^j)$ are needed for covering B? In other words we are interested in the number

$$N(d, X) := \min\{n \mid \exists x^1, \dots, x^n : B_X \subset \cup_{j=1}^n B_X^o(x^j)\}. \tag{3}$$

This problem is related to Borsuk's conjecture. Borsuk [3] raised the following question.

Problem B. Is it true that every set of diameter one in $\mathbb{R}^d$ can be partitioned into $d+1$ sets of diameter smaller than one?

Settings with *diameter* as in Problem B and with *covering* as in (3) are related but different. It is clear that if A is covered by n balls $B(x^j, r)$, $j = 1, \dots, n$, then A can be partitioned into n sets of diameter $\le 2r$. It is known (see, for instance, [4]) that the fact that A is of diameter $2r$ does not imply that A can be covered by a single $B(x^0, r)$. Kahn and Kalai [5] gave an example showing that for all sufficiently large d the answer to Problem B is negative (see [1] for further discussion and [2] for the latest results). Problem B is formulated in terms of diameter of a set as a characteristic of its size. One can reformulate Problem B in terms of radius of a set

$$r(A) := \inf\{r \mid \exists x : A \subset B(x, r)\}.$$

Problem Br. Is it true that every set of radius one in $\mathbb{R}^d$ can be partitioned into $d+1$ sets of radius smaller than one?

Problem Br is equivalent to the question: Does equality $N(d, X) = d+1$ hold? In the case of the Hilbert space ℓ_2^d it is a classical problem of discrete geometry. It is well known that in this case $N(d, X) = d+1$. This result is a sort of folklore result in discrete geometry. The author learnt it from Misha Gromov in 1997 at the conference Foundations of Computational Mathematics held in Brazil.

We proved in [8] that if X is a uniformly smooth Banach space then $N(d, X) = d+1$. Thus, contrary to Problem B its analog Problem Br has a positive answer in a very general situation. With this result in hands we discuss the problem: How small ϵ can be for the relation $N_\epsilon(B) = d+1$ to hold? The left inequality in Proposition 1 gives the lower bound for such ϵ: $\epsilon \ge 1 - \frac{\ln(d+1)}{d}$. It is known in discrete geometry (see, for instance, [6]) that for the Hilbert space ℓ_2^d we have $\epsilon \le 1 - Cd^{-2}$. In [8] we give a very simple analytic construction of such a cover.

Simple explicit analytic constructions of coverings are important in applications. We explain in [8] how explicit constructions of coverings with radius close to 1 can be used for building coverings with arbitrarily small radius. Keeping in mind importance of simple analytic constructions we present in [8] a construction based on Hadamard matrices. We discuss in detail Hadamard matrices because their application provides a simple construction with a nice property of the vectors of the corresponding ϵ-net: they have the same absolute values of all coordinates. This construction also provides an upper bound $\epsilon \le 1 - Cd^{-2}$ for those d for which the Hadamard matrices exist.

Thus, we show in [8] that the best known upper bound $\epsilon \le 1-Cd^{-2}$ follows from two different constructions. In both constructions we use a system $\mathcal{D} := \{g^j\}_{j=1}^{d+1}$ of vectors and build a covering of B_2 in the form $\cup_{j=1}^{d+1} B_2^o(ag^j, r)$ with an appropriate r. In [8] we apply this idea with $\mathcal{D}$ being an incoherent dictionary for covering in the Hilbert space ℓ_2^d. We prove the following bound in [8]. For $r = (1-\mu^2)^{1/2}$, $\mu \in [(2d)^{-1/2}, 1/2]$, we have

$$N_r(B_2) \le 2\exp(C_1 d\mu^2 \ln(2/\mu)). \tag{4}$$

Bound (4) is based on known results on the maximal size of incoherent dictionaries (see [7], section 5.7).

In [8] we use incoherent dictionaries in a smooth Banach space X to build a good covering for B_X. We prove in [8] a new result on the maximal size of incoherent dictionaries in Banach spaces. Then we use it to obtain an upper bound on $N_r(B_X)$ (see (5) below). Let $\rho(u)$ denote the modulus of smoothness of X and $a(\mu)$ be a solution (actually, it is a unique solution) to the equation

$$a\mu = 4\rho(2a).$$

We prove the following bound in [8]. For $r = 1 - \frac{1}{2}\mu a(\mu)$, $\mu \le 1/2$, we have

$$N_r(B_X) \le 2\max(C_2 d, \exp(C_2 d\mu^2 \ln(2/\mu))). \tag{5}$$

It is interesting to note that in the case $X := \ell_p^d$, $p \in [2,\infty)$, we have $1 - r = \frac{1}{2}\mu a(\mu) \asymp \mu^2$ as in the case $X = \ell_2^d$.

In [8] we consider several specific examples of X and make a conclusion that the technique based on extremal incoherent dictionaries works well and provides either optimal or close to optimal bounds in the sense of order of $\ln N_\epsilon(B_X)$.

References

[1] N. Alon, Discrete mathematics: methods and challenges, Proceedings of the International Congress of Mathematicians, Beijing 2002, **1**, 119–135.

[2] A. Bondarenko, On Borsuk's conjecture for two-distance sets, arxiv:1305.2584v3 [math.MG] 29 Aug 2013, 1–9.

[3] K. Borsuk, Drei Sätze über die n-dimensionale euclidische Sphäre, Fund. Math. **20** (1933), 177–190.

[4] J. Bourgain and J. Lindenstrauss, On covering a set in $\mathbb{R}^d$ by balls of the same diameter, Geometric Aspects of Functional Analysis, Lecture Notes in Mathematics, **1469** (1991), 138–144, Springer-Verlag, Berlin.

[5] J. Kahn and G. Kalai, A counterexample to Borsuk's conjecture, Bulletin of the AMS **29** (1993), 60–72.

[6] A.M. Raigorodskii, The Kneser conjecture and the topological method in combinatorics, Moscow Centre for Continuous Mathematical Education, Moscow, Russia, 2011.

[7] V.N. Temlyakov, Greedy approximation, Cambridge University Press, 2011.

[8] V.N. Temlyakov, A remark on covering, arxiv:1301.3043v1 [math.MG] 10 Jan 2013.

Recent results on $(t, \mathbf{e}, s)$-sequences

Shu Tezuka

Recently [1], a notion of $(t, \mathbf{e}, s)$-sequences in base b was introduced, where $\mathbf{e} = (e_1, ..., e_s)$ is a positive integer vector, and their discrepancy bounds were obtained based on the signed splitting method. In this extended abstract, we present the latest results on $(t, \mathbf{e}, s)$-sequences.

First, we introduce the definition of discrepancy. For a point set $P_N = \{X_0, X_1, \ldots, X_{N-1}\}$ of N points in $[0,1]^s$ and an interval $J \subseteq [0,1]^s$, we define $A_N(J)$ as the number of n, $0 \le n \le N-1$, with $X_n \in J$ and $\mu(J)$ is the volume of J. Then the star discrepancy of P_N is defined by

$$D_N^* = \sup_J \left| \frac{A_N(J)}{N} - \mu(J) \right|,$$

where the supremum is taken over all intervals J of the form $\prod_{i=1}^s [0, \alpha_i)$ for $0 < \alpha_i \le 1$. The (unanchored) discrepancy D_N is obtained when the supremum is taken over all intervals J of the form $\prod_{i=1}^s [\alpha_i, \beta_i)$ for $0 \le \alpha_i < \beta_i \le 1$.

Let $b \ge 2$ be an integer. An elementary interval in base b, which is a key concept of the net theory, is an interval of the form

$$E(\mathbf{l}; \mathbf{a}) = \prod_{i=1}^s \left[\frac{a_i}{b^{l_i}}, \frac{a_i + 1}{b^{l_i}} \right),$$

where a_i and l_i are integers with $0 \le a_i < b^{l_i}$ and $l_i \ge 0$ for $i = 1, ..., s$. Denote a subset of nonnegative integer vectors by $\mathcal{E} \subseteq \mathbb{N}_0^s$, where $\mathrm{card}(\mathcal{E}) = \infty$. Define a set of elementary intervals as follows:

$$\mathbb{E}(\mathcal{E}) = \bigcup_{\mathbf{l} \in \mathcal{E}} E(\mathbf{l}),$$

where

$$E(\mathbf{l}) = \{E(\mathbf{l}; \mathbf{a}) |\ 0 \le a_i < b^{l_i},\ (1 \le i \le s)\}.$$

Define $|\mathbf{l}| = l_1 + \cdots + l_s$ for a nonnegative integer vector $\mathbf{l} = (l_1, ..., l_s)$. Denote a set of nonnegative integers by $\mathbb{N}(\mathcal{E}) = \{|\mathbf{l}| \mid \mathbf{l} = (l_1, ..., l_s) \in \mathcal{E}\}$. We first give the definition of $(t, m, \mathcal{E}, s)$-nets as follows:

