FOCM 2017 - Workshop C1
Computational Harmonic Analysis and Compressive Sensing

C1 - July 17, 14:30 – 15:20

Enhancing Resolution in Undersampled Physical Imaging

Robert Calderbank
Duke University, USA
robert.calderbank@duke.edu

Compressed sensing MRI was first proposed by Lustig and Donoho, and it is now starting to appear in commercial systems. Candes, Tao and Romberg used Incoherence between measurements and representations of signals to develop performance guarantees for compressed sensing. We demonstrate here that fidelity and resolution in physical imaging systems can be improved by taking advantage of shared asymptotic characteristics of signals and measurement devices. We present new experimental results, explaining why and how improvements are possible.

Joint work with Bogdan Roman, University of Cambridge, UK, Ben Adcock, Simon Fraser University, Canada, Daniel Nietlispach, University of Cambridge, UK, Mark Bostock, University of Cambridge, UK, Irene Calvo-Almazan, University of Cambridge, UK, Martin Graves, University of Cambridge, UK and Anders Hansen, University of Cambridge, UK.

C1 - July 17, 15:30 – 15:55

The Sample Complexity of Multi-Reference Alignment

Afonso Bandeira
New York University, USA
bandeira@cims.nyu.edu

Many problems in signal and image processing amount to estimating a signal, or image, from corrupted measurements. A particularly challenging form of measurement corruption are latent transformations of the underlying signal to be recovered.

A fundamental example of this problem is the multi-reference alignment problem: in its simplest form the goal is to estimate a signal from noisy cyclically shifted copies. We will show that the number of observations needed has a surprising dependency on the signal-to-noise ratio (SNR). Computationally efficient methods that achieve optimal sample complexity will also be described.

We will also discuss the sample complexity of the heterogeneous multi-reference alignment problem where the samples come from a mixture of signals, and provide the first known procedure that provably achieves signal recovery in the low SNR regime. A related problem is heterogenous reconstruction in Cryo-Electron Microscopy (Cryo-EM) imaging where projections of a mixture of molecule densities are taken from unknown rotations. Our work provides a first step towards a complete statistical theory of heterogenous Cryo-EM.

Joint work with Jonathan Weed (MIT, USA), Amelia Perry (MIT, USA), Philippe Rigollet (MIT, USA) and Amit Singer (Princeton University, USA).

C1 - July 17, 16:00 – 16:25

Robustness to unknown error in sparse regularization

Ben Adcock
Simon Fraser University, Canada
ben_adcock@sfu.ca

Sparse regularization techniques have become increasingly popular in the last decade with the advent of techniques such compressed sensing and matrix completion. Most theoretical guarantees for sparse regularization assume that an upper bound for the noise is known in advance. However, such a bound is absent in most, if not all, practical scenarios. While estimation of the noise level may be possible, e.g. via cross validation, there are few theoretical results which explain the effect of unknown or estimated noise levels on the overall reconstruction. In this talk I will present new results on the performance of several popular sparse regularization techniques when subject to unknown noise. These results cover both Gaussian random matrices, and large classes of structured random matrices with heavy-tailed rows. Time permitting, I will give several applications of this work, including high-dimensional function approximation, infinite-dimensional sparse regularization for inverse problems, and fast algorithms for non-Cartesian Magnetic Resonance Imaging.

Joint work with Simone Brugiapaglia (SFU).

