Mathematisches Forschungsinstitut Oberwolfach

Report No. 38/2015

DOI: 10.4171/OWR/2015/38

Applied Harmonic Analysis and Sparse Approximation

Organised by
Ingrid Daubechies, Durham
Gitta Kutyniok, Berlin
Holger Rauhut, Aachen
Thomas Strohmer, Davis

16 August – 22 August 2015

Abstract. Efficiently analyzing functions, in particular multivariate functions, is a key problem in applied mathematics. The area of applied harmonic analysis has a significant impact on this problem by providing methodologies both for theoretical questions and for a wide range of applications in technology and science, such as image processing. Approximation theory, in particular the branch of the theory of sparse approximations, is closely intertwined with this area with a lot of recent exciting developments in the intersection of both. Research topics typically also involve related areas such as convex optimization, probability theory, and Banach space geometry. The workshop was the continuation of a first event in 2012 and intended to bring together world leading experts in these areas, to report on recent developments, and to foster new developments and collaborations.

Mathematics Subject Classification (2010): 42-XX, 65Txx, 94Axx, 65K05, 15A52.

Introduction by the Organisers

The workshop *Applied Harmonic Analysis and Sparse Approximation* was organized by Ingrid Daubechies (Durham), Gitta Kutyniok (Berlin), Holger Rauhut (Aachen) and Thomas Strohmer (Davis). This meeting was attended by 57 participants from 11 countries and 3 continents.

Applied Harmonic Analysis provides one key approach towards the problem of efficiently representing, decomposing, and analyzing univariate and multivariate functions. Its applications range from theoretical ones such as the decomposition of specific operators to more practical ones such as image and signal processing as well as inverse problems. Research is typically driven by real-world applications leading to mathematically highly challenging questions, thereby also significantly advancing the mathematical understanding of harmonic analysis itself and in turn impacting the respective application.

The area of sparse approximation with its daughter compressed sensing constitutes an even more recent development which is closely intertwined with applied harmonic analysis, but also has roots in other areas such as statistics, optimization, Banach space geometry and random matrix theory. Despite its young age, this field has already reached a mature state and is nowadays considered a mathematical discipline of its own. Some parts of the core theory are nowadays established, which led to the fact that research in this area has gained even more momentum with many exciting new research directions such as structured dictionary learning, matrix completion, novel methods for phaseless reconstruction, and high-dimensional function reconstruction.

One key focus of research in applied harmonic analysis is the introduction and analysis of representation systems which are designed according to their Fourier domain behavior with constraints such as to derive optimally sparse approximations of curvilinear singularities. Examples are **Gabor systems**, **wavelets**, and also the novel representation systems of **curvelets** and **shearlets**. Such systems are typically utilized for and have impacted both theoretically oriented questions such as sparse expansions of Fourier integral operators and application oriented areas such as image processing. A very recent development just in this year was a first approach to providing a unifying mathematical framework called **parabolic molecules** for all directional representation systems based on parabolic scaling such as curvelets and shearlets. But even for the single systems such as shearlets, many key questions are far from being solved, such as the introduction and characterization of associated function spaces. Also associated functional analytic properties are often far from being well understood, since most of such systems do not constitute orthonormal bases but form redundant systems for which a natural concept with additional stability properties is the notion of a **frame**. Interestingly, **frame theory** – which might be even considered a theory of its own, studying various aspects of redundancy as a mathematical concept – provides another link to the area of sparse approximations elaborated upon below. Besides structured representation systems, one main novel focus is also on **dictionary learning**, i.e., the generation of data-dependent sparsifying systems. Due to the highly complex nature of this problem, and with it the highly unstructured systems which are generated, the mathematical theory is at its beginning. Recently, **structured dictionary learning** has become one new focus of research in an attempt to bridge the gap between structured systems and dictionary learning, and so-called **α-shearlets** might be regarded as one such development.

The paradigms of **sparsity** and **sparse approximations** have had a tremendous impact on various areas in applied mathematics such as imaging sciences and signal processing. It states that functions and signals which come from applications typically exhibit the property of admitting an (approximate) representation in a suitable orthonormal basis or frame. Suitable representation systems were and are still being developed in the areas of applied harmonic analysis such as Gabor systems, wavelets, curvelets, or shearlets. Although compression schemes such as JPEG as well as denoising via thresholding might be considered the first breakthroughs of this general approach, quite recently, the new area of compressed sensing revealed another use of sparse representations with tremendous impact. Roughly speaking, it showed that signals exhibiting a sparse approximation can be recovered efficiently from what would previously have been considered highly incomplete measurements. This discovery has led to a canon of fundamentally new approaches for various previously considered almost insolvable problems, for instance, for signal and image recovery problems. Remarkably, good strategies for designing the measurement process known so far are based on randomness, and the mathematical research in compressed sensing uses also tools, sometimes quite sophisticated, from probability theory and the geometry of Banach spaces.

Several new directions have emerged on the heels of compressed sensing: **Low-rank matrix recovery** aims at recovering a matrix with small rank from incomplete data. In particular, **matrix completion** recovers the matrix from only a small fraction of its entries. Since low-rank structures arise in numerous applications, one can expect an enormous impact. **Phase retrieval** aims at reconstruction of signals from measurements where the phase information is missing. While previous methods both lack stability and provable reconstruction guarantees, recent reconstruction algorithms based on ideas from matrix completion do provide both of these features. Such methods are expected to have a major impact in applications such as X-ray cristallography. Many challenging mathematical problems remain open in these areas.

Interesting recent developments have also occurred at the intersection of harmonic analysis, **high-dimensional manifolds**, and **large data sets**. Traditional reconstruction methods suffer from the curse of dimension which predicts that the required number of samples scales exponentially in the number of variables. Based on similar ideas as in sparse approximation, recent approaches are able to circumvent this phenomenon and work with much fewer samples. In a different direction, certain data processing applications model the data as being elements of a high-dimensional manifold. Tools from harmonic analysis are very useful in describing and processing such manifolds. Yet another very promising direction in this context is to apply concepts from harmonic analysis to graphs and complex data sets.

The workshop featured 29 talks, thereof 9 longer overview talks. Moreover, a session of short presentations of 3 minutes took place on Monday, which we called the *3 Minutes of Fame* (following Andy Warhol's concept of 15 minutes of fame). Every participant had the possibility to contribute to this session, and this session worked out very well. It provided a quick overview on what the participants are presently working or would like to discuss with other participants.

Some highlights of the program included:

- **Multiscale geometric methods for high-dimensional data analysis:** Mauro Maggioni gave an overview on multiscale methods for building up efficient representations for data in high dimensions, where the assumption consists in the data being (almost) contained in a low dimensional manifold. He presented various interesting applications such as learning metastable dynamical systems from short time simulations, dictionary learning via multiscale SVDs.

- **Variable bandwidth:** Several mathematical approaches for formulating the concept of a function which has a bandwidth that varies over time have been formulated in the last decades. All of these have certain shortcomings. Karlheinz Gröchenig presented an elegant new approach to this topic and showed that, in contrast to previous approaches, natural analogs of the sampling theorem hold in this context. This work represents a promising new direction in applied harmonic analysis.
- **Sparse approximation:** Several contributions to sparse approximation were presented: A first analysis of ℓ_1-support vector machine (Jan Vybiral), a deterministic sparse Fourier algorithm (Gerlind Plonka), sparse analysis for art conservation (Ingrid Daubechies), advances on sparse dictionary learning (Karin Schnass),
- **Low rank recovery in quantum mechanics:** David Gross offered a nice introduction to quantum mechanics and the use of low rank approximation techniques for quantum state tomography. Building on this, Richard Küng presented a brand new application in the design of quantum optical curcuits, where techniques of low rank matrix recovery (extensions of compressed sensing) may be very important.

The organizers would like to take the opportunity to thank MFO for providing support and a very inspiring environment for the workshop. The magic of the place and the pleasant atmosphere contributed greatly to the success of the workshop.

Acknowledgement: The MFO and the workshop organizers would like to thank the National Science Foundation for supporting the participation of junior researchers in the workshop by the grant DMS-1049268, “US Junior Oberwolfach Fellows”.

Workshop: Applied Harmonic Analysis and Sparse Approximation

Table of Contents

Abstracts

Non-unique games over compact groups

Afonso S. Bandeira
(joint work with Yutong Chen and Amit Singer)

Let $\mathcal{G}$ be a compact group and let $f_{ij} \in L^2(\mathcal{G})$. The *Non-Unique Games* (NUG) problem is defined as

$$\begin{aligned} &\underset{g_1,\dots,g_n}{\text{minimize}} && \sum_{i,j=1}^{n} f_{ij}\left(g_i g_j^{-1}\right) \\ &\text{subject to} && g_i \in \mathcal{G}, \end{aligned} \tag{1}$$

Many inverse problems can be solved as instances of (1). A simple example is angular synchronization [18, 5], where one is tasked with estimating angles $\{\theta_i\}_i$ from information about their offsets $\theta_i - \theta_j \mod 2\pi$. The problem of estimating the angles can then be formulated as an optimization problem depending on the offsets, and thus be written in the form of (1). In general, many inverse problems, where the goal is to estimate multiple group elements from information about group offsets, can be formulated as (1).

One of the simplest instances of (1) is the `Max-Cut` problem, where the objective is to partition the vertices of a graph as to maximize the number of edges (the *cut*) between the two sets. In this case, $\mathcal{G} \cong \mathbb{Z}_2$, the group of two elements $\{\pm 1\}$, and f_{ij} is zero if (i,j) is not an edge of the graph and

$$f_{ij}(1) = 0 \text{ and } f_{ij}(-1) = -1,$$

if (i,j) is an edge. In fact, the semidefinite programming based approach towards (1) discussed here is inspired by — and can be seen as a generalization of— the semidefinite relaxation for the `Max-Cut` problem by Goemans and Williamson [12].

Another important source of inspiration was the semidefinite relaxation of the `Max-2-Lin`$(\mathbb{Z}_L)$ problem, proposed in [9], for the *Unique Games* problem, a central problem in theoretical computer science [14, 15]. Given integers n and L, an Unique-Games instance is a system of linear equations over $\mathbb{Z}_L$ on n variables $\{x_i\}_{i=1}^n$. Each equation constraints the difference of two variables. More precisely, for each (i,j) in a subset of the pairs, we associate a constraint

$$x_i - x_j = b_{ij} \mod L.$$

The objective is then to find $\{x_i\}_{i=1}^n$ in $\mathbb{Z}_L$ that satisfy as many equations as possible. This can be easily described within our framework by taking, for each constraint,

$$f_{ij}(g) = -\delta_{g \equiv b_{ij}},$$

and $f_{ij} = 0$ for pairs not corresponding to constraints.

The semidefinite relaxation for the unique games problem proposed in [9] was investigated in [6] in the context of the signal alignment problem, where the f_{ij}

are not forced to have a special structure (but $\mathcal{G} \cong \mathbb{Z}_L$). This framework can be seen as a generalization of the approach in [6] to other compact groups $\mathcal{G}$.

Besides the signal alignment problem treated in [6] the semidefinite relaxation to the NUG problem coincides with other effective relaxations. When $\mathcal{G} \cong \mathbb{Z}_2$ it coincides with the semidefinite relaxations for `Max-Cut` [12], little Grothendieck problem over $\mathbb{Z}_2$ [3, 17], recovery in the stochastic block model [2, 4], and Synchronization over $\mathbb{Z}_2$ [1, 4, 11]. When $\mathcal{G} \cong SO(2)$ and the functions f_{ij} are linear with respect to the representation $\rho_1 : SO(2) \to \mathbb{C}$ given by the $\rho_1(\theta) = e^{i\theta}$, it coincides with the semidefinite relaxation for angular synchronization [18]. Similarly, when $\mathcal{G} \cong O(d)$ and the functions are linear with respect to the natural d-dimensional representation, then the NUG problem essentially coincides with the little Grothendieck problem over the orthogonal group [8, 16]. Other examples include the shape matching problem in computer graphics for which $\mathcal{G}$ is a permutation group (see [13, 10]).

We refer the reader to the manuscript [7] for more on the Non-Unique Games framework and for a description of its application to the problem of orientation estimation in Cryo-Electron Microscopy.

References

[1] E. Abbe, A. S. Bandeira, A. Bracher, and A. Singer. Decoding binary node labels from censored edge measurements: Phase transition and efficient recovery. *Network Science and Engineering, IEEE Transactions on*, 1(1):10–22, Jan 2014.

[2] E. Abbe, A. S. Bandeira, and G. Hall. Exact recovery in the stochastic block model. *Available online at arXiv:1405.3267 [cs.SI]*, 2014.

[3] N. Alon and A. Naor. Approximating the cut-norm via Grothendieck's inequality. In *Proc. of the 36 th ACM STOC*, pages 72–80. ACM Press, 2004.

[4] A. S. Bandeira. Random Laplacian matrices and convex relaxations. *Available online at arXiv:1504.03987 [math.PR]*, 2015.

[5] A. S. Bandeira, N. Boumal, and A. Singer. Tightness of the maximum likelihood semidefinite relaxation for angular synchronization. *Available online at arXiv:1411.3272 [math.OC]*, 2014.

[6] A. S. Bandeira, M. Charikar, A. Singer, and A. Zhu. Multireference alignment using semidefinite programming. *5th Innovations in Theoretical Computer Science (ITCS 2014)*, 2014.

[7] A. S. Bandeira, Y. Chen, and A. Singer. Non-unique games over compact groups and orientation estimation in cryo-em. *Available online at arXiv:1505.03840 [cs.CV]*, 2015.

[8] A. S. Bandeira, C. Kennedy, and A. Singer. Approximating the little grothendieck problem over the orthogonal group. *Available online at arXiv:1308.5207 [cs.DS]*, 2013.

[9] M. Charikar, K. Makarychev, and Y. Makarychev. Near-optimal algorithms for unique games. *Proceedings of the 38th ACM Symposium on Theory of Computing*, 2006.

[10] Y. Chen, Q.-X. Huang, and L. Guibas. Near-optimal joint object matching via convex relaxation. *Proceedings of the 31st International Conference on Machine Learning*, 2014.

[11] M. Cucuringu. Synchronization over Z2 and community detection in signed multiplex networks with constraints. *Journal of Complex Networks*, 2015.

[12] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefine programming. *Journal of the Association for Computing Machinery*, 42:1115–1145, 1995.

[13] Q.-X. Huang and L. Guibas. Consistent shape maps via semidefinite programming. *Computer Graphics Forum*, 32(5):177–186, 2013.

[14] S. Khot. On the power of unique 2-prover 1-round games. *Thiry-fourth annual ACM symposium on Theory of computing*, 2002.

[15] S. Khot. On the unique games conjecture (invited survey). In *Proceedings of the 2010 IEEE 25th Annual Conference on Computational Complexity*, CCC '10, pages 99–121, Washington, DC, USA, 2010. IEEE Computer Society.

[16] A. Naor, O. Regev, and T. Vidick. Efficient rounding for the noncommutative Grothendieck inequality. In *Proceedings of the 45th annual ACM symposium on Symposium on theory of computing*, STOC '13, pages 71–80, New York, NY, USA, 2013. ACM.

[17] Y. Nesterov. Semidefinite relaxation and nonconvex quadratic optimization. *Optimization Methods and Software*, 9(1-3):141–160, 1998.

[18] A. Singer. Angular synchronization by eigenvectors and semidefinite programming. *Appl. Comput. Harmon. Anal.*, 30(1):20 – 36, 2011.

Solving equations using nonlinear approximations

Gregory Beylkin

The usual approach to solving partial differential and integral equations is to select a basis (possibly a multiresolution basis) or a grid, project equations onto such basis and solve the resulting discrete equations. An alternative is to look for the solution within a large class of functions (larger than any basis) by constructing optimal or near optimal approximations at every step of an (iterative) algorithm for solving the equations. We present two examples of such solvers, one for the viscous Burgers' equation [6] and, another, for solving the Hartree-Fock equations of quantum chemistry [2], and discuss the merits of the approach.

The choice of the Burgers' equation allows us to thoroughly test algorithms for constructing rational approximations with (near) optimally small L^∞ error. When the viscosity ν is small, solutions of Burgers' equation develop sharp (moving) transition regions of width $\mathcal{O}(\nu)$, which presents a challenge for standard numerical methods.

In solving the Hartree-Fock equations, we present a new approach [2] to electronic structure calculations based on recently developed algorithms for computing near optimal approximations [5, 1, 4]. We maintain a functional form for the spatial orbitals consisting of linear combinations of products of decaying exponentials and spherical harmonics centered at the nuclear cusps. While such representations are similar to the classical Slater-type orbitals, in the course of computation we optimize both the exponents and the coefficients in order to achieve an efficient representation of solutions and to obtain guaranteed error bounds. In this way, we combine the efficiency of traditional Slater-type representations with the adaptivity of current multiresolution methods [3, 7, 8].

References

[1] C. Ahrens and G. Beylkin. *Rotationally Invariant Quadratures for the Sphere* Proceedings of the Royal Society A, **465** (2009), 3103–3125.

[2] G. Beylkin and T. S. Haut. *Nonlinear approximations for electronic structure calculations* Proceedings of the Royal Society A, **469** (2013), 20130231

[3] R.J. Harrison, G.I. Fann, T. Yanai, Z. Gan, and G. Beylkin. *Multiresolution quantum chemistry: basic theory and initial applications* J. Chem. Phys., **121** (2004), 11587–11598.

[4] G. Beylkin and L. Monzón. *On approximation of functions by exponential sums* Appl. Comput. Harmon. Anal., **19** (2005), 17–48.

[5] T. S. Haut and G. Beylkin. *Fast and accurate con-eigenvalue algorithm for optimal rational approximations* SIAM J. Matrix Anal. Appl., **33** (2012), 1101–1125.

[6] T. S. Haut, G. Beylkin, and L. Monzón. *Solving Burgers' equation using optimal rational approximations* Appl. Comput. Harmon. Anal., **34** (2013), 83–95.

[7] T. Yanai, G.I. Fann, Z. Gan, R.J. Harrison, and G. Beylkin. *Multiresolution quantum chemistry: Analytic derivatives for Hartree-Fock and density functional theory* J. Chem. Phys., **121** (2004), 2866–2876.

[8] T. Yanai, G.I. Fann, Z. Gan, R.J. Harrison, and G. Beylkin. *Multiresolution quantum chemistry: Hartree-Fock exchange* J. Chem. Phys., **121** (2004), 6680–6688.

Robust and stable compressive phase retrieval

Bernhard G. Bodmann

(joint work with Nathaniel Hammen)

This abstract is concerned with the problem of recovering a signal from magnitude measurements. In our formulation of this so-called phase retrieval problem, we represent the signal as a d-dimensional vector $x \in \mathbb{C}^d$ and take noisy measurements as $b_i = |\langle x, f_i\rangle|^2 + \epsilon_i$ for some set of measurement vectors $\{f_i\}_{i=1}^M$ and measurement noise $\{\epsilon_i\}_{i=1}^M$. In this setting, the objective is to stably recover an approximation to the vector x from the measurements $\{b_i\}_{i=1}^M$, up to an overall unimodular factor.

Often, there are fewer measurements that are feasibly available than the dimension of the signal to be recovered. Phase retrieval allows recovery of a vector from fewer linear measurements than the dimension of the vector, if the vector is known to be sparse. Thus, we would like to combine phase retrieval results with compressive sensing results. To do this, we represent noisy measurements as $b_i = |\langle Ax, f_i\rangle|^2 + \epsilon_i$ for some matrix A that allows underdetermined recovery of a sparse vector and some set of measurement vectors $\{f_i\}_{i=1}^M$ and measurement noise $\{\epsilon_i\}_{i=1}^M$. In this case, a phase retrieval algorithm can recover the vector Ax, and then a compressive sensing algorithm can recover the vector x. This type of procedure has been shown to have an error bound that is linear in terms of the input noise [2].

The recovery result presented here is stable with respect to input noise and also allows the assumption of sparsity to be only approximately satisfied. The estimates are an improvement over prior work [1] and include the treatment of sparsity. To define the notion of approximate sparsity, we recall the following definitions.

For any vector $x \in \mathbb{C}^N$, we define the error of best s-term approximation to x by

$$\sigma_s(x)_1 = \min_{z \in \mathbb{C}^N, \|z\|_0 \leq s} \|x - z\|_1$$

and a best s-term approximation to x is given by

$$H_s(x) = \arg \min_{z \in \mathbb{C}^N, \|z\|_0 \leq s} \|x - z\|_1 .$$

Note that this best s-term approximation is not necessarily unique for a given x, but that the error $\sigma_s(x)$ and the norm $\|H_s(x)\|_2$ are independent of the choice of $H_s(x)$.

Next, we state the main result. Let $N \geq d \geq s$. Let $x \in \mathbb{C}^M$, let $\eta_0, \eta_1, \eta_2 \in \mathbb{R}^{2N-1}$, and let $\Phi \in \mathbb{C}^{d \times M}$ satisfy the ℓ_2-robust null space property of order s with constants $0 < \rho < 1$ and $\tau > 0$. If $\upsilon = e^{\frac{2i\pi}{2N-1}}$ and $\nu = e^{\frac{2i\pi}{d}}$, let $B \in \mathbb{C}^{(6N-3)\times d}$ be given by

$$B_{j,k} = \begin{cases} \upsilon^{j(k-1)} & \text{if } 1 \leq j \leq 2N-1, \\ \upsilon^{j(k-1)} - (\upsilon^j \nu)^{k-1} & \text{if } 2N \leq j \leq 4N-2, \\ \upsilon^{j(k-1)} - i(\upsilon^j \nu)^{k-1} & \text{if } 4N-1 \leq j \leq 6N-3 \end{cases}$$

and let $\epsilon \in \mathbb{R}^{6N-3}$ be given by

$$\epsilon_j = \begin{cases} (\eta_0)_j & \text{if } 1 \leq j \leq 2N-1, \\ (\eta_1)_{j-(2N-1)} & \text{if } 2N \leq j \leq 4N-2, \\ (\eta_2)_{j-(4N-2)} & \text{if } 4N-1 \leq j \leq 6N-3\,. \end{cases}$$

Let $r = \sin(\frac{2\pi}{(d-1)d^2})$, $0 < \alpha < 1$, and $\beta = \frac{r^{\frac{(d-1)d}{2}}\left(\frac{d-1}{2d}\right)^d \frac{2}{d-1}}{\left(\prod_{k=1}^{d-1}(r^k+1)\right)}$. If the noise is bounded by $\|\eta_0\|_2 \leq \frac{\sqrt{2N-1}}{\sqrt{2d-1}}\alpha\beta^2\|p\|_2^2$ and x satisfies the approximate sparsity requirement

$$\rho\sigma_s(x)_1 < \sqrt{s}\|H_s(x)\|_2$$

then an approximation $x^{\#}$ for x can be reconstructed from the vector $|B\Phi x|^2 + \epsilon$ (where $|\bullet|^2$ is taken component-wise), such that

$$\|c_0 x - x^{\#}\|_2 \leq \frac{C_1}{\sqrt{s}}\sigma_s(x)_1 + \frac{C_2}{\sqrt{2N-1}} \frac{\|\epsilon\|_2}{\|H_s(x)\|_2 - \frac{\rho}{\sqrt{s}}\sigma_s(x)_1}$$

where $C_1 = \frac{(1+\rho)^2}{1-\rho}$, and $C_2 = \frac{\sqrt{2}(3+\rho)\tau^2}{1-\rho}\left\|w\left(C, \frac{1}{d}\beta^2(1-\alpha)\right)\right\|_2$ with

$$C = \frac{(1+\sqrt{2})\frac{\sqrt{2d-1}}{\sqrt{2N-1}}\max_{0\leq l\leq 2}\{\|\eta_l\|_2\} + d\|x\|_2^2}{\beta^2(1-\alpha)\|x\|_2^2}\sqrt{d}$$

and

$$w(C,m)_j = \begin{cases} \sum_{k=2}^{d} \frac{C^{k-2}(1+\frac{\sqrt{2}}{2})}{m_1} + \sum_{k=1}^{d} \frac{C^{k-1}}{2\sqrt{m_1}} & \text{if } j = 1 \\ \sum_{k=1}^{d-j} \frac{C^{k-1}(1+\frac{\sqrt{2}}{2})}{m_j} + \sum_{k=0}^{d-j} \frac{C^k \frac{\sqrt{2}}{2}}{m_{j-1}} & \text{if } 2 \leq j \leq d \\ \sum_{k=1}^{2d-j} \frac{C^{k-1}}{2m_{j-d}} & \text{if } d+1 \leq j \leq 2d-1 \\ \sum_{k=1}^{3d-1-j} \frac{C^{k-1}}{2m_{j+1-2d}} & \text{if } 2d \leq j \leq 3d-2 \end{cases}$$

and $c_0 \in \mathbb{C}, |c_0| = 1$ is a remaining undetermined unimodular factor, .

References

[1] B. G. Bodmann and N. Hammen, Algorithms and error bounds for noisy phase retrieval with low-redundancy frames, preprint, arXiv:1412.6678.

[2] M. Iwen, Mark, A. Viswanathan, and Y. Wang, Robust sparse phase retrieval made easy, preprint, arXiv:1410.5295.

Adaptive signal processing, Hilbert transform, and a problem of Ul'yanov

Holger Boche
(joint work with Volker Pohl)

Let $\mathrm{T} : \mathcal{B}_1 \to \mathcal{B}_2$ be a bounded linear operator between Banach spaces $\mathcal{B}_1$ and $\mathcal{B}_2$. An important problem in analysis with many applications in engineering and science is the approximation of T by a sequence $\{\mathrm{T}_N\}_{N\in\mathbb{N}}$ of linear, bounded operators $\mathrm{T}_N : \mathcal{B}_1 \to \mathcal{B}_2$. Practical problems often imply more or less stringent restrictions on T_N. In digital signal processing for example, one naturally requires that the calculation of $\mathrm{T}_N f$ is based on a finite number $\{f(\lambda_{n,N})\}_{n=1}^{M_N}$ of samples of f. The question is then whether the sequence $\{\mathrm{T}_N f\}_{N\in\mathbb{N}}$, with the required structure of T_N, converges to $\mathrm{T}f$ for every $f \in \mathcal{B}_1$. For many important problems, $\mathrm{T}_N f$ converges to $\mathrm{T}f$ for f in a dense subset of $\mathcal{B}_1$ but fails to converge for all $f \in \mathcal{B}_1$. Such negative results are often stated as

$$\limsup_{N\to\infty} \|\mathrm{T}_N f_* - \mathrm{T} f_*\|_{\mathcal{B}_2} = \infty \quad \text{for some} \quad f_* \in \mathcal{B}_1\ . \tag{1}$$

A sequence $\{\mathrm{T}_N\}_{N\in\mathbb{N}}$ satisfying (1) has bad subsequences $\{\mathrm{T}_{N_k}\}_{k\in\mathbb{N}}$ such that $\mathrm{T}_{N_k} f_*$ does not converge to $\mathrm{T}f_*$ for some $f_* \in \mathcal{B}_1$. However, this behavior does not exclude the existence of good subsequences such that $\{\mathrm{T}_{N_k} f\}_{k\in\mathbb{N}}$ converges to $\mathrm{T}f$ for all $f \in \mathcal{B}_1$. More precisely, (1) does not exclude

$$\liminf_{N\to\infty} \|\mathrm{T}_N f - \mathrm{T} f\|_{\mathcal{B}_2} < \infty \quad \text{or even} \quad \liminf_{N\to\infty} \|\mathrm{T}_N f - \mathrm{T} f\|_{\mathcal{B}_2} = 0$$

for all $f \in \mathcal{B}_1$. If a convergent subsequence exists, it depends generally on the actual $f \in \mathcal{B}_1$. So the selection of a good subsequence $\{N_k(f)\}_{k\in\mathbb{N}}$ such that

$$\lim_{k\to\infty} \|\mathrm{T}_{N_k(f)} f - \mathrm{T} f\|_{\mathcal{B}_2} = 0 \tag{2}$$

can be regarded as an *adaption* of $\{\mathrm{T}_N\}_{N\in\mathbb{N}}$ to the actual function $f \in \mathcal{B}_1$.