Definition 1. *Let t and m be integers with $0 \le t \le m$ such that $m - t \in \mathbb{N}(\mathcal{E})$. A $(t, m, \mathcal{E}, s)$-net in base b is a point set of b^m points in $[0,1]^s$ such that $A_{b^m}(E) = b^t$ for every elementary interval $E \in \mathbb{E}(\mathcal{E})$ with $\mu(E) = b^{t-m}$.*

Let $\mathbf{T}_\mathcal{E}$ be a mapping from $\mathbb{N}(\mathcal{E})$ to $\mathbb{N}_0$, where $0 \le \mathbf{T}_\mathcal{E}(m) \le m$ for $m \in \mathbb{N}(\mathcal{E})$, such that there are infinitely many m satisfying $m - \mathbf{T}_\mathcal{E}(m) \in \mathbb{N}(\mathcal{E})$. Then $(\mathbf{T}_\mathcal{E}, \mathcal{E}, s)$-sequences are defined as follows:

Definition 2. *A $(\mathbf{T}_\mathcal{E}, \mathcal{E}, s)$-sequence in base b is an infinite sequence, $X = (X_n)_{n \ge 0}$, of points in $[0,1]^s$ such that for all integers $k \ge 0$ and all $m \ge \mathbf{T}_\mathcal{E}(m)$ satisfying $m - \mathbf{T}_\mathcal{E}(m) \in \mathbb{N}(\mathcal{E})$, the point set $\{[X_{kb^m}]_{b,m}, ..., [X_{(k+1)b^m - 1}]_{b,m}\}$ is a $(\mathbf{T}_\mathcal{E}(m), m, \mathcal{E}, s)$-net, where $[X_n]_{b,m}$ means the coordinate-wise b-ary m-digit truncation of a point X_n.*

It is easy to obtain the following propositions.

Proposition 1. *When $\mathcal{E} = \mathbb{N}_0^s$, a $(\mathbf{T}_\mathcal{E}, \mathcal{E}, s)$-sequence in base b is identical to a $(\mathbf{T}, s)$-sequence in base b.*

Proposition 2. *Let $\mathbf{e} = (e_1, ..., e_s)$ be a positive integer vector. When the mapping $\mathbf{T}_\mathcal{E}$ is constant, i.e., $\mathbf{T}_\mathcal{E} \equiv t$, and $\mathcal{E} = \{\mathbf{l} \mid e_i$ divides l_i $(1 \le i \le s)\}$, then a $(\mathbf{T}_\mathcal{E}, \mathcal{E}, s)$-sequence in base b is identical to a $(t, \mathbf{e}, s)$-sequence in base b.*

The main result [2] is given as follows:

Theorem 1. *Let $b \ge 2$ be an arbitrary integer. The star discrepancy for the first $N > b^t$ points of a $(t, \mathbf{e}, s)$-sequence in base b is bounded as follows:*

$$ND_N^* \le \frac{b^t}{s!} \prod_{i=1}^{s} \left(\frac{b^{e_i} - 1}{2e_i} (\log_b N - t) + s \right)$$

$$+ \frac{b^{t+e_s} |\mathbf{e}|}{2} \prod_{i=1}^{s-1} \left(\frac{\lfloor b^{e_i}/2 \rfloor}{e_i} (\log_b N - t) + \lfloor b^{e_i}/2 \rfloor \right)$$

$$+ \sum_{k=0}^{s-1} \frac{b^{t+e_{k+1}}}{k!} \prod_{i=1}^{k} \left(\frac{\lfloor b^{e_i}/2 \rfloor}{e_i} (\log_b N - t) + k \right).$$

Thus, we arrive at the leading constant for $(t, \mathbf{e}, s)$-sequences in base b as follows:

$$c_s^*(new) = \frac{b^t}{s!} \prod_{i=1}^{s} \frac{b^{e_i} - 1}{2e_i \log b}.$$

In comparison with the previous constant

$$c_s^* = \frac{b^t}{s!} \prod_{i=1}^{s} \frac{\lfloor b^{e_i}/2 \rfloor}{e_i \log b},$$

the new constant yields improvement for the case of even bases.

The leading constant currently known as the best for (t, s)-sequences in base b, which was recently obtained by Faure and Kritzer based on an improvement of the double recursion method, is given as

$$c_s^*(t, s) = \begin{cases} \dfrac{b^2}{2(b^2-1)} \dfrac{b^t}{s!} \left(\dfrac{b-1}{2 \log b} \right)^s & \text{if } b \text{ is even,} \\ \dfrac{1}{2} \dfrac{b^t}{s!} \left(\dfrac{b-1}{2 \log b} \right)^s & \text{otherwise.} \end{cases}$$

Since (t, s)-sequences in base b are equivalent to $(t, \mathbf{e}, s)$-sequences in base b with $\mathbf{e} = (1, ..., 1)$, we can compare the above constants to conclude that the new constant is slightly bigger than $c_s^*(t, s)$ by a factor smaller than 2. In the most practical case of $b = 2$, the factor is 1.5 for the new constant, while it is 1.5×2^s for the previous constant. Therefore, our improvement is significant in particular for large dimensions s.

References

[1] S. Tezuka, On the discrepancy of generalized Niederreiter sequences, *Journal of Complexity*, **29** (2013), 240-247.

[2] S. Tezuka, Improvement on the discrepancy of $(t, \mathbf{e}, s)$-sequences, preprint, (2013).

Limit Theorems for Discrepancy Functions

Robert Tichy

We consider sequences $(n_k x)$ where x is a real number and (n_k) a lacunary sequences of positive integers, i.e. $n_{k+1}/n_k \geq q > 1$. It is the aim of this lecture to explore, how the distribution properties of such deterministic sequences differ from the ones of "true" random sequences. Let D_N denote the usual discrepancy function defined by

$$D_N(x) = D_N(n_k x) = \sup_J \left| \frac{1}{N} \neq \{k \leq N : n_k x \bmod 1 \ \in J\} - \mid J \mid \right|,$$

where the sup is extended over all intervals of length $\mid J \mid$ contained in the unit interval $[0, 1)$. By a general theorem of R. Baker, for arbitrary increasing (n_k) the bound ($\varepsilon > 0$)

$$D_N(x) = O(\frac{(\log N)^{3/2+\varepsilon}}{N})$$

holds for almost all $x \in \mathbb{R}$ (in the sense of usual Lebesgue measure). For lacunary sequences (n_k) a much sharper result due to W. Philipp (1975) is

$$C_2 \leq \limsup_{N\to\infty} \frac{N D_N(x)}{\sqrt{N \log\log N}} \leq C_1 \tag{1}$$

for almost all x. In the present lecture we report on recent developments in metric discrepancy theory, most of the results are contained in joint papers of C. Aistleitner, I. Berkes and R.F. Tichy.

We are mainly interested in computing the lim sup if possible and distinguish three case where in each of the cases the results are quite different. If

$$\lim_{k\to\infty} \frac{n_{k+1}}{n_k} = \infty$$

("strong lacunarity") the sequence of random variables $D_N(x)$ behaves like a sum of i.i.d randomvariables, i.e. a "strong" law of the iterated logarithm holds

$$\limsup_{N\to\infty} \frac{N D_N(x)}{\sqrt{N \log\log N}} = \frac{1}{2} \qquad a.e. \tag{2}$$

And even more holds in this case: the result is permutation invariant as in the i.i.d. case, i.e. it is also true for sequences $(n_{\sigma(k)} x)$, where σ is a permutation of the positive integers. In the case when the growth rate of (n_k) is exponential

("proper lacunarity") there is an old example of Erdős and Fortet, $n_k = 2^k - 1$ where no LIL (law of the iterated logarithm) holds. Recently, Fukuyama computed the corresponding values $\sum_a$ of the lim sup for sequences $n_k = a^k$ (a an integer ≥ 2):

$$\sum_a \begin{cases} \sqrt{42}/9 & \text{for} \quad a = 2 \\ \frac{\sqrt{(a+1)a(a-2)}}{2\sqrt{(a-3)^3}} & \text{for} \quad a \geq 4 \quad \text{even} \\ \frac{\sqrt{a+1}}{2\sqrt{a-1}} & \text{for} \quad a \geq 3 \quad \text{odd} \end{cases}$$

In the sublacunary case in general a LIL does not hold. However, in special situations a LIL is still true. For instance, an Erdős gap condition

$$\frac{n_k + 1}{n_k} \geq 1 + k^{-\alpha}, \quad (0 > \alpha < 1) \tag{3}$$

is sufficient for a non-permutation invariant LIL. For $\alpha = \frac{1}{2}$ it becomes false as shown by Berkes and Philipp. The permutation, invariant law does not hold even if (3) is satisfied: For any positive sequence $\varepsilon_k \to 0$ there exists (n_k) with $n_{k+1}/n_k \geq 1 + \varepsilon_k$ and a permutation $\sigma : \mathbb{N} \to \mathbb{N}$ such that the LIL fails. In the case of the sublacunary Hardy-Littlewood-Polya sequence a LIL of type (1) remains true. These sequences n_k are the numbers (in increasing oder) of the multiplicative semigroup which is generated by coprime integers $q_1, \ldots, q_t$ ($t \geq 2$). It should be remarked that similar results also hold for weighted sums

$$\frac{1}{A(N)} \sum_{k=1}^{N} a_k f(n_k x), \quad A(N) = \sum_{k=1}^{N} a_k \tag{4}$$

where f satisfies a Lipschitz condition. For trigonometric functions $f(x)$ this is classical work of Salem and Zygmund (1947) who showed a central limit theorem and a LIL in this case. If (n_k) satisfies $\frac{n_{k+1}}{n_k} \to \infty$ a permutation invariant central limit theorem and a permutation invariant LIL hold.