C1 - July 17, 17:00 – 17:25

Energy Propagation in Deep Convolutional Neural Networks

philipp grohs
University of Vienna, austria
philipp.grohs@univie.ac.at

Many practical machine learning tasks employ very deep convolutional neural networks. Such large depths pose formidable computational challenges in training and operating the network. It is therefore important to understand how many layers are actually needed to have most of the input signal's features be contained in the feature vector generated by the network. This question can be formalized by asking how quickly the energy contained in the feature maps decays across layers. In addition, it is desirable that none of the input signal's features be "lost" in the feature extraction network or, more formally, we want energy conservation in the sense of the energy contained in the feature vector being proportional to that of the corresponding input signal. This paper establishes conditions for energy conservation for a wide class of deep convolutional neural networks and characterizes corresponding feature map energy decay rates. Specifically, we consider general scattering networks, and find that under mild analyticity and high-pass conditions on the filters (which encompass, inter alia, various constructions of Weyl-Heisenberg filters, wavelets, ridgelets, (α)-curvelets, and shearlets) the feature map energy decays at least polynomially fast. For broad families of wavelets and Weyl-Heisenberg filters, the guaranteed decay rate is shown to be exponential. Our results yield handy estimates of the number of layers needed to have at least $((1 - \varepsilon) \cdot 100)\%$ of the input signal energy be contained in the feature vector.

Joint work with Helmut Bölcskei (ETH Zurich) and Thomas Wiatowski (ETH Zurich).

C1 - July 17, 17:30 – 17:55

Adaptive regression on low-dimensional sets in high-dimensions

Mauro Maggioni

Johns Hopkins Univeristy, U.S.A.

mauro.maggioni@jhu.edu

We discuss the problem of regressing a function of unknown regularity on an unknown low-dimensional set (e.g. a manifold) embedded in high-dimensional space, given samples on the set and corresponding noisy function values. We introduce a multiscale approximation scheme that adapts to the unknown smoothness, has strong theoretical guarantees, and is implemented by fast algorithms.

Joint work with Wenjing Liao (Johns Hopkins University) and Stefano Vigogna (Johns Hopkins University).

C1 - July 17, 18:00 – 18:25

A polynomial-time relaxation of the Gromov-Hausdorff distance

Soledad Villar

New York University, United States

soledad.villar@nyu.edu

The Gromov-Hausdorff distance provides a metric on the set of isometry classes of compact metric spaces. Unfortunately, computing this metric directly is believed to be computationally intractable. Motivated by applications in shape matching and point-cloud comparison, we study a semidefinite programming relaxation of the Gromov-Hausdorff metric. This relaxation can be computed in polynomial time, and somewhat surprisingly is itself a pseudometric. We describe the induced topology on the set of compact metric spaces. Finally, we demonstrate the numerical performance of various algorithms for computing the relaxed distance and apply these algorithms to several relevant data sets. In particular we propose a greedy algorithm for finding the best correspondence between finite metric spaces that can handle hundreds of points.

Joint work with Afonso Bandeira (New York University), Andrew Blumberg (University of Texas at Austin) and Rachel Ward (University of Texas at Austin).

C1 - July 17, 18:30 – 18:55

TBA

Gilad Lerman

University of Minnesota , USA

lerman@umn.edu

C1 - July 18, 14:30 – 14:55

Low-rank matrix recovery from Clifford orbits

David Gross

University of Cologne, Germany

david.gross@thp.uni-koeln.de

We prove that low-rank matrices can be recovered efficiently from a small number of measurements that are sampled from orbits of a certain matrix group. We focus in particular on the phase retrieval problem - a special case of low-rank recovery for rank-1 matrices. Variants of the group in question have appeared under different names in many areas of mathematics. In coding theory and quantum information, it is the complex Clifford group; in time-frequency analysis the oscillator group; and in mathematical physics the metaplectic group. It affords one particularly small and highly structured orbit that includes and generalizes the discrete Fourier basis: While the Fourier vectors have coefficients of constant modulus and phases that depend linearly on their index, the vectors in said orbit have phases with a quadratic dependence. In quantum information, the orbit is used extensively and is known as the set of stabilizer states. We argue that due to their rich geometric structure and their near-optimal recovery properties, stabilizer states form an ideal model for structured measurements for phase retrieval.

Joint work with R. Kueng and H. Zhu.