The following interesting problem arises: Given a sequence of approximation operators $\{\mathrm{T}_N\}_{N\in\mathbb{N}}$ satisfying (1) and which converges on a dense subset of $\mathcal{B}_1$. Is it possible to find for every $f \in \mathcal{B}_1$ a subsequence $\{N_k(f)\}_{k\in\mathbb{N}}$ such that (2) holds? To investigate this problem, the notion of strong divergence was introduced in [1].

Definition 1 (Strong divergence)**.** *Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be Banach spaces, and let $\{\mathrm{T}_N\}_{N\in\mathbb{N}}$ be a sequence of bounded linear operators $\mathrm{T}_N : \mathcal{B}_1 \to \mathcal{B}_2$. We say that* $\{\mathrm{T}_N\}_{N\in\mathbb{N}}$ diverges strongly *if*

$$\lim_{N\to\infty} \|\mathrm{T}_N f_*\|_{\mathcal{B}_2} = \infty \quad \textit{for some} \quad f_* \in \mathcal{B}_1\ .$$

If $\{\mathrm{T}_N\}_{N\in\mathbb{N}}$ diverges strongly then no convergent subsequence exists. It is an interesting problem to investigate sequences which satisfy (1) whether they diverge strongly i.e. whether all subsequences diverge. This contribution investigates this problem for the so-called Hilbert transform.

Strong divergence of Hilbert transform approximations

Let $\mathbb{T}$ be the additive group of real numbers modulus 2π. For any $f \in L^1(\mathbb{T})$, its *Hilbert transform* is defined by

$$(3) \qquad \big(\mathrm{H}f\big)(t) = \lim_{\epsilon\to 0} \frac{1}{2\pi} \int_{\epsilon\leq|\tau|\leq\pi} \frac{f(\tau+t)}{\tan(\tau/2)}\,\mathrm{d}\tau\ ,$$

where the limit on the right hand side exists for almost all $t \in \mathbb{T}$. This transformation plays a very important role in many different areas of science and engineering [5]. We consider H on the Banach space $\mathcal{B} := \big\{f \in \mathcal{C}(\mathbb{T}) : \mathrm{H}f \in \mathcal{C}(\mathbb{T})\big\}$ equipped with the norm $\|f\|_{\mathcal{B}} := \max\{\|f\|_\infty, \|\mathrm{H}f\|_\infty\}$.

Our goal is to approximate (3) by a sequence $\{\mathrm{H}_N f\}_{N=1}^{\infty}$, where each H_N belongs to $\mathcal{L}(\mathcal{B})$, the set of bounded linear operators on $\mathcal{B}$. Moreover, each H_N should be computational feasible, i.e. defined on a finite number of samples of f. More precisely, we require that $\{\mathrm{H}_N\}_{N\in\mathbb{N}} \subset \mathcal{L}(\mathcal{B})$ has the following natural properties:

(A) *Concentration on a finite sampling set:* For every $N \in \mathbb{N}$ there is a finite set $\Lambda_N = \{\lambda_{n,N} : n = 1,\dots,M_N\} \subset \mathbb{T}$ such that for all $f, g \in \mathcal{B}$ we have: $f(\lambda) = g(\lambda)$ for all $\lambda \in \Lambda_N$ implies $(\mathrm{H}_N f)(t) = (\mathrm{H}_N g)(t)$ for all $t \in \mathbb{T}$.

(B) *Convergence on a dense subset:* We have $\lim_{N\to\infty} \big\|\mathrm{H}_N f - \mathrm{H}f\big\|_\infty = 0$ for all $f \in \mathcal{C}^\infty(\mathbb{T})$, i.e. for all infinitely differentiable functions.

(C) *Generation by a sampling series:* There is $\{\mathrm{A}_N\}_{N\in\mathbb{N}} \subset \mathcal{L}(\mathcal{B})$ such that $\lim_{N\to\infty} \big\|\mathrm{A}_N f - f\big\|_\infty = 0$ for all $f \in \mathcal{B}$ and such that $\mathrm{H}_N f = \mathrm{H}\mathrm{A}_N f$.

The following known result [2] shows that every sequence $\{\mathrm{H}_N\}_{N\in\mathbb{N}}$ with properties (A), (B), (C) diverges weakly on $\mathcal{B}$.

Theorem 1. *Let $\{\mathrm{H}_N\}_{N\in\mathbb{N}} \subset \mathcal{L}(\mathcal{B})$ be a sequence with properties (A), (B), (C). There is a residual set $\mathcal{D}_{\mathrm{w}} \subset \mathcal{B}$ so that $\limsup_{N\to\infty} \|\mathrm{H}_N f\|_\infty = \infty$ for all $f \in \mathcal{D}_{\mathrm{w}}$.*

We strongly believe that all sequences $\{\mathrm{H}_N\}_{N\in\mathbb{N}} \subset \mathcal{L}(\mathcal{B})$ with properties (A), (B), and (C) not only diverge weakly but diverge strongly.

Conjecture 1. *Let $\{\mathrm{H}_N\}_{N\in\mathbb{N}} \subset \mathcal{L}(\mathcal{B})$ be a sequence with properties (A), (B), (C). There exists an $f_* \in \mathcal{B}$ such that $\lim_{N\to\infty} \|\mathrm{H}_N f_*\|_\infty = \infty$.*

This conjecture is supported by two results. First it is shown that the *sampled Fejér means* diverge strongly. The corresponding operators are defined by

$$(4) \qquad \big(\mathrm{H}_N^{\mathcal{F}} f\big)(t) = \sum_{n=0}^{N-1} f\big(n\,\tfrac{2\pi}{N}\big)\,\widetilde{\mathcal{F}}_N\big(t - n\,\tfrac{2\pi}{N}\big) \qquad (t \in \mathbb{T})\ ,$$

where the *conjugate Fejér kernel* is given by

$$\widetilde{\mathcal{F}}_N(\tau) = \frac{N\sin(\tau) - \sin(N\tau)}{2\big[N\sin(\tau/2)\big]^2} = \frac{1}{N}\left(\frac{1}{\tan(\tau/2)} - \frac{\sin(N\tau)}{2N\sin^2(\tau/2)}\right)\ .$$

It is easy to see that $\{\mathrm{H}_N^{\mathcal{F}}\}_{N\in\mathbb{N}} \subset \mathcal{L}(\mathcal{B})$ has properties (A), (B), and (C).

Theorem 2. *Let $\{\mathrm{H}_N^{\mathcal{F}}\}_{N\in\mathbb{N}}$ be the sequence of sampled conjugate Fejér means as defined in (4). There exists an $f_* \in \mathcal{B}$ such that $\lim_{N\to\infty} \big(\mathrm{H}_N^{\mathcal{F}} f_*\big)(\pi) = \infty$.*

The second result in support of our conjecture shows that for any $\{\mathrm{H}_N\}_{N\in\mathbb{N}}$ with properties (A), (B), (C) there is an $f_* \in \mathcal{B}$ such that $\|\mathrm{H}_N f_*\|_\infty$ exceeds any given bound for any given number of sufficiently large consecutive indices N.

Theorem 3. *Let $\{\mathrm{H}_N\}_{N\in\mathbb{N}} \subset \mathcal{L}(\mathcal{B})$ be a sequence with properties (A), (B), (C). Then there exists a function $f_* \in \mathcal{B}$ with $\|f_*\|_{\mathcal{B}} \leq 1$ and the following property: For all $M, N_0 \in \mathbb{N}$ and every $\delta \in (0,1)$ there exists two integers $N^{(1)} = N^{(1)}(M, N_0, \delta)$ and $N^{(2)} = N^{(2)}(M, N_0, \delta)$ with*

$$N^{(2)} > N^{(1)} \geq N_0 \qquad \text{and} \qquad \big[N^{(2)} - N^{(1)}\big]/N^{(2)} > 1 - \delta$$

such that $\|\mathrm{H}_N f_*\|_\infty > M$ for all $N \in [N^{(1)}, N^{(2)}]$.

A problem of Ul'yanov

Theorem 3 has a close relation to a problem on Fourier series due to Ul'yanov [4, 6]. This problem may be generalized and reformulated to the approximation of arbitrary operators $\mathrm{T} \in \mathcal{L}(\mathcal{B}_1, \mathcal{B}_2)$ where $\mathcal{L}(\mathcal{B}_1, \mathcal{B}_2)$ is the set of bounded linear operators between the Banach spaces $\mathcal{B}_1$ and $\mathcal{B}_2$.

Problem 1 (Ul'yanov-Type Problem)**.** *Let $\{\mathrm{T}_N\}_{N\in\mathbb{N}} \subset \mathcal{L}(\mathcal{B}_1, \mathcal{B}_2)$ be an approximation method of $\mathrm{T} \in \mathcal{L}(\mathcal{B}_1, \mathcal{B}_2)$. Does there exist a strictly increasing sequence $\{N_k\}_{k\in\mathbb{N}} \subset \mathbb{N}$ such that for every $f \in \mathcal{B}_1$ there is a strictly increasing sequence $\{\widehat{N}_k\}_{k\in\mathbb{N}} \subset \mathbb{N}$ so that for all $k \in \mathbb{N}$*

$$\widehat{N}_k \leq N_k \qquad \text{and} \qquad \sup_{k\in\mathbb{N}} \big\|\mathrm{T}_{\widehat{N}_k} f - \mathrm{T} f\big\|_{\mathcal{B}_2} < \infty\ ?$$

For sampling based approximations of the Hilbert transform (3) the answer is negative and we can prove the following result [3].

Theorem 4. *Let $\{\mathrm{H}_N\}_{N\in\mathbb{N}} \subset \mathcal{L}(\mathcal{B})$ be a sequence with properties (A), (B), (C), and let $\{N_k\}_{k\in\mathbb{N}} \subset \mathbb{N}$ be an arbitrary strictly increasing sequence. There exists a function $f_* \in \mathcal{B}$ such that*

$$\limsup_{k\to\infty} \min_{N\in(N_k, N_{k+1}]} \big\|\mathrm{H}_N f_*\big\|_\infty = \infty\ ,$$

and the Ul'yanov-Type Problem has no positive solution.

References

[1] H. Boche, B. Farrell, *Strong divergence of reconstruction procedures for the Paley-Wiener space $\mathcal{PW}_\pi^1$ and the Hardy space $\mathcal{H}^1$*, J. Approx. Theory **183** (2014), 98–117.

[2] H. Boche, V. Pohl, *On the calculation of the Hilbert transform from interpolated data*, IEEE Trans. Inform. Theory **54** (2008), 2358–2366.

[3] H. Boche, V. Pohl, *On the strong divergence of Hilbert transform approximations and a problem of Ul'yanov*, J. Approx. Theory (2015), submitted for publication.

[4] S. V. Konyagin, *Almost everywhere convergence and divergence of Fourier series*, Proc. Intern. Congress of Mathematicians, Madrid, Spain, Aug. 2006, 1393–1403.

[5] V. Pohl, H. Boche, *Advanced topics in system and signal theory: A mathematical approach*, Foundations in Signal Processing, Communications and Networking, Springer, Berlin, 2009.

[6] P. L. Ul'yanov, *Solved and unsolved problems in the theory of trigonometric and orthogonal series*, Russian Math. Surveys **19** (1964), 1–62.

Blind spikes deconvolution with lifting

Yuejie Chi

In many applications, the goal is to estimate the set of delays and amplitudes of point sources contained in a sparse spike signal $x(t)$ from its convolution with a band-limited or diffraction-limited point spread function (PSF) $g(t)$, which is either determined by the nature or designed by the practitioners. This describes the problem of estimating target locations in radar and sonar, firing times of neurons, direction-of-arrivals in array signal processing, etc.

When the PSF is assumed perfectly known, many algorithms have been developed to retrieve the spike signal, ranging from subspace methods [1] to total variation minimization [2]. However, in many applications, the PSF is not known a priori, and must be estimated together with the spike model, referred to as *blind spikes deconvolution.* A related problem is *blind calibration* of uniform linear arrays [3], where it is desirable to calibrate the gains of the array antennas in a blind fashion.

In this work, we study the problem of blind spikes deconvolution, where we want to joint estimate the PSF and the spike signal composed of a small number of delayed and scaled Dirac functions. We start by sampling the Fourier transform of the convolution which gives a measurement vector $\boldsymbol{y} = \boldsymbol{g} \odot \boldsymbol{x} \in \mathbb{C}^N$, where $\odot$ denotes point-wise product, $\boldsymbol{g} \in \mathbb{C}^N$ is the sampled Fourier transform of the PSF, $\boldsymbol{x} \in \mathbb{C}^N$ is the sampled Fourier transform of the spike signal, which is a sum of K complex sinusoids with frequencies determined by the corresponding delays, K is the number of spikes.

Motivated by [4], we assume that the PSF $\boldsymbol{g}$ lies in a known low-dimensional subspace $\boldsymbol{B} \in \mathbb{C}^{N\times L}$, i.e. $\boldsymbol{g} = \boldsymbol{B}\boldsymbol{h}$, $\boldsymbol{h} \in \mathbb{C}^L$, where the orientation of $\boldsymbol{g}$ in the subspace, given by $\boldsymbol{h}$, still needs to be estimated. This assumption is quite flexible and holds, at least approximately, in a sizable number of applications [4]. We introduce a novel application of the lifting trick, i.e. $y_i = x_i \cdot g_i = (\boldsymbol{e}_i^T \boldsymbol{x})(\boldsymbol{b}_i^T \boldsymbol{h}) = \boldsymbol{e}_i^T(\boldsymbol{x}\boldsymbol{h}^T)\boldsymbol{b}_i$, where y_i, x_i and g_i are the ith entry of $\boldsymbol{y}$, $\boldsymbol{x}$ and $\boldsymbol{g}$, $\boldsymbol{e}_i$ and $\boldsymbol{b}_i$ are the ith row of the identity matrix $\boldsymbol{I}$ and $\boldsymbol{B}$, respectively. It is now obvious we can translate the measurement vector into a set of linear measurements with respect to the matrix $\boldsymbol{Z}^\star = \boldsymbol{x}\boldsymbol{h}^T \in \mathbb{C}^{N\times L}$, i.e. $\boldsymbol{y} = \mathcal{X}(\boldsymbol{Z}^\star)$. While it is tempting to directly recover $\boldsymbol{Z}^\star$ from the above linear system of equations, it is under-determined no matter how large N is since we have more unknowns, NL, than the number of observations, N. Fortunately, note that the columns of $\boldsymbol{Z}^\star$ can be regarded as an ensemble of spectrally-sparse signals with the same spectral support, it is therefore possible to motivate this structure in the solution using the recently proposed atomic norm for spectrally-sparse ensembles [5, 6]. Specifically, we seek the matrix with minimum atomic norm that matches the set of linear measurements. The proposed algorithm is referred to as *AtomicLift.* AtomicLift can be efficiently implemented via semidefinite programming using off-the-shelf solvers. Moreover, the spikes can be localized by identifying the peaks of a dual polynomial constructed from the dual solution of AtomicLift.

To establish rigorous performance guarantees of AtomicLift, we assume that each row of $\boldsymbol{B}$ is identically and independently drawn from a distribution that obeys a simple isotropy property and an incoherence property, which is motivated by Candès and Plan in their development of a RIPless theory of compressed sensing [7]. This implies the PSF to have certain "spectral-flatness" property, so that the PSF has on average the same energy at different frequencies. Moreover, this assumption is flexible to allow the entries in each row of $\boldsymbol{B}$ to be correlated. On the other hand, we assume the minimum separation between spikes is at least $1/M$, where $N = 4M + 1$. This condition is the same as the requirement in [2, 8] even when the PSF is known perfectly. Under these conditions, we show that, with high probability, AtomicLift recovers the spike signal model up to a scaling factor as soon as N is on the order of $O(K^2L)$ up to logarithmic factors. Our result do not make randomness assumptions on the spike signal $\boldsymbol{x}$ nor the orientation of the PSF in the subspace $\boldsymbol{h}$. Recall that when the PSF is known exactly, it is capable to resolve K spikes as soon as N is on the order of $O(K)$. Therefore, when both K and L are not too large, AtomicLift is capable of blind spikes deconvolution at a price of slightly more samples.

Our proof is based on constructing a valid *vector-valued* dual polynomial that certifies the optimality of the proposed convex optimization algorithm with high probability. Our construction is inspired by [2, 8], where the squared Fejer's kernel is an essential building block in the construction. Nonetheless, significant, and nontrivial, modifications are necessary since our dual polynomial is vector-valued rather than scalar-valued as in the existing works, which is additionally complicated by the special linear operator induced from lifting. The details of the proof can be found in [9].

Our approach is inspired by the pioneering work of [4, 3], which applied the lifting trick to bilinear inverse problems such as blind deconvolution. Unfortunately, the results therein do not apply to our setting since the locations of the spikes do not necessarily lie on any a priori defined grid, $\boldsymbol{x}$ cannot be approximated by a sparse signal. An interesting future direction is to consider blind spikes deconvolution without subspace constraints on the PSF, by allowing more measurement vectors, e.g. $\boldsymbol{y}_i = \text{diag}(\boldsymbol{g})\boldsymbol{x}_i$, $i = 1, \dots, p$, which draws an interestingly connection to dictionary learning with translation-invariant constraints.

References

[1] P. Stoica and R. L. Moses, *Introduction to spectral analysis.* New Jersey: Prentice Hall, 1997, vol. 1.

[2] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Communications on Pure and Applied Mathematics*, vol. 67, no. 6, pp. 906–956, 2014.

[3] S. Ling and T. Strohmer, "Self-calibration and biconvex compressive sensing," *arXiv preprint arXiv:1501.06864*, 2015.

[4] A. Ahmed, B. Recht, and J. Romberg, "Blind deconvolution using convex programming," *Information Theory, IEEE Transactions on*, vol. 60, no. 3, pp. 1711–1732, 2014.

[5] Y. Chi, "Joint sparsity recovery for spectral compressed sensing," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, 2014, pp. 3938–3942.

[6] Y. Li and Y. Chi, "Off-the-grid line spectrum denoising and estimation with multiple measurement vectors," *arXiv preprint arXiv:1408.2242*, 2014.
[7] E. Candes and Y. Plan, "A probabilistic and RIPless theory of compressed sensing," *IEEE Transactions on Information Theory*, vol. 57, no. 11, pp. 7235–7254, 2011.
[8] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE transactions on information theory*, vol. 59, no. 11, pp. 7465–7490, 2013.
[9] Y. Chi, "Guaranteed blind sparse spikes deconvolution via lifting and convex optimization," *arXiv preprint arXiv:1506.02751*, 2015.

Forecasting with high-dimensional time series

Christine De Mol
(joint work with Domenico Giannone, Lucrezia Reichlin)

We consider the problem of predicting a given time series, representing for example a macroeconomic or financial variable, based on the information contained in a large ensemble of time series, strongly correlated with the series to forecast.

Let X_t be a high-dimensional time series, with (cross-sectional) dimension N, composed e.g. of N macroeconomic or financial variables, observed at discrete time intervals $t = 1, 2, \ldots$, e.g. every day, month, quarter or year. Each individual time series in X_t is assumed to be a stationary process, having zero mean and unit variance. The aim is to forecast a given economic variable y_t (in general included in X_t), e.g. inflation, unemployment or GDP growth, based on the information contained in the whole set of series, and not only based on the past of y_t. At a given time T, we want an estimate of y_{T+h}, i.e. a forecast at some given horizon h, based on the information available at time T. This information consists of the input data (predictors) $\{x_{n\,t}\}$, for $n = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$, ranged in a T by N matrix X, and of the observed response or dependent variable y_{t+h} for each $t = 1, 2, \ldots, T - h$, ranged in a vector $\mathrm{y} = (y_{1+h}, y_{2+h}, \ldots, y_T)'$ (the prime denoting the transpose). We assume a linear dependence between y_{t+h} and X_t, namely $y_{t+h} = \sum_n \beta_n x_{n\,t}$ for $t = 1, 2, \ldots, T - h$, or, in matrix form, $\mathrm{y} = \mathrm{X}\beta$. Let us remark that to model the possible dependence of y_{t+h} on past values of the predictors, we can augment X_t with the p lagged time series $X_{t-1}, X_{t-2}, \ldots, X_{t-p}$, as done in Vector Autoregressive (VAR) models. Hence, to simplify, and without loss of generality, we can set $p = 0$. We can also set $h = 0$ by using a proper redefinition of the time labels. To cope with noisy observations, we include a zero-mean error term e in the model,

$$\mathrm{y} = \mathrm{X}\beta + \mathrm{e}\ ,$$

and we reformulate the problem as a classical linear regression problem, which amounts to the minimization of the quadratic loss function $\mathrm{L}(\beta) = \|\mathrm{y} - \mathrm{X}\beta\|^2$, where $\|\mathrm{y}\|^2 = \sum_t |y_t|^2$ is the squared L_2-norm of y. When $\mathrm{X}'\mathrm{X}$ is full-rank, the minimizer is the well-known ordinary least-squares (OLS) solution $\hat{\beta}_{ols} = (\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'\mathrm{y}$, which, typically, becomes numerically unstable for large N and T due to ill-conditioning and, moreover, is not feasible as soon as N is larger than T, a common instance for macroeconomic data. The standard remedy to this problem, proposed in the econometric literature, is Principal Component Regression (PCR), or equivalently truncated singular value decomposition (TSVD) where only the K largest eigenvalues of $\mathrm{X}'\mathrm{X}$ are taken into account (see e.g. [4], and also [3] for dynamic principal components, i.e. principal components in the Fourier domain). These papers address the question of whether the accumulation of time samples and of series can help forecasting the target variable, and, under a factor model assumption, derive asymptotic convergence rates for both $T \to \infty$ and $N \to \infty$.

In [2], as an alternative to the PCR paradigm, we have considered other types of regularization of the problem, namely penalized least-squares such as ridge and lasso regression. We recall that in ridge regression, the estimator for β is given by

$$\hat{\beta} = \arg\min_{\beta} \left\{ \frac{1}{NT} \|\mathrm{y} - \mathrm{X}\beta\|^2 + \lambda \|\beta\|^2 \right\} \quad \text{i.e.} \quad \hat{\beta} = \left(\frac{\mathrm{X}'\mathrm{X}}{NT} + \lambda I \right)^{-1} \frac{\mathrm{X}'\mathrm{y}}{NT} \ ,$$

where λ is a positive regularization parameter. From this expression, we get the following "bias-variance" decomposition:

$$\hat{\beta} - \beta = -\lambda \left(\frac{\mathrm{X}'\mathrm{X}}{NT} + \lambda I \right)^{-1} \beta + \left(\frac{\mathrm{X}'\mathrm{X}}{NT} + \lambda I \right)^{-1} \frac{\mathrm{X}'\mathrm{e}}{NT} \ .$$

Under fairly general assumptions, standard for time series, one can show that $\frac{\|\mathrm{X}'\mathrm{e}\|}{\sqrt{NT}} = O(1)$ as $N, T \to \infty$, whence the following bound for the estimation of β:

$$\|\hat{\beta} - \beta\| \leq \|\beta\| + O\left(\frac{1}{\lambda} \frac{1}{\sqrt{NT}} \right) \ ,$$

a bound which will not vanish asymptotically unless the norm of β vanishes. More relevant for prediction is the Mean Square Forecast Error (MSFE), bounded by

$$\frac{1}{\sqrt{T}} \|\mathrm{X}\hat{\beta} - \mathrm{X}\beta\| \leq \frac{1}{\sqrt{T}} \sqrt{\lambda} \sqrt{NT} \|\beta\| + \frac{1}{\sqrt{T}} \frac{1}{\sqrt{\lambda}} O(1) \ .$$

The value of λ which minimizes this bound, equally balancing the two terms, is $\lambda \sim \frac{1}{\sqrt{T}\sqrt{N}\|\beta\|}$ and the resulting asymptotic rate for the MSFE is

$$\frac{1}{\sqrt{T}} \|\mathrm{X}\hat{\beta} - \mathrm{X}\beta\| \leq \frac{N^{1/4}}{T^{1/4}} \sqrt{\|\beta\|} \ ,$$

which increases with N. Therefore consistency results seem hard to obtain without further assumptions about the data-generating model.

A way of modelling strong comovement is provided by so-called "factor models", assuming that the high-dimensional time series is driven by a small number of factors spanning a subspace of (fixed) dimension K and can be written as the sum of a common component governing the global evolution and of an idiosyncratic component proper to each individual series:

$$X_t = \Lambda F_t + \xi_t \ ,$$

where the factors F_t are a K-dimensional stationary process, with covariance matrix $\mathrm{E}F_tF_t' = I_K$ and the idiosyncratic components ξ_t are a N-dimensional stationary process, orthogonal to the factors, with covariance matrix $\mathrm{E}\xi_t\xi_t' = \Psi$, of full rank for every N. The matrix Λ loading the factors is a non-random matrix of dimension $N \times K$, assumed to be of full rank K for every N. Moreover, all the eigenvalues of $\Lambda'\Lambda$ are supposed to grow asymptotically as N, which means that all predictors are informative on the factors. As a consequence, the spectrum of the population covariance matrix given by $\Sigma_{XX} = \mathrm{E}(X_tX_t') = \Lambda\Lambda' + \Psi$ (here we take $\Psi = I_N$, for simplicity) presents two clusters of eigenvalues separated with a spectral gap which increases with N. Under the previous assumptions, one can show [2] that $\|\beta\| \sim \frac{1}{\sqrt{N}}$, yielding a decay rate proportional to $T^{-1/4}$ for the MSFE and to $N^{-1/2}$ for the norm of the estimation error on β.

In recent work in progress, we have been able to improve the asymptotic rates derived in [2]. The principle of the somewhat technical proof is to take into account the spectral gap and to use Weyl's perturbation lemma to control the difference between the eigenvalues of the population and sample covariance matrices. In such a way, when setting $\lambda \sim \frac{1}{\sqrt{T}}$, we can show that

$$|X_t'\hat{\beta} - X_t'\beta| \leq O(\frac{1}{\sqrt{N}}) + O(\frac{1}{\sqrt{T}}) \ .$$

This asymptotic bound establishes consistency for $N \to \infty$ and $T \to \infty$, along any path in (N, T). The rate for ridge regression is the same as the rate which would be obtained with principal component regression, when K, the number of factors, is known. However, the estimation of K is a hard problem, widely discussed in the literature.

The framework described above can be generalized in various ways. A first extension allows to deal with "approximate" factor models in which the idiosyncratic components are mildly correlated ($\Psi \neq I_N$). Next, weights can be introduced in the L_2-norm of the residuals and of the penalty. The setting can also be extended to other linear regularization methods based on "spectral filtering". Finally, let us mention that a closely related methodology has been used to deal with large "Bayesian VAR" models [1].

References

[1] M. Banbura, D. Giannone, L. Reichlin, *Large Bayesian vector autoregressions*, Journal of Applied Econometrics **25** (2010), 71–92.

[2] C. De Mol, D. Giannone, L. Reichlin, *Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components?*, Journal of Econometrics **146** (2008), 318–328.

[3] M. Forni, M. Hallin, M. Lippi, L. Reichlin, *The generalized dynamic-factor model: identification and estimation*, The Review of Economics and Statistics **82** (2000), 540–554.

[4] J.H. Stock, M.W. Watson, *Forecasting using principal components from a large number of predictors*, Journal of the American Statistical Association **97** (2002), 1167–1179.

Dimensionality reduction with sparse Johnson-Lindenstrauss transforms

Sjoerd Dirksen
(joint work with Jean Bourgain, Jelani Nelson)

In a wide variety of disciplines, such as statistics, machine learning, numerical linear algebra and compressed sensing, one is faced with computational tasks on data sets that are not only large, but also high-dimensional. Unfortunately, the high-dimensionality leads to a large storage consumption and, since many algorithms have a running time depending at least linearly on the dimension, a large computational burden. For this reason, one would like to embed the data into a lower-dimensional space. Depending on the computational task, the embedding needs to preserve certain properties in order to be able to make inference about the original data. We consider the situation where we want to embed the data with a map that is linear and preserves inter-point Euclidean distances. We also require that the map is generated independently of the data (i.e., non-adaptively).