The proofs of all mentioned results depend a combinations of probabilistic tools (martingale inequalities) and methods from diophantine analysis such as the following result: For any fixed integers $a \neq 0, b \neq 0, c$ the number of solutions of the diophantine equation $an_k + bn_l = c$ is bounded by a constant $K(a, b)$ independent of c, where for $c = 0$ also $k \neq l$ is required. If this general bound for the number of solutions is satisfied, permutation invariant limit laws hold for discrepancies and Lipschitz functions (4).

Well-separated spherical designs

Maryna Viazovska
(joint work with Andriy Bondarenko, Danylo Radchenko)

In this talk we will discuss the interrelation between several classical optimization problems on spheres S^d such as minimal equal-weight quadratures (spherical designs), best packing problems, and minimal energy problems. For $d = 1$, a regular polygon is an optimal configuration for all of these problems. However, for $d \geq 2$ exact solutions are known in very few cases. Even asymptotically optimal configurations are sometimes very hard to obtain (see for example Smale's 7th Problem [13]).

In this talk we will prove the existence of certain configurations in S^d which are spherical t-designs with asymptotically minimal number of points and that simultaneously have asymptotically the best separation property. These configurations also provide approximate solutions for several other optimization problems.

Let $S^d = \left\{x \in \mathbb{R}^{d+1} : |x| = 1\right\}$ be the unit sphere in $\mathbb{R}^{d+1}$ equipped with the Lebesgue measure μ_d normalized by $\mu_d(S^d) = 1$. A set of points $x_1, \dots, x_N \in S^d$ is called a *spherical t-design* if

$$\int_{S^d} P(x)\, d\mu_d(x) = \frac{1}{N} \sum_{i=1}^{N} P(x_i)$$

for all polynomials in $d+1$ variables, of total degree at most t. The concept of a spherical design was introduced by Delsarte, Goethals, and Seidel [8]. For each $d, t \in \mathbb{N}$ denote by $N(d,t)$ the minimal number of points in a spherical t-design in S^d. The following lower bound

$$N(d,t) \geq \begin{cases} \binom{d+k}{d} + \binom{d+k-1}{d} & \text{if } t = 2k, \\ 2\binom{d+k}{d} & \text{if } t = 2k+1, \end{cases} \tag{1}$$

is proved in [8] (see also the classical monograph [7]). On the other hand, it follows from the general result by Seymour and Zaslavsky [12] that spherical designs exist for all positive integers d and t. The method of proof used in [12] was not constructive and the authors did not indicate an upper bound for $N(d,t)$ in terms of d and t. First feasible upper bounds were given by Wagner [14] ($N(d,t) \leq C_d t^{Cd^4}$) and Bajnok [2] ($N(d,t) \leq C_d t^{Cd^3}$). Korevaar and Meyers [10] have improved these inequalities by showing that $N(d,t) \leq C_d t^{(d^2+d)/2}$. They have also conjectured that $N(d,t) \leq C_d t^d$. Note that (1) implies $N(d,t) \geq c_d t^d$. Here and in what follows we use the notations C_d, L_d, etc. (c_d, λ_d, etc.) for sufficiently large (small) constants depending only on d.

Korevaar and Meyers were motivated by the following problem coming from potential theory: How to choose N equally charged points $x_1, \dots, x_N$ in S^2 to

minimize the value

$$U_r(x_1,\dots,x_N) := \sup_{|x|=r}\left|\frac{1}{N}\sum_{i=1}^{N}\frac{1}{|x-x_i|}-1\right|, \quad r\in(0,1)?$$

The classical Faraday cage phenomenon states that any stable charge distribution on the compact closed surface cancels the electric field inside the surface. According to this model the minimal value of U_r should rapidly decay to 0, when N grows.

It was shown in [10] that if the set of points $x_1,\dots,x_N$ is a spherical t-design for some $t > cN^{1/2}$ then

$$U_r(x_1,\dots,x_N) \le r^{\alpha N^{1/2}}.$$

The estimate is optimal up to the constant in the power.

Recently we have suggested a nonconstructive approach to obtain an optimal asymptotic bound for $N(d,t)$ based on the application of the topological degree theory; see [4, 5]. We have proved the following

Theorem 1. *For each $N \ge C_d t^d$ there exists a spherical t-design in S^d consisting of N points.*

This implies the Korevaar-Meyers conjecture.

Now we will give the definition of a well-separated sequence of configurations. A sequence of N-point configurations $X_N = \{x_{1N},\dots,x_{NN}\}$ in S^d is called well-separated if

$$\min_{1\le i<j\le N}|x_{iN}-x_{jN}| \ge \lambda_d N^{-1/d} \tag{2}$$

for some constant λ_d and all $N \ge 2$. The inequality (2) is optimal up to the constant λ_d. That is, there exists a constant L_d such that for any N-point configuration $\{x_1,\dots,x_N\}$

$$\min_{1\le i<j\le N}|x_i-x_j| < L_d N^{-1/d}.$$

Many authors have predicted the existence of well-separated spherical t-designs in S^d of asymptotically minimal cardinality $O(t^d)$ as $t\to\infty$ (see, e.g. [1] and [9]). Moreover, in [9] it was shown that if such spherical designs exist then they have asymptotically minimal Riesz s-energy. In our recent paper [3] we prove the existence of above mentioned spherical designs. Namely, we show the following:

Theorem 2. *For each $d \ge 2$ there exist positive constants C_d and λ_d depending only on d such that for each $t\in\mathbb{N}$ and each $N > C_d t^d$ there exists a spherical t-design in S^d consisting of N points $\{x_i\}_{i=1}^N$ with $|x_i-x_j| \ge \lambda_d N^{-1/d}$ for $i\ne j$.*

References

[1] C. An, X. Chen, I. H. Sloan, and R. S. Womersley, *Well conditioned spherical designs for integration and interpolation on the two-sphere*, SIAM Journal on Numerical Analysis 48 (2010), pp. 2135-2157.

[2] B. Bajnok, *Construction of spherical t-designs*, Geom. Dedicata, 43 (1992), pp. 167–179.

[3] A. Bondarenko, D. Radchenko, and M. Viazovska, *Well-separated spherical designs*, arXiv:1303.5991 (2013).
[4] A. Bondarenko, D. Radchenko, and M. Viazovska, *Optimal asymptotic bounds for spherical designs*, Annals of Mathematics, 178 (2013), pp. 443-452.
[5] A. Bondarenko, M. Viazovska, *Spherical designs via Brouwer fixed point theorem*, SIAM J. Discrete Math., 24 (2010), pp. 207–217.
[6] X. Chen, R. S. Womersley, *Existence of solutions to systems of underdetermined equations and spherical designs*, SIAM Journal on Numerical Analysis, 44 (2006), pp. 2326 - 2341.
[7] J. H. Conway, N. J. A. Sloane, *Sphere packings, lattices and groups.*, 3rd ed., Springer, New York, 1999.
[8] P. Delsarte, J. M. Goethals, and J. J. Seidel, *Spherical codes and designs*, Geom. Dedicata, 6 (1977), pp. 363–388.
[9] K. Hesse, P. Leopardi, *The Coulomb energy of spherical designs*, Adv. Comp. Math., 28 (2008), pp. 331–354.
[10] J. Korevaar, J. L. H. Meyers, *Spherical Faraday cage for the case of equal point charges and Chebyshev-type quadrature on the sphere*, Integral Transforms Spec. Funct. 1 (1993), pp. 105–117.
[11] H. N. Mhaskar, F. J. Narcowich, and J. D. Ward, *Spherical Marcinkiewicz-Zygmund inequalities and positive quadrature*, Math. Comp., 70 (2001), pp. 1113–1130.
[12] P. D. Seymour and T. Zaslavsky, *Averaging sets*, Adv. Math., 52 (1984), pp. 213–240.
[13] S. Smale, *Mathematics: frontiers and perspectives*, Amer. Math, Soc., Providence, (2000), pp. 271-294.
[14] G. Wagner, *On averaging sets*, Monatsh. Math., 111 (1991), pp. 69–78.