C1 - July 18, 15:00 – 15:25

Computational Harmonic Analysis meets Deep Learning

Gitta Kutyniok
Technische Universität Berlin, Germany
kutyniok@math.tu-berlin.de

Deep learning has shown impressive results in real-world applications. However, most of the research on deep neural networks is based on heuristics, lacking a fundamental mathematical theory. In this talk, we will focus on the relation between the complexity of a neural network in terms of sparse connectivity and its approximation properties. We will derive a fundamental lower bound on this sparsity for a prescribed approximation accuracy. Using methods from computational harmonic analysis, in particular, α-shearlets, we then prove that this bound is in fact optimal. Finally, our numerical experiments do not only show that this bound is achieved when training a deep neural network by the standard backpropagation algorithm, but by using a carefully chosen architecture, the trained network surprisingly even exhibits elements of representation systems from computational harmonic analysis.

Joint work with Helmut Bölcskei (ETH Zürich, Switzerland), Philipp Grohs (Universität Wien, Austria) and Philipp Petersen (Technische Universität Berlin, Germany).

C1 - July 18, 15:30 – 15:55

Statistics, Computation and Learning with Graph Neural Networks

Joan Bruna
New York University, USA
bruna@cims.nyu.edu

Deep Learning, thanks mostly to Convolutional architectures, has recently transformed computer vision and speech recognition. Their ability to encode geometric stability priors, while offering enough expressive power, is at the core of their success. In such settings, geometric stability is expressed in terms of local deformations, and it is enforced thanks to localized convolutional operators that separate the estimation into scales.

Many problems across applied sciences, from particle physics to recommender systems, are formulated in terms of signals defined over non-Euclidean geometries, and also come with strong geometric stability

priors. In this talk, I will present techniques that exploit geometric stability in general geometries with appropriate graph neural network architectures. We will show that these techniques can all be framed in terms of local graph generators such as the graph Laplacian. We will present some stability certificates, as well as applications to computer graphics, particle physics and graph estimation problems. In particular, we will describe how graph neural networks can be used to reach statistical detection thresholds in community detection, and attack hard combinatorial optimization problems, such as the Quadratic Assignment Problem.

Joint work with Xiang Li (UC Berkeley), Soledad Villar (NYU), Alex Nowak (NYU) and Afonso Bandeira (NYU).

C1 - July 18, 16:00 – 16:25

What Can Deep Learning Learn from Compressive Sensing?

Benjamin Recht
University of California, Berkeley, USA
brecht@berkeley.edu

Recent successes in neural networks have demonstrated that models with an excessive numbers parameters can achieve tremendous improvements in pattern recognition. Moreover, empirical evidence demonstrates that such performance is achievable without any obvious forms of regularization or capacity control. These findings at first glance suggest that traditional learning theory fails to explain why large neural networks generalize. In this talk, I will discuss possible mechanisms to explain generalization in such large models, appealing to insights from linear predictors. I will discuss how many of the observations can be understood by direct analogies to the linear case. I will close by proposing some possible directions of future research that connect a decade of work in compressive sensing with the contemporary phenomenology in large-scale machine learning.

Joint work with Samy Bengio (Google Brain), Moritz Hardt (Google Brain/University of California Berkeley), Oriol Vinyals (DeepMind), Chiyuan Zhang (Massachusetts Institute of Technology).

C1 - July 18, 17:00 – 17:25

Vandermonde matrices and the large sieve

Helmut Boelcskei
ETH Zurich, Switzerland
boelcskei@nari.ee.ethz.ch

Vandermonde matrices arise in many fields of applied mathematics and engineering, e.g., subspace methods—such as MUSIC and ESPRIT—for the estimation of cisoid parameters, super-resolution, line spectral estimation, compressed sensing, interpolation and approximation theory, sampling theory, differential equations, and control theory. In this talk, we establish a systematic connection between Vandermonde matrices and the large sieve, a set of inequalities developed in analytic number theory by Linnik, Rényi, Roth, and Bombieri. Based on this relationship, we present new bounds on the extremal singular values and the condition number of Vandermonde matrices with nodes in the unit disk. We then build on these bounds to develop a deterministic, finite-SNR, finite sample-size performance analysis of MUSIC, ESPRIT, and the matrix pencil method. These results demonstrate that polynomial root finding (e.g.,

Berlekamp-Massey and Peterson-Gorenstein-Zierler) algorithms over the reals as used widely in the sparse signal recovery literature do not exhibit an intrinsic lack of numerical stability.