Formally, consider a set X of vectors in a high-dimensional space $\mathbb{R}^n$. We want to find a matrix $\Phi \in \mathbb{R}^{m\times n}$ such that

$$(1) \qquad (1-\varepsilon)\|x-y\|_2^2 \leq \|\Phi x - \Phi y\|_2^2 \leq (1+\varepsilon)\|x-y\|_2^2, \quad \text{for all}\, x,y \in X.$$

If $T = \{(x-y)/\|x-y\|_2 \ : \ x,y \in X\}$ is the set of normalized differences, then the constant

$$\varepsilon_T = \sup_{x\in T} |\|\Phi x\|_2^2 - 1|,$$

which is called the *restricted isometry constant* of Φ on T, is exactly the smallest possible ε that one can take in (1).

A classical theorem due to Gordon [1] states that if Φ is an $m \times n$ matrix filled with i.i.d. mean-zero Gaussian entries with variance $1/m$ and $T \subset S^{n-1}$, then $\varepsilon_T \leq \varepsilon$ with probability at least $1-\eta$ if $m \gtrsim \varepsilon^{-2}(w^2(T) + \log(\eta^{-1}))$, where $\gtrsim$ hides an absolute constant. The parameter $w(T) := \mathbb{E}\sup_{x\in T}\langle x, g\rangle$, where g is an n-dimensional standard Gaussian vector, is known as the *Gaussian width* and describes the ℓ_2-geometric complexity of T. Gordon's theorem can be considered as an instance-optimal version of the Johnson-Lindenstrauss lemma, since $w^2(T) \lesssim \log|T|$, where $|T|$ is the cardinality of T, but the Gaussian width can be much smaller. For example, if T is the set of all k-sparse vectors in S^{n-1}, then $w^2(T) \lesssim k\log(n/k)$.

Thanks to Gordon's result, good dimensionality reduction guarantees are known for Gaussian matrices. However, a clear downside of these matrices is that they are densely populated and therefore matrix-vector multiplication is slow. In many applications of ε-isometries it is desirable or even necessary that the embedding matrix Φ supports fast matrix-vector multiplication. One 'fast' construction is the sparse Johnson-Lindenstrauss transform (SJLT) [2, 3]. One possible way to construct the SJLT (cf. [3]) is as follows: start with a random sign matrix. For each column independently, pick exactly s entries uniformly at random without replacement and put the rest to zero. Finally, rescale the matrix by $1/\sqrt{s}$. Since the resulting matrix Φ has exactly s non-zeros per column, one can compute Φx in time $O(s\|x\|_0)$ instead of $O(m\|x\|_0)$ for a dense matrix.

We consider the following general question: given $T \subset S^{n-1}$, how should we set m and s to guarantee that the restricted isometry constant ε_T of the SJLT is small in expectation? Our answer can be considered as a sparse analogue of Gordon's theorem.

Theorem 1. [4, 5] *Let $T \subset S^{n-1}$ and Φ be an SJLT with column sparsity s. Define the complexity parameter*

$$\kappa(T) = \max_{q \leq \frac{m}{s} \log s} \Big\{ \frac{1}{\sqrt{qs}} \Big(\mathbb{E}_\eta \Big(\mathbb{E}_g \sup_{x \in T} \Big| \sum_{j=1}^{n} \eta_j g_j x_j \Big| \Big)^q \Big)^{1/q} \Big\},$$

where $(g_j)_{j=1}^n$ are i.i.d. standard gaussian and $(\eta_j)_{j=1}^n$ i.i.d. Bernoulli with mean $qs/(m \log s)$. Suppose that $s \gtrsim_* \varepsilon^{-2}$, where $\gtrsim_*$ hides a polylog-factor. Then $\mathbb{E}\varepsilon_T \leq \varepsilon$ holds as long as s, m are such that $\kappa(T) \lesssim_* \varepsilon$.

The parameter $\kappa(T)$ can be seen as a new complexity measure that replaces the Gaussian width featuring in Gordon's theorem. It may look daunting at first, but it can in fact be controlled using standard tools from high-dimensional probability for all data structures of interest. In particular, Theorem 1 qualitatively (i.e. up to log-factors) unifies all known results for specific data sets $T \subset S^{n-1}$, and yields new sparse dimensionality reduction results for (possibly infinite) unions of subspaces. Moreover, we find the first (and in a certain sense optimal) guarantees for the SJLT to preserve geodesic distances on a manifold up to a multiplicative error. In addition, we obtain novel results for using the SJLT to accelerate approximately solving convexly constrained least squares programs such as the Lasso.

References

[1] Y. Gordon, "On Milman's inequality and random subspaces which escape through a mesh in $\mathbf{R}^n$," in *Geometric aspects of functional analysis (1986/87)*, ser. Lecture Notes in Math. Berlin: Springer, 1988, vol. 1317, pp. 84–106.

[2] A. Dasgupta, R. Kumar, and T. Sarlós, "A sparse Johnson-Lindenstrauss transform," in *Proceedings of the 42nd ACM Symposium on Theory of Computing (STOC)*, 2010, pp. 341–350.

[3] D. M. Kane and J. Nelson, "Sparser Johnson-Lindenstrauss transforms," *Journal of the ACM*, vol. 61, no. 1, art. 4, 2014.

[4] J. Bourgain, S. Dirksen, and J. Nelson, "Toward a unified theory of sparse dimensionality reduction in Euclidean space," in *Proceedings of the 47th ACM Symposium on Theory of Computing (STOC)*, 2015, pp. 499–508.

[5] ——, "Toward a unified theory of sparse dimensionality reduction in Euclidean space," *Geometric and Functional Analysis*, vol. 25, no. 4, pp. 1009–1088, 2015.

Consistency of probability measure quantization by means of power repulsion-attraction potentials

Massimo Fornasier

In this talk we present the study of the consistency of a variational method for probability measure quantization, deterministically realized by means of a minimizing principle, balancing power repulsion and attraction potentials. The proof of consistency is based on the construction of a target energy functional whose unique minimizer is actually the given probability measure to be quantized. Then we show that the discrete functionals, defining the discrete quantizers as their minimizers, actually Gamma -converge to the target energy with respect to the narrow topology on the space of probability measures. A key ingredient is the reformulation of the target functional by means of a Fourier representation, which extends the characterization of conditionally positive semi-definite functions from points in generic position to probability measures. As a byproduct of the Fourier representation, we also obtain compactness of sublevels of the target energy in terms of uniform moment bounds, which already found applications in the asymptotic analysis of corresponding gradient flows. To model situations where the given probability is affected by noise, we additionally consider a modified energy, with the addition of a regularizing total variation term and we investigate again its point mass approximations in terms of Gamma -convergence. We show that such a discrete measure representation of the total variation can be interpreted as an additional nonlinear potential, repulsive at a short range, attractive at a medium range, and at a long range not having effect, promoting a uniform distribution of the point masses. Inspired by these results we eventually sketch a model of social inclusion/exclusion and redistribution of resourced for balancing social and wealth inequality in societies.

References

[1] M. Di Francesco, M. Fornasier, J.-C. Htter, and D. Matthes. Asymptotic behavior of gradient flows driven by nonlocal power repulsion and attraction potentials in one dimension. *SIAM J. Math. Anal.*, Vol. 46, No. 6, 2014, pp. 3814-3837.

[2] M. Fornasier, J. Haskovec and G. Steidl. Consistency of variational continuous-domain quantization via kinetic theory. *Applicable Analysis*, Vol. 92, No. 6, 2013, pp. 1283-1298.

[3] M. Fornasier and J.-C. Htter. Consistency of probability measure quantization by means of power repulsion-attraction potentials. *J. Fourier Anal. Appl*, to appear.

Diamond norm as regularizer for low rank matrix recovery

David Gross, Richard Kueng
(joint work with Martin Kliesch, Jens Eisert)

In the common approach to low-rank matrix recovery [1, 2, 3], one uses the nuclear norm as a convex surrogate for rank. Geometric proof techniques like Tropp's Bowling scheme [4] or Mendelson's small ball method [5, 6] bound the reconstruction error in terms of the descent cone of the norm at the matrix that is to be recovered. Moreover, these arguments suggest that the error would decrease if another convex function be used, which has a smaller descent cone at the set of matrices of interest. Here, we construct such an improved convex function based on the *diamond norm* [7, 8]. We characterize those low-rank matrices for which we expect improved recovery, demonstrate the increased performance numerically, and point out applications to learning *quantum channels*. This is an extended abstract of [9].

The objects we aim to recover are linear maps on a tensor product of Hilbert spaces $L(V \otimes V)$ with $V \simeq \mathbb{C}^n$ for some dimension n. On this space, we introduce the norm

$$\|Z\| := n \sup_{\|A\|_F = \|B\|_F = 1} \|(A \otimes \mathbb{I})\, Z\, (B \otimes \mathbb{I})\|_* \,. \tag{1}$$

Here, $\|\cdot\|_F$ is the Frobenius norm and $\|\cdot\|$ the nuclear norm. Up to an elementary but notoriously confusing shuffling of indices (the *partial transpose*, which we shall not introduce here), the resulting function is the *diamond norm* [8], which in turn is dual to the *norm of complete boundedness* studied in operator theory [10]. We point out that

$$\|Z\|_* \leq \|Z\| \quad \forall Z, \tag{2}$$

because setting $A = B = \frac{1}{\sqrt{n}}\mathbb{I}$ is a feasible point in the optimization on the right hand side of (1), which recovers exactly the nuclear norm. Moreover, although the definition relies on an optimization over two continuous sets of matrices, the function can be cast into the form of an SDP [8, Section 3.2]. Hence (1) can be evaluated computationally efficiently.

It is the main message of this work that for certain low-rank recovery problems, replacing the nuclear norm regularizer by the norm function (1) can improve performance. To see why, note that the norm is a supremum over many different convex functions $n\|\,(A \otimes \mathbb{I}) \cdot (B \otimes \mathbb{I})\,\|_*$, one of them being the nuclear norm itself. Elementary convex analysis [9] then reveals that the norm's descent cone obeys

$$\mathcal{D}\left(\|\cdot\|, Z\right) \subset \bigcap_{(A,B)\text{ active}} \mathcal{D}\left(\|\,(A \otimes \mathbb{I}) \cdot (B \otimes \mathbb{I})\,\|_*, Z\right) \quad \forall Z \in L\left(V \otimes V\right), \tag{3}$$

where we call (A, B) *active* if the supremum in (1) is attained for A, B.

If now Z is such that $\|Z\| = \|Z\|_*$ holds (i.e. the pair $\left(\frac{1}{\sqrt{n}}\mathbb{I}, \frac{1}{\sqrt{n}}\mathbb{I}\right)$ is active), then (3) implies

$$\mathcal{D}\left(\|\cdot\|, Z\right) \subset \mathcal{D}\left(\|\cdot\|_*, Z\right). \tag{4}$$

Consequently, the norm's descent cone at Z is at most as large as the corresponding one of the nuclear norm.

A main result of [9] is the characterization of the set of matrices for which equality holds. To state it, we need the *partial trace*

$$\mathrm{tr}_2 : L(V \otimes V) \mapsto L(V)$$

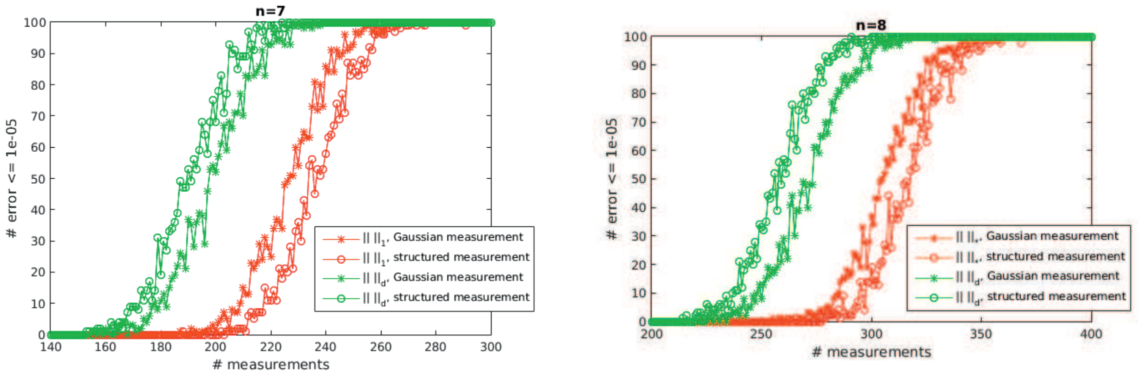

FIGURE 1. Number of successful low rank matrix reconstructions out of 100 trials by means of nuclear norm minimization (red) and diamond norm minimization (green) in the absence of noise. As test matrices, we have considered matrices of the form $X = (U \otimes \mathbb{I}) \sum_{i,j=1}^{n} e_i e_j^* \otimes e_i e_j^* (V \otimes \mathbb{I})$, where $U, V \in U(n)$ are independent Haar-random unitaries. Test matrices of this form have rank one and approximately obey $\|X\| = \|X\|_*$ by construction (if $V = U^*$, they obey this relation exactly). We considered a single recovery to be successful, if the Frobenius distance of minimizer of the respective optimization problem to the original target matrix was closer than 10^{-5}. As measurements we considered both independent Gaussian measurements and also certain structured measurements that are well motivated by quantum mechanical applications. The two plots illustrate such experiments for $n = 7$ and $n = 8$. In both cases and for the two different types of measurement matrices (Gaussian and structured ones), the diamond norm minimization clearly outperforms its well-established nuclear norm counterpart.

which is defined by linearly extending its action on tensor products:

$$\operatorname{tr}_2(X \otimes Y) = X \operatorname{tr} Y.$$

We then have:

Theorem 1 ([9]). *The equality $\|Z\| = \|Z\|_*$ holds if and only if both $\operatorname{tr}_2 \sqrt{ZZ^*}$ and $\operatorname{tr}_2 \sqrt{Z^*Z}$ are proprtional to the identity matrix.*

It is natural to ask whether maps satisfying this extremality condition play any role in practice? They do – at least for practitioning quantum physicsits. Indeed, the so-called "Choi matrix of a quantum channel" [13, 12] (see also [14, Lecture 5] for a concise introduction) fulfills the condition of Theorem 1. (*Channels* are the quantum analogue of stochastic maps in classical probability theory. The condition above amounts to requiring the quantum channels preserve the normalization of quantum probability distributions, see e.g. [14, Lecture 5: Theorem 5.4]).

The discussion above suggests that our norm might outperform the recovery of quantum channels with an associated low-rank Choi matrix (which are important in practice). Numerical simulations suggest that this is indeed the case. We have summarized the results of two such studies in Figure 1. Further appications will be presented in [9].

References

[1] B. Recht, M. Fazel, P.A. Parrilo, *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Rev. **52** (2010), 471–501

[2] D. Gross, *Recovering low-rank matrices from few coefficients in any basis*, IEEE Trans. Inform. Theory **57** (2011), 1548–1566

[3] M. Kabanava, R. Kueng, H. Rauhut, U. Terstiege *Stable low-rank matrix recovery via null space properties*, preprint arXive:1507.07184 (2015)

[4] J. Tropp, *Convex recovery of a structured signal from independent random linear measurements*, to appear in "Sampling Theory, a Renaissance", preprint arXiv:1405.1102 (2014)

[5] G. Lecué, S. Mendelson, *Sparse recovery under weak moment assumptions*, preprint arXiv:1401.2188 (2014)

[6] R. Kueng, H. Rauhut, U. Terstiege, *Low rank matrix recovery from rank one measurements*, Appl. Comput. Harmon. Anal., doi:10.1016/j.acha.2015.07.007 (2015)

[7] A.Y. Kitaev, A. Shen, M.N. Vyalyi, *Classical and quantum computation*, Providence: American Mathematical Society, **47** (2002)

[8] J. Watrous, *Simpler semidefinite programs for completely bounded norms*, Chicago Journal on Theoretical Computer Science (2013), article 8

[9] M. Kliesch, R. Kueng, D. Gross, J. Eisert, *Convex recovery of structured rank-4 tensors from independent random linear measurements*, in preparation (2015)

[10] V. Paulsen, *Completely bounded maps and operator algebras*, Cambridge University Press, Cambridge, UK, 2002.

[11] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information.* Cambridge University Press (2010).

[12] M.D. Choi, *Completely positive maps on complex matrices*, Linear Algebra Appl. **10** (1975), 285

[13] A. Jamiolkowski, *Linear transformations which preserve trace and positive semidefiniteness of operators*, Rep. Math. Phys. *3* (4) (1972), 275

[14] J. Watrous, *Theory of quantum information*, lecture notes (2011), available online: https://cs.uwaterloo.ca/˜watrous/LectureNotes.html

What is variable bandwidth?

Karlheinz Gröchenig
(joint work with Andreas Klotz)

A function or signal $f \in L^2(\mathbb{R})$ has bandwidth $\Omega > 0$, if its Fourier transform $\hat{f}(\xi) = \int_{\mathbb{R}} f(x)e^{-ix\xi}\,dx$ vanishes outside the interval $[-\Omega, \Omega]$. The number Ω is the maximal frequency contributing to f and is called the bandwidth of f. According to Shannon the bandwidth is also an information-theoretic quantity.

In the context of time-frequency analysis, it is perfectly plausible to assign different local bandwidths to different segments of a signal. This becomes even more obvious in the often cited metaphor of music: the highest frequency of musical piece is time-varying. However, a rigorous definition of variable bandwidth is difficult, perhaps even elusive, because bandwidth is global by definition and the assignment of a local bandwidth is in contradiction with the uncertainty principle.

So what is a function of variable bandwidth?

Before giving a precise definition, we need to single out the distinctive features of bandlimited functions. In our view the essence of bandwidth is encapsulated in three fundamental theorems about bandlimited functions:

(1) the Shannon-Whittaker-Kotelnikov sampling theorem and its variations,
(2) the existence of a critical density (Nyquist rate in engineering terms), and
(3) some inherent analyticity.

We propose a new notion of variable bandwidth. It is based on the spectral theory of the differential operator

$$A_p = -\tfrac{d}{dx}\big((p(x)\tfrac{d}{dx}\big)\,, \tag{1}$$

where $p > 0$ is the *bandwidth-parametrizing function.* By imposing mild assumptions on p and choosing a suitable domain, A_p becomes a positive, unbounded, self-adjoint operator on $L^2(\mathbb{R})$. Its spectral representation enables us to make the following definition.

Definition 1. *Let $\Lambda \subseteq \mathbb{R}^+$ be a fixed compact spectrum. A function is A_p-bandlimited with spectrum Λ, if $f \in c_\Lambda(A_p)L^2(\mathbb{R})$. The range of the spectral projection $c_\Lambda(A_p)L^2(\mathbb{R})$ is called the Paley-Wiener space with respect to A_p and spectrum Λ and will be denoted by $PW_\Lambda(A_p)$.*

If $p \equiv 1$ and $A_p = -\frac{d^2}{dx^2}$, then $PW_{[0,\Omega]}(-D^2)$ consists exactly of the classical bandlimited functions with bandwidth $\sqrt{\Omega}$.

We will show that this definition is indeed a meaningful notion of variable bandwidth. First, functions of variable bandwidth admit sampling theorems.

Theorem 1 (Sampling theorem for $PW_\Lambda(A_p)$)**.** *Fix $\Lambda \subseteq \mathbb{R}^+$ compact and set $\Omega = \max \Lambda$. Assume that $0 < c \leq p(x)$ for all $x \in \mathbb{R}$. Let $X = (x_i)_{i\in\mathbb{Z}}$ be an increasing sequence with $\lim_{i\to\pm\infty} x_i = \pm\infty$ and $\inf_i(x_{i+1} - x_i) > 0$. If*

$$\delta = \sup_{i\in\mathbb{Z}} \frac{x_{i+1} - x_i}{\inf_{x\in[x_i,x_{i+1}]} \sqrt{p(x)}} < \frac{\pi}{\sqrt{\Omega}}\,, \tag{2}$$

then there exist $A, B > 0$ such that, for all $f \in PW_\Lambda(A_p)$,

$$A\|f\|_2^2 \leq \sum_{i\in\mathbb{Z}} |f(x_i)|^2 \leq B\|f\|_2^2 \tag{3}$$

Theorem 1 supports our interpretation that $p(x)^{-1/2}$ is a measure for the local bandwidth. If p is constant on an interval I, $p|_I = p_0$, then the local gap condition (2) reads as $x_{i+1} - x_i \leq \delta\sqrt{p_0} < \frac{\pi\sqrt{p_0}}{\sqrt{\Omega}}$ for $x_i \in I$. This is precisely the sufficient condition on the maximal gap that arises for bandlimited functions with bandwidth $(\Omega/p_0)^{1/2}$. In other words, $f \in PW_{[0,\Omega]}(A_p)$ behaves like a $(\Omega/p_0)^{1/2}$-bandlimited function on I.

The second result is a necessary density condition for sampling in the style of Landau [4]. For the formulation we need an adaptation of the Beurling density to variable bandwidth. As in (2) we impose a new measure or distance on $\mathbb{R}$ determined by the bandwidth parametrization p, namely $\mu_p(I) = \int_I p^{-1/2}(u)\,du$ and define the Beurling density of a set $X \subseteq \mathbb{R}$ as

$$D_p^-(X) = \varliminf_{r\to\infty} \inf_{\mu_p(I)=r} \frac{\#(X \cap I)}{r}\,,$$

where the infimum runs over all bounded intervals $I \subseteq \mathbb{R}$.

Theorem 2. *Assume that $p \in C^2$ and that, for some $a > 0$, $p(x) = p_-$ for $x \leq -a$ and $p(x) = p_+$ for $x \geq a$. Fix $\Lambda \subseteq \mathbb{R}^+$ with finite (Lebesgue) measure. If $X \subseteq \mathbb{R}$ is a separated set such that the sampling inequality*

$$A\|f\|_2^2 \leq \sum_{i \in I} |f(x_i)|^2 \leq B\|f\|_2^2$$

holds for all $f \in PW_\Lambda(A_p)$, then $D_p^-(X) \geq \frac{|\Lambda|^{1/2}}{\pi}$.

Theorem 2 is again consistent with our interpretation of $PW_\Lambda(A_p)$ as a space of functions with variable bandwidth. If p is constant on an interval I, $p|_I = p_0$, and $\Lambda = [0, \Omega]$, then $\mu_p(I) = |I|/\sqrt{p_0}$ and we obtain

$$\#(X \cap I) \geq \frac{\Omega^{1/2}|I|}{\pi\sqrt{p_0}}\,.$$

Comparing with Landau's classical result for bandlimited function, this is exactly the minimum number of samples in I required for a bandlimited function with bandwidth $(\Omega/p_0)^{1/2}$. Again, $f \in PW_{[0,\Omega]}(A_p)$ behaves like a $(\Omega/p_0)^{1/2}$-bandlimited function on I.

Methods. Whereas the formulation of these theorems is almost the same as the standard theorems for classical bandlimited functions, the proofs require input from two areas, namely the applied harmonic analysis of sampling theory and the detailed spectral analysis of Sturm-Liouville operators and Schrödinger operators. The methodical input from sampling theory is the oscillation method from [2] for the proof of Theorem 1, and the proof of Theorem 2 follows the outline of Nitzan and Olevski [5] in which a (discrete) frame of reproducing kernels is compared to a continuous resolution of the identity. The second methodical input is from the theory of Sturm-Liouville problems and of (one-dimensional) Schrödinger operators. In the proof of Theorem 1 we need a version of the Wirtinger-Poincaré inequality. The main effort will be devoted to finding appropriate estimates and cancellation properties of the reproducing kernel of $PW_\Lambda(A_p)$. These are derived by means of spectral theory of Sturm-Liouville operators. The detailed analysis of the spectral measure of A_p yields a representation of functions in $PW_\Lambda(A_p)$ as

$$f(x) = \int_\Lambda F(\lambda) \cdot \Phi(\lambda, x)\, d\rho(\lambda)$$

where $\Phi(\lambda, x) = \big(\Phi_1(\lambda, x), \Phi_2(\lambda, x)\big)^T$ is a set of fundamental solutions of $-DpD\Phi = \lambda\Phi$, ρ is the 2×2-matrix-valued spectral measure, and $F \in L^2(\Lambda, \mathbb{R}^2, \rho)$. Though not as explicit as the Fourier transform, this spectral representation of functions of variable bandwidth will enable us to derive the essential properties of $PW_\Lambda(A_p)$. For the proof of the density theorem we will switch to an equivalent Schrödinger equation and use the scattering theory of the Schrödinger equation.

Related work and other notions of variable bandwidth. In the literature one finds several approaches to variable bandwidth. Among them are time-warping (e.g., [7]), the time-frequency methods of Aceska and Feichtinger [1], and the procedural concept of Kempf [3]. Perhaps closest to our approach is Pesenson's deep and original work on abstract bandlimitedness [6]. Given an unbounded, self-adjoint operator on a Hilbert space H, the spectral subspaces $c_\Lambda(A)H$ are considered abstract spaces of bandlimited vectors. If A is the Laplace-Beltrami operator on a manifold and thus the corresponding Paley-Wiener spaces are concrete function spaces, for which Pesenson and Zayed have already shown the existence of qualitative sampling theorems.

References

[1] R. Aceska and H. G. Feichtinger. Functions of variable bandwidth via time-frequency analysis tools. *J. Math. Anal. Appl.*, 382(1):275–289, 2011.

[2] K. Gröchenig. Reconstruction algorithms in irregular sampling. *Math. Comp.*, 59(199):181–194, 1992.

[3] A. Kempf. Black holes, bandwidths and Beethoven. *J. Math. Phys.*, 41(4):2360–2374, 2000.

[4] H. J. Landau. Necessary density conditions for sampling and interpolation of certain entire functions. *Acta Math.*, 117:37–52, 1967.

[5] S. Nitzan and A. Olevskii. Revisiting Landau's density theorems for Paley-Wiener spaces. *C. R. Math. Acad. Sci. Paris*, 350(9-10):509–512, 2012.

[6] I. Pesenson and A. I. Zayed. Paley-Wiener subspace of vectors in a Hilbert space with applications to integral transforms. *J. Math. Anal. Appl.*, 353(2):566–582, 2009.

[7] D. Wei and A. Oppenheim. Sampling based on local bandwidth. In *Signals, Systems and Computers, 2007. ACSSC 2007. Conference Record of the Forty-First Asilomar Conference on*, pages 1103–1107, Nov 2007.

Certifying linear optical circuits via phaseless estimation techniques

Richard Kueng

(joint work with Daniel Suess, David Gross)

The problem of retrieving a complex-valued signal from measurements that are ignorant towards phase information has a long history in many different branches of science. In a discrete (digital) setting, this problem is usually phrased as the task of inferring a complex signal $x \in \mathbb{C}^n$ from measurements of the form

$$y_i = |\langle a_i, x\rangle|^2 + \epsilon_i \quad i = 1, \ldots, m, \tag{1}$$

where $a_1, \ldots, a_m \in \mathbb{C}^n$ are measurement vectors and $\epsilon_1, \ldots, \epsilon_m \in \mathbb{R}$ models additive noise. Clearly, all phase information is lost in such a measurement process which is why the inverse problem is ill-posed. Recently, the mathematical structure of such a problem has received considerable interest in its own right and many efficient recovery procedures for certain types of measurements have been proposed [1, 2, 3, 4]. In this work we shall focus on *PhaseLift* [5, 6] – one such approach that is based on recasting the phase retrieval problem as a particular instance of low rank matrix recovery [10, 11]. For Gaussian measurement vectors, the results

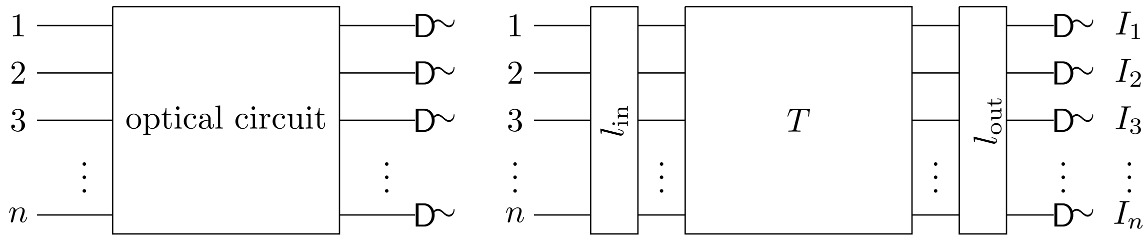

Figure 1. Sketch of a linear optical circuit with n incoming and n outgoing wires followed by n light intensity detectors. On the right hand side, the incoming and outgoing light configurations are represented as vectors $l_{\mathrm{in}}, l_{\mathrm{out}} \in \mathbb{C}^n$. Also, the n corresponding detector measurements are listed as $I_1, \ldots, I_n$.

in [7, 8, 9] assure that $m = Cn$ measurements[1] of the form (1) suffice with high probability (w.h.p.) to reconstruct any signal $x \in \mathbb{C}^n$ via a convex optimization problem. In addition to being uniform (one choice of measurement vectors w.h.p. suffices to reconstruct *any* signal) this recovery is stable towards additive noise in the measurement process.