Tractability of approximation of ∞-variate functions with bounded mixed partial derivatives

Grzegorz W. Wasilkowski

There are many practical problems dealing with ∞-variate functions. These involve stochastic differential equations, partial differential equations with random coefficients, and path integrals, see, e.g., [1, 2, 4, 5] and papers cited therein. This is why the study of complexity and tractability of such problems has become a popular field of study. In this talk, that is based on [7], we assume that the space $\mathcal{F}$ of functions f to be approximated admits the unique representation

$$f(\boldsymbol{x}) = \sum_{\mathfrak{u}} f_{\mathfrak{u}}(\boldsymbol{x}),$$

where $\mathfrak{u}$ enumerate the finite subsets of $\mathbf{N}_+$ listing the so-called *active variables* upon which $f_{\mathfrak{u}}$ depends. More precisely, each $f_{\mathfrak{u}}$ belongs to a normed space $F_{\mathfrak{u}}$ that is a $|\mathfrak{u}|$-fold tensor product of a space F of univariate functions. Previous work on this topic assumed that F is a Hilbert space, with $F_{\mathfrak{u}}$ being Hilbert spaces obtained via the usual tensor product. In this paper, we let $F_{\mathfrak{u}}$ be Banach spaces of functions with mixed partial derivatives of order r bounded in a ψ-weighted L_p norm for some $p \in [1, \infty]$. The space $\mathcal{F}$ is endowed with the norm

$$\|f\|_{\mathcal{F}} = \left[\sum_{\mathfrak{u}} \left(\frac{\|f_{\mathfrak{u}}\|_{F_{\mathfrak{u}}}}{\gamma_{\mathfrak{u}}} \right)^q \right]^{1/q}$$

for $q \in [1,\infty]$ and a given family $\gamma = \{\gamma_{\mathfrak{u}}\}_{\mathfrak{u}}$ of non-negative numbers, called *weights*. The goal is to approximate $f \in \mathcal{F}$ with the error measured in a ω-weighted L_s-norm for some $s \in [1,\infty]$. Both ψ and ω are probability density functions.

The role of weights $\gamma_{\mathfrak{u}}$ has been explained in many papers; roughly speaking, they quantify the importance of interactions among variables listed in $\mathfrak{u}$. Previous work has only considered the case $q = 2$, resulting in $\mathcal{F}$ also being a Hilbert space. In this talk, we consider arbitrary q to study the tradeoff between the size of q and the rate of the decay of the weights γ. In particular, we show that with $q = 1$, we have positive results even if $\gamma_{\mathfrak{u}}$ converges to zero very slowly. On the other hand, we need quickly-converging $\gamma_{\mathfrak{u}}$ for $q = \infty$. For more, see a short discussion at the end.

The role of ψ is to control the resulting space $\mathcal{F}$, since the faster the decay of ψ at $\pm\infty$ the larger the space $\mathcal{F}$. More precisely, the spaces $F_{\mathfrak{u}}$ are completions of $|\mathfrak{u}|$-times (algebraic) tensor products of the following space F of univariate functions. The domain of functions f from F is an arbitrary interval $D \subseteq \mathbf{R}$ and $f^{(r-1)}$ are absolutely (locally) continuous with bounded

$$\left[\int_D \left|f^{(r)}(x)\right|^p \psi(x)\,dx\right]^{1/p} < \infty.$$

We present the so-called *multivariate decomposition methods* (or *MDM* for short). These are modifications of methods introduced in [5], *changing dimension algorithms*. Roughly speaking, these methods identify a *set of important variable interactions* $\mathfrak{U}(\varepsilon)$ with the following desirable properties:

- **(i):** the cardinality of $\mathfrak{U}(\varepsilon)$ is polynomial in $1/\varepsilon$,
- **(ii):** elements $f_{\mathfrak{u}}$ for $\mathfrak{u} \notin \mathfrak{U}(\varepsilon)$ can be neglected,
- **(iii):** it is enough to approximate $f_{\mathfrak{u}}$ for $\mathfrak{u} \in \mathfrak{U}(\varepsilon)$,
- **(iv):** the number $|\mathfrak{u}|$ of active variables is sub-logarithmic in $1/\varepsilon$ for all $\mathfrak{u} \in \mathfrak{U}(\varepsilon)$.

In other words, an MDM replaces one problem with ∞-many variables by a number of problems, each with at most $O(\ln(1/\varepsilon)/\ln(\ln(1/\varepsilon)))$ variables, which is very small.

In general, the terms $f_{\mathfrak{u}}$ in $f = \sum_{\mathfrak{u}} f_{\mathfrak{u}}$ are not available. Hence, to take advantage of **(iii)** and **(iv)**, the space $\mathcal{F}$ has to be such that values of $f_{\mathfrak{u}}$ could be obtained by evaluating f at certain points. Moreover, we need efficient algorithms for multivariate problems with relatively small number $|\mathfrak{u}|$ of variables. When the $F_{\mathfrak{u}}$ are Hilbert spaces, Smolyak's construction provides such algorithms, see [6] for the idea and [8] for specific results. Such algorithms are often called *sparse grid algorithms*.

Fortunately, due to a very special nature of the spaces $F_{\mathfrak{u}}$ and $G_{\mathfrak{u}} = L_s(D^{|\mathfrak{u}|})$, it follows from [3] that Smolyak's construction and the results of [8] can be applied to the setting of this talk.

Among a number of results, we have the following one. Suppose that the univariate functions $f \in F$ have $D = [0,1]$ as the domain and vanish at 0, with the

norm in F given by

$$\|f\|_F = \left[\sum_{\ell=1}^{r-1} |f^{(\ell)}(0)|^p + \|f^{(r)}\|^p_{L_p([0,1])}\right]^{1/p}.$$

Here the regularity degree r is a positive integer. We measure the errors by L_s-norms for some $s \in [1,\infty]$. Suppose that the weights have the special product order-dependent form

$$\gamma_{\mathfrak{u}} = (|\mathfrak{u}|!)^\alpha \cdot \prod_{j\in\mathfrak{u}} \frac{c}{j^\beta} \quad \text{for} \quad \beta > \max(0,\alpha).$$

If $\beta > 1 - 1/q = 1/q^*$ then the minimal cost of approximating the ∞-variate functions with errors not exceeding $\varepsilon \cdot \|f\|_{\mathcal{F}}$ is bounded from above by

$$O\left(\varepsilon^{-\theta-\delta}\right) \quad \text{for all } \delta > 0, \tag{1}$$

where

$$\theta = \max\left(\frac{1}{r + \min(1/s - 1/p, 0)}\,,\, \frac{1}{\beta - 1/q^*}\right).$$

It is well-known that the complexity is proportional to $\varepsilon^{-1/(r+\min(1/s-1/p,0))}$ for the univariate problem. Hence the result given above is optimal as far as the order of convergence is concerned. Also the dependence of θ on β and q is sharp. Note that if $q = 1$, then we only need $\beta > \max(0,\alpha)$. Since $1 - 1/q$ increases with q, we require the stronger condition $\beta > 1 - 1/q$ for $q > 1$. In particular, we need $\beta > 1$ if $q = \infty$. Hence there is a tradeoff between β and q.

As it is well-known, algorithms for approximation in the L_s-norm with $s = 1$ can be used to derive cubatures for approximating ∞-variate integrals. The worst case errors of such cubatures are bounded by the errors of the L_1-norm approximation problem. Hence, for the function spaces considered in this paper, these cubatures provide ε approximation with the cost bounded as in (1), where now

$$\theta = \max\left(\frac{1}{r}\,,\, \frac{1}{\beta - 1/q^*}\right).$$

Again, this is a nearly optimal result, since the complexity of the corresponding univariate problem is $O\left(\varepsilon^{-1/r}\right)$, and the dependence on $1/(\beta - 1/q^*)$ is sharp.

References

[1] J. Creutzig, S. Dereich, T. Müller-Gronbach, and K. Ritter, *Infinite-dimensional quadrature and approximation of distributions*, Found. Comput. Math. **9** (2009), 391–429.

[2] M. Gnewuch, *Infinite-dimensional integration on weighted Hilbert spaces,* Math. Comp. **81** (2012), 2175–2205.

[3] Fengbo Hang and Youming Li, *On applicability of the sparse grid method in the worst case setting*, Numerical Algorithms **42** (2006), 95-105.

[4] F. Y. Kuo, C. Schwab, and I. H. Sloan, *Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients*, SIAM J. Numer. Anal. **50** (2012), 3351–3374.

[5] F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, and H. Woźniakowski, *Liberating the dimension*, J. of Complexity **26** (2010), 422–454. doi: 10.1016/j.jco.2009.12.003.

[6] S. A. Smolyak, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Dokl. Acad. Nauk SSSR **4** (1963), 240-243.
[7] G. W. Wasilkowski, *Tractability of approximation of ∞-variate functions with bounded mixed partial derivatives*, in review.
[8] G. W. Wasilkowski and H. Woźniakowski, *Explicit cost bounds for multivariate tensor product problems*, J. of Complexity **11** (1995), 1–56.