Joint work with Céline Aubel.

C1 - July 18, 17:30 – 17:55

Packings in real projective spaces

Dustin Mixon
Air Force Institute of Technology, USA
dustin.mixon@gmail.com

This talk discusses techniques from algebraic and differential geometry to determine how to best pack points in real projective spaces. We present a computer-assisted proof of the optimality of a particular 6-packing in $\mathbb{R}\mathbf{P}^3$, we introduce a linear-time constant-factor approximation algorithm for packing in the so-called Gerzon range, and we provide local optimality certificates for two infinite families of packings. Finally, we present perfected versions of various putatively optimal packings from Sloane's online database, along with a handful of infinite families they suggest, and we prove that these packings enjoy a certain weak notion of optimality.

Joint work with Matthew Fickus (Air Force Institute of Technology, USA) and John Jasper (University of Cincinnati, USA).

C1 - July 18, 18:00 – 18:25

Multireference alignment and clustering

Amit Singer
Princeton University, United States
amits@math.princeton.edu

Multi-reference alignment consists of estimating a signal from its many noisy shifted copies, where the shifts are unknown. The number of observations needed for accurate estimation, efficient recovery algorithms, extension to recovery of multiple signals, and application to cryo-electron microscopy will be discussed.

Joint work with Tamir Bendory (Princeton University, USA), Nicolas Boumal (Princeton University, USA), Roy Lederman (Princeton University, USA), William Leeb (Princeton University, USA), Nir Sharon (Princeton University, USA), Emmanuel Abbe (Princeton University, USA), Joao Pereira (Princeton University, USA) and Afonso Bandeira (NYU, USA).

C1 - July 19, 14:30 – 14:55

Extrapolation from sparsity

Laurent Demanet
MIT, USA
demanet@gmail

This talk considers the basic question of frequency extrapolation of sparse signals observed over some frequency band, such as scattered bandlimited waves. How far, and how stably can we extend? I will review recent progress on the mathematical aspects of this question, which are tied to the notion of super-resolution. I will also discuss the robust "phase tracking" algorithmic approach, which is suitable for seismic imaging where the bandlimiting model is far from accurately known.

Joint work with Nam Nguyen and Yunyue Elita Li.

C1 - July 19, 15:00 – 15:25

The Usefulness of a Modified Restricted Isometry Property

Simon Foucart
Texas A&M University, USA
foucart@tamu.edu

The restricted isometry property is arguably the most prominent tool in the theory of compressive sensing. In its classical version, it features ℓ_2-norms as inner and outer norms. The modified version considered in this talk features the ℓ_1-norm as the inner norm, while the outer norm depends a priori on the distribution of the random entries populating the measurement matrix. The modified version holds for a wider class of random matrices and still accounts for the success of sparse recovery via basis pursuit and via iterative hard thresholding. In the special case of Gaussian matrices, the outer norm actually reduces to an ℓ_2-norm. This fact allows one to retrieve results from the theory of one-bit compressive sensing in a very simple way. Extensions to one-bit matrix recovery are then straightforward.