In this note we propose to employ this recently established phaseless recovery method to a seemingly very different problem: certifying the performance of linear optical circuits. In quantum mechanics, these devices have received considerable attention over the past years since they might function as elementary circuits of a future universal quantum computer [12]. Unlike other physical realizations of these elementary quantum logical gates, optical devices have the prospect of scalability by including many elementary gates in a single integrated chip. Doing so and accurately certifying that such a device performs the way it is supposed to is an important milestone in quantum optics [13] and motivated the work presented here.

In classical optics[2] a linear optical circuit is an optical device which maps a light configuration localized in n optical fibres linearly onto a different n-wire outcome configuration. In order to characterize such a circuit, one measures the outputted light configuration with light intensity detectors positioned at each wire. This setup is illustrated in Figure 1. It is quite straightforward to formulate such a configuration mathematically: input and output configurations can be represented by vectors $l_{\mathrm{in}}, l_{\mathrm{out}} \in \mathbb{C}^n$, where the k-th component of l_{in} encapsulates intensity and phase of the incoming light configuration at the k-th wire and an analogous identification holds for l_{out}. Since the optical circuit is linear, it can be described by a complex $n \times n$ matrix T and, in particular, $l_{\mathrm{out}} = T l_{\mathrm{in}}$ is true for any input configuration l_{in}. Finally, the individual detectors simply measure the intensity of l_{out} at the different wires. So, upon feeding in a certain light configuration l_{in}, the

[1] Here, C denotes an absolute constant of sufficient size.

[2] For the sake of simplicity, here we shall content ourselves with certifying classical optical devices. Although less general than the full quantum problem, classical circuits still make up the bulk of state-of-the art quantum optical circuits [13].

the k-th detector outcome corresponds to

(2)
$$I_k\left(l_{\text{in}}\right)=\left|\langle e_k,l_{\text{out}}\rangle\right|^2=\left|\langle e_k,Tl_{\text{in}}\rangle\right|^2=\left|\langle T^*e_k,l_{\text{in}}\rangle\right|^2=\left|\langle\bar{t}_k,l_{\text{in}}\rangle\right|^2=\left|\langle\bar{l}_{\text{in}},t_k\rangle\right|^2,$$

where t_k denotes T's k-th row t_k and $\bar{t}_k$ its complex conjugate. Consequently, the different detectors single out the corresponding rows of T, up to complex conjugation. Moreover, the mathematical structure of the k-th detector measurement (2) exactly resembles a "traditional" phaseless measurement (1) of $\bar{t}_k\in\mathbb{C}^n$, where the input configuration l_{in} assumes the role of a single measurement vector a.

Since we can choose the input light configuration l_{in} at will and are able to perform such an experiment multiple times for different input configurations, the structure of (2) in combination with PhaseLift suggests the following protocol for estimating T:

(1) Prepare $m=Cn$ light configurations $l_{\text{in}}^{(1)},\ldots,l_{\text{in}}^{(m)}$ that resemble independently chosen complex standard Gaussian vectors.
(2) Successively feed the complex conjugate light configurations into the optical circuit and record the corresponding measurement outcomes $I_k\left(\bar{l}_{\text{in}}^{(i)}\right)=\left|\langle l_{\text{in}}^{(i)},t_k\rangle\right|^2$ for $1\leq i\leq m$ separately in n different data arrays (one for each detector $1\leq k\leq n$).
(3) Perform n independent instances of PhaseLift to reconstruct the different rows t_k individually from these data arrays up to a global phase each.

Stacking the recovered rows $t_1,\ldots,t_n$ back together yields the sought for matrix T up to left-multiplication with a diagonal unitary matrix. Note that such a degree of freedom is inherent in the problem's structure and cannot be avoided. For such a procedure to work, it is crucial that the recovery for PhaseLift is both uniform and stable. This in turn assures that recovering all n rows t_k from a common set of m Gaussian input configurations indeed works w.h.p. Moreover, estimating T in such a way assures stability towards noisy detector measurements. This is arguably the greatest advantage of our proposed recovery scheme over more traditional estimation methods for linear optical circuits [14] which, by and large, require interactive measurements and are very prone to error propagation.

Our protocol overcomes both these issues at the cost of a higher number of Gaussian input light configurations. Indeed, note that the constant C in the sampling rate of PhaseLift is not explicitly specified. However, numerical experiments suggest that a number of $m=4n-4$ different input configurations suffices for our protocol to work properly. The results of these studies can be found in Figure 2 and underline the prospect of our novel protocol to estimate linear optical circuits. We conclude with mentioning an ongoing collaboration with J.O. Brien's quantum photonics group in Bristol with the aim of employing such a PhaseLift-based protocol to certify the performance of spick-and-span universal optical circuit [13]. The results of this experimental collaboration will be presented elsewhere.

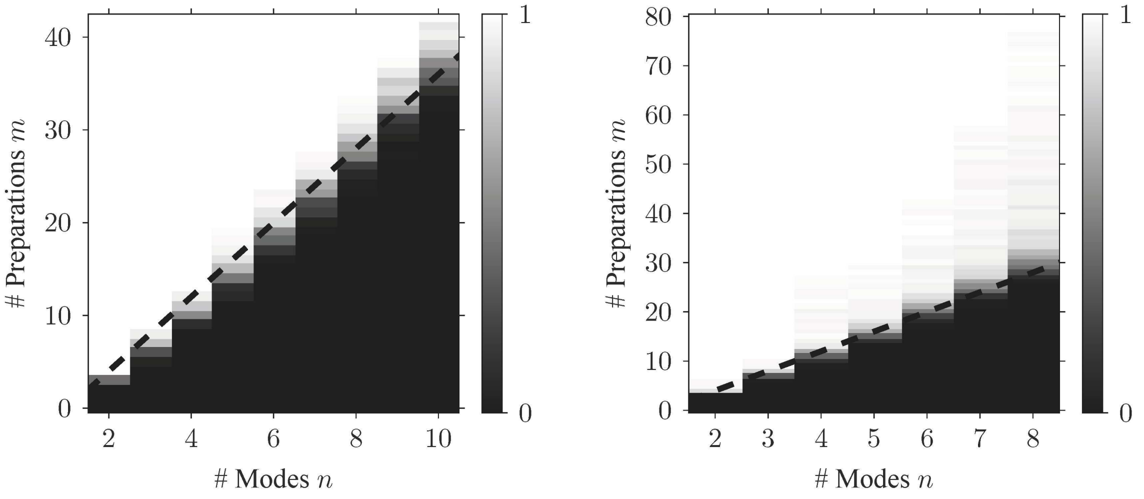

FIGURE 2. Phase transition diagrams for the recovery of Haar random unitary test matrices T via our protocol. The x-axis indicates the problem's dimension n (i.e. the number of incoming and outgoing wires), while the y-axis denotes the number of different Gauss random input light configurations. The frequency of successful recovery over 100 independent runs appears color-coded from black (zero) to white (one). The plot on the right hand side considers recovery in the absence of noise, while on the right hand side additive Gaussian noise with mean zero and variance 0.001 has been added to each detector measurement. We have considered an individual recovery to be successful, if the reconstructed matrix is – up to left multiplication with a diagonal unitary matrix – sufficiently close in Frobenius distance to the original test matrix (10^{-5} in the noiseless case and 10^{-2} in the noisy setting). In both plots, the dashed line indicates $m = 4n - 4$.

References

[1] R. Balan, P. Casazza, D. Edidin, *On signal reconstruction without phase*, Appl. comput. Harmon. Anal **20** (2006), 345–356.

[2] B. Alexeev, A.S. Bandeira, M. Fickus, D.G. Mixon, *Phase retrieval with polarization*, SIAM J. Imaging Sci. **7** (2014), 35–66.

[3] B.G. Bodmann, N. Hammen *Stable phase retrieval with low redundancy frames*, Adv. Comput. Math, **41** (2015), 317–331.

[4] E.J. Candès, X. Li, M. Soltanolkotabi, *Phase Retrieval via Wirtinger Flow: Theory and Algorithms*, IEEE Trans. Inform. Theory, **61** (2015), 1985–2007.

[5] E.J. Candès, Y.C. Eldar, T. Strohmer, V. Voroninski, *Phase retrieval via matrix completion*, SIAM Rev., **57(2)** (2015), 225–251.

[6] E.J. Candès, T. Strohmer, V. Voroninski, *PhaseLift: Exact and Stable Signal Recovery from Magnitude Measurements via Convex Programming*, Commun. Pure Apll. Math., **66** (2013), 1241–1274

[7] E.J. Candès, X. Li, *Solving Quadratic Equations via PhaseLift When There Are About as Many Equations as Unknowns*, Found. Comput. Math. **14** (2014), 1017–1026

[8] R. Kueng, H. Rauhut, U. Terstiege, *Low rank matrix recovery from rank one measurements*, Appl. Comput. Harmon. Anal., doi:10.1016/j.acha.2015.07.007 (2015)

[9] M. Kabanava, R. Kueng, H. Rauhut, U. Terstiege, *Stable low-rank matrix recovery via null space properties*, preprint arXiv:1507.07184 (2015)

[10] B. Recht, M. Fazel, P.A. Parrilo, *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Rev. **52** (2010), 471–501

[11] D. Gross, *Recovering low-rank matrices from few coefficients in any basis*, IEEE Trans. Inform. Theory **57** (2011), 1548–1566

[12] E. Knill, R. Laflamme, G.J. Milburn, *A scheme for efficient quantum computation with linear optics*, Nature **409** (2000), 46–52

[13] J. Carolan, C. Harrold, C. Sparrow, E. Martín-López, N.J. Russell, J.W. Silverstone, P.J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh, G.D. Marshall, M.G. Thompson, J.C.F. Matthews, T. Hashimoto, J.L. O'Brien, A. Laing, *Universal linear optics*, preprint arXiv:1505.01182 (2015)

[14] S. Rahimi-Keshari, M.A. Broome, R. Fickler, A. Fedrizzi, T.C. Ralph, A.G. White, *Direct characterization of linear-optical networks*, Opt. Express **21** (2013), 13450–13458

Integral geometry and phase transitions in conic optimization

Martin Lotz

(joint work with Dennis Amelunxen, Michael B. McCoy, Joel A. Tropp)

More than 10 years ago, foundational work in compressive sensing established that sparse or compressible signals could be reconstructed efficiently, using convex optimization, from a number of observations nearly proportional to the sparsity of the signal, rather than the ambient dimension. In addition, an interesting phase transition phenomenon has been observed: the probability that the optimization problem $\min_x \|x\|_1$ subject to $Ax = b$ recovers an s-sparse, or all s-sparse, solutions of $Ax = b$ for random A jumps from almost 0 to almost 1 as the number of rows of A passes a certain threshold. Donoho and Tanner accurately described the phase transition location [9], and their work was confirmed using an (at the time) seemingly unrelated approach by Stojnic [20]. The observed phase transitions are not unique to ℓ_1 minimization; they are a feature of convex optimization problems with random constraints. More generally, the optimality conditions in convex optimization can usually be formulated as intersection conditions of certain subdifferential cones associated to the problem; for example, the problem $\min f(x)$ subject to $Ax = b$ has $\hat{x}$ as unique solution if and only if the cone spanned by the subdifferential $\partial f(\hat{x})$ intersects the image $\mathrm{im} A^\top$ nontrivially (or the descent cone of f at $\hat{x}$ does not intersect the kernel of A nontrivially). The general geometric question is then: given two or more cones, randomly rotated according to the Haar measure on the orthogonal group, what is the probability that they intersect?

Classical integral geometry, going back to the work of Santaló and Blaschke, is tailor made to address such questions (see [19] for a modern treatment). Via the principal kinematic formula, it provides exact expressions for the probability that various randomly moved geometric objects intersect in terms of geometric invariants, the *intrinsic volumes* of the objects involved. For the case of two cones C and D and random orthogonal Q it takes the form

$$\mathrm{P}\{C \cap QD \neq \{0\}\} = 2 \sum_{k \text{ odd}} \sum_{i+j=d+k} v_i(C) v_j(D). \tag{1}$$

Not surprisingly, integral geometry has played a role in many fields such as the probabilistic analysis of condition numbers [8, 4, 16], complexity of optimization problems [21, 5], statistics [1], or the analysis of randomized algorithms [17], to name a few. A special case of the kinematic formula, applied to descent cones of the ℓ_1 norm at sparse vectors, also formed the basis for the asymptotic estimates by Donoho and Tanner [9]. In [3], it was shown that the integral-geometric approach naturally leads to an explanation of the phase transition phenomenon in convex optimization. The key is the observation that the intrinsic volumes of convex cones form a probability distribution, and that the intersection probability, given by the kinematic formula (1), is described by a convolution of this discrete distribution. Using an extension of the classical Steiner formula, it was shown that the intrinsic volume distribution concentrates around its mean, the statistical dimension $\delta(C)$, from which the existence of the phase transitions, and their location, immediately follows. More precisely, given cones C and D (for example, C descent cone of a norm and D the kernel of a Gaussian matrix),

$$\begin{aligned}\delta(C)+\delta(D) \lesssim d \qquad &\Longrightarrow \qquad \mathrm{P}\big\{C\cap QD=\{0\}\big\}\approx 1\\ \delta(C)+\delta(D) \gtrsim d \qquad &\Longrightarrow \qquad \mathrm{P}\big\{C\cap QD=\{0\}\big\}\approx 0.\end{aligned}$$

The statistical dimension is the unique, orthogonal invariant and continuous valuation on the set of convex cones that coincides with the dimension on linear subspaces. Moreover, it coincides with the expected squared projected length of a Gaussian vector on the cone, a quantity that has appeared in many contexts (for example, in [6]), and is closely related to the Gaussian width. In the case when one of the cones is the kernel of a Gaussian matrix, then accurate bounds on the intersection probability are given by Gordon's escape through the mesh argument, and the corresponding lower bounds can be derived from a duality argument. The integral-geometric approach, however, does not make reference to the representation of the random transformation operator as a Gaussian matrix, and can be seen as a direct generalization of the basic fact from linear algebra that subspaces intersect nontrivially (almost surely) if an only if their dimensions add up to more than the ambient dimension, without reference to their representation as matrices. When dealing with robustness questions, however, an analysis of the cone-restricted smallest singular value of a random matrix using probabilistic tools becomes important, see [7] or [11, Chapters 8-9] for an analysis of linear inverse problems based on the Gaussian width and Gordon's inequality or [2] for some discussion of links to integral geometry and the analysis of condition numbers.

Despite the importance of the intrinsic volumes for understanding the statistics of random convex cones, many of their features are still not fully understood. An important contribution has been the recent work by Goldstein, Peccati and Nourdin [12], who established a central limit theorem, including a Berry-Esseen type bound, for the intrinsic volume distribution. A consequence is that the phase transitions can be understood in terms of Gaussian distributions associated to the individual cones. A natural question is whether the conic intrinsic volumes are log-concave, i.e., satisfy an Alexandrov-Fenchel inequality $v_i(C)^2 \geq v_{i-1}(C)v_{i+1}(C)$. It is known [15] that the average of the i-th intrinsic volume of the chambers of a hyperplane arrangement is determined by the i-th coefficient of the characteristic polynomial of the arrangement, the value of a Möbius function. Recent work by June Huh has shown that the (absolute) coefficients of the characteristic polynomial of a hyperplane arrangement are log-concave [13], solving a long-standing open problem by Rota. Another question is whether, among cones with a fixed statistical dimension, the intrinsic volumes of circular cones have the largest variance. Finally, there is the question of universality: do the phase transitions, and their location, persist for more general distributions, as extensive experiments [10, 14] suggest? The answer appears to be yes [18].

References

[1] Rober J. Adler and Jonathan E. Taylor. *Random fields and geometry.* Springer Monographs in Mathematics. Springer, New York, 2007.

[2] Dennis Amelunxen and Martin Lotz. Gordon's inequality and condition numbers in conic optimization. *arXiv preprint arXiv:1408.3016*, 2014.

[3] Dennis Amelunxen, Martin Lotz, Michael B. McCoy, and Joel A. Tropp. Living on the edge: phase transitions in convex programs with random data. *Information and Inference*, 2014.

[4] Peter Bürgisser. Smoothed analysis of condition numbers. In *Proceedings of the International Congress of Mathematicians, Hyderabad*, volume IV, pages 2609–2633. World Scientific, 2010.

[5] Peter Bürgisser, Felipe Cucker, and Martin Lotz. Coverage processes on spheres and condition numbers for linear programming. *Ann. Probab.*, 38(2):570–604, 2010.

[6] Venkat Chandrasekaran and Michael I. Jordan. Computational and statistical tradeoffs via convex relaxation. *arXiv preprint arXiv:1211.1073*, 2012.

[7] Venkat Chandrasekaran, Benjamin Recht, Pablo A. Parrilo, and Alan S. Willsky. The convex geometry of linear inverse problems. *Found. Comput. Math.*, 12(6):805–849, 2012.

[8] James Demmel. The probability that a numerical analysis problem is difficult. *Math. Comp.*, 50:449–480, 1988.

[9] David L. Donoho and Jared Tanner. Counting faces of randomly projected polytopes when the projection radically lowers dimension. *J. Amer. Math. Soc.*, 22(1):1–53, 2009.

[10] David L. Donoho and Jared Tanner. Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing. *Phil. Trans. R. Soc. A*, 367(1906):4273–4293, 2009.

[11] Simon Foucart and Holger Rauhut. *A mathematical introduction to compressive sensing*, volume 336 of *Applied and Numerical Harmonic Analysis.* Birkhäuser, Basel, 2013.

[12] Larry Goldstein, Ivan Nourdin, and Giovanni Peccati. Gaussian phase transitions and conic intrinsic volumes: Steining the steiner formula. *arXiv preprint arXiv:1411.6265*, 2014.

[13] June Huh. Milnor numbers of projective hypersurfaces and the chromatic polynomial of graphs. *Journal of the American Mathematical Society*, 25(3):907–927, 2012.

[14] Jakob Sauer Jørgensen and EY Sidky. How little data is enough? phase-diagram analysis of sparsity-regularized x-ray computed tomography. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 373(2043):20140387, 2015.

[15] Caroline J Klivans and Ed Swartz. Projection volumes of hyperplane arrangements. *Discrete & Computational Geometry*, 46(3):417–426, 2011.

[16] Martin Lotz. On the volume of tubular neighborhoods of real algebraic varieties. *Proceedings of the American Mathematical Society*, 143(5):1875–1889, 2015.

[17] Oren Mangoubi and Alan Edelman. Integral geometry for markov chain monte carlo: overcoming the curse of search-subspace dimensionality. *arXiv preprint arXiv:1503.03626*, 2015.

[18] Samet Oymak and Joel Tropp. Private communication, 2015.

[19] Rolf Schneider and Wolfgang Weil. *Stochastic and Integral Geometry.* Springer series in statistics: Probability and its applications. Springer, 2008.

[20] Mihailo Stojnic. Various thresholds for ℓ_1-optimization in compressed sensing. *preprint*, 2009. arXiv:0907.3666.

[21] Anatoly M Vershik and Piotr V Sporyshev. An asymptotic estimate of the average number of steps of the parametric simplex method. *USSR Computational Mathematics and Mathematical Physics*, 26(3):104–113, 1986.

Signal decomposition and analysis via extraction of frequencies

Hrushikesh N. Mhaskar
(joint work with Charles K. Chui)

Time-frequency analysis is central to signal processing, with standard adaptation to higher dimensions for imaging applications, and beyond. However, although the theory, methods, and algorithms for stationary signals are well developed, mathematical analysis of non-stationary signals is almost nonexistent. For a real-valued signal defined on the time-domain $\mathbb{R}$, a classical approach to compute its instantaneous frequency (IF) is to consider the amplitude- frequency modulated (AM–FM) formulation of its complex (or analytic) signal extension, via the Hilbert transform.

In a popular paper by Huang et.al. [1], the so-called empirical mode decomposition (EMD) scheme is introduced to separate such a signal as a sum of finitely many intrinsic mode functions (IMF's), with a slowly oscillating signal as the remainder, so that more than one IF's of the given signal can be computed by extending each IMF to an AM-FM signal component. Based on the continuous wavelet transform (CWT), the notion of synchrosqueezing transform (SST), introduced by Daubechies and Mae in 1996, and further developed by Daubechies, Lu, and Wu (DLW) in a 2011 paper [2], provides another approach to extract more than one IF's of the signal on R. Furthermore, by introducing a list of fairly restrictive conditions on the adaptive harmonic (AHM) signal model, the DLW paper also derives a theory for estimating the signal components according to this model, by using the IF's with estimates The objective of our present paper is to introduce another mathematical theory, along with rigorous methods and computational schemes, to achieve a more ambitious goal than the SST approach, first to extract the polynomial-like trend from the source signal, then to compute the exact number of signal components according to a less restrictive AHM model, then to obtain better estimates of the IF's and instantaneous amplitudes (IA's) of the signal components, and finally to separate the signal components from the (blind) source signal. Furthermore, our computational scheme can be realized in near real-time, and our mathematical theory has direct extension to the multivariate setting.

References

[1] Norden E. Huang, Zheng Shen, Steven R. Long, Manli C. Wu, Hsing H. Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung and Henry H. Liu. The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis. *Proceedings: Mathematical, Physical and Engineering Sciences*, Vol. 454, No. 1971 (Mar. 8, 1998), pp. 903–995.

[2] Daubechies, Ingrid, Jianfeng Lu, and Hau-Tieng Wu. Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool. *Applied and computational harmonic analysis* 30.2 (2011): 243–261.

Recent advances in mathematical data science

Dustin G. Mixon

(joint work with Afonso S. Bandeira, Takayuki Iguchi, Jesse Peterson, Benjamin Recht, Soledad Villar)

This talk describes recent work on three different problems of interest in mathematical data science, namely, compressive classification, k-means clustering, and deep learning.

First, compressive classification is a problem that comes on the heels of compressive sensing. In compressive sensing, one exploits the underlying structure of a signal class in order to exactly reconstruct any signal from the class given very few linear measurements of the signal. However, many applications do not require an exact reconstruction of the image, but rather a classification of that image (for example, is this a picture of a cat, or of a dog?). As such, it makes intuitive sense that the classification task might succeed given far fewer measurements than are necessary for compressive sensing.

Much like compressive sensing, compressive classification must exploit some notion of simplicity of the data set. For this talk, we consider data sets which are linearly separable, that is, one may distinguish two classes of points $A, B \subseteq \mathbb{R}^n$ by thresholding an inner product:

$$\begin{aligned} x \in A \quad &\Longrightarrow \quad \langle x, v \rangle < \theta, \\ x \in B \quad &\Longrightarrow \quad \langle x, v \rangle > \theta. \end{aligned}$$

In particular, given sets A and B and tolerance η, we seek the smallest m such that PA and PB are linearly separable for a random $m \times n$ matrix P with probability $\geq 1 - \eta$. This establishes a minimal number of measurements necessary to classify in the compressed domain. To approach this problem, we note that PA and PB are linearly separable if and only if the null space of P trivially intersects the cone generated by the Minkowski difference $A - B$. As such, we may leverage Gordon's theory of escape through a mesh [5], or more recently, the approximate kinematic formula from conic integral geometry [1]. With these tools, we find estimates in the special cases where A and B are Euclidean balls, and more generally, when they are ellipsoids (see [3] for details).

The second problem we discuss is k-means clustering. Here, given a point cloud $P \subseteq \mathbb{R}^n$, one is asked to partition the points into k clusters $C_1, \ldots, C_k$ such that the corresponding cluster centers exhibit the smallest possible sum of squared error:

$$\text{minimize} \quad \sum_{t=1}^{k} \sum_{x \in C_t} \bigg\| x - \frac{1}{|C_t|} \sum_{y \in C_t} y \bigg\|^2 \quad \text{subject to} \quad C_1 \sqcup \cdots \sqcup C_k = P$$

Unfortunately, minimizing the so-called k-means objective is NP-hard in general [9]. However, Lloyd's algorithm, which alternates between finding cluster centers for a proto-partition and reassigning points to the nearest cluster center, performs well in practice. Unfortunately, there is currently no guarantee that such an algorithm finds the k-means-optimal clustering.

In this talk, we consider a semidefinite relaxation of the k-means problem. Denoting $p = |P|$, we define the $p \times p$ matrix D by $D_{ij} := \|x_i - x_j\|^2$. Next, letting 1_{C_t} denote the p-dimensional indicator vector of C_t, then a straightforward manipulation gives

$$\sum_{t=1}^{k} \sum_{x \in C_t} \left\| x - \frac{1}{|C_t|} \sum_{y \in C_t} y \right\|^2 = \frac{1}{2} \operatorname{Tr}\left(D \sum_{t=1}^{k} \frac{1}{|C_t|} 1_{C_t} 1_{C_t}^\top \right).$$

Writing $X := \sum_{t=1}^{k} \frac{1}{|C_t|} 1_{C_t} 1_{C_t}^\top$, we identify several convex constraints that X satisfies, and we may optimize subject to these constraints in order to produce a polytime-solvable program:

$$\begin{array}{ll} \text{minimize} & \operatorname{Tr}(DX) \\ \text{subject to} & \operatorname{Tr}(X) = k \\ & X1 = 1 \\ & X \geq 0 \\ & X \succeq 0 \end{array}$$

In general, we can expect the value of this relaxed program to be smaller than the value of the original k-means program. However, if the relaxed optimizer happens to be integral, meaning the optimal X has the form $\sum_{t=1}^{k} \frac{1}{|C_t|} 1_{C_t} 1_{C_t}^\top$, then we may conclude that the corresponding clustering $C_1, \ldots, C_k$ is k-means optimal. Amazingly, when the data is drawn randomly from a reasonable model (the so-called stochastic ball model, introduced in [13]), one may show that the relaxed optimizer is integral with high probability (see [2] and [8]). It remains to find faster-than-SDP algorithms which enjoy a similar performance guarantee.

The last problem we consider is deep learning. Today, deep learning is the state-of-the-art technique for several important classification tasks [15, 7, 6]. For each of these problems, given a labeled training set, one is tasked with producing a labeling function that not only matches the training set, but is also simple enough to generalize well to a test set. For deep learning, the labeling function is implemented by a deep neural network, which amounts to a large circuit of neurons, where each neuron linearly combines the outputs of its parent neurons, and then outputs a nonlinear function of this combination. To learn this labeling function, practitioners locally optimize the function parameters so as to fit the training set. To date, there is little theory to explain why this should perform as well as it does in practice.