On lower bounds for integration of multivariate permutation-invariant functions

Markus Weimar

The talk is based on the recent paper [5]. We consider the integration problem

$$\mathrm{Int}_d\colon E_{d,\alpha} \to \mathbb{C}, \qquad \mathrm{Int}_d(f) = \int_{[0,1]^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},$$

for periodic, complex-valued functions in the Korobov class

$$E_{d,\alpha} := \left\{ f \in L_1([0,1]^d) \,\middle|\, \|f \,|\, E_{d,\alpha}\| := \sup_{\boldsymbol{k}\in\mathbb{Z}^d} \left|\widehat{f}(\boldsymbol{k})\right| \left(\overline{k_1}\cdot\ldots\cdot\overline{k_d}\right)^{\alpha} < \infty \right\}$$

where $d \in \mathbb{N}$ and $\alpha > 1$. Here $\overline{k_m} := \max\{1, |k_m|\}$ and $\widehat{f}(\boldsymbol{k})$ denotes the $\boldsymbol{k}$th Fourier coefficient of f.

To approximate $\mathrm{Int}_d(f)$, $d \in \mathbb{N}$, without loss of generality, we consider algorithms from the class of all linear cubature rules

$$\mathcal{A}_{N,d}(f) := \sum_{n=1}^{N} w_n\, f\left(\boldsymbol{t}^{(n)}\right), \qquad N \in \mathbb{N}_0,$$

that use at most N values of the input function f at some points $\boldsymbol{t}^{(n)} \in [0,1]^d$, $n = 1\ldots, N$. The weights w_n can be arbitrary complex numbers.

In [3] Sloan and Woźniakowski showed that for every $d \in \mathbb{N}$ the *Nth minimal worst case error*,

$$e(N,d;\mathrm{Int}_d, E_{d,\alpha}) := \inf_{\mathcal{A}_{N,d}} \sup_{\|f\,|\,E_{d,\alpha}\|\le 1} \left|\mathrm{Int}_d(f) - \mathcal{A}_{N,d}(f)\right|,$$

equals the initial error $e(0,d;\mathrm{Int}_d, E_{d,\alpha}) = 1$ provided that $N < 2^d$. In other words, the integration problem on the full spaces $(E_{d,\alpha})_{d\in\mathbb{N}}$ suffers from the *curse of dimensionality*, since for every fixed $\varepsilon \in (0,1)$ its *information complexity* grows exponentially with the dimension d:

$$n(\varepsilon,d) := n(\varepsilon,d;\mathrm{Int}_d, E_{d,\alpha}) := \min\left\{N \in \mathbb{N}_0 \,\middle|\, e(N,d;\mathrm{Int}_d, E_{d,\alpha}) \le \varepsilon\right\} \ge 2^d.$$

We generalize this result to the case of permutation-invariant subspaces in the sense of [4]. To this end, for $d \in \mathbb{N}$ let $I_d \subseteq \{1,\ldots,d\}$ be some subset of coordinates and consider the integration problem restricted to the subspace $\mathfrak{S}_{I_d}(E_{d,\alpha})$ of all *I_d-permutation-invariant* functions $f \in E_{d,\alpha}$. That is, in dimension d we restrict ourselves to functions f that satisfy

$$f(\boldsymbol{x}) = f(\sigma(\boldsymbol{x})) \quad \text{for all} \quad \boldsymbol{x} \in [0,1]^d$$

and any permutation σ from

$$\left\{\sigma\colon \{1,\dots,d\}\to\{1,\dots,d\}\,\middle|\,\sigma \text{ bijective and } \sigma\big|_{\{1,\dots,d\}\setminus I_d}=\mathrm{id}\right\}.$$

Observe that in the case $I_d=\emptyset$ we clearly have $\mathfrak{S}_{I_d}(E_{d,\alpha})=E_{d,\alpha}$.

We prove the following

Theorem 1. *Let*

$$N^*:=N^*(d,I_d):=(\#I_d+1)\cdot 2^{d-\#I_d},\qquad d\in\mathbb{N}.$$

Then, for every $N<N^*$,

$$e(N,d;\mathrm{Int}_d,\mathfrak{S}_{I_d}(E_{d,\alpha}))=1$$

and

$$e(N^*,d;\mathrm{Int}_d,\mathfrak{S}_{I_d}(E_{d,\alpha}))\le\left(1+\frac{\zeta(\alpha)}{2^{\alpha-1}}\right)^d-1$$

for all $d\in\mathbb{N}$ and $\alpha>1$. Consequently,

$$\lim_{\alpha\to\infty} e(N^*,d;\mathrm{Int}_d,\mathfrak{S}_{I_d}(E_{d,\alpha}))=0$$

for all $d\in\mathbb{N}$.

Remember that a problem is called *polynomially tractable* if its information complexity $n(\varepsilon,d)$ is bounded from above by some polynomial in d and ε^{-1}, i.e.,

$$n(\varepsilon,d)\le C\,\varepsilon^{-p}\,d^q\quad\text{for some}\quad C,p>0,\ q\ge 0\quad\text{and all}\quad\varepsilon\in(0,1],\ d\in\mathbb{N}.$$

If the last inequality remains valid even for $q=0$ then we say that the problem is *strongly polynomially tractable*.

Consequently Theorem 1 implies that the integration problem for permutation-invariant functions in the above sense can *never* be strongly polynomially tractable, independent of the size of the sets I_d which describes the number of imposed permutation-invariance conditions. The question under which conditions polynomial tractability holds is posed as an open problem. Apart from that implications for other types of tractability (such as weak, (s,t)-weak, and uniform weak tractability) are presented. For details and further information we refer to [1, 2, 5].

References

[1] E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems. Vol. I: Linear Information*. EMS Tracts in Mathematics 6. European Math. Soc. (EMS), Zürich. 2008.

[2] P. Siedlecki, *Uniform weak tractability*, J. Complexity **29** (2013), 438-453.

[3] I.H. Sloan and H. Woźniakowski, *An intractability result for multiple integration*, Math. Comp. **66** (1997), 1119–1124.

[4] M. Weimar, *The complexity of linear tensor product problems in (anti)symmetric Hilbert spaces*, J. Approx. Theory **164** (2012), 1345–1368.

[5] M. Weimar, *On lower bounds for integration of multivariate permutation-invariant functions*, to appear in J. Complexity.

A survey on recursive nonlinear pseudorandom number generators

Arne Winterhof

Let p be a prime, r a positive integer, $q = p^r$ and denote by $\mathbb{F}_q$ the finite field of q elements. Given a polynomial $f(X) \in \mathbb{F}_q[X]$ of degree $d \geq 2$, we define the *recursive nonlinear pseudorandom number generator* (μ_n) of elements of $\mathbb{F}_q$ by the recurrence relation

$$\mu_{n+1} = f(\mu_n), \qquad n = 0, 1, \ldots, \tag{1}$$

with some *initial value* $\mu_0 \in \mathbb{F}_q$. This sequence is eventually periodic with some period $T \leq q$. We assume that the sequence (μ_n) is purely periodic.

In this survey we mention recent results on different features of these sequences in view of possible applications for Monte Carlo methods or cryptography.

In [15, 17, 20], a method has been presented to study the additive character sums

$$S_{\mathbf{a},N}(f) = \sum_{n=0}^{N-1} \chi\left(\sum_{j=0}^{s-1} \alpha_j \mu_{n+j}\right), \quad 1 \leq N \leq T,$$

and thus the distribution of such sequences for arbitrary polynomials $f(X)$ where χ is a nontrivial additive character of $\mathbb{F}_q$ and $\mathbf{a} = (\alpha_0, \ldots, \alpha_{s-1}) \in \mathbb{F}_q^s \setminus \mathbf{0}$, see also the recent surveys [19, 20, 22]. More precisely, under some necessary restrictions, say $\gcd(d, p) = 1$, we can prove:

$$S_{\mathbf{a},N}(f) \ll N\left(\log \frac{2q}{N}\right)^{1/2} (\log d)^{1/2}/(\log q)^{1/2}, \quad 1 \leq N \leq T. \tag{2}$$

Unfortunately, the general bound (2) is only nontrivial if $d = q^{o(1)}$. In two special cases, nonlinear generators with small *p-weight degree* [10] and *inversive generators*, see [14] and references therein, the method in the proof of (2) leads to stronger bounds. For other special classes of polynomials, namely for *monomials* and *Dickson polynomials*, an alternative approach, producing much stronger bounds has been proposed in [1, 5, 6, 7]. Related results for sequences produced by *Rédei functions* are obtained in [9].