C1 - July 19, 15:30 – 15:55

Joint Blind Deconvolution and Blind Demixing via Nonconvex Optimization

Shuyang Ling
University of California Davis, USA
syling@math.ucdavis.edu

We study the question of extracting a sequence of functions $\{f_i, g_i\}_{i=1}^s$ from observing only the sum of their convolutions, i.e., from $y = \sum_{i=1}^s f_i * g_i$. While convex optimization techniques are able to solve this joint blind deconvolution-demixing problem provably and robustly under certain conditions, for medium-size or large-size problems we need computationally faster methods without sacrificing the benefits of mathematical rigor that come with convex methods. We present a non-convex algorithm which guarantees exact recovery under conditions that are competitive with convex optimization methods, with the additional advantage of being computationally much more efficient. Our two-step algorithm converges to the global minimum linearly and is also robust in the presence of additive noise. While the derived performance bounds are suboptimal in terms of the information-theoretic limit, numerical simulations show remarkable performance even if the number of measurements is close to the number of degrees of freedom. We will discuss blind deconvolution as a special case and an application of the proposed framework in wireless communications.

Joint work with Thomas Strohmer (University of California Davis, USA), Xiaodong Li (University of California Davis, USA) and Ke Wei (University of California Davis, USA).

C1 - July 19, 16:00 – 16:25

Multi-taper estimators with geometric constraints

Jose Luis Romero

Acoustics Research Institute, Austrian Academy of Sciences, Austria

jlromero@kfs.oeaw.ac.at

We consider the problem of estimating a stationary point process from samples taken on a set with possibly general geometry. The multi-taper approach consists in weighting the data with adequate masks to then perform an average that reduces variance. We describe the performance of such estimators, and the computational challenges in their implementation. We consider workarounds for the main computational obstacles and quantify their effectiveness.

Joint work with Luis Daniel Abreu (Acoustics Research Institute, Austrian Academy of Sciences).

C1 - July 19, 17:00 – 17:50

TBA

Ingrid Daubechies

Duke University, USA

ingrid@math.duke.edu

C1 - Poster

The Beurling-Selberg Box Minorant Problem

Jacob Carruth

University of Texas at Austin, United States

jcarruth@math.utexas.edu

Let F be a function on $\mathbb{R}^d$ which is bandlimited to the unit cube and which minorizes the characteristic function of the unit cube. How large can the integral of F be? We show that for a sufficiently large dimension d^* the answer to this question is zero for $d \geq d^*$. Furthermore, we discuss a numerical method which gives an explicit upper bound on d^*. We also show that the integral of F can be larger than zero when $d \leq 5$ by explicitly constructing non-trivial minorants. Selberg first asked this question to solve a problem in number theory but it has since found many applications including to signal recovery, crystallography, and sphere packing. We highlight some of the applications to signal recovery.

Joint work with Noam Elkies (Harvard University), Felipe Goncalves (University of Alberta) and Michael Kelly (Center for Communications Research).

C1 - Poster

Exact support recovery in unmixing problems by an altered Lasso-path algorithm for multi-penalty functionals

Timo Klock

Simula Research Laboratory, Norway

timo@simula.no

Inverse problems of unmixing type arise in many real-life applications such as audio processing or medical image analysis. In such problems additive noise directly affects a sparse signal before being measured through a sampling matrix. Consequently, the noise in the measurement is amplified through the sampling process and the so-called noise folding phenomenon occurs. This amplification worsens the results on support identification by means of "classical" sparse recovery techniques based on the ℓ_1-penalised Lasso functional. Several recent works suggest to apply a multi-penalty framework for a correct modeling and separation of the original signal in such type of problems. Theoretical results and numerical experiments with oracle-given regularisation parameters have shown that this approach allows to recover the correct support in a larger number of problems than its single-penalty counterpart. Admittedly, the parameter choice in such multi-penalty functionals becomes more involved compared to the Lasso or other single-penalty methods.

In this poster, we introduce the multi-penalty regularisation for the unmixing problem. It has been shown that the resulting functional can be rewritten as a parameterised Lasso such that we can solve the multi-penalty minimisation with known techniques for the Lasso. In this spirit, we provide an extension of the Lasso-path algorithm for an efficient calculation of large parts of the multi-penalty solution space without performing extensive grid-searches over regularisation parameters. Finally, by using a naive support selection heuristic based on signal-to-noise ratios, we identify a unique support that is compared to results from conventional sparse recovery techniques. Such experiments confirm improved results of the multi-penalty function, especially when compared to single-penalty counterparts.