Our approach to deep learning is motivated by an analogy with Boolean circuits. These circuits are similar to neural nets, except neurons are replaced by

Boolean gates (such as AND, OR, or threshold gates). Interestingly, Boolean circuits of simple structure tend to implement Boolean functions $f\colon \{\pm1\}^n \to \{\pm1\}$ of concentrated spectra [11, 14], that is, $\{a_S\}_{S\subseteq[n]}$ is nearly sparse, where

$$f(x_1,\ldots,x_n) = \sum_{S\subseteq[n]} a_S \prod_{i\in S} x_i \qquad \forall (x_1,\ldots,x_n)\in\{\pm1\}^n.$$

Passing through the analogy, we hypothesize that learnable neural nets are necessarily simple, and furthermore, Boolean functions that are well approximated by such neural nets necessarily have concentrated spectra. If this hypothesis is true, then the deep learning problem can be relaxed to a sparse approximation problem. We test this hypothesis by classifying the zeros and ones in the MNIST database of handwritten digits [10] by way of sparse approximation (indeed, deep neural nets currently hold the record for classifying MNIST digits [4]). In our experiment, we managed to obtain a misclassification rate of 0.74%, thereby proving the concept of classification by sparse approximation (see [12] for details). Unfortunately, our naïve method of sparse approximation does not scale well, and so it remains to find a scalable alternative to implement on other instances of binary classification.

References

[1] D. Amelunxen, M. Lotz, M. B. McCoy, J. A. Tropp, Living on the edge: Phase transitions in convex programs with random data, Inform. Inference 3 (2014) 224–294.

[2] P. Awasthi, A. S. Bandeira, M. Charikar, R. Krishnaswamy, S. Villar, R. Ward, Relax, no need to round: integrality of clustering formulations, Available online: arXiv:1408.4045

[3] A. S. Bandeira, D. G. Mixon, B. Recht, Compressive classification and the rare eclipse problem, Available online: arXiv:1404.3203

[4] D. Ciresan, U. Meier, J. Schmidhuber, Multi-column deep neural networks for image classification, CVPR (2012) 3642–3649.

[5] Y. Gordon, On Milman's inequality and random subspaces which escape through a mesh in $\mathbb{R}^n$, Geometric aspects of functional analysis, Israel Seminar 1986–87, Lecture Notes in Mathematics 1317 (1988) 84–106.

[6] A. Hannun, C. Case, J. Casper, B. Catanzaro, G. Diamos, E. Elsen, R. Prenger, S. Satheesh, S. Sengupta, A. Coates, A. Y. Ng, Deep Speech: Scaling up end-to-end speech recognition, Available online: arXiv:1412.5567

[7] K. He, X. Zhang, S. Ren, J. Sun, Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification, Available online: arXiv:1502.01852

[8] T. Iguchi, D. G. Mixon, J. Peterson, S. Villar, On the tightness of an SDP relaxation of k-means, Available online: arXiv:1505.04778

[9] K. Jain, M. Mahdian, A. Saberi, A new greedy approach for facility location problems, STOC (2002) 731–740.

[10] Y. LeCun, C. Cortes, C. J. C. Burges, The MNIST database of handwritten digits, Available online: `yann.lecun.com/exdb/mnist/`

[11] N. Linial, Y. Mansour, N. Nisan, Constant depth circuits, Fourier transform, and learnability, J. ACM 40 (1993) 607–620.

[12] D. G. Mixon, J. Peterson, Learning Boolean functions with concentrated spectra, Proc. SPIE 9597 (2015) 95970C/1–8.

[13] A. Nellore, R. Ward, Recovery guarantees for exemplar-based clustering, Available online: arXiv:1309.3256

[14] A. Shpilka, A. Tal, B. lee Volk, On the structure of Boolean functions with small spectral norm, Available online: arXiv:1304.0371

[15] Y. Taigman, M. Yang, M. Ranzato, L. Wolf, DeepFace: Closing the gap to human-level performance in face verification, CVPR (2014) 1701–1798.

Shearlet systems on bounded domains

Philipp Petersen

(joint work with Philipp Grohs, Gitta Kutyniok, Jackie Ma)

Driven by an overwhelming amount of applications, numerical approximation of partial differential equations was established as one of the core areas in applied mathematics. During the last decades a trend for the solution of PDEs emerged, that focuses on employing systems from applied harmonic analysis for the adaptive solution of these equations. Most notably wavelet systems have been used, which lead for instance to provably optimal solvers for elliptic PDEs, see [1]. Inspired by this success story also other systems with various advantages in different directions should be employed in various discretization problems. For instance, ridgelets where recently successfully used in the discretization of linear transport equations, see [4].

Another famous system is that of shearlets, [7], which admits optimal representations of functions that have singularities along smooth curves and, perhaps more importantly for the solution of PDEs, they also provide drastically improved approximation rates when compared with wavelets of functions that have first or higher order cartoon-like derivatives, see [8]. The main bottleneck in developing shearlet, or ridgelet-based PDE solvers is the fact that originally these systems are constructed as representation systems, or frames, for functions defined on $\mathbb{R}^d$. On the other hand, most PDEs are defined on a finite domain $\Omega \subset \mathbb{R}^d$ which implies that the development of effective PDE solvers crucially depends on the construction of anisotropic representation systems on finite domains, satisfying various boundary conditions. Hence it is necessary to have a system on a bounded domain Ω, which fulfills the following desiderata

- **[D1]** yields a frame for $L^2(\Omega)$,
- **[D2]** is able to incorporate boundary conditions,
- **[D3]** gives rise to optimal approximation rates for functions with anisotropic structures,
- **[D4]** characterizes Sobolev spaces by weighted ℓ^2 norms.

Although there have been first approaches to construct shearlet systems for the solution of PDEs on bounded domains, e.g. [6], they fail to satisfy all the desiderata above.

1. Construction

In this talk we proposed a new construction of a shearlet system for the adaptive solution of PDEs. We start with a compactly supported shearlet frame for $\mathbb{R}^2$, as constructed in [5]. Let for $j, k \in \mathbb{Z}$

$$A_j := \text{ diag } (2^j, 2^{\frac{j}{2}}) := \begin{pmatrix} 2^j & 0 \\ 0 & 2^{\frac{j}{2}} \end{pmatrix}, \text{ and } S_k := \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix},$$

denote the *parabolic scaling matrix* and *shearing matrix.* Let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$, $c = [c_1, c_2]^T \in \mathbb{R}^2$ with $c_1, c_2 > 0$. Then the cone-*adapted shearlet system* is defined by $(\psi_{j,k,m,\iota})_{(j,k,m,\iota)\in\Lambda}$, where

$$\Lambda := \left\{(j,k,m,\iota) \,:\, \iota \in \{-1,0,1\}, |\iota j| \geq j \geq 0,\ |k| \leq |\iota| \left\lceil 2^{\frac{j}{2}} \right\rceil,\ m \in \mathbb{Z}^2 \right\}.$$

and

$$\begin{aligned} \psi_{0,0,m,0} &:= \phi(\cdot - c_1 m), \\ \psi_{j,k,m,1} &:= 2^{\frac{3j}{4}} \psi(S_k A_j \cdot - M_c m), \\ \psi_{j,k,m,-1} &:= 2^{\frac{3j}{4}} \tilde{\psi}(S_k^T \tilde{A}_j \cdot - M_{\tilde{c}} m), \end{aligned}$$

with $M_c := \text{diag}(c_1, c_2)$, $M_{\tilde{c}} = \mathrm{diag}(\mathrm{c_2}, \mathrm{c_1})$, and $\tilde{A}_{2^j} = \mathrm{diag}(2^{\frac{\mathrm{j}}{2}}, 2^{\mathrm{j}})$.

Of this shearlet system $(\psi_{j,k,m,\iota})_{j,k,m,\iota\in\Lambda}$ we only keep the frame elements whose support is fully contained in Ω. Clearly this system is not complete in $L^2(\Omega)$. For this reason we augment the aforementioned shearlet subsystem by boundary wavelets as constructed for instance in [2]. To be more precise we have the following definition:

Definition 1. [3] *For $t \in \mathbb{Z}$ we denote by Γ_t the part of Ω that has distance less than $2^{-\frac{t}{2}}$ from $\partial\Omega$, i.e. $\Gamma_t := \{x \in \Omega : d(x, \partial\Omega) < 2^{-\frac{t}{2}}\}$.*

Let $(\psi_{j,k,m,\iota})_{j,k,m,\iota\in\Lambda}$ be a shearlet system. Further, let $t \in \mathbb{N}$ and $\mathcal{W}$ be an orthonormal bases of wavelets on $L^2(\Omega)$ and $\mathcal{W}_t := \{\omega_{j,m} \in \mathcal{W} :\ supp\ \omega_{j,m} \cap \Gamma_{j-t} \neq \emptyset\}$. Then, the boundary shearlet system with offset t *is defined as*

$$(\varphi_n)_{n\in\mathbb{N}} := \{\psi_{j,k,m,\iota} :\ supp\ \psi_{j,k,m,\iota} \subseteq \Omega\} \cup \mathcal{W}_t.$$

2. Properties

We can describe some properties of the boundary shearlet system $(\varphi_n)_{n\in\mathbb{N}}$. In particular it fulfills all desiderata **[D1]** - **[D4]** mentioned in the introduction. Under appropriate assumptions on smoothness and vanishing moments of the wavelet ONB and the shearlet system $(\psi_{j,k,m,\iota})_{j,k,m,\iota\in\Lambda}$ we get

Theorem 1. [3] *There exists $t \in \mathbb{N}$ such that the boundary shearlet system $(\varphi_n)_{n\in\mathbb{N}}$ with offset $t' > t$ has the property, that there exist $0 < A \leq B < \infty$ such that for all $s \leq N$ we have*

$$A\|f\|^2_{H^s(\Omega)} \leq \sum_{n\in\mathbb{N}} 2^{2j_n s} |\langle f, \varphi_n\rangle|^2 \leq B\|f\|^2_{H^s(\Omega)}, \ \textit{for all } f \in H^s(\Omega)$$

In particular this holds for $s = 0$.

This yields [**D1**] and [**D4**]. The incorporation of boundary conditions, [**D2**], is possible due to the fact, that this can be done by the wavelet bases used in the construction. Concerning the optimal approximation rates of [**D3**] we get the following result which we state slightly informal, since we do not want to introduce to much notation. Under appropriate assumptions on smoothness and vanishing moments of the wavelet ONB and the shearlet system $(\psi_{j,k,m,\iota})_{j,k,m,\iota} \in \Lambda$ we get.

Theorem 2. [3] *Let $(\varphi_n)_{n\in\mathbb{N}}$ be a boundary shearlet system and let f be a C^2 function apart from finitely many disjoint C^2 discontinuity curves, that touch the boundary of Ω only finitely many times. Then we have*

$$\|f - f_N\|^2_{L^2(\Omega)} \leq CN^{-2}\log(N)^3 \text{ for } N \to \infty,$$

where f_N is the best N-term approximation of f with respect to the boundary shearlet system $(\varphi_n)_{n\in\mathbb{N}}$ and C is a constant.

References

[1] Albert Cohen, Wolfgang Dahmen, and Ronald DeVore. Adaptive wavelet methods for elliptic operator equations: convergence rates. *Math. Comp.*, 70(233):27–75, 2001.

[2] Albert Cohen, Ingrid Daubechies, and Pierre Vial. Wavelets on the interval and fast wavelet transforms. *Appl. Comput. Harmon. Anal.*, 1(1):54–81, 1993.

[3] Philipp Grohs, Gitta Kutyniok, Jackie Ma, Philipp Petersen, Multiscale anisotropic directional systems on bounded domains, *in preparation*, 2015.

[4] Philipp Grohs and Axel Obermeier. Optimal adaptive ridgelet schemes for linear transport equations. Technical Report 2014-21 (revised), Seminar for Applied Mathematics, ETH Zürich, 2014.

[5] Pisamai Kittipoom, Gitta Kutyniok, and Wang-Q Lim. Construction of compactly supported shearlet frames. *Constr. Approx.*, 35(1):21–72, 2012.

[6] Gitta Kutyniok and Wang-Q Lim. Shearlets on bounded domains. Approximation theory XIII: San Antonio 2010, 187–206, 2012.

[7] Demetrio Labate, Wang-Q. Lim, Gitta Kutyniok, and Guido Weiss. Sparse multidimensional representation using shearlets. *Wavelets XI.*

[8] Philipp Petersen Shearlet approximation of functions with discontinuous derivatives, arXiv:1508.00409 (2015).

A deterministic sparse FFT algorithm for vectors with short support

Gerlind Plonka
(joint work with Katrin Wannenwetsch)

We consider sparse signals $\mathbf{x} \in \mathbb{C}^N$ which are known to vanish outside a support interval of length bounded by $m < N$. For the case that m is known, we propose a deterministic algorithm of complexity $\mathcal{O}(m \log m)$ for reconstruction of $\mathbf{x}$ from its discrete Fourier transform $\widehat{\mathbf{x}} \in \mathbb{C}^N$.

Introduction. Fast algorithms for the computation of the discrete Fourier transform of a vector of length N have been known for many years. These FFT algorithms have an arithmetical complexity of $\mathcal{O}(N \log N)$. Recently, there has been a strong interest in Fourier algorithms for sparse vectors with sublinear complexity. Randomized sparse Fourier algorithms achieving a complexity of $\mathcal{O}(m \log N)$ resp.

$\mathcal{O}(m \log m)$ for m-sparse vectors can e.g. be found in [3] resp. [5], [6]. An overview of the methods of randomized sparse Fourier transforms is given in [2]. Previous deterministic approaches (see e.g. [1, 4]) also lead to sublinear algorithms having polynomials costs in m and $\log N$.

Here, we present a deterministic FFT algorithm and restrict ourselves to vectors with a short support interval. Such vectors occur in different applications, such as in X-ray microscopy, where compact support is a frequently used a-priori condition in phase retrieval, as well as in computer tomography reconstructions.

Let $\mathbf{x} = (x_k)_{k=0}^{N-1} \in \mathbb{C}^N$. We define the *support length* $m = |\operatorname{supp} \mathbf{x}|$ of $\mathbf{x}$ as the minimal integer m for which there exists a $\mu \in \{0, \ldots, N-1\}$ such that the components x_k of $\mathbf{x}$ vanish for all $k \notin I := \{(\mu + r) \mod N, \quad r = 0, \ldots, m-1\}$. The index set I is called *support interval* of $\mathbf{x}$. We always have $x_\mu \neq 0$ and $x_{\mu+m-1} \neq 0$, but there may be components of $\mathbf{x}$ equal to zero within the support interval. Observe that if $m \leq \frac{N}{2}$, the support interval and hence the first support index μ of $\mathbf{x}$ is uniquely determined.

We define the discrete Fourier transform of a vector $\mathbf{x} \in \mathbb{C}^N$ by $\widehat{\mathbf{x}} = \mathbf{F}_N \mathbf{x}$, where the Fourier matrix $\mathbf{F}_N$ is given by $\mathbf{F}_N := (\omega_N^{jk})_{j,k=0}^{N-1}$, $\omega_N := \mathrm{e}^{-\frac{2\pi \mathrm{i}}{N}}$. In the following, we describe a deterministic algorithm for the reconstruction of $\mathbf{x}$ of length $N = 2^J$ from Fourier data $\widehat{\mathbf{x}} \in \mathbb{C}^N$. The algorithm is based on the idea that the (at most) m nonzero components of $\mathbf{x}$ can already be identified from a periodization of $\mathbf{x}$ of length $2^L \geq m$. Hence for the complete reconstruction it remains to determine the support interval (i.e., the first support index) of $\mathbf{x}$.

Reconstruction of x with short support interval. Let $N := 2^J$ for some $J > 0$. We define the periodizations $\mathbf{x}^{(j)} \in \mathbb{C}^{2^j}$ of $\mathbf{x}$ by

$$\mathbf{x}^{(j)} = (x_k^{(j)})_{k=0}^{2^j-1} = \left(\sum_{\ell=0}^{2^{J-j}-1} x_{k+2^j \ell} \right)_{k=0}^{2^j-1} \tag{1}$$

for $j = 0, \ldots, J$. The discrete Fourier transform of the vectors $\mathbf{x}^{(j)}$, $j = 0, \ldots, J$, can be described in terms of $\widehat{\mathbf{x}}$. According to the following lemma, it can be obtained by just picking suitable components of $\widehat{\mathbf{x}}$.

Lemma 1 (see [7]). *For the vectors $\mathbf{x}^{(j)} \in \mathbb{C}^{2^j}$, $j = 0, \ldots, J$, in* (1), *we have the discrete Fourier transform*

$$\widehat{\mathbf{x}}^{(j)} := \mathbf{F}_{2^j} \mathbf{x}^{(j)} = (\widehat{x}_{2^{J-j}k})_{k=0}^{2^j-1},$$

where $\widehat{\mathbf{x}} = (\widehat{x}_k)_{k=0}^{N-1} = \mathbf{F}_N \mathbf{x}$ is the Fourier transform of $\mathbf{x} \in \mathbb{C}^N$.

Assume that the Fourier data $\widehat{\mathbf{x}} = \mathbf{F}_N \mathbf{x} \in \mathbb{C}^N$ and $|\operatorname{supp} \mathbf{x}| \leq m$ for some given m. Choose L such that $2^{L-1} < m \leq 2^L$. By Lemma 1 we have $\widehat{\mathbf{x}}^{(L+1)} = (\widehat{x}_{2^{J-(L+1)}k})_{k=0}^{2^{L+1}-1}$. Thus, we can compute $\mathbf{x}^{(L+1)}$ using inverse FFT of length 2^{L+1}.

The resulting vector $\mathbf{x}^{(L+1)}$ has already the same support length as $\mathbf{x}$, since $|\operatorname{supp}\mathbf{x}| \le m \le 2^L$, and for each $k \in \{0, \dots, 2^{L+1}-1\}$ the sum in

$$x_k^{(L+1)} = \sum_{\ell=0}^{2^{J-L-1}-1} x_{k+2^{L+1}\ell} \tag{2}$$

contains at most one nonvanishing term. Therefore, the support of $\mathbf{x}^{(L+1)}$ and its first index $\mu^{(L+1)}$ are uniquely determined. For reconstruction of the complete vector $\mathbf{x}$ it is now sufficient to determine the first support index $\mu^{(J)} = \mu$ of the support interval of $\mathbf{x}$. Then the components of $\mathbf{x}$ are given by

$$x_{(\mu^{(J)}+k) \bmod N} = \begin{cases} x^{(L+1)}_{(\mu^{(L+1)}+k) \bmod 2^{L+1}} & k = 0, \dots, m-1, \\ 0 & k = m, \dots, N-1. \end{cases} \tag{3}$$

In order to find $\mu^{(J)}$, we observe that $\mu^{(J)} = \mu^{(L+1)} + \nu 2^{L+1}$ for some $\nu \in \{0, \dots, 2^{J-L-1}\}$ since $\mathbf{x}^{(L+1)}$ is a periodization of $\mathbf{x}$. In order to find ν, we first construct $\tilde{\mathbf{x}} \in \mathbb{C}^N$ from $\mathbf{x}^{(L+1)}$ by setting $\nu = 0$. Then the desired vector $\mathbf{x}$ is obtained from $\tilde{\mathbf{x}}$ by a (periodic) shift of all components by $\nu 2^{L+1}$. Using the properties of the DFT of a shifted vector, we can now obtain ν by comparing a suitable Fourier component of $\tilde{\mathbf{x}}$ with the corresponding given Fourier component of $\mathbf{x}$.

Theorem 1 (see [7]). *Let $\mathbf{x} \in \mathbb{C}^N$, $N = 2^J$, have support length m (or a support length bounded by m) with $2^{L-1} < m \le 2^L$. For $L < J-1$, let $\mathbf{x}^{(L+1)}$ be the 2^{L+1}-periodization of $\mathbf{x}$. Then $\mathbf{x}$ can be uniquely recovered from $\mathbf{x}^{(L+1)}$ and one nonzero component of the vector $(\widehat{x}_{2k+1})_{k=0}^{N/2-1}$.*

Sparse FFT Algorithm. We summarize the reconstruction of $\mathbf{x}$ from Fourier data $\widehat{\mathbf{x}}$ in the following algorithm.

Algorithm (see [7]) (Sparse FFT for vectors with short support)

Input: $\widehat{\mathbf{x}} \in \mathbb{C}^N$, $N = 2^J$, $|\operatorname{supp}\mathbf{x}| \le m < N$.

- Compute L such that $2^{L-1} < m \le 2^L$, i.e., $L := \lceil \log_2 m \rceil$.
- If $L = J$ or $L = J-1$, compute $\mathbf{x} = \mathbf{F}_N^{-1}\widehat{\mathbf{x}}$ using an FFT of length N.
- If $L < J-1$:
  1. Choose $\widehat{\mathbf{x}}^{(L+1)} := (\widehat{x}_{2^{J-(L+1)}k})_{k=0}^{2^{L+1}-1}$ and compute $\mathbf{x}^{(L+1)} := \mathbf{F}_{2^{L+1}}^{-1}\widehat{\mathbf{x}}^{(L+1)}$ using an FFT of length 2^{L+1}.
  2. Determine the first support index $\mu^{(L+1)} \in \{0, \dots, 2^{L+1}-1\}$ of $\mathbf{x}^{(L+1)}$ such that $x^{(L+1)}_{\mu^{(L+1)}} \neq 0$ and $x_k^{(L+1)} = 0$ for $k \notin \{(\mu^{(L+1)} + r) \bmod 2^{L+1},\ r = 0, \dots, m-1\}$.
  3. Choose a Fourier component $\widehat{x}_{2k_0+1} \neq 0$ of $\widehat{\mathbf{x}}$ and compute the sum

$$a := \sum_{\ell=0}^{m-1} x^{(L+1)}_{(\mu^{(L+1)}+\ell) \bmod 2^{L+1}}\, \omega_N^{(2k_0+1)(\mu^{(L+1)}+\ell)}.$$

(4) Compute $b := \widehat{x}_{2k_0+1}/a$ that is by construction of the form $b = \omega^p_{2^{J-L-1}}$ for some $p \in \{0, \dots, 2^{J-L-1} - 1\}$, and find $\nu \in \{0, \dots, 2^{J-L-1} - 1\}$ such that $(2k_0 + 1)\,\nu = p \bmod 2^{J-L-1}$.

(5) Set $\mu^{(J)} := \mu^{(L+1)} + 2^{L+1}\nu$, and $\mathbf{x} := (x_k)_{k=0}^{N-1}$ with entries

$$x_{(\mu^{(J)}+\ell) \bmod N} := \begin{cases} x^{(L+1)}_{(\mu^{(L+1)}+\ell) \bmod 2^{L+1}} & \ell = 0, \dots, m-1, \\ 0 & \ell = m, \dots, N-1. \end{cases}$$

Output: $\mathbf{x}$.

Our algorithm has an arithmetical complexity of $\mathcal{O}(m \log m)$. This can be seen as follows: In the first step, an FFT algorithm of this complexity is performed. All further steps require at most $\mathcal{O}(m)$ operations. Moreover, the algorithm needs less than $4m$ Fourier values. More detailed results can be found in [7] where we also propose an algorithm for noisy input data. In this case the evaluation of $\mu^{(L+1)}$ and of $\mu^{(J)}$ is stabilized using additional Fourier values and $\mathcal{O}(m \log N)$ arithmetical operations. Furthermore, if the nonzero entries of $\mathbf{x}$ are assumed to be real and positive, then a similar sparse FFT algorithm can be derived that does not require a priori knowledge of the support length m.

Acknowledgement. We gratefully acknowledge the funding of this work by the DFG in the project PL 170/16-1.

References

[1] A. Akavia, Deterministic sparse Fourier approximation via approximating arithmetic progressions, IEEE Trans. Inform. Theory **60**(3) (2014), 1733–1741.

[2] A. Gilbert, P. Indyk, M.A. Iwen, and L. Schmidt, *Recent developments in the sparse Fourier transform*, IEEE Signal Process. Magazine **31**(5) (2014), pp. 91–100.

[3] H. Hassanieh, P. Indyk, D. Katabi, and E. Price, *Near-optimal algorithm for sparse Fourier transform*, Proc. 44th annual ACM symposium on Theory of Computing, 2012, pp. 563–578.

[4] M.A. Iwen, Improved approximation guarantees for sublinear-time Fourier algorithms, Appl. Comput. Harmon. Anal. **34** (2013), 57–82.

[5] D. Lawlor, Y. Wang, and A. Christlieb, *Adaptive sub-linear time Fourier algorithms*, Adv. Adapt. Data Anal. **5**(1) (2013), 1350003.

[6] S. Pawar and K. Ramchandran, *Computing a k-sparse n-length discrete Fourier transform using at most $4k$ samples and $\mathcal{O}(k \log k)$ complexity*, IEEE International Symposium on Information Theory, 2013, pp. 464–468.

[7] G. Plonka, K. Wannenwetsch, A deterministic sparse FFT algorithm for vectors with small support, Numer. Alg., 2015, DOI: 10.1007/s11075-015-0028-0.

Multiscale basis dictionaries on graphs and some of their applications

Naoki Saito
(joint work with Jeff Irion)

1. Introduction

We previously introduced two multiscale transforms for signals on graphs: the *Hierarchical Graph Laplacian Eigen Transform* (HGLET) [1] and the *Generalized Haar-Walsh Transform* (GHWT) [2], both of which utilize a recursive partitioning of a graph to generate overcomplete dictionaries of orthonormal bases and corresponding expansion coefficients. The HGLET and GHWT can be viewed as generalizations of the hierarchical block Discrete Cosine Transform (DCT) and the Haar-Walsh Wavelet Packets, respectively, to the setting of signals on graphs.

Since these overcomplete dictionaries contain a huge number of possible orthonormal bases, it is important for us to be able to select the most suitable one (as well as the corresponding expansion coefficients of an input signal) for our task at hand. To do so, we generalized the classical best-basis algorithm [3] to the setting of our graph-based transforms. In [2], we have demonstrated how our transforms can be used to achieve good denoising of signals on graphs. Here, we apply our graph-based transforms to tackle two problems: simultaneous segmentation and denoising of classical 1-D signals and matrix data analysis.

2. Simultaneous Segmentation and Denoising of Classical 1-D Signals

Given a 1-D noisy signal sampled on regular grids, our goal here is three-fold: 1) to divide the signal into segments of similar characteristics; 2) to reduce the noise in the signal; and 3) to achieve better approximation and compression of the underlying signal. Our proposed method below can be viewed as a significant improvement over our previous attempt [4] where we used conventional tools without the graph setting. Here we fully utilize our graph-based transforms by viewing such a 1-D signal as data on a path graph. Doing so affords us more flexibility in our segmentation, as we no longer have to work within a dyadic constraint on the segment lengths.

The first step in our algorithm is to recursively partition the unweighted path graph. Next, we use the three HGLET variations (using the unnormalized, random-walk-normalized, and symmetrically-normalized graph Laplacian matrices [5]) to analyze each segment of the signal. We note that there is no need to compute the eigenvectors of these matrices because on unweighted path graphs they are known to be three different types of DCT, i.e., the DCT Type II, the weighted version of the DCT Type I, and the DCT Type I, respectively. From these three sets of expansion coefficients, we select a *hybrid* best basis using the minimum description length (MDL) criterion [6, 7] as our cost functional, which was also used in [4]. We note that we quantize the expansion coefficients with a threshold for denoising, but the appropriate quantization resolution and the threshold are automatically selected by the MDL criterion. In addition, the model parameters to be determined are: 1) the segmentation configuration of the signal (i.e., the set of disjoint intervals, which we also quantize via the levels list description method [8]) and 2) a flag to specify the HGLET variation used for each segment. Thus, by using the MDL cost functional to perform the best basis search, we are searching for the segmentation whose structure is choosable from the current partitioning tree that allows us to most efficiently approximate the underlying signal in the noisy input data.