We derive from the sequence (μ_n) defined by (1) a nonlinear method for pseudorandom number generation defined as follows. Let $\{\beta_1, \ldots, \beta_r\}$ be an ordered basis of $\mathbb{F}_q$ over $\mathbb{F}_p$ and identify $\mathbb{F}_p$ with the set of integers $\{0, 1, \ldots, p-1\}$. If

$$\mu_n = u_{n,1}\beta_1 + \ldots + u_{n,r}\beta_r, \quad \text{with } u_{n,i} \in \mathbb{F}_p,$$

then we derive *digital nonlinear pseudorandom numbers* in the unit interval $[0, 1)$ by putting $y_n = \sum_{j=1}^{r} u_{n,j} p^{-j}$.

The *discrepancy* is a measure for the deviation from uniform distribution and thus suitability for Monte Carlo methods. Suitable general discrepancy bounds reduce the discrepancy of the points $\mathbf{y}_n = (y_n, y_{n+1}, \ldots, y_{n+s-1})$, $n = 0, \ldots, N-1$, of consecutive digital nonlinear pseudorandom numbers to the additive character

sums mentioned above, see for example the bound of [12, Theorem 3.12] or for $r = 1$ the *Erdős-Turán-Koksma inequality*, see [4, Theorem 1.21].

For $N \geq 1$ and a sequence (μ_n) over $\mathbb{F}_q$, its *Nth linear complexity* $L(\mu_n, N)$ over $\mathbb{F}_q$ is the smallest integer L such that there exist $\alpha_0, \ldots, \alpha_{L-1} \in \mathbb{F}_q$ such that

$$\mu_{n+L} = \alpha_{L-1}\mu_{n+L-1} + \ldots + \alpha_0\mu_n \quad \text{for } 0 \leq n < N - L,$$

with the conventions that $L(\mu_n, N) = 0$ if $\mu_0 = \cdots = \mu_{N-1} = 0$ and $L(\mu_n, N) = N$ if $\mu_0 = \cdots = \mu_{N-2} = 0$ but $\mu_{N-1} \neq 0$. Its *linear complexity* is $L(\mu_n) = \sup_{N \geq 1} L(\mu_n, N)$. Note that for a T-periodic sequence we have $L(\mu_n) \leq T$. The linear complexity is a measure for the unpredictability and thus suitability in cryptography. For recent surveys on linear complexity and related measures see [13, 21].

For the sequence (μ_n) defined in (1) we know the general lower bound

$$L(\mu_n, N) \geq \frac{\min\{\log(N - \log N/\log d), \log T\}}{\log d}, \quad N \geq 1,$$

of [8, Theorem 4]. Specially tailored results have been proved for the following generators, see [22] and references therein: the inversive generator, power generator, Dickson generator and Rédei generator.

Whereas linear complexity comes from cryptography, a closely related concept called *lattice test* has its origin in the area of Monte Carlo methods. The following lattice test was introduced in [18]. Let (μ_n), $n = 0, 1, \ldots$, be a T-periodic sequence over $\mathbb{F}_q$. For given integers $s \geq 1$, $0 \leq d_0 < d_1 < \ldots < d_{s-1} < T$, and $N \geq 2$, we say that (μ_n) passes the *s-dimensional N-lattice test with lags $d_0, \ldots, d_{s-1}$* if the vectors $\{\vec{u}_n - \vec{u}_0 : 0 \leq n \leq N-1\}$ span $\mathbb{F}_q^s$, where

$$\vec{u}_n = (\mu_{n+d_0}, \mu_{n+d_1}, \ldots, \mu_{n+d_{s-1}}), \quad 0 \leq n \leq N-1.$$

In the case $d_i = i$ for $0 \leq i < s$, this test coincides essentially with the lattice test introduced in [2]. The latter lattice test is closely related to the concept of the linear complexity profile, see [2, 3, 16]. If additionally $N \geq T$, this special lattice test was proposed by Marsaglia [11].

If d_{s-1} is small, the same methods as for estimating the linear complexity profile apply. However, for arbitrary lags little is known [18]. The only good bounds known are for a modification of the inversive generator introduced in [14] and further analyzed in [18].

References

[1] J. Bourgain. Mordell's exponential sum estimate revisited. *J. Amer. Math. Soc.*, 18:477–499, 2005.

[2] G. Dorfer and A. Winterhof. Lattice structure and linear complexity profile of nonlinear pseudorandom number generators. *Appl. Algebra Engrg. Comm. Comput.*, 13:499–508, 2003.

[3] G. Dorfer and A. Winterhof. Lattice structure of nonlinear pseudorandom number generators in parts of the period. In H. Niederreiter (ed.). *Monte Carlo and Quasi-Monte Carlo Methods 2002*, 199–211. Springer, Berlin, 2004.

[4] M. Drmota and R.F. Tichy. *Sequences, discrepancies and applications. Lecture Notes in Mathematics*, 1651. Springer-Verlag, Berlin, 1997.

[5] J.B. Friedlander, J. Hansen, and I.E. Shparlinski. Character sums with exponential functions. *Mathematika*, 47:75–85, 2000.
[6] J.B. Friedlander and I.E. Shparlinski. On the distribution of the power generator. *Math. Comp.*, 70:1575–1589, 2001.
[7] D. Gomez-Perez, J. Gutierrez, and I.E. Shparlinski. Exponential sums with Dickson polynomials. *Finite Fields Appl.*, 12:16–25, 2006.
[8] J. Gutierrez, I.E. Shparlinski, and A. Winterhof. On the linear and nonlinear complexity profile of nonlinear pseudorandom number-generators. *IEEE Trans. Inform. Theory*, 49:60–64, 2003.
[9] J. Gutierrez and A. Winterhof. Exponential sums of nonlinear congruential pseudorandom number generators with Rédei functions. *Finite Fields Appl.*, 14:410–416, 2008.
[10] A. Ibeas and A. Winterhof. Exponential sums and linear complexity of nonlinear pseudorandom number generators with polynomials of small p-weight degree. *Unif. Distrib. Theory*, 5:79–93, 2010.
[11] G. Marsaglia. The structure of linear congruential sequences. In S.K. Zaremba (ed.). *Applications of Number Theory to Numerical Analysis*, 249–285. New York: Academic Press 1972.
[12] H. Niederreiter. *Random number generation and quasi-Monte Carlo methods*, CBMS-NSF Regional Conference Series in Applied Mathematics, 63. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
[13] H. Niederreiter. Linear complexity and related complexity measures for sequences. In *Progress in cryptology—INDOCRYPT 2003. Lecture Notes in Comput. Sci.*, 2904:1–17, 2003.
[14] H. Niederreiter and J. Rivat. On the correlation of pseudorandom numbers generated by inversive methods. *Monatsh. Math.*, 153:251–264, 2008.
[15] H. Niederreiter and I.E. Shparlinski. On the distribution and lattice structure of nonlinear congruential pseudorandom numbers. *Finite Fields Appl.*, 5:246–253, 1999.
[16] H. Niederreiter and A. Winterhof. Lattice structure and linear complexity of nonlinear pseudorandom numbers. *Appl. Algebra Engrg. Comm. Comput.* 13:319–326, 2002.
[17] H. Niederreiter and A. Winterhof. Exponential sums for nonlinear recurring sequences. *Finite Fields Appl.*, 14:59–64, 2008.
[18] H. Niederreiter and A. Winterhof. On the structure of inversive pseudorandom number generators. In S. Boztas and H.-F. Lu. *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes AAECC 17 (Bangalore, 2007)*, *Lecture Notes Comput. Science*, 4851:208–216, 2007.
[19] I.E. Shparlinski. On some dynamical systems in finite fields and residue rings. *Discrete Contin. Dyn. Syst.*, 17: 901–917, 2007.
[20] A. Topuzoğlu and A. Winterhof. Pseudorandom sequences. In *Topics in geometry, coding theory and cryptography*, volume 6 of *Algebr. Appl.*, pages 135–166. Springer, Dordrecht, 2007.
[21] A. Winterhof. Linear complexity and related complexity measures. In *Selected topics in information and coding theory*, pages 3–40. World Scientific, 2010.
[22] A. Winterhof. Recent results on recursive nonlinear pseudorandom number generators. Sequences and their applications–SETA 2010, 113–124, Lecture Notes in Comput. Sci., 6338, Springer, Berlin, 2010.

Well distributed points on the sphere – What do you want?