Joint work with Valeriya Naumova (Simula Research Laboratory) and Markus Grasmair (Norwegian University of Science and Technology).

C1 - Poster

Harmonic Mean Iteratively Reweighted Least Squares for Low-Rank Matrix Recovery

Christian Kümmerle
Technische Universität München, Germany
christian.kuemmerle@ma.tum.de

We propose a new iteratively reweighted least squares (IRLS) algorithm for the recovery of a matrix $X \in \mathbb{C}^{d_1 \times d_2}$ of rank $r \ll \min(d_1, d_2)$ from incomplete linear observations, solving a sequence of low complexity linear problems. The easily implementable algorithm, which we call harmonic mean iteratively reweighted least squares (HM-IRLS), optimizes a non-convex Schatten-p quasi-norm penalization to promote low-rankness and carries three major strengths, in particular for the matrix completion setting. First, the algorithm converges globally to the low-rank matrix for relevant, interesting cases, for which any other (non-)convex state-of-the-art optimization approach fails the recovery. Secondly, HM-IRLS exhibits an empirical recovery probability close to 1 even for a number of measurements very close to the theoretical lower bound $r(d_1 + d_2 - r)$, i.e., already for significantly fewer linear observations than any other tractable approach in the literature. Thirdly, HM-IRLS exhibits a locally superlinear rate of convergence (of order $2 - p$) if the linear observations fulfill a suitable null space property. While for the first two properties we have so far only strong empirical evidence, we prove the third property as our main theoretical result.

Joint work with Juliane Sigl (Technische Universität München, Germany).

C1 - Poster

Novel aspects of approximating Hilbert Schmidt operators via Gabor Multipliers and Spline-type Spaces

Darian Onchis

University of Vienna, Austria

darian.onchis@univie.ac.at

In this work, we introduce two new aspects to improve the approximation of Hilbert Schmidt operators over locally compact Abelian groups via generalized Gabor multipliers (GGM). One aspect is to formulate both as particular examples of pseudodifferential operators and to consider the approximation of the symbol of an Hilbert-Schmidt operator in the spline-type space associated to a Gabor multiplier. This gives the possibility to employ a selection procedure of the analysis and synthesis function, interpreted as time-frequency lag; hence, with the related algorithm it is possible to handle both underspread and overspread operators. In the numerical section, we exploit the case of approximating overspread operators having compact spreading function and time-varying systems. For the latter, the approximation of discontinuities in the symbol is not straightforward achievable in the GGM setting. For this reason, another aspect is to further process the symbol through a Hough transform, to detect discontinuities and to smooth them using a new class of approximants. This procedure creates a bridge between features detections techniques and harmonic analysis methods and in specific cases it almost doubles the accuracy of approximation.

Joint work with Simone Zappala (University of Vienna, Austria).

C1 - Poster

Estimation of block sparsity in compressive sensing

Zhiyong Zhou

Umeå University, Sweden

zhiyong.zhou@umu.se

Explicitly using the block structure of the unknown signal can achieve better recovery performance in compressive censing. An unknown signal with block structure can be accurately recovered from under-determined linear measurements provided that it is sufficiently block sparse. However, in practice, the block sparsity level is typically unknown. In this paper, we consider a soft measure of block sparsity, $k_\alpha(\mathbf{x}) = (\|\mathbf{x}\|_{2,\alpha}/\|\mathbf{x}\|_{2,1})^{\frac{\alpha}{1-\alpha}}, \alpha \in [0, \infty]$ and propose a procedure to estimate it by using multivariate isotropic symmetric α-stable random projections without sparsity or block sparsity assumptions. The limiting distribution of the estimator is given. Some simulations are conducted to illustrate our theoretical results.

Joint work with Jun Yu (Umeå University, Sweden).