The MDL-guided best basis search yields two outputs: a segmentation of the signal and the corresponding set of quantized expansion coefficients. Utilizing the segmentation obtained by the best basis algorithm, we modify the edge weights of the graph. The purpose of doing so is to encourage edges between regions of similar characteristics to be preserved in the next iteration and to encourage those edges between regions of different characteristics to be cut. Whereas we began with an unweighted path graph, we now have a weighted one. We then iterate this process. We generate a new recursive partitioning of the signal, which will differ from the previous recursive partitioning due to the modified edge weights. We analyze the signal again using the three HGLET variations although we treat the graph as being unweighted. This is because the purpose of modifying the edge weights is to influence the partitioning while preserving the relationship between the HGLET on a path graph and the block DCTs. As the recursive partitioning of the signal is different, the expansion coefficients will be different as well. We then find a new best basis and corresponding segmentation, and we modify the edge weights as before. We repeat this process until it converges to a particular basis, which gives us both a segmentation of the signal and a set of quantized and thresholded coefficients from which we can reconstruct the denoised signal. Empirically, we have observed that convergence occurs between 6 and 15 iterations for all the signals (both synthetic and real) that we have examined so far.

3. Matrix Data Analysis

Next, we want to analyze and efficiently approximate or denoise scrambled matrices that are typical for ratings/reviews databases using our graph-based tools. As in [9], our first step is to discover the underlying structure of the input matrix. To do this, we generate recursive partitioning trees on both the rows and the columns of the matrix based on appropriately defined affinities (e.g., the regularized inverse Euclidean distances) among rows and columns.

Equipped with these recursive partitionings, we analyze the matrix using the GHWT. Specifically, we first expand each column of the matrix into the GHWT best basis computed on the rows, and then we expand those coefficients into the GHWT best basis on the columns. Thus, our orthonormal best basis is the tensor product of the best basis on the row partitioning and the best basis on the column partitioning.

Using the same Science News database reported in [9] as an input data matrix, we compared the performance in data approximation of our GHWT best basis with that of the classical Haar wavelet basis and the classical Haar-Walsh wavelet packet best basis using the same sparsity criterion (i.e., ℓ^1-norm minimization) as in our GHWT best basis. Whereas the classical Haar wavelet basis is a non-adaptive transform, meaning that the basis is fixed, both the classical Haar-Walsh wavelet packet best basis and the GHWT best basis are data-adaptive. The difference between the latter two is that the classical Haar-Walsh wavelet packet dictionary considers only homogeneous dyadic partitions of rows and columns, whereas the GHWT does not impose such a constraint on row and column partitions. As such, the GHWT does a better job of capturing the underlying structure of the matrix and achieves better approximation than the classical transforms. With only 2% of the coefficients retained, the relative ℓ^2 errors were 16.3%, 15.6%, and 8.2% for the Haar basis, Haar-Walsh wavelet packet best basis, and GHWT best basis, respectively. Retaining 10% of the coefficients, those numbers became 4.5%, 4.4%, and 2.1%.

Acknowledgment

This research was partially supported by ONR grant N00014-12-1-0177 and NSF grant DMS-1418779, and was conducted with Government support under contract FA9550-11-C-0028 and awarded by the Department of Defense, Air Force Office of Scientific Research, National Defense Science and Engineering Graduate (NDSEG) Fellowship, 32 CFR 168a.

References

[1] J. Irion and N. Saito, *Hierarchical graph Laplacian eigen transforms*, JSIAM Letters, **6** (2014), 21–24.

[2] J. Irion and N. Saito, *The generalized Haar-Walsh transform*, in: Proc. 2014 IEEE Workshop on Statistical Signal Processing, pp. 472–475, 2014.

[3] R. R. Coifman and M. Wickerhauser, *Entropy-based algorithms for best basis selection*, IEEE Trans. Inform. Theory, **38** (1992), 713–718.

[4] N. Saito and E. Woei, *Simultaneous segmentation, compression, and denoising of signals using polyharmonic local sine transform and minimum description length criterion*, in: Proc. 13th IEEE Workshop on Statistical Signal Processing, pp. 315–320, 2005.

[5] U. von Luxburg, *A tutorial on spectral clustering*, Stat. Comput. **17**(4) (2007), 395–416.

[6] J. Rissanen, *Stochastic Complexity in Statistical Inquiry*, World Scientific, Singapore, 1989.

[7] P. D. Grünwald, *The Minimum Description Length Principle*, The MIT Press, Cambridge, MA, 2007.

[8] M. W. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software*, A K Peters, Ltd., Wellesley, MA, 1994.

[9] R. R. Coifman and M. Gavish, *Harmonic analysis of digital data bases*, in: Wavelets and Multiscale Analysis (J. Cohen and A. I. Zayed, eds.), Birkhäuser, pp. 161–197, 2011.

Dictionary learning - fast and dirty

Karin Schnass

In this talk we gave a small introduction to fast dictionary learning algorithms with local convergence guarantees. Many problems in signal/data processing and analysis can be efficiently solved if the signals are sparse in a dictionary. Therefore it is desirable to have an automated way to learn this sparsifying dictionary directly from a few training signals of the data class of interest, that is, a dictionary learning algorithm. To be more precise we have two requirements on a good dictionary learning algorithm, first that it is fast or computationally cheap and second that we have some guarantees that the algorithm will recover an underlying dictionary Φ if the data is known to be sparse in the dictionary. Currently there are mainly two promising directions. On one hand there are graph clustering algorithms and sums of squares methods, [2, 5, 6], which have global convergence guarantees but are computationally very costly. On the other hand there are (alternating) optimisation schemes, [8, 3, 10, 1, 11], which are computationally efficient, experimentally globally convergent but in the overcomplete case only have local convergence guarantees. For a recent solution to the basis learning case see [17], and for a more comprehensive introduction into dictionary learning see [12, 16]. The starting point for the probably most famous alternating optimisation dictionary learning algorithm - K-SVD, [3], is the following optimisation programme,

$$\min_{\Psi \in \mathcal{D}, X \in \mathcal{X}_S} \|Y - \Psi X\|_F^2, \tag{1}$$

where $Y = (y_1 \dots y_N)$ collects the N training signals on its columns, $\mathcal{D} := \{\Psi = (\psi_1 \dots \psi_K), \psi_k \in \mathbb{R}^d, \|\psi_k\|_2 = 1\}$ is the set of admissable dictionaries and $\mathcal{X}_S := \{X = (x_1 \dots x_N), x_n \in \mathbb{R}^K, \|x_n\|_0 \leq S\}$ the set of columnwise S-sparse coefficient matrices. If the data are generated from an S-sparse random coefficient model in combination with a unit norm tight frame Φ it can be shown that Φ is at/near a local minimiser of (1), [14]. In the special case S=1 this means that K-SVD locally converges to the generating dictionary. The fact that the K-SVD principle seems to have difficulties in recovering non-tight frames together with the fact that K-SVD itself is still quite computationally costly led to the development of a simpler optimisation criterion, which can be interpreted as a generalisation of the K-means criterion,

$$\max_{\Psi \in \mathcal{D}} \sum_n \max_I \|\Psi_I^\star y_n\|_1, \tag{2}$$

Given enough training samples, this criterion locally identifies also a non-tight dictionary Φ up to arbitrary precision, provided the sparsity level scales as $S = O(\mu^{-1})$, where μ is the coherence of Φ, that is, $\mu := \max_{i \neq j} |\langle \phi_i, \phi_j \rangle|$. This local identification property is stable for $S = O(\mu^{-2})$, non exact sparsity and large (random) noiselevels, [13]. With the criterion the following very simple alternating optimisation algorithm, called Iterative Thresholding and K (signal) Means, can be associated.

ITKsM Algorithm (one iteration)
Given an input dictionary Ψ and N training signals y_n do:

- For all n find $I^t_{\Psi,n} = \arg\max_{I:|I|=S} \|\Psi^\star_I y_n\|_1$.
- For all k calculate
$$\bar{\psi}_k = \frac{1}{N} \sum_{n:k\in I^t_{\Psi,n}} y_n \cdot \operatorname{sign}(\langle \psi_k, y_n\rangle).$$
- Output $\bar{\Psi} = (\bar{\psi}_1/\|\bar{\psi}_1\|_2, \dots, \bar{\psi}_K/\|\bar{\psi}_K\|_2)$.

The main advantage of ITKsM over K-SVD is that it is computationally much cheaper, of the order $O(dKN)$ corresponding to the matrix vector calculations $\Psi^\star y_n$. The signals can be processed sequentially, so the algorithm can be used online or in parallel. The disadvantage is that it is not globally convergent and with a random initialisation will only recover most but not all atoms of the generating dictionary.

To remedy this problem ITKsM was modified to use residual rather than signal means, leading to an algorithm called Iterative Thresholding and K (residual) Means.

ITKrM Algorithm (one iteration)
Given an input dictionary Ψ and N training signals y_n do:

- For all n find $I^t_{\Psi,n} = \arg\max_{I:|I|=S} \|\Psi^\star_I y_n\|_1$.
- For all k calculate
$$\bar{\psi}_k = \sum_{n:k\in I^t_n} \big[\mathbb{I} - P(\Phi_{I^t_n}) + P(\psi_k)\big] y_n \cdot \operatorname{sign}(\langle \psi_k, y_n\rangle),$$
where $P(M)$ denotes the orthogonal projection onto the rowspan of a matrix M.
- Output $\bar{\Psi} = (\bar{\psi}_1/\|\bar{\psi}_1\|_2, \dots, \bar{\psi}_K/\|\bar{\psi}_K\|_2)$.

ITKrM can be regarded as a hybrid between K-SVD and ITKsM, as it uses residuals like K-SVD and means like ITKsM. As a hybrid it inherits all the desirable properties of its parents. Given $O(K \log K \varepsilon^{-2})$ training samples it can locally recover a generating dictionary up to precision ε if the sparsity level scales as $S = O(\mu^{-1})$ and up to precision $\varepsilon = K^{-\ell}$ if the sparsity scales as $S = O(\mu^{-2}/(\ell \log K))$. It is still computationally very cheap $O(dN(K + S^2))$ and sequential and on top of that experimentally globally convergent, [15].

Another advantage of ITKrM as well as a currently investigated weighted variant, where $\operatorname{sign}(\langle \psi_k, y_n\rangle)$ is replaced by $\langle \psi_k, y_n\rangle$, is that it is very easy to incorporate additional information. Thus in ongoing work with V. Naumova we are currently investigating how to learn dictionaries from corrupted/missing data, which can for instance be applied to inpainting.

Several other questions are also under scrutiny at the moment. Is the generating dicitionary the global optimum of (1/2)? Are there efficient initialisation strategies for $S = O(\mu^{-2})$? (For $S = O(\mu^{-1})$ recently a polynomial time algorithm has been found, [4].) Can we extend the convergence radius of ITKsM resp. ITKrM from $O(1/\log K)$ resp. $O(1/\sqrt{S})$ to $O(1)$, as suggested by experiments? How do we transfer the local convergence proof for ITKM to a local convergence proof for K-SVD, especially in the non-tight but experimentally stable case. An interesting question that arises in this context is the calculation/estimation of the expectation

$$\mathbb{E}_{I:k\in I,|I|=S}\left[(\mathbb{I}-\Phi_I\Phi_I^\dagger)\Phi\Phi^\star\phi_k\right].$$

Finally to improve our results to signals with large dynamic range of the coefficients we need to redive into the investigation of performance guarantees for average case sparse approximation including stability under perturbation of the dictionary, the best candidates being iterative thresholding, [7], or hard thresholding pursuit, [9].

References

[1] A. Agarwal, A. Anandkumar, P. Jain, P. Netrapalli, and R. Tandon. Learning sparsely used overcomplete dictionaries via alternating minimization. *COLT 2014 (arXiv:1310.7991)*, 2014.

[2] A. Agarwal, A. Anandkumar, and P. Netrapalli. Exact recovery of sparsely used overcomplete dictionaries. *COLT 2014 (arXiv:1309.1952)*, 2014.

[3] M. Aharon, M. Elad, and A.M. Bruckstein. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing.*, 54(11):4311–4322, November 2006.

[4] S. Arora, R. Ge, T. Ma, and A. Moitra. Simple, efficient, and neural algorithms for sparse coding. *arXiv:1503.00778*, 2015.

[5] S. Arora, R. Ge, and A. Moitra. New algorithms for learning incoherent and overcomplete dictionaries. *COLT 2014 (arXiv:1308.6273)*, 2014.

[6] B. Barak, J.A. Kelner, and D. Steurer. Dictionary learning and tensor decomposition via the sum-of-squares method. *arXiv:1407.1543*, 2014.

[7] T. Blumensath and M.E. Davies. Iterative thresholding for sparse approximation. *Journal of Fourier Analysis and Applications*, 14(5-6):629–654, 2008.

[8] K. Engan, S.O. Aase, and J.H. Husoy. Method of optimal directions for frame design. In *ICASSP99*, volume 5, pages 2443–2446, 1999.

[9] S. Foucart. Hard thresholding pursuit: an algorithm for compressive sensing. *SIAM Journal on Numerical Analysis*, 49(6):2543–2563, 2011.

[10] R. Gribonval and K. Schnass. Dictionary identifiability - sparse matrix-factorisation via l_1-minimisation. *IEEE Transactions on Information Theory*, 56(7):3523–3539, July 2010.

[11] R. Jenatton, F. Bach, and R. Gribonval. Sparse and spurious: dictionary learning with noise and outliers. *arXiv:1407.5155*, 2014.

[12] R. Rubinstein, A. Bruckstein, and M. Elad. Dictionaries for sparse representation modeling. *Proceedings of the IEEE*, 98(6):1045–1057, 2010.

[13] K. Schnass. Local identification of overcomplete dictionaries. *accepted to Journal of Machine Learning Research (arXiv:1401.6354)*, 2014.

[14] K. Schnass. On the identifiability of overcomplete dictionaries via the minimisation principle underlying K-SVD. *Applied Computational Harmonic Analysis*, 37(3):464–491, 2014.

[15] K. Schnass. Convergence radius and sample complexity of ITKM algorithms for dictionary learning. *arXiv:1503.07027*, 2015.

[16] K. Schnass. A personal introduction to theoretical dictionary learning. *Internationale Mathematische Nachrichten*, 228:5–15, 2015.

[17] J. Sun, Q. Qu, and J. Wright. Complete dictionary recovery over the sphere. *arXiv:1504.06785*, 2015.

Dynamic optimal transport for RGB image processing

Gabriele Steidl
(joint work with Jan Henrik Fitschen, Friederike Laus)

In this paper we deal with the interpolation between two color images in the RGB space using the concept of dynamic optimal transport. We mention that the problem can be also tackled by other techniques such as metamorphoses, cf. [8] which are not within the scope of our paper. RGB images of size $N_1 \times N_2$ are usually given by three $N_1 \times N_2$ matrices with values in $\{0, \dots, 255\}$. To explain the idea behind dynamic optimal transport let us consider the images as continuous three-dimensional density functions f_0 and f_1 of probability measures $\mu_0 = f_0 dx$, $\mu_1 = f_1 dx$ which are absolutely continuous with respect to the Lebesgue measure. In particular, we have $\int_{\mathbb{R}^d} f_0 dx = \int_{\mathbb{R}^d} f_1 dx = 1$. We assume that the density functions are finitely supported in the two spatial dimensions. Due to the human color perception the functions should be periodic in the third dimension. More generally, let us consider d-dimensional distributions, where $d = 3$ in our specific application. We assume that the measures μ_1 and μ_2 belong to a Wasserstein space

$$\mathcal{P}_p(\mathbb{R}^d) := \{\mu \in \mathcal{P}(\mathbb{R}^d) : \int_{\mathbb{R}^d} |x|^p d\mu(x) < +\infty\}, \quad p \in [1, \infty)$$

equipped with a distance function, the so-called Wasserstein distance

$$W_p(\mu_0, \mu_1) := \min_{\nu \in \Pi(\mu_0, \mu_1)} \int_{\mathbb{R}^d} |x - y|^p \, d\nu(x, y).$$

Indeed, the joint probability measure ν which minimizes the Wasserstein distance exists for $p \in [1, \infty)$ and is uniquely determined for $p > 1$. It is called optimal transport map between μ_0 and μ_1. Wasserstein spaces $(\mathcal{P}_p, W_p)$ are geodesic space. In particular there exists for any $\mu_0, \mu_1 \in \mathcal{P}_p(\mathbb{R}^d)$ a geodesic $\gamma : [0, 1] \to \mathcal{P}_p(\mathbb{R}^d)$ with $\gamma(0) = \mu_0$, $\gamma(1) = \mu_1$. For interpolating our images we ask for $\mu_t = \gamma(t)$, $t \in [0, 1]$.

At least theoretically there are several ways to compute μ_t. If the optimal transport map ν is known, then $\mu_t = \mathcal{L}_{t\#}\nu := \nu \circ \mathcal{L}_t^{-1}$, where $\mathcal{L}_t : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ is the linear interpolation map $\mathcal{L}_t(x, y) := (1 - t)x + ty$. This requires the knowledge of the optimal transport map and of $\mathcal{L}_t^{-1}$. At present there are efficient ways for computing optimal transport map for one-dimensional distributions by an ordering procedure and for Gaussian distributions in the case $p = 2$. For $p = 2$ one can also use the fact that ν is indeed induced by a transport plan $T : \mathbb{R}^d \to \mathbb{R}^d$, i.e., $\nu = (\mathrm{id}, T)$ having a potential ψ, i.e., $T = \nabla \psi$ which fulfills the Monge-Ampere equation. However, this is a second order nonlinear elliptic PDE which is numerically hard to solve. For the analysis of the Monge-Ampere equation we refer to papers of Caffarelli, e.g. [3]. Moreover there exit numerical results for the (simplified) semi-geostrophic equation. For other numerical approaches to compute optimal transport maps, see, e.g., [1, 5]. Another approach relaxes the condition of minimizing a Wasserstein distance by using instead an entropy regularized Wasserstein distance. Such distances can be computed more efficiently by the Sinkhorn algorithm and were applied within a barycentric approach by Cuturi et al. [4].

In the talk we will apply the fluid dynamic formulation of the dynamic optimal transport problem for $p \in (1,2]$, i.e., we minimize:

$$\int_0^1 \int_{\mathbb{R}^d} \frac{1}{p} |\mathrm{v}(t,x)|^p f(t,x)\, dx dt$$

subject to

$$\begin{aligned} \partial_t f(t,x) + \nabla_x \cdot (\mathrm{v}(t,x) f(t,x)) &= 0, \\ f(0,\cdot) = f_0,\ f(1,\cdot) &= f_1. \end{aligned}$$

Further we have to suppose $\cup_{t\in[0,1]} \mathrm{supp} f(t,\cdot) \subseteq [0,1]^d$ with appropriate boundary conditions. Benamou and Brenier [2] suggested to substitute the momentum $m = fv$ which makes the problem convex in f, m. We provide a discrete model based on a staggered grid discretization as it was also proposed by Papadakis et al. [7] for $p = 2$. In particular, we provide a sound matrix-vector notation of the problem by using the tensor product notation. Moreover, we modify the model by penalizing the continuity constraint:

Constrained Transport Problem:

$$\begin{aligned} &\mathrm{argmin}_{m,f,u,v} \| J_p(u,v) \|_1, \\ &\text{subject to} \quad S_{\mathrm{m}} m = u, \quad S_{\mathrm{f}} f + f^+ = v, \\ &\qquad (D_{\mathrm{m}} | D_{\mathrm{f}}) \begin{pmatrix} m \\ f \end{pmatrix} = f^-, \quad f \geq 0. \end{aligned}$$

Penalized Transport Problem ($\lambda > 0$):

$$\begin{aligned} &\mathrm{argmin}_{m,f,u,v} \left\{ \| J_p(u,v) \|_1 + \lambda \| (D_{\mathrm{m}} | D_{\mathrm{f}}) \begin{pmatrix} m \\ f \end{pmatrix} - f^- \|_2^2 \right\}, \\ &\text{subject to} \quad S_{\mathrm{m}} m = u, \quad S_{\mathrm{f}} f + f^+ = v, \quad f \geq 0. \end{aligned}$$

Here D_* are appropriate difference matrices and S_* are averaging matrices. Moreover,

$$J_p(x,y) := \begin{cases} \frac{1}{p} \frac{|x|^p}{y^{p-1}} & \text{if } y > 0, \\ 0 & \text{if } (x,y) = (0,0), \\ +\infty & \text{otherwise} \end{cases}$$

and $|x| := (\sum_{i=1}^d x_i^2)^{\frac{1}{2}}$.

We suggest to solve the minimization problem by primal-dual minimization algorithms. It turns out that one step of the algorithm requires the solution of a 4D Poisson equation which includes simultaneously zero-, mirror- and periodic boundary conditions. This can be efficiently realized by fast Sine-, Cosine- and Fourier transforms. Another step involves the finding of a somehow uniquely determined root of a function which by a Newton method. With an appropriately defined initialization quadratic convergence the method is ensured. We provide interesting numerical results which demonstrate the good performance of our algorithms.

References

[1] S. Angenent, S. Haker, and A. Tannenbaum. Minimizing flows for the Monge-Kantorovich problem. *SIAM Journal of Mathematical Analysis*, 35:61–97, 2003.

[2] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

[3] L. A. Caffarelli. Interior $W^{2,p}$ estimates for solutions of the Monge-Ampère equation. *Annals of Mathematics*, 131(1):135–150, 1990.

[4] M. Cuturi and A. Coucet. Fast computation of Wasserstein barycenters. *Proceedings of the 31st International Conference on Machine Learning*, 32, 2014.

[5] E. Haber, T. Rehman, and A. Tannenbaum. An efficient numerical method for the solution of the l_2 optimal mass transfer problem. *SIAN Journal of Scientific Computing*, 32(1):197–211, 2010.

[6] J. Maas, M. Rumpf, C.-B. Schönlieb, and S. Simon. A generalized model for optimal transport of images including dissipation and density modulation. *Preprint*, 2014.

[7] N. Papadakis, G. Peyré, and E. Oudet. Optimal transport with proximal splitting. *SIAM Journal on Imaging Sciences*, 7(1):212–238, 2014.

[8] A. Trouvé and L. Younes. Metamorphoses through Lie group action. *Foundations of Computational Mathematics*, 5(2):173–198, 2005.

Nonlinear phase unwinding of functions

Stefan Steinerberger
(joint work with Ronald R. Coifman)

Blaschke factorization is a classical tool in complex analysis: any ‘sufficiently nice’ holomorphic function $F : \mathbb{C} \to \mathbb{C}$ can be written as $F = B \cdot G$, where G has no roots inside the unit disk $\mathbb{D}$ and B is a Blaschke product, i.e. of the form

$$B(z) = z^m \prod_j \frac{z - \alpha_j}{1 - \overline{\alpha_j} z}$$

for some $m \in \mathbb{N}$ and $\alpha_i \in \mathbb{D}$. Note that

$$|B(z)| = 1 \qquad \text{whenever } |z| = 1$$

and hence $|F(z)| = |G(z)|$ on $\partial\mathbb{D}$. A rough interpretation would be that

$$B \sim \text{phase} \qquad \text{and} \qquad G \sim \text{amplitude}.$$

In the mid-1990s it was observed by one of the authors (Coifman) that Blaschke factorization could be iteratively applied if we subtract a suitable constant (i.e. the value at the origin) after each step. This gives rise to a nonlinear analogue of Fourier series, an unwinding series of the form

$$F \sim \alpha_0 B_0 + \alpha_1 B_0 B_1 + \alpha_2 B_0 B_1 B_2 + \dots$$

Fast computation of the Blaschke factorization can be accomplished without explicitely computing the roots via a method first discussed by Guido & Mary Weiss [13] in 1962. Extensive numerical investigation by Michel Nahon [6] suggested

- that the formal series converges
- that this seems to (at least generically) happen with an exponential rate
- and that the method is very stable.

The method was further studied by Letelier & Saito [5] who applied it to underwater acoustics problems and studied regularization methods and Healy [3, 4]. Recently, the method has been independently discovered by Tao Qian and collaborators [7, 8, 9, 10, 11, 12]. However, so far the convergence of the purely formal series has not been studied from a rigorous mathematical viewpoint.

Our main result in the most general form is as follows. Let $0 = \gamma_0 \leq \gamma_1 \leq \dots$ be an arbitrary monotonically increasing sequence of real numbers and let X be the subspace of $L^2(\mathbb{T})$ for which

$$\left\| \sum_{n \geq 0} a_n z^n \right\|_X^2 := \sum_{n \geq 0} \gamma_n |a_n|^2 < \infty.$$

We define a norm Y (merely a semi-norm whenever γ is not strictly increasing)

$$\left\| \sum_{n \geq 0} a_n z^n \right\|_Y^2 := \sum_{n \geq 0} (\gamma_{n+1} - \gamma_n) |a_n|^2.$$

Our main statement is that the Blaschke factorization acts nicely on these spaces. The first part of our statement is known (being ascribed to Digital Signal Processing in [8]) and can be equivalently phrased as follows: given a Blaschke decomposition $F = B \cdot G$ and assuming both functions are expanded into a Fourier series

$$F(z) = \sum_{n=0}^{\infty} f_n z^n \qquad \text{and} \qquad G(z) = \sum_{n=0}^{\infty} g_n z^n,$$

then, for every $N \in \mathbb{N}$

$$\sum_{n \geq N}^{\infty} |g_n|^2 \leq \sum_{n \geq N}^{\infty} |f_n|^2.$$

Phrased differently, inner outer factorization shifts the energy to lower frequencies in a strictly monotonous way. Our main tool will be a refinement of that inequality.

Theorem 1. *If $F \in \mathcal{H}^2$ has a Blaschke factorization $F = B \cdot G$, then*

$$\|G(e^{i\cdot})\|_X \leq \|F(e^{i\cdot})\|_X.$$

Moreover, if $F(\alpha) = 0$ for some $\alpha \in \mathbb{D}$, we even have

$$\|G(e^{i\cdot})\|_X^2 \leq \|F(e^{i\cdot})\|_X^2 - (1 - |\alpha|^2) \|G(e^{i\cdot})\|_Y^2.$$

An iterative application combined with telescoping yields that the sequence converges in Y for initial values in X. The case $\gamma_n = n$ recovers the Dirichlet space $\mathcal{D}$ and allows to reprove a formula of Carleson [1]. Our recent paper [2] also discuss other new phenomena, for example the following curious stability property: when doing Blaschke factorization $F = BG$ numerically, we will introduce some roundoff errors; even though we never actually compute the roots of the functions, this roundoff error can be imagined as perturbing the roots a little bit. We have the following curious and purely algebraic *pointwise* stability statement.

Theorem 2. *Suppose $F_1, F_2 : \mathbb{C} \to \mathbb{C}$ are polynomials having the same roots outside of $\mathbb{D}$ and the same number of roots inside $\mathbb{D}$. Then the Blaschke factorizations*

$$F_1 = B_1 G_1 \qquad \text{and} \qquad F_2 = B_2 G_2,$$

satisfy

$$|G_1(z) - G_2(z)| = |F_1(z) - F_2(z)| \qquad \text{for all} \quad z \in \partial\mathbb{D}.$$

References

[1] L. Carleson, A representation formula for the Dirichlet integral. Math. Z. 73 1960 190-196.

[2] Ronald R. Coifman and S. Steinerberger, Nonlinear Phase Unwinding of Functions, arXiv:1508.01241

[3] D. Healy Jr., Multi-Resolution Phase, Modulation, Doppler Ultrasound Velocimetry, and other Trendy Stuff, talk, slides via personal communication

[4] D. Healy, Phase analysis, Talk given at the University of Maryland, slides as private communication

[5] J. Letelier and N. Saito, Amplitude and Phase Factorization of Signals via Blaschke Product and Its Applications, talk given at JSIAM09, HTTPS://WWW.MATH.UCDAVIS.EDU/ SAITO/TALKS/JSIAM09.PDF

[6] M. Nahon, *Phase Evaluation and Segmentation*, PhD Thesis, Yale, 2000.

[7] W. Mi, T. Qian and F. Wan, A Fast Adaptive Model Reduction Method Based on Takenaka-Malmquist Systems, Systems & Control Letters. Volume 61, Issue 1, January 2012, Pages 223–230.

[8] T. Qian, Adaptive Fourier Decomposition, Rational Approximation, Part 1:Theory, invited to be included in a special issue of International Journal of Wavelets, Multiresolution and Information Processing.