Robert S. Womersley

Consider the problem of generating a set $\mathcal{X}_N = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ of N well distributed points on the unit sphere $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ for numerical integration or approximation. The geometric quality can be measured by the separation $\delta(\mathcal{X}_N)$ (which is twice the packing radius) and the mesh norm $\rho(\mathcal{X}_N)$ (or covering radius) defined by

$$\begin{aligned} \delta(\mathcal{X}_N) &= \min_{i \neq j} d(\mathbf{x}_i, \mathbf{x}_j), \\ \rho(\mathcal{X}_N) &= \max_{\mathbf{x} \in \mathbb{S}^d} \min_{j=1,\ldots,N} d(\mathbf{x}, \mathbf{x}_j), \end{aligned}$$

where $d(\mathbf{x}, \mathbf{y})$ can be either the geodesic distance $\cos^{-1}(\mathbf{x} \cdot \mathbf{y})$ or the Euclidean distance $|\mathbf{x} - \mathbf{y}|$. The separation affects aspects related to stability, while the mesh norm affects the quality of the approximation or numerical integration rule. The mesh ratio

$$\gamma(\mathcal{X}_N) = \frac{\rho(\mathcal{X}_N)}{\delta(\mathcal{X}_N)} \geq \frac{1}{2}$$

combine these. Many applications and results assume that a sequence of point sets is quasi-uniform, that is $\gamma(\mathcal{X}_N)$ is uniformly bounded, with good separation $\delta(\mathcal{X}_N) \geq c_d^{\text{sep}} N^{-1/d}$ and good covering $\rho(\mathcal{X}_N) \leq c_d^{\text{cov}} N^{-1/d}$.

Spherical t-designs are equal weight numerical integration rules that are exact for all spherical polynomials of degree at most t. Spherical designs were introduced by Delsarte, Goethals and Seidel [10], who gave lower bounds $N^*(t, d)$ on the number of points N required for a spherical t-design. Yudin [21] improved these lower bounds. The existence of spherical t-designs for N sufficiently large has been known for some time (see, for example, the survey Bannai and Bannai [1] and [8]). However only recently Bondarenko, Radchenko and Viazovska [3] have shown that spherical t-designs exist for all $N \geq c_d t^d$. It is the spherical t-designs with $N = O(t^d)$ that are the most interesting and useful.

Yudin [20] has shown that spherical t-designs have a covering radius that is bounded by $\bar{c}_d t^{-1}$. Thus spherical t-designs with $N = O(t^d)$ have good covering. On the other hand it is well known that two spherical t designs are still a spherical t-design, so spherical t-designs with $N = O(t^d)$ points can have arbitrarily bad separation. Bondarenko, Radchenko and Viazovska [4] have recently also shown that there exist spherical t-designs with $N = O(t^d)$ points and good separation.

Grabner and Sloan in an earlier talk at this workshop have shown that if $N/N^*(t, d) < 2$ then the design must have good separation. Examples of computed spherical t-designs for $\mathbb{S}^2$ and $t \leq 180$ and symmetric spherical t-designs for $t \leq 250$, both with $\gamma(\mathcal{X}_N) \leq 1$ and $N/N^*(t, 2) < 2$ (however the difference goes to 0 like $1/t$) are available from the author.

A standard measure of the uniformity of a point set $\mathcal{X}_N$ on the unit sphere is the spherical cap discrepancy

$$D^C(\mathcal{X}_N) = \sup_{\mathcal{C}} \left| \frac{\#(\mathcal{X}_N \cap \mathcal{C})}{N} - \frac{|\mathcal{C}|}{|\mathbb{S}^d|} \right|,$$

where the supremum is taken over all spherical caps. Grabner and Tichy [12] give Erdös-Turan style upper bounds on the spherical cap discrepancy, and for spherical t-designs with $N = O(t^d)$ points show that $D^C(\mathcal{X}_N) \leq O(N^{-1/d})$. For the computed spherical designs these bounds give $D^C(\mathcal{X}_N) \leq N^{-1/2}$, while the lower bound on the spherical cap discrepancy is $N^{\frac{1}{2}-\frac{1}{2d}}$, which for $d = 2$ is $N^{-3/4}$. If the L_2 norm of the discrepancy is used, then the discrepancy can be related to the (generalised) sum of distances using Stolarsky's invariance principle, as discussed in the talk by Brauchart at this workshop and [6, 7].

The Cui and Freeden discrepancy [9] has a simple closed form representation and can be interpreted as the worst case error for the Sobolev space $\mathbb{H}^{\frac{3}{2}}(\mathbb{S}^2)$ [19]. The concept of QMC designs as sequences of point sets which have the optimal order $O(N^{-s/d})$ of convergence of the worst case error for the Sobolev space $\mathbb{H}^s(\mathbb{S}^d)$, was introduced in [7]. For a particular sequence of point sets, the supremum of the values of $s > d/2$ for which it is a QMC design is referred to as the design strength. This gives another criterion associated with the order of convergence of the error for numerical integration for measuring the quality of sequences of point sets. Spherical t-designs with $N = O(t^d)$ provide QMC designs for all $s > d/2$.

Another desirable property is good (minimal) Riesz s-energy, which for $s > 0$ is defined by

$$E^s(\mathcal{X}_N) = \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|^s}.$$

For $s > d$, minimal energy points are asymptotically uniformly distributed [13] and have good separation for quite general manifolds. The limit as $s \to \infty$ gives best packing. Moreover there exist quasi-uniform minimal energy point sets [14]. For $0 < s < d$, the leading term in the asymptotic expansion of the Riesz energy is determined by the continuous integral and the quality of the point set is determined by the coefficient of the next term in the expansion.

Recently Erdélyi and Saff [11] considered the dual concept of polarisation where point sets are chosen to maximize

$$M^s(\mathcal{X}_N) = \min_{\mathbf{x} \in \mathbb{S}^d} \sum_{j=1}^{N} \frac{1}{|\mathbf{x} - \mathbf{x}_j|^s}.$$

In the limit as $s \to \infty$ such point sets give the best covering.

All these criteria, and many others, can be use to choose well-distributed points sets on the sphere $\mathbb{S}^d$. The interesting aspect is that for $d > 1$ they give different point sets (except for a few special cases [8]), while for $d = 1$ equally spaced points on the circle are optimal for many criteria.

In some cases combinations of characteristics may be satisfied, for example quasi-uniform spherical designs or minimal energy points. For all these optimization criteria, it is difficult to find global optimizers, as the problems typically have many local optimum, often very close together. Ideally points which are fast to generate, such as the generalized spiral points [17, 2, 15] or equal area points [17], or transformations of low discrepancy point sets [5] could be used. However many of these schemes are specific to $\mathbb{S}^2$, which is the primary area of interest for applications. Just which point set is best really depends on your particular application.

References

[1] E. Bannai, E. Bannai, *A survey on spherical designs and algebraic combinatorics on spheres*, European J. Comb. **30** (2009) 1392–1425.

[2] R. Bauer, *Distribution of points on a sphere with application to star catalogs*, J. Guid. Control **23**, 1 (2000) 130–137.

[3] A. Bondarenko, D. Radchenko, M. Viazovska, *Optimal asymptotic bounds for spherical designs*, Annal. Math. **178** (2013) 443–452.

[4] A. Bondarenko, D. Radchenko, M. Viazovska, *Well separated spherical designs*, arXiv:1303.5991 [math.MG] July 2013.

[5] J.S. Brauchart, J. Dick, *Quasi-Monte Carlo rules for numerical integration over the unit sphere* $\mathbb{S}^2$, Numer. Math. **121** (2012) 473–502.

[6] J.S. Brauchart, J. Dick, *A characterization of Sobolev spaces on the sphere and an extension of Stolarsky's invariance principle to arbitrary smoothness*, Constr. Approx. (to appear).

[7] J.S. Brauchart, E.B. Saff, I.H. Sloan, R.S. Womersley, *QMC designs: optimal order Quasi Monte Carlo integration schemes on the sphere*, Math. Comp. (to appear).

[8] H. Cohn, A. Kumar, *Universally optimal distribution of points on spheres*, J. Amer. Math. Soc. **20** (2007) 99–148.

[9] J. Cui, W. Freeden, *Equidistribution on the sphere*, SIAM J. Sci. Comput. **18**, 2 (1997) 595–609.

[10] P. Delsarte, J.M. Goethals, J.J. Seidel, *Spherical codes and designs*, Geom. Dedicata **6** (1977), 363–388.

[11] T. Erdély, E.B. Saff, *Riesz polarization inequalities in higher dimensions*, J. Approx. Th. **171** (2013), 128–147.

[12] P.J. Grabner, R.F. Tichy, *Spherical designs, discrepancy and numerical integration*, Math. Comp. **60** (1993), 327–336.

[13] D.P. Hardin, E.B.. Saff, *Discretizing manifolds via minimum energy points*, Notices Amer. Math. Soc. **51**, 10 (2004), 1186–1194.

[14] D.P. Hardin, E.B.. Saff, J.T. Whitehouse, *Quasi-uniformity of minimal weighted energy points on compact metric spaces*, J. Complexity **28**, 2, (2012), 177–191.

[15] C.G. Koay, *Analytically exact spiral scheme for generating uniformly distributed points on the unit sphere*, J. Comp. Sci. **2**, 2, (2011), 88–91.

[16] A.B.J. Kuijlaars, E.B. Saff, *Asymptotics for minimal discrete energy on the sphere*, Trans. Amer. Math. Soc. **350**, (1998), 523–538.

[17] E.A. Rakhmanov, E.B. Saff, Y.M. Zhou, *Minimal discrete energy on the sphere*, Math. Res. Lett. **1** (1994), 647–662.