[9] T. Qian, I. T. Ho, I. T. Leong and Y. B. Wang, Adaptive decomposition of functions into pieces of non-negative instantaneous frequencies, International Journal of Wavelets, Multiresolution and Information Processing, 8 (2010), no. 5, 813-833.

[10] T. Qian, L.H. Tan and Y.B. Wang, Adaptive Decomposition by Weighted Inner Functions: A Generalization of Fourier Serie, Journal of Fourier Analysis and Applications, 2011, 17(2): 175-190.

[11] T. Qian and L. Zhang, Mathematical theory of signal analysis vs. complex analysis method of harmonic analysis, Appl. Math. J. Chinese Univ, 2013, 28(4): 505-530.

[12] T. Qian, L. Zhang and Z. Li, Algorithm of Adaptive Fourier Decomposition, IEEE Transactions on Signal Processing, Issue Date: Dec. 2011 Volume: 59 Issue:12 On page(s): 5899 - 5906.

[13] G. Weiss and M. Weiss, A derivation of the main results of the theory of H^p-spaces. Rev. Un. Mat. Argentina (20) 63–71, 1962.

Tensor theta norms

Željka Stojanac
(joint work with Holger Rauhut)

We are interested in the problem of low rank tensor recovery via small number of measurements. In other words, we want to recover a low rank dth order tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ from a linear measurement map $\mathbf{\Phi} : \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d} \to \mathbb{R}^m$ and measurement vector $\mathbf{b} \in \mathbb{R}^m$, where $\mathbf{b} = \mathbf{\Phi}(\mathbf{X})$ and $m \ll n_1 n_2 \cdots n_d$. We consider a generalization of the matrix singular value decomposition called canonical decomposition (or CP-decomposition) and the corresponding notion of rank and norm (tensor nuclear norm). A dth order tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is a rank one tensor if and only if there exist d vectors $\mathbf{u}_j \in \mathbb{R}^{n_j}$, for $j \in [d] = \{1, 2, \ldots, d\}$ such that

$$\mathbf{X}(i_1, i_2, \ldots, i_d) = \mathbf{u}_1(i_1)\,\mathbf{u}_2(i_2) \cdots \mathbf{u}_d(i_d), \quad \text{for all } i_p \in [n_p],\, p \in [d].$$

The rank of a dth order tensor is the smallest number of rank one tensors that sum up to the original tensor. The tensor nuclear norm is a generalization of the matrix nuclear/trace norm, i.e., for a tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$

$$\|\mathbf{X}\|_* = \inf\Big\{ \sum_{k=1}^{r} |c_k| : \mathbf{X} = \sum_{k=1}^{r} c_k\, \mathbf{u}_1^k \otimes \mathbf{u}_2^k \otimes \cdots \otimes \mathbf{u}_d^k, \\ r \in \mathbb{N},\ \|\mathbf{u}_i^k\|_{\ell_2} = 1,\ \text{for } i \in [d], k \in [r] \Big\}.$$

Unfortunately, the set of rank-r tensors (for $r > 1$) is not closed and thus determining the rank as well as the nuclear norm of a given tensor is in general NP-hard, see [4, 5].

To tackle this problem, we suggest an approach based on theta bodies which were recently introduced in real algebraic geometry. As a result, we obtain new tensor norms (called theta tensor norms) that can be computed via semidefinite programming. This idea was first proposed in paper [2].

Next, we explain the idea behind the theta bodies. In the following, $\mathbb{R}[\mathbf{x}] = \mathbb{R}[x_1, x_2, \ldots, x_n]$ denotes the set of all real polynomials in variables $x_1, x_2, \ldots, x_n$ and $\mathbb{R}[\mathbf{x}]_k$ denotes the set of all real polynomials of degree at most k in the same variables. The central problem in optimization is finding a maximum of a linear functional over a given set $\mathcal{S}$, i.e., solving the problem

$$\max_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle \ \text{ s.t. } \mathbf{x} \in \mathcal{S} \tag{1}$$

which is equivalent to solving

$$\max_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle \ \text{ s.t. } \mathbf{x} \in \overline{\operatorname{conv}(\mathcal{S})},$$

where $\overline{\operatorname{conv}(\mathcal{S})}$ denotes the closure of the convex hull of the set $\mathcal{S}$. For example, in linear programming the set $\mathcal{S}$ is of the form $\mathcal{S} = \{\mathbf{x} : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$. We are interested in the case where the set $\mathcal{S}$ is the real algebraic variety of the polynomial ideal $I \in \mathbb{R}[\mathbf{x}]$, i.e., in the set

$$\nu_{\mathbb{R}}(I) = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = 0, \text{ for all } f \in I\}.$$

Recall, the set $\overline{\operatorname{conv}(\nu_{\mathbb{R}}(I))}$ can be described as

$$\overline{\operatorname{conv}(\nu_{\mathbb{R}}(I))} = \left\{\mathbf{x} \in \mathbb{R}^n : \ell(\mathbf{x}) \geq 0, \text{ for all } \ell \text{ affine s.t. } \ell\,|_{\nu_{\mathbb{R}}(I)} \geq 0\right\}.$$

However, checking only for one fixed affine polynomial ℓ whether it is nonnegative on the set $\mathcal{S} = \nu_{\mathbb{R}}(I)$ can be a tedious task. The idea is to find a relaxation $\mathcal{T}$ of the set $\overline{\operatorname{conv}(\nu_{\mathbb{R}}(I))}$ such that the corresponding optimization problem (1) over the set $\mathcal{T}$ (instead of $\mathcal{S}$) can be solved via semidefinite programming. Rather than considering all affine polynomials ℓ which are nonnegative on the set $\mathcal{S}$, we restrict our search to its subset. That is, we consider only affine polynomials ℓ which can be written as

$$\ell(\mathbf{x}) = \sum_{i=1}^{t} h_i^2(\mathbf{x}) + g(\mathbf{x}), \quad h_i \in \mathbb{R}[\mathbf{x}], g \in I, t \in \mathbb{N}. \tag{2}$$

Clearly, every polynomial ℓ defined as in (2) is nonnegative on the set $\mathcal{S} = \nu_{\mathbb{R}}(I)$ since $h_i^2(\mathbf{x}) \geq 0$ and $g(\mathbf{x}) = 0$, for every $\mathbf{x} \in \nu_{\mathbb{R}}(I)$. Theta bodies form a hierarchy of sets and were introduced first by Lovász in [6] and later analyzed in [3]. The k-th theta body takes into account only the affine polynomials ℓ which are k-sos[1] mod I, i.e., affine polynomials ℓ as in (2) with

$$\deg(h_i) \leq k, \text{ for all } i \in [t], \text{ and } g \in I.$$

It results in the following definition of the k-th theta body (for $k \in \mathbb{N}$)

$$\mathrm{TH}_k(I) := \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \geq 0, \text{for every } f \text{ affine and } k\text{-sos mod I}\}.$$

Theta bodies are closed, convex sets and satisfy the following nestedness property

$$\mathrm{TH}_1(I) \supseteq \mathrm{TH}_2(I) \supseteq \cdots \supseteq \overline{\operatorname{conv}(\nu_{\mathbb{R}}(I))}.$$

The idea is to define an ideal I such that

$$\nu_{\mathbb{R}}(I) = \{\text{all rank-one, Frobenius norm-one tensors}\}. \tag{3}$$

Then, for every $k \in \mathbb{N}$, the k-th theta body defines a new tensor unit θ_k-norm ball which is a relaxation of the tensor unit nuclear norm ball and can be computed for a fixed tensor $\mathbf{X}$ via

$$\|\mathbf{X}\|_{\theta_k} = \left\{\inf_t t \text{ s.t. } \mathbf{X} \in t\,\mathrm{TH}_k(I)\right\}.$$

Computing theta norms relies heavily on computing a Groebner basis of the corresponding ideal I. We have used the so called grevlex ordering (graded reverse lexicographic ordering, see [1]) and after some tedious computations, we end up with the combinatorial moment matrix $\mathbf{M}_{\mathcal{B}_k}(\mathbf{X}, \mathbf{y})$, see [7] for details.

[1] sos = sum of squares

Given the combinatorial moment matrix $\mathbf{M}_{\mathcal{B}_k}(\mathbf{X},\mathbf{y})$, computing the θ_k-norm of a given tensor $\mathbf{X}$ is given by the semidefinite program

$$\min_{\mathbf{y}} t \quad \text{subject to} \quad \mathbf{M}_{\mathcal{B}_k}(\mathbf{X},\mathbf{y}) \succeq 0, y_0 = t$$

and the tensor recovery via θ_k-norm minimization is given by the following semidefinite program

$$\arg\min_{\mathbf{y},\mathbf{Z}} t \quad \text{subject to } \mathbf{M}_{\mathcal{B}_k}(\mathbf{Z},\mathbf{y}) \succeq 0,\ y_0 = t \quad \text{and} \quad \mathbf{\Phi}(\mathbf{Z}) = \mathbf{b}.$$

For a matrix case, we need to find an ideal I_M such that its real algebraic variety is $\nu_{\mathbb{R}}(I_M) = \{\mathbf{X} \in \mathbb{R}^{n_1 \times n_2} : \operatorname{rank}(\mathbf{X}) = 1, \|\mathbf{X}\|_F = 1\}$. Here, we work in the set of all real polynomials in variables which correspond to the entries of a matrix $\mathbf{X}$ denoted as $\mathbb{R}[\mathbf{x}] = \mathbb{R}[X_{11}, X_{12}, \ldots, X_{n_1 n_2}]$. Since a matrix is rank one if and only if all its 2×2 minors vanish, we define the corresponding ideal I_M through its basis B_M

$$B_M = \left\{ \bigcup_{i<k,j<l} \{X_{il}X_{kj} - X_{ij}X_{kl}\} \cup \left\{ \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} X_{ij}^2 - 1 \right\} \right\}.$$

In this case, all θ_k-norms are equal to the matrix nuclear norm.

For a third order tensor case, we define the polynomial ideal $I \in \mathbb{R}[\mathbf{x}] = \mathbb{R}[X_{111}, X_{112}, \ldots, X_{n_1 n_2 n_3}]$ satisfying (3) through its reduced Groebner basis B

$$\begin{aligned} B = \Big\{ f_1^{ijk\hat{i}\hat{j}\hat{k}} &= -X_{ijk}X_{\hat{i}\hat{j}\hat{k}} + X_{ij\hat{k}}X_{\hat{i}\hat{j}k}, i < \hat{i}, j \le \hat{j}, k < \hat{k}, \\ f_2^{ijk\hat{i}\hat{j}\hat{k}} &= -X_{ijk}X_{\hat{i}\hat{j}\hat{k}} + X_{i\hat{j}k}X_{\hat{i}j\hat{k}}, i \le \hat{i}, j < \hat{j}, k < \hat{k}, \\ f_3^{ijk\hat{i}\hat{j}\hat{k}} &= -X_{ijk}X_{\hat{i}\hat{j}\hat{k}} + X_{i\hat{j}\hat{k}}X_{\hat{i}jk}, i < \hat{i}, j < \hat{j}, k \le \hat{k}, \\ g &= \sum_{i,j,k} X_{ijk}^2 - 1,\ f_1^{ijk\hat{i}\hat{j}\hat{k}}, f_2^{ijk\hat{i}\hat{j}\hat{k}}, f_3^{ijk\hat{i}\hat{j}\hat{k}}, g \in \mathbb{R}[\mathbf{X}] \Big\}. \end{aligned}$$

Every matricization of a rank one tensor is a rank one matrix. Thus, the ideal I should contain all 2×2 minors of every matricization. Since B is the reduced Groebner basis of I, it contains only the subset of all 2×2 minors. However, every other 2×2 minor not contained in B can be obtained as a difference of two polynomials in B and thus is contained in the ideal I.

Finally, in Table 1 we present some numerical results for third order tensor recovery via θ_1-norm minimization from a random Gaussian measurement map $\mathbf{\Phi} : \mathbb{R}^{n_1 \times n_2 \times n_3} \to \mathbb{R}^m$. For fixed tensor dimensions, rank and number of measurements m we performed 200 trials. We say that a tensor $\mathbf{X}$ is recovered if the entry-wise difference between the original tensor $\mathbf{X}$ and the tensor $\mathbf{X}^* = \arg\min_{\mathbf{Z}:\mathbf{\Phi}(\mathbf{Z})=\mathbf{\Phi}(\mathbf{X})} \|\mathbf{Z}\|_{\theta_1}$ is at most 10^{-6}. With $m_{\max}$ we denote the maximal number of measurements m for which we did not manage to recover any out of 200 tensors and with $m_{\min}$ we denote the minimal number of measurements m for which we managed to recover all 200 tensors. The last column contains the

tensor dimensions	rank	$m_{\max}$	$m_{\min}$	deg. of freedom
$2 \times 2 \times 3$	1	4	12	12
$3 \times 3 \times 3$	1	7	21	27
$3 \times 4 \times 5$	1	10	31	60
$4 \times 4 \times 4$	1	12	34	64
$4 \times 5 \times 6$	1	18	42	120
$5 \times 5 \times 5$	1	18	43	125
$3 \times 4 \times 5$	2	38	47	60
$4 \times 4 \times 4$	2	31	51	64
$4 \times 5 \times 6$	2	41	85	120

Table 1. Numerical results of third order tensor recovery

minimal number of independent measurements which would always be enough for tensor recovery.

References

[1] D. A. Cox, J. Little, D. O'Shea. Ideals, Varieties and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra. *Springer*, 1991.

[2] V. Chandrasekaran, B. Recht, P. A. Parrilo, A. S. Willsky. The Convex Geometry of Linear Inverse Problems. *Foundations of Computational Mathematics*, 12(6):805–849, 2012.

[3] J. Gouveia, P. A. Parrilo, R. R. Thomas. Theta Bodies for Polynomial Ideals. *SIAM Journal on Optimization*, 20(4):2097–2118, 2010.

[4] J. Håstad. Tensor rank is NP-complete. J. Algorithms, 11(4):644–654, 1990.

[5] C. Hillar, L.-H. Lim. Most tensor problems are NP-hard. Journal of the ACM, 60(6):45:1–45:39, 2013.

[6] L. Lovász. On the Shannon capacity of a graph. IEEE Trans. Inform. Theory, 25(1):1–7, 1979.

[7] H. Rauhut, Ž. Stojanac. Tensor theta norms and low rank recovery. arXiv:1505.05175, 2015.

Compressed sensing meets harmonic analysis

Vladimir Temlyakov

A generic problem of mathematical and numerical analysis is to approximately represent a given function. It is a classical problem that goes back to the first results on Taylor's and Fourier's expansions of a function.

The first step in solving the representation problem is to choose a representation system. Traditionally, a representation system has natural features such as minimality, orthogonality, simple structure and nice computational characteristics. The most typical representation systems are the trigonometric system $\{e^{ikx}\}$, the algebraic system $\{x^k\}$, the spline system, the wavelet system and their multivariate versions. In general we may speak of a basis $\Psi = \{\psi_k\}_{k=1}^{\infty}$ in a Banach space X.

The second step in solving the representation problem is to choose a form of an approximant that is built on the base of the chosen representation system Ψ. In a classical way that was used for centuries, an approximant a_m is a polynomial with respect to Ψ:

$$a_m := \sum_{k=1}^{m} c_k \psi_k. \tag{1}$$

The complexity of the approximant a_m is characterized by the order m of the polynomial. It is well known in approximation theory that approximation by polynomials is closely related to smoothness properties of a function being approximated. Approximation of this type is referred to as *linear approximation theory* because, for a fixed m, approximants come from a linear subspace spanned by $\psi_1, \dots, \psi_m$.

It was understood in numerical analysis and approximation theory that in many problems from signal/image processing it is more beneficial to use an m-term approximant with respect to Ψ than a polynomial of order m. This means that for $f \in X$ we look for an approximant of the form

$$a_m(f) := \sum_{k \in \Lambda(f)} c_k \psi_k \tag{2}$$

where $\Lambda(f)$ is a set of m indices which is determined by f. The complexity of this approximant is characterized by the cardinality $|\Lambda(f)| = m$ of $\Lambda(f)$. Approximation of this type is referred to as *nonlinear approximation theory* because, for a fixed m, approximants $a_m(f)$ come from different linear subspaces spanned by ψ_k, $k \in \Lambda(f)$, which depend on f. The cardinality $|\Lambda(f)|$ is a fundamental characteristic of $a_m(f)$ called *sparsity* of $a_m(f)$ with respect to Ψ. It is now well understood that we need to study nonlinear sparse representations in order to significantly increase our ability to process (compress, denoise, etc.) large data sets. Sparse representations of a function are not only a powerful analytic tool but they are utilized in many applications in image/signal processing and numerical computation.

The third step in solving the representation problem is to choose a method of construction of an approximant of desired form. The fundamental question of nonlinear approximation is how to devise good constructive methods (algorithms) of approximation. This problem has two levels of nonlinearity. The first level of nonlinearity is m-term approximation with regard to bases. In this problem one can use the unique function expansion with regard to a given basis to build an approximant. Nonlinearity enters by looking for m-term approximants with terms (i.e. basis elements in approximant) allowed to depend on a given function. On the second level of nonlinearity, we replace a basis by a more general system which is not necessarily minimal (for example, redundant system, dictionary). This setting is much more complicated than the first one (bases case), however, there is a solid justification of importance of redundant systems in both theoretical questions and in practical applications.

Recent results have established that greedy type algorithms are suitable methods of nonlinear approximation in both m-term approximation with regard to bases and m-term approximation with regard to redundant systems. It turns out that there is one fundamental principal that allows us to build good algorithms both for arbitrary redundant systems and for very simple well structured bases like the Haar basis. This principal is the use of a greedy step in searching for a new element to be added to a given m-term approximant. By a *greedy step*, we mean one which maximizes a certain functional determined by information from the previous steps of the algorithm. We obtain different types of greedy algorithms by varying the above mentioned functional and also by using different ways of constructing (choosing coefficients of the linear combination) the m-term approximant from the already found m elements of the dictionary.

In the case of nonlinear approximation with respect to a basis the *Thresholding Greedy Algorithm* is the simplest and the most studied one. The following question is very natural and fundamental. Which bases are suitable for the use of the Thresholding Greedy Algorithm (TGA)? Answering this question researchers introduced several new concepts of bases of a Banach space X: *greedy bases*, *quasi-greedy bases*, *almost greedy bases*. The greedy bases are the best for application of the TGA for sparse approximation – for any $f \in X$ the TGA provides after m iterations approximation with the error of the same order as the best m-term approximation of f. If a basis Ψ is a quasi-greedy basis then it merely guarantees that for any $f \in X$ the TGA provides approximants that converge to f but does not guarantee the rate of convergence. It turns out that the wavelet type bases are very good for the TGA. However, it is known that the TGA does not work well for the trigonometric system.

It was discovered recently, that the Weak Chebyshev Greedy Algorithm (WCGA) works much better than the TGA for the trigonometric system. We discuss and compare approximation by the TGA and the WCGA. We present some Lebesgue-type inequalities for the Weak Chebyshev Greedy Algorithm. The main message of the talk is that it is time to conduct a deep and thorough study of the WCGA with respect to bases in a style of the corresponding study of the TGA.

Non-asymptotic analysis of ℓ_1-SVM

Jan Vybíral

(joint work with Anton Kolleck – TU Berlin)

Support vector machines (SVM) are a group of popular classification methods in machine learning. Their input is a set of data points $x_1, \ldots, x_m \in \mathbb{R}^d$, each equipped with a label $y_i \in \{-1, +1\}$, which assigns each of the data points to one of two groups. SVM aims for binary linear classification based on separating hyperplane between the two groups of training data, choosing a hyperplane with separating gap as large as possible.

Since their introduction by Vapnik and Chervonenkis [9], the subject of SVM was studied intensively. We will concentrate on the so-called soft margin SVM [2], which allow also for misclassification of the training data.

In its most common form (and neglecting the bias term), the soft-margin SVM is a convex optimization program

$$\min_{\substack{w\in\mathbb{R}^d\\ \xi\in\mathbb{R}^m}} \frac{1}{2}\|w\|_2^2+\lambda\sum_{i=1}^m \xi_i \quad \text{subject to} \quad y_i\langle x_i,w\rangle \ge 1-\xi_i \quad \text{and} \quad \xi_i\ge 0 \tag{1}$$

for some tradeoff parameter $\lambda>0$ and so called slack variables ξ_i. It will be more convenient for us to work with the following equivalent reformulation of (1)

$$\min_{w\in\mathbb{R}^d} \sum_{i=1}^m [1-y_i\langle x_i,w\rangle]_+ \quad \text{subject to} \quad \|w\|_2\le R, \tag{2}$$

where $R>0$ gives the restriction on the size of w. We refer to monographs [7, 10, 11] and references therein for more details on SVM and to [4, Chapter B.5] and [3, Chapter 9] for a detailed discussion on dual formulations.

As the classical SVM (1) and (2) do not use any pre-knowledge about w, one typically needs to have more training data than the underlying dimension of the problem, i.e. $m\gg d$. Especially in analysis of high-dimensional data, this is usually not realistic and we typically deal with much less training data, i.e. with $m\ll d$. On the other hand, we can often assume some structural assumptions on w, in the most simple case that it is *sparse*, i.e. that most of its coordinates are zero. Motivated by the success of LASSO [8] in sparse linear regression, it was proposed in [1] that replacing the ℓ_2-norm $\|w\|_2$ in (2) by its ℓ_1-norm $\|w\|_1=\sum_{j=1}^d|w_j|$ leads to sparse classifiers $w\in\mathbb{R}^d$. This method was further popularized in [12] by Zhu, Rosset, Hastie, and Tibshirani, who developed an algorithm that efficiently computes the whole solution path (i.e. the solutions of (2) for a wide range of parameters $R>0$).

ℓ_1-SVM (and its variants) found numerous applications in high-dimensional data analysis, most notably in bioinformatics for gene selection and microarray classification. Finally, ℓ_1-SVM's are closely related to other popular methods of data analysis, like elastic nets, or sparse principal components analysis.

For the non-asymptotic analysis of ℓ_1-SVM, we shall make the following

Standing assumptions:

(i) $a \in \mathbb{R}^d$ is the true (nearly) sparse classifier with $\|a\|_2 = 1, \quad \|a\|_1 \le R$, $R \ge 1$, which we want to approximate;

(ii) $x_i = r\tilde{x}_i, \quad \tilde{x}_i \sim \mathcal{N}(0, id), i = 1, \ldots, m$ are i.i.d. training data points for some constant $r > 0$;

(iii) $y_i = \operatorname{sgn}(\langle x_i, a\rangle), \quad i = 1, \ldots, m$ are the labels of the data points;

(iv) $\hat{a}$ is the minimizer of the ℓ_1-SVM

$$\min_{w \in \mathbb{R}^d} \sum_{i=1}^{m} [1 - y_i\langle x_i, w\rangle]_+ \quad \text{subject to} \quad \|w\|_1 \le R. \tag{3}$$

(v) Furthermore, we denote

$$K = \{w \in \mathbb{R}^d \mid \|w\|_1 \le R\},$$

$$f_a(w) = \frac{1}{m}\sum_{i=1}^{m} [1 - y_i\langle x_i, w\rangle]_+,$$

where the subindex a denotes the dependency of f_a on a (via y_i).

In order to estimate the difference between a and $\hat{a}$ we adapt the ideas of [6]. First we observe

$$\begin{aligned} 0 \le f_a(a) - f_a(\hat{a}) &= \big(\mathbb{E}f_a(a) - \mathbb{E}f_a(\hat{a})\big) + \big(f_a(a) - \mathbb{E}f_a(a)\big) - \big(f_a(\hat{a}) - \mathbb{E}f_a(\hat{a})\big) \\ &\le \mathbb{E}(f_a(a) - f_a(\hat{a})) + 2\sup_{w \in K} |f_a(w) - \mathbb{E}f_a(w)|, \end{aligned}$$

i.e.

$$\mathbb{E}(f_a(\hat{a}) - f_a(a)) \le 2\sup_{w \in K} |f_a(w) - \mathbb{E}f_a(w)|. \tag{4}$$

Hence, it remains

- to bound the right hand side of (4) from above and
- to estimate the left hand side in (4) by the distance between a and $\hat{a}$ from below.

Both these task can be done with success by standard concentration arguments, cf. [5]. In this way, we obtain the following

Theorem 1. *Let $d \ge 2$, $0 < \varepsilon < 0.18$, $r > \sqrt{2\pi}(0.57 - \pi\varepsilon)^{-1}$ and $m \ge C\varepsilon^{-2}r^2R^2\ln(d)$ for some constant C. Under the "Standing assumptions" it holds*

$$\frac{\left\| a - \frac{\hat{a}}{\|\hat{a}\|_2} \right\|_2}{\langle a, \frac{\hat{a}}{\|\hat{a}\|_2}\rangle} \le C'\left(\varepsilon + \frac{1}{r}\right)$$

with probability at least

$$1 - \gamma\exp\left(-C''\ln(d)\right)$$

for some positive constants γ, C', C''.

References

[1] P.S. Bradley and O.L. Mangasarian, *Feature selection via concave minimization and support vector machines*, In Proceedings of the 13th International Conference on Machine Learning, pp. 82–90, 1998.

[2] C. Cortes and V. Vapnik, *Support-vector networks*, Machine Learning, vol. 20, no.3, 1995. 273–297.

[3] F. Cucker and D. X. Zhou, *Learning Theory: An Approximation Theory Viewpoint*, Cambridge University Press, 2007.

[4] S. Foucart and H. Rauhut, *A mathematical introduction to compressive sensing*, Applied and Numerical Harmonic Analysis, Birkhäuser, Boston, 2013.

[5] A. Kolleck, J. Vybíral, *Non-asymptotic analysis of ℓ_1-Support Vector Machines*, preprint.

[6] Y. Plan and R. Vershynin, *Robust 1-bit compressed sensing and sparse logistic regression: a convex programming approach*, IEEE Trans. Inform. Theory, vol. 59, pp. 482–494, 2013

[7] I. Steinwart and A. Christmann, *Support Vector Machines*, Springer, Berlin, 2008.

[8] R. Tibshirani, *Regression shrinkage and selection via the Lasso*, J. Royal Stat. Soc. Ser. B, vol. 58, no. 1, pp. 267–288, 1996.

[9] V. Vapnik and A. Chervonenkis, *A note on one class of perceptrons*, Automation and Remote Control, vol. 25, 1964.

[10] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, Berlin, 1995.

[11] V. Vapnik, *Statistical Learning Theory*, Wiley, Chichester, 1998.

[12] J. Zhu, S. Rosset, T. Hastie, and R. Tibshirani, *1-norm support vector machines*, In Proc. Advances in Neural Information Processing Systems, vol. 16, pp. 49–56, 2004.

Wavelet transform modulus: phase retrieval and scattering

Irène Waldspurger
(joint work with Stéphane Mallat)

This talk presents the results described in my PhD thesis about the wavelet transform modulus operator, in particular its inversion and its use in the deep representation named *scattering transform.*

A wavelet family $(\psi_j)_{j\in\mathbb{Z}}$ is defined by:

$$\forall j \in \mathbb{Z}, t \in \mathbb{R}, \qquad \psi_j(t) = 2^{-j}\psi(2^{-j}t)$$

where $\psi \in L^1 \cap L^2(\mathbb{R},\mathbb{C})$ is any function such that $\int_{\mathbb{R}} \psi(t)dt = 0$. We call *wavelet transform* the operator:

$$W : f \in L^2(\mathbb{R},\mathbb{C}) \quad \to \quad \{f \star \psi_j\}_{j\in\mathbb{Z}} \in (L^2(\mathbb{R}))^{\mathbb{Z}}$$

The *wavelet transform modulus* is the composition of W and a pointwise modulus:

$$|W| : f \in L^2(\mathbb{R},\mathbb{C}) \quad \to \quad \{|f \star \psi_j|\}_{j\in\mathbb{Z}} \in (L^2(\mathbb{R}))^{\mathbb{Z}}$$

Under mild conditions on ψ, W is invertible and its inverse is uniformly continuous. Do the same properties hold for $|W|$? To what extent is it possible to reconstruct $f \in L^2(\mathbb{R},\mathbb{C})$ from $|W|f$, up to multiplication by a global phase?

This is an inverse problem whose main motivations come from audio processing. Indeed, the wavelet transform modulus (or *scalogram*) is a widespread representation of acoustic signals, similar to the *spectrogram.* It possesses the essential property that two signals have almost identical wavelet transform modulus if and only if they are indistinguishable for a human ear. Initiated in the eighties [7, 13], the study of the inverse problem provides a theoretical framework to this empirical property. Moreover, many sound processing applications (like blind source separation [14]) operate in the scalogram domain. In order to obtain audible results, it is thus necessary to invert the wavelet transform modulus operator.

This problem belongs to the class of *phase retrieval problems*, where one aims at reconstructing an unknown object from the modulus of linear measurements. Main theoretical questions raised by these problems are the uniqueness of reconstruction and its stability to measurement noise. They are well-understood in the case where measurements are randomly chosen according to some probability distributions [3, 2]. However, they are in general difficult when the measurements are deterministic and imposed by practical considerations.

For the wavelet transform, they can be precisely answered if the wavelets are assumed to be Cauchy wavelets, that is:

$$\forall \omega \in \mathbb{R} \qquad \hat{\psi}(\omega) = \omega^p e^{-\omega} 1_{\omega \geq 0} \qquad \text{for some } p > 0.$$

In this case, we prove that any analytic function of $L^2(\mathbb{R}, \mathbb{C})$ is uniquely determined by its wavelet transform modulus, up to a global phase [12]. The corresponding inverse operator is continuous, but not uniformly continuous: the reconstruction is not stable to measurement noise in a strong sense. Nevertheless, it satisfies a local stability property: if two wavelet transforms are approximately equal in modulus, then they are approximately equal up to a global phase in a neighborhood of each point $(t, j) \in \mathbb{R} \times \mathbb{Z}$ of the time-frequency plane, except maybe around points where the modulus is close to zero.

The proof of this result is driven by the same techniques as in [1, 8].

From a numerical point of view, as phase retrieval problems are non-convex, most generic reconstruction algorithms suffer from a local optima phenomenon; even with no measurement noise, they do not always return the correct solution. Methods by convexification [3] seem more robust to this problem, but their computational cost is prohibitive.

In the case of the wavelet transform, we propose an algorithm avoiding these drawbacks. It relies on a reformulation of the phase retrieval problem involving the holomorphic extension of the wavelet transform. Additionnally, the reconstruction is performed in a multiscale manner, from low to high frequencies; each wavelet scale begins to be reconstructed after the reconstruction for the coarser scales has converged. The resulting algorithm is accurate and sufficiently fast to be applied to audio signals.

In the second part of the talk, we discuss the integration of the wavelet transform modulus in a deep representation, the scattering transform, defined by Mallat [11].

The scattering transform is a cascade of wavelet transform modulus, followed by convolutions with a low-pass filter ϕ_J. To a signal $f \in L^2(\mathbb{R})$, it associates *scattering coefficients*, of the form:

$$|...||f \star \psi_{j_1}| \star \psi_{j_2}| ... \star \psi_{j_n}| \star \phi_J$$

The index n is called the *order* of the coefficient.

Since its introduction, the scattering transform has been applied to many data analysis tasks [5]. It performs on par with or better than deep learned transformations. But compared to these transformations, the scattering transform does not require any learning. It is thus easier to analyze mathematically, and can provide interesting insights on the behavior of deep representations.

Mallat proved that the scattering transform is stable to small deformations and translations, and preserves the norm, provided that the wavelets satisfy a so-called *admissibility* condition [11].

We explain that, in the norm preservation theorem, the admissibility condition can be removed. Moreover, we can give an upper bound of the energy contained in the n-th order scattering coefficients of f, as a function of the decay of $|\hat{f}|$. For band-limited signals, this result implies that the energy decays exponentially with the order. It matches empirical observations according to which, in most applications, scattering coefficients of order $n \geq 3$ carry a negligible amount of information.

It is then natural to focus our analysis on coefficients with order $n = 1$ or $n = 2$.

First-order scattering coefficients are simply the wavelet transform modulus, convolved with the low-pass filter, as studied in the first part of this talk. At this level, the scattering transform is close to many widely-used representations. It corresponds to *dynamic features* [6] in prior audio processing work. In computer vision, Histograms Of Gradients (HOG) [4] and, to some extent, Scale-Invariant Feature Transforms (SIFT) [10], behave similarly to a wavelet transform modulus. The first layer of most convolutional neural networks also consists in wavelet-like filters [9], followed by a nonlinearity comparable to a complex modulus.

A better understanding of second-order coefficients, and their relation to other deep representations, needs to be addressed in future work.

References

[1] E. J. Akutowicz, *On the determination of the phase of a Fourier integral, I*, Transactions of the American Mathematical Society, **83** (1956) 179–192

[2] E. J. Candès, X. Li, *Solving quadratic equations via PhaseLift when there are about as many equations as unknowns*, Foundations of computational mathematics **14** (2012) 1017–1026

[3] E. J. Candès, T. Strohmer, V. Voroninski, *PhaseLift: exact and stable signal recovery from magnitude measurements via convex programming*, Communications in Pure and Applied Mathematics **66** (2011) 1241–1274

[4] N. Dalal, B. Triggs, *Histograms of oriented gradients for human detection*, Conference on computer vision and pattern recognition (2005) 886–893

[5] DATA team, École Normale Supérieure, *http://www.di.ens.fr/data/publications/*

[6] S. Furui, *Speaker-independent isolated work recognition using dynamic features of speech spectrum*, IEEE Transactions on Acoustics, Speech and Signal Processing **ASSP-34** (1986) 52–59

[7] D. Griffin, J. S. Lim, *Signal estimation from modified short-time Fourier transform*, IEEE Transactions on acoustics, speech and signal processing **32** (1984) 236–243

[8] P. Jaming, *Uniqueness results in an extension of Pauli's phase retrieval problem*, Applied and computational harmonic analysis, **37** (2014) 413–441

[9] A. Krizhevsky, I. Sutskever, G. E. Hinton, *ImageNet classification with deep convolutional neural networks* Advances in neural information processing systems (2012) 1097–1105

[10] D. G. Lowe, *Object recognition from local scale-invariant features*, International conference on computer vision (1999) 1150–1157

[11] S. Mallat, *Group invariant scattering*, Communications in pure and applied mathematics **65** (2012) 1331-1398

[12] S. Mallat, I. Waldspurger, *Phase retrieval for the Cauchy wavelet transform*, To appear in the Journal of Fourier Analysis and Applications (2014)

[13] S. H. Nawab, T. F. Quatieri, J. S. Lim, *Signal reconstruction from short-time Fourier transform magnitude*, IEEE Transactions on acoustics, speech and signal processing **ASSP-31** (1983) 986–998

[14] T. Virtanen, *Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria*, IEEE Transactions on acoustics, speech and signal processing **15** (2007) 1066–1074

Global solutions to k-means and k-median clustering objectives

Rachel Ward

The k-means clustering objective aims to partition a set of n points into k clusters in such a way that each observation belongs to the cluster with the nearest mean, and such that the sum of squared distances from each point to its nearest mean is minimal. In general, this is a hard optimization problem, requiring an exhaustive search over all possible partitions of the data into k clusters in order to find the optimal clustering. At the same time, fast heuristic algorithms for the k-means optimization problem are often applied in many data processing applications, despite having few guarantees on the clusters they produce. In this talk, we will introduce a semidefinite programming relaxation of the k-means optimization problem, along with geometric conditions on a set of data such that the algorithm is guaranteed to find the optimal k-means clustering for the data. For points drawn randomly within separated balls, the important quantities are the distances between the centers of the balls compared to the relative densities of points within them, and at sufficient density, the SDP relaxation is guaranteed to resolve such clusters at arbitrarily small separation distance. We will also discuss certain convex relaxations and recovery guarantees for another geometric clustering objective, k-median clustering. We will conclude by discussing several open questions related to this work. References are [1, 2, 3].

References

[1] Nellore, A, and Ward, R. *Recovery guarantees for exemplar-based clustering*, Accepted, Information and Computation, 2015.

[2] Awasthi, P., Bandeira, A., Charikar, M. and Krishnaswamy, R. and Villar, S., and Ward, R. *Relax, no need to round: integrality of clustering formulations*, Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science. ACM, 2015.

[3] Iguchi, T. and Mixon, D. and Peterson, J. and Villar, S., *On the tightness of an SDP relaxation of k-means*, arXiv preprint arXiv:1505.04778, 2015.

Data assimilation

Przemyslaw Wojtaszczyk
(joint work with Peter Binev, Albert Cohen, Wolfgang Dahmen, Ron DeVore, Guergana Petrova)

We tried to understand and put in a more general framework the work of Yvon Maday and coworkers on data assimilation; in particular the paper [3]. The general framework behind this work is the following:

We are given $V_1 \subset V_2 \subset \cdots \subset V_N \subset \mathcal{H}$ a sequence of finite dimensional subspaces and sequence of numbers $\epsilon_1 \geq \epsilon_2 \geq \cdots \geq \epsilon_N > 0$ a Hilbert space $\mathcal{H}$. We are also given a linear measurement map $M : \mathcal{H} \to \mathbb{R}^m$. We always assume that M maps $\mathcal{H}$ onto $\mathbb{R}^m$. We want to approximate $u \in \mathcal{H}$ given the information

(1) $\operatorname{dist}(u, V_j) \leq \epsilon_j$ for $j = 1, 2, \dots, N$
(2) $M(u) = w$ where $w \in \mathbb{R}^m$ is known.

We also consider the case when instead of a Hilbert space $\mathcal{H}$ we have a Banach space $\mathcal{X}$. The case when $N = 1$ (one-space case) is special and receives a different treatment.

Such a setup and a name comes from the following considerations: We have a natural process which is describe by a model but we do not know the parameters of the models. Quite often e.g. using a reduced basis method our information which comes from the model is in the form (1). On top of this we have t he additional information coming from measurements, this is modeled by w. We want to use them together.

Thus we want to have an algorithm $A : \mathbb{R}^m \to \mathcal{H}$ such that

- $A(w)$ satisfy 1. and 2. or it tells us that such a u does not exists (i.e. something is wrong).
- We want a'priori and a'posteriori error bounds for best possible algorithm.
- We want A to be the best or almost the best.

We denote $\mathcal{N} =: \ker M$ and $\mathcal{N}_w = M^{-1}(w)$ and $\mathcal{K} = \{h \in \mathcal{H} \,:\, \operatorname{dist}(h, V) \leq \epsilon\}$.

$$\mathcal{K}_w = \mathcal{K} \cap \mathcal{N}_w$$

If $\dim V \cap \mathcal{N} \geq 1$ than either $\mathcal{K}_w = \emptyset$ or $\mathcal{K}$ contains a line so we always assume $V \cap \mathcal{N} = \{0\}$. This forces $m \geq \dim V$. For a set $S \subset \mathcal{H}$ we define

$$\operatorname{diam}(S) = sup_{x,y \in S} \|x - y\|$$

$$\operatorname{rad}(S) = \inf_{y \in S} \sup_{x \in S} \|y - x\|.$$

The center of S is any y for which inf is attained.

The best algorithm is:

$A(w)$ equals the center of $\mathcal{K}_w$ and the error is $\operatorname{rad}(\mathcal{K}_w)$.

More geometry: We put $\mu(\mathcal{N}, V) = \sup_{x \in \mathcal{N};\, y \in V} \frac{\|x\|}{\|x - y\|}$. Also $\operatorname{diam}(\mathcal{K}_w) \leq \operatorname{diam}(\mathcal{K}_0)$.

In Hilbert and Banach one space case we have $\mathrm{rad}(\mathcal{K}_0) = \epsilon\mu(\mathcal{N}, V)$. *In the Hilbert case* $\mathrm{rad}(\mathcal{K}_w) \leq \epsilon\mu(\mathcal{N}, V)$ *and in the Banach case* $\mathrm{rad}(\mathcal{K}_w) \leq 2\epsilon\mu(\mathcal{N}\mathcal{H}, V)$

Using only SVD and Gram-Schmidt orthogonalisation we define an orthonormal basis in $\mathcal{H}$. Using this basis we describe $\mathcal{K}_w$ explicitely. It is an elipsoid in $\mathcal{N}_w$, we have the center and the radius equals $\epsilon\mu(\mathcal{N}, V)\Phi(w)$. *This gives a numerical algorithm in one space Hilbert case.*

In the multispace case the set $\mathcal{K}_w^{\mathrm{multi}}$ of all $x \in \mathcal{H}$ or $\mathcal{X}$ such that 1. and 2. holds equals

$$\mathcal{K}_w^{\mathrm{multi}} = \bigcap_{j=1}^{N} \left(\{x \ : \ \mathrm{dist}(x, V_j) \leq \epsilon_j\} \cap \mathcal{N}_w\right).$$

We have $\mathrm{diam}(\mathcal{K}_w^{\mathrm{multi}}) \leq 2\mathrm{rad}(\mathcal{K}_0^{\mathrm{multi}})$. Clearly

The best algorithm possible is $A(w)$ *equals the center of* $\mathcal{K}_w^{\mathrm{multi}}$

Unfortunately in this case we do not know how to calculate the center of $\mathcal{K}_w^{\mathrm{multi}}$ or its radius even in the Hilbert case. It is known to be NP hard. $\mathcal{K}_w^{\mathrm{multi}}$ is a finite intersection of convex sets so alternating projection algorithm generally works. In Banach space it is rather theoretical.

In Hilbert space the closest point projection CP_j onto $\{x \ : \ \mathrm{dist}(x, V_j) \leq \epsilon_j\}$ is defined as

$$CP_j(x) = P_{V_j}(x) + \alpha(x - P_{V_j}(x))$$

where $\alpha = \min\{1, \epsilon_j \|x - P_j(x)\|_2^{-1}\}$ and CP_w onto $\mathcal{N}_w$ is the affine orthogonal projection. We put

$$u^{k+1} =: CP_N CP_{N-1} \ldots CP_1 CP_w(u^k).$$

If $\mathcal{K}_w^{\mathrm{multi}} \neq \emptyset$ then $u^k \to u \in \mathcal{K}_w^{\mathrm{multi}}$ and $\|u - u^k\| = O(k^{-1/2})$. This u is not a center but the error $\leq \mathrm{diam}\mathcal{K}_w^{\mathrm{multi}}$

The obvious estimate $\mathrm{diam}\mathcal{K}_w^{\mathrm{multi}} \leq \min_j \mathrm{diam}\mathcal{K}_w^j = 2\min_j \epsilon_j \mu(\mathcal{N}, V_j)$ *is far from optimal.*

Now we discuss the Banach space case: On $\mathbb{R}^m$ we introduce the new norm (quotient norm) as

$$\|w\|_M = \inf\{\|x\|_{\mathcal{X}} \ : \ M(x) = w\}.$$

We fix a lifting $\Delta : \mathbb{R}^m \to \mathcal{X}$ i.e. a map such that $M(\Delta(w) = w$ for all $w \in \mathbb{R}^m$ and $\Delta(tw) = t\Delta(w)$ for $t \geq 0$. We put

$$\|\Delta\| = \sup_{\|w\|_M \leq 1} \|\Delta(w)\|.$$

Note that $\|\Delta\| < \infty$ does not implies that Δ is continuous. We define $L \subset \mathbb{R}^m$ as $L = M(V)$. $M|V$ is 1-1 from V onto L so we calculate the inverse map M^{-1}.

Given $w \in \mathbb{R}^m$ we find $\Lambda(w) \in L$ which is (almost) the best approximation to w in L in $\|.\|_M$, say

$$\|w - \Lambda(w)\|_M \leq \lambda \inf\{\|w - \ell\|_M \ : \ \ell \in L\}.$$

We define

$$A(w) := M^{-1}(\Lambda(w)) + \Delta(w - \Lambda(w)).$$

Let $w = M(x)$. We have

(1) $A(w) \in \mathcal{N}_w$ *i.e.* $M(A(w)) = w$

(2) $\mathrm{dist}(A(M(x)), V) \leq \lambda \|\Delta\| \mathrm{dist}(x, V)$

so if $\lambda = 1 = \|\Delta\|$ *and* $x \in \mathcal{K}_w$ *then* $A(w) \in \mathcal{K}_w$ *and* $\sup_{x \in \mathcal{K}_w} \|x - A(w)\| \leq 2\mathrm{rad}(\mathcal{K}_w)$.

- Norm $\|.\|_M$ generally is difficult to compute. Formulas exists in exceptional cases. To compute $\|w\|_M$ up to given accuracy may require exponential in m number of functionals. Finding approximation Λ is a convex minimization problem but in $\|.\|_M$. Theoretically Λ with $\lambda = 1$ exists.
- Lifting Δ with $\|\Delta\| = 1$ may not exists. Continuous Δ with $\|\Delta\| \leq 1 + \eta$ exists for every $\eta > 0$ (Bartle-Graves theorem) but are not linear and with bad modulus of continuity. In nice Banach spaces (uniformly convex e.g. L_p with $1 < p < \infty$) continuous Δ with $\|\Delta\| = 1$ exists.
- When $\mathcal{X}$ is a Hilbert space, $\|.\|_M$ is euclidean norm so Λ is an orthogonal projection and $\lambda = 1$. Δ is linear unitary map and we get the optimal algorithm discussed earlier.
- When $\mathcal{X} = C(S)$ and $M(f) = (f(s_1), \ldots, f(s_m))$ then $\|(w_j)\|_M = \max_j |w_j|$. We fix functions $(\phi_j)_{j=1}^m$ such that $\phi_j(s_j) = 1$ and $\sum_{j=1}^n |\phi_j(s)| \leq 1$ for all $s \in S$. Then $\Delta(w) =: \sum_{j=1}^m w_j \phi_j$ is a linear lifting with $\|\Delta\| = 1$. We have
$$\Lambda(w) = \mathrm{Argmin}_{\ell \in L} \max_j |l_j - w_j|$$
which gives
$$M^{-1}(\Lambda(w) = \mathrm{Argmin}_{v \in V} \max_j |v(s_j) - w_j|.$$

References

[1] P.Binev, A.Cohen, W.Dahmen, R.DeVore, G.Petrova, P. Wojtaszczyk, *Data assimilation in reduced modelling*, (submitted)

[2] R.DeVore, G.Petrova, P. Wojtaszczyk, *Data Assimilation in Banach spaces* (in preparation)

[3] Y. Maday, A.T. Patera, J.D. Penn and M. Yano, *A parametrized-background data-weak approach to variational data assimilation: Formulation, analysis, and application to acoustics* (submitted).

Reporter: Jean-Luc Bouchot

Participants

Rima Alaifari
Departement Mathematik
ETH-Zentrum
Rämistrasse 101
8092 Zürich
SWITZERLAND

Afonso S. Bandeira
Department of Mathematics
MIT, 2-246 C
77, Massachusetts Ave.
Cambridge, MA 02139
UNITED STATES

Dr. Dmitry Batenkov
Department of Computer Science
Technion-Israel Institute of Technology
Haifa 32000
ISRAEL

Prof. Dr. Gregory Beylkin
Department of Applied Mathematics
University of Colorado at Boulder
Campus Box 526
Boulder, CO 80309-0526
UNITED STATES

Prof. Dr. Holger Boche
LST für Theoretische
Informationstechnik
Technische Universität München
Theresienstr. 90/IV (LTI)
80333 München
GERMANY

Prof. Dr. Bernhard G. Bodmann
Department of Mathematics
University of Houston
Houston TX 77204-3008
UNITED STATES

Dr. Jean-Luc Bouchot
Lehrstuhl für Mathematik C (Analysis)
RWTH Aachen
Pontdriesch 10
52062 Aachen
GERMANY

Claire Boyer
Institut de Mathématiques de Toulouse
Université Paul Sabatier
118, route de Narbonne
31062 Toulouse Cedex 9
FRANCE

Prof. Dr. A. Robert Calderbank
Pratt School of Engineering
Duke University
Durham, NC 27708
UNITED STATES

Prof. Dr. Yuejie Chi
Department of Electrical & Computer
Engineering
The Ohio State University
205 Dreese Labs
2015 Neil Ave.
Columbus OH 43210
UNITED STATES

Prof. Dr. Stephan Dahlke
FB Mathematik & Informatik
Philipps-Universität Marburg
Hans-Meerwein-Strasse (Lahnbg.)
35032 Marburg
GERMANY

Prof. Dr. Ingrid Daubechies
Department of Mathematics
Duke University
P.O.Box 90320
Durham, NC 27708-0320
UNITED STATES

Prof. Dr. Christine De Mol
Department of Mathematics
Université Libre de Bruxelles
CP 217 Campus Plaine
Bd. du Triomphe
1050 Bruxelles
BELGIUM

Dr. Sjoerd Dirksen
Lehrstuhl für Mathematik C (Analysis)
RWTH Aachen
Templergraben 55
52062 Aachen
GERMANY

Kostiantyn D. Drach
School of Mathematics & Mechanical Engin.
V.N. Karazin Kharkiv National University
Maidan Svobody 4
Kharkiv 61022
UKRAINE

Dr. Martin Ehler
Fakultät für Mathematik
Universität Wien
Oskar-Morgenstern-Platz 1
1090 Wien
AUSTRIA

Markus Faulhuber
Fakultät für Mathematik
Universität Wien
Nordbergstrasse 15
1090 Wien
AUSTRIA

Prof. Dr. Hans Georg Feichtinger
Fakultät für Mathematik
Universität Wien
Oskar-Morgenstern-Platz 1
1090 Wien
AUSTRIA

Jonathan Fell
Lehrstuhl für Mathematik C (Analysis)
RWTH Aachen
Pontdriesch 10
52062 Aachen
GERMANY

Prof. Dr. Carlos Fernandez-Granda
Department of Mathematics
Courant Institute of Mathematical Sciences
New York University
251, Mercer Street
New York, NY 10012-1110
UNITED STATES

Prof. Dr. Massimo Fornasier
Zentrum Mathematik
Technische Universität München
Boltzmannstr. 3
85748 Garching bei München
GERMANY

Prof. Dr. Hartmut Führ
Lehrstuhl A für Mathematik
RWTH Aachen
52056 Aachen
GERMANY

Prof. Dr. Karlheinz Gröchenig
Fakultät für Mathematik
Universität Wien
Oskar-Morgenstern-Platz 1
1090 Wien
AUSTRIA

Prof. Dr. Philipp Grohs
Seminar for Applied Mathematics
ETH Zürich
Rämistrasse 101
8092 Zürich
SWITZERLAND

Prof. Dr. David Groß
Institut für Theoretische Physik
Universität Köln
50937 Köln
GERMANY

Dr. Maryia Kabanava
Lehrstuhl C für Mathematik (Analysis)
RWTH Aachen
Pontdriesch 10
52062 Aachen
GERMANY

Sandra Keiper
Institut für Mathematik
Sekr. MA 5-4
Technische Universität Berlin
Straße des 17. Juni 136
10623 Berlin
GERMANY

Prof. Dr. Felix Krahmer
Zentrum Mathematik
Lehr- u. Forschungseinheit M 15
Technische Universität München
Boltzmannstr. 3
85748 Garching bei München
GERMANY

Richard Küng
Physikalisches Institut
Universität Freiburg
Hermann-Herder-Str. 3a
79104 Freiburg i. Br.
GERMANY

Dr. Jan Lellmann
Centre for Mathematical Sciences
University of Cambridge
Wilberforce Road
Cambridge CB3 0WA
UNITED KINGDOM

Lizao Li
Department of Mathematics
University of Minnesota
504 Vincent Hall
206 Church Street S. E.
Minneapolis, MN 55455
UNITED STATES

Shuyang Ling
Department of Mathematics
University of California, Davis
1, Shields Avenue
Davis, CA 95616-8633
UNITED STATES

Dr. Martin Lotz
Department of Mathematics
The University of Manchester
Manchester M13 9PL
UNITED KINGDOM

Jackie Ma
Institut für Mathematik
Sekr. MA 5-4
Technische Universität Berlin
Straße des 17. Juni 136
10623 Berlin
GERMANY

Dr. Mauro Maggioni
Department of Computer Science
Duke University
117 Physics Building
P.O. Box 90320
Durham, NC 27708-0320
UNITED STATES

Prof. Dr. Madhusudan Manjunath
Department of Mathematics
University of California, Berkeley
970 Evans Hall
Berkeley CA 94720-3840
UNITED STATES

Prof. Dr. Shahar Mendelson
Department of Mathematics
Technion
Haifa 32000
ISRAEL

Prof. Dr. Hrushikesh N. Mhaskar
Institute for Mathematical Sciences
Claremont Graduate University
Claremont, CA 91125
UNITED STATES

Dr. Dustin G. Mixon
Air Force Institute of Technology
2950 Hobson Way
Wright-Patterson, OH 45433
UNITED STATES

Prof. Dr. Kasso Okoudjou
Department of Mathematics
University of Maryland
College Park, MD 20742-4015
UNITED STATES

Philipp Petersen
Institut für Mathematik
Sekr. MA 4-1
Technische Universität Berlin
Straße des 17. Juni 136
10623 Berlin
GERMANY

Prof. Dr. Gerlind Plonka-Hoch
Institut f. Numerische & Angew.
Mathematik
Universität Göttingen
Lotzestrasse 16-18
37083 Göttingen
GERMANY

Prof. Dr. Holger Rauhut
Lehrstuhl für Mathematik C (Analysis)
RWTH Aachen
Pontdriesch 10
52062 Aachen
GERMANY

Dr. Jose Luis Romero
Fakultät für Mathematik
Universität Wien
Oskar-Morgenstern-Platz 1
1090 Wien
AUSTRIA

Prof. Dr. Naoki Saito
Department of Mathematics
University of California, Davis
1, Shields Avenue
Davis, CA 95616-8633
UNITED STATES

Dr. Karin Schnass
Institut für Mathematik
Universität Innsbruck
Technikerstr. 13
6020 Innsbruck
AUSTRIA

Prof. Dr. Gabriele Steidl
Fachbereich Mathematik
Technische Universität Kaiserslautern
67653 Kaiserslautern
GERMANY

Dr. Stefan Steinerberger
Department of Mathematics
Yale University, Rm. 456 L
P.O. Box 208283
New Haven CT 06520
UNITED STATES

Zeljka Stojanac
Hausdorff Center for Mathematics
Institute for Numerical Simulation
Endenicher Allee 60
53115 Bonn
GERMANY

Prof. Dr. Thomas Strohmer
Department of Mathematics
University of California, Davis
1, Shields Avenue
Davis, CA 95616-8633
UNITED STATES

Prof. Dr. Vladimir N. Temlyakov
Department of Mathematics
University of South Carolina
Columbia, SC 29208
UNITED STATES

Dr. Ulrich Terstiege
Lehrstuhl C für Mathematik
RWTH Aachen
Pontdriesch 10
52062 Aachen
GERMANY

Felix Voigtlaender
Lehrstuhl A für Mathematik
RWTH Aachen
Templergraben 55
52062 Aachen
GERMANY

Dr. Jan Vybiral
Department of Mathematical Analysis
Faculty of Mathematics & Physics
Charles University
Sokolovská 83
186 75 Praha 8 - Karlin
CZECH REPUBLIC

Dr. Abdul Wahab
Department of Mathematics
COMSATS Institute of Information
Technology
G. T. Road
47040 Wah Cantt.
PAKISTAN

Irene Waldspurger
Dept. de Mathématiques et Applications
École Normale Superieure
45, rue d'Ulm
75005 Paris Cedex
FRANCE

Prof. Rachel Ward
Department of Mathematics
University of Texas at Austin
2515 Speedway
Austin, TX 78712
UNITED STATES

Prof. Dr. Przemek Wojtaszczyk
Interdisciplinary Centre for
Mathematical
and Computational Modelling
University of Warsaw
ul. Prosta 69
00-838 Warszawa
POLAND