[18] E.B. Saff, A.B.J. Kuijlaars, *Distributing many points on a sphere*, Math. Intell. **19**, 1 (1997), 7–11.

[19] I.H. Sloan, R.S. Womersley, *Extremal systems of points and numerical integration on the sphere*, Adv. Comp. Math. **21** (2004), 107–125.

[20] V.A. Yudin, *Covering a sphere and extremal properties of orthogonal polynomials*, Discrete Math. Appl. **5** (1995), 371–379.

[21] V.A. Yudin, *Lower bounds for spherical designs*, Izv. Math. **61** (1997), 673–683.

Reporter: Michael Gnewuch

Participants

Dr. Christoph Aistleitner
Department of Mathematics
Kobe University
Rokko-dai, 1-1 Nada-ku
Kobe 657-8501
JAPAN

Prof. Dr. Dmytro Bilyk
School of Mathematics
University of Minnesota
127 Vincent Hall
Minneapolis MN 55455-0436
UNITED STATES

Dr. Johann S. Brauchart
School of Mathematics and Statistics
University of New South Wales
Sydney NSW 2052
AUSTRALIA

Prof. Dr. Ronald Cools
Departement Computerwetenschappen
Katholieke Universiteit Leuven
Celestijnenlaan 200A
3001 Heverlee
BELGIUM

Prof. Dr. Steffen Dereich
Institut f. Mathematische Statistik
Universität Münster
Einsteinstr. 62
48149 Münster
GERMANY

Dr. Josef Dick
School of Mathematics & Statistics
The University of New South Wales
Sydney NSW 2052
AUSTRALIA

Prof. Dr. Benjamin Doerr
Laboratoire d'Informatique
École Polytechnique de Paris
Plateau de Palaiseau
91126 Palaiseau Cedex
FRANCE

Dr. Carola Doerr
LIP 6 Laboratory
Université PMC - Paris VI
Case 169
4, Place Jussieu
75252 Paris Cedex 05
FRANCE

Prof. Dr. Henri Faure
Institut de Mathématiques de Luminy
Case 907
163 Avenue de Luminy
13288 Marseille Cedex 9
FRANCE

PD Dr. Michael Gnewuch
Fachbereich Mathematik
T.U. Kaiserslautern
Postfach 3049
67618 Kaiserslautern
GERMANY

Dr. Takashi Goda
Grad. School of Information Science & Techn.
The University of Tokyo
2-11-16 Yayoi, Bunkyo-ku
Tokyo
JAPAN

Prof. Dr. Peter J. Grabner
Institut für Analysis und Computational Number Theory
Technische Universität Graz
Steyrergasse 30
8010 Graz
AUSTRIA

Mario Hefter
Technische Universität Kaiserslautern
Fachbereich Mathematik
Postfach 3049
67653 Kaiserslautern
GERMANY

Prof. Dr. Stefan Heinrich
Fachbereich Informatik
Universität Kaiserslautern
Postfach 3049
67653 Kaiserslautern
GERMANY

Andre Herzwurm
Fachbereich Mathematik
T.U. Kaiserslautern
Postfach 3049
67653 Kaiserslautern
GERMANY

Prof. Dr. Aicke Hinrichs
Institut für Mathematik
Universität Rostock
18057 Rostock
GERMANY

Dr. Roswitha Hofer
Institut für Finanzmathematik
Universität Linz
Altenbergerstr. 69
4040 Linz
AUSTRIA

Prof. Dr. Stephen Joe
Department of Mathematics
University of Waikato
Private Bag 3105
3240 Hamilton
NEW ZEALAND

Dr. Peter Kritzer
Institut für Finanzmathematik
Universität Linz
Altenbergerstr. 69
4040 Linz
AUSTRIA

Prof. Dr. Thomas Kühn
Mathematisches Institut
Universität Leipzig
Augustusplatz 10
04109 Leipzig
GERMANY

Dr. Frances Y. Kuo
School of Mathematics and Statistics
The University of New South Wales
Sydney NSW 2052
AUSTRALIA

Prof. Dr. Michael T. Lacey
School of Mathematics
Georgia Institute of Technology
686 Cherry Street
Atlanta, GA 30332-0160
UNITED STATES

Prof. Dr. Gerhard Larcher
Institut für Finanzmathematik
Universität Linz
Altenbergerstr. 69
4040 Linz
AUSTRIA

Prof. Dr. Pierre L'Ecuyer
Dept. of Mathematics and Statistics
University of Montreal
CP 6128, succ. Centre Ville
Montreal, QC H3C 3J7
CANADA

Dr. Gunther Leobacher
Institut für Finanzmathematik
Universität Linz
Altenbergerstr. 69
4040 Linz
AUSTRIA

Sangmeng Li
Institut f. Mathematische Statistik
Universität Münster
48149 Münster
GERMANY

Sebastian Mayer
Hausdorff Center for Mathematics
Institute for Numerical Simulation
Endenicher Allee 62
53115 Bonn
GERMANY

Prof. Dr. Thomas Müller-Gronbach
Fakultät f. Mathematik & Informatik
Universität Passau
Innstr. 33
94032 Passau
GERMANY

Prof. Dr. Harald Niederreiter
Johann Radon Institute for Comput. and Applied Mathematics (RICAM)
Austrian Academy of Sciences
Altenbergerstr. 69
4040 Linz
AUSTRIA

Prof. Dr. Erich Novak
Mathematisches Institut
Friedrich-Schiller-Universität
07743 Jena
GERMANY

Prof. Dr. Dirk Nuyens
Departement Computerwetenschappen
Katholieke Universiteit Leuven
Celestijnenlaan 200A
3001 Heverlee
BELGIUM

Prof. Dr. Friedrich Pillichshammer
Institut für Finanzmathematik
Universität Linz
Altenbergerstr. 69
4040 Linz
AUSTRIA

Prof. Dr. Leszek Plaskota
Institute of Applied Mathematics
University of Warsaw
ul. Banacha 2
02-097 Warszawa
POLAND

Prof. Dr. Daniel Potts
Fakultät für Mathematik
Technische Universität Chemnitz
09126 Chemnitz
GERMANY

Prof. Dr. Klaus Ritter
Fachbereich Mathematik
T.U. Kaiserslautern
Erwin-Schrödinger-Straße
67653 Kaiserslautern
GERMANY

Dr. Daniel Rudolf
Fakultät f. Mathematik & Informatik
Friedrich-Schiller-Universität
Ernst-Abbe-Platz 1-4
07743 Jena
GERMANY

Dr. Wolfgang Christian Schmid
Fachbereich Mathematik
Universität Salzburg
Hellbrunnerstr. 34
5020 Salzburg
AUSTRIA

Prof. Dr. Ian H. Sloan
School of Mathematics and Statistics
The University of New South Wales
Sydney NSW 2052
AUSTRALIA

Stefan Steinerberger
Mathematisches Institut
Universität Bonn
Endenicher Allee 60
53115 Bonn
GERMANY

Kousuke Suzuki
Department of Mathematical Sciences
University of Tokyo
3-8-1 Komaba, Meguro-ku
Tokyo 153-8914
JAPAN

Prof. Dr. Vladimir N. Temlyakov
Department of Mathematics
University of South Carolina
Columbia, SC 29208
UNITED STATES

Prof. Dr. Shu Tezuka
Faculty of Mathematics
Kyushu University
744, Motooka, Nishi-ku
Fukuoka 819-0395
JAPAN

Prof. Dr. Robert F. Tichy
Institute for Analysis and
Computational Number Theory
Technische Universität Graz
Steyrergasse 30
8010 Graz
AUSTRIA

Prof. Dr. Vilmos Totik
Department of Mathematics
University of South Florida
Tampa, FL 33620-5700
UNITED STATES

Prof. Dr. Giancarlo Travaglini
Dipt. Statistica & Metodi Quantitativi
Università Milano Bicocca
Via Bicocca degli Arcimboldi 8
20126 Milano
ITALY

Maryna Viazovska
Max-Planck-Institut für Mathematik
Vivatsgasse 7
53111 Bonn
GERMANY

Prof. Dr. Grzegorz W. Wasilkowski
Computer Science Department
University of Kentucky
Davis Marksbury Building
329 Rose St.
Lexington KY 40506-0633
UNITED STATES

Dr. Markus Weimar
Fachbereich Mathematik
Universität Marburg
35032 Marburg
GERMANY

Prof. Dr. Arne Winterhof
Johann Radon Institute
Austrian Academy of Sciences
Altenberger Straße 69
4040 Linz
AUSTRIA

Prof. Dr. Robert S. Womersley
School of Mathematics and Statistics
The University of New South Wales
Sydney NSW 2052
AUSTRALIA

Prof. Dr. Henryk Wozniakowski
Department of Computer Science
Columbia University
450 Computer Science Building
1214 Amsterdam Avenue
New York, NY 10027-7003
UNITED STATES