SampTA 2019

13th International conference on Sampling Theory and Applications

université de BORDEAUX July 8-12

Tuesday July 9, 1.45-8.00 pm

Wegener	1.45-2.45	*Combinatorial compressed sensing with expanders* Bubacarr Bah Chair: Ana Gilbert
		Deep learning (invited session) Chair: Misha Belkin & Mahdi Soltanolkotabi
A9 Amphi 1	2:55-3:20	*Reconciling modern machine learning practice and the classical bias-variance trade-off* Mikhail Belkin, Daniel Hsu, Siyuan Ma & Soumik Mandal
A9 Amphi 1	3.20-3.45	*Overparameterized Nonlinear Optimization with Applications to Neural Nets* Samet Oymak
A9 Amphi 1	3.45-4.10	*General Bounds for 1-Layer ReLU approximation* Bolton R. Bailey & Matus Telgarsky
A9 Amphi 1	4.10-4.20	Short Break
A9 Amphi 1	4.20-4.45	*Generalization in deep nets: an empirical perspective* Tom Goldstein
A9 Amphi 1	4.45-5.10	*Neuron birth-death dynamics accelerates gradient descent and converges asymptotically* Joan Bruna
	5.45-8.00	**Poster simposio**
	Domaine du Haut-Carré	

SampTA 2019

13th International conference on Sampling Theory and Applications

université de BORDEAUX July 8-12

Tuesday July 9, 1.45-8.00 pm

Wegener	1.45-2.45	*Combinatorial compressed sensing with expanders* Bubacarr Bah Chair: Anna Gilbert
		Frame Theory Chair: Ole Christensen
A29 Amphi 2	2.55-3.20	*Banach frames and atomic decompositions in the space of bounded operators on Hilbert spaces* Peter Balazs
	3.20-3.45	*Frames by Iterations in Shift-invariant Spaces* Alejandra Aguilera, Carlos Cabrelli, Diana Carbajal & Victoria Paternostro
	3.45-4.10	*Frame representations via suborbits of bounded operators* Ole Christensen & Marzieh Hasannasabjaldehbakhani
	4.10-4.20	Short Break
	4.20-4.45	*Sum-of-Squares Optimization and the Sparsity Structure of Equiangular Tight Frames* Dmitriy Kunisky & Afonso Bandeira
	4.45-5.10	*Frame Potentials and Orthogonal Vectors* Josiah Park
	5.10-5.35	*Compactly Supported Tensor Product Complex Tight Framelets with Directionality* Xiaosheng Zhuang & Bin Han
	5.45-8.00	**Poster simposio**
	Domaine du Haut-Carré	

Tuesday July 9, 1.45-8.00 pm

Wegener	1.45-2.45	*Combinatorial compressed sensing with expanders* Bubacarr Bah Chair: Anna Gilbert
		Phase retrieval Chair: Joseph Lakey
A29 Amphi 3	2.55-3.20	*Phase Estimation from Noisy Data with Gaps* Yitong Huang, Clark Bowman, Olivia Walch & Daniel Forger
A29 Amphi 3	3.20-3.45	*Phase retrieval from local correlation measurements with fixed shift length* Oleh Melnyk, Frank Filbir & Felix Krahmer
A29 Amphi 3	3.45-4.10	*Ill-conditionedness of discrete Gabor phase retrieval and a possible remedy* Matthias Wellershoff & Rima Alaifari
	17.10-17.20	Short break
		Quantization Chair: Ozgur Yilmaz
A29 Amphi 3	4.30-4.55	*Higher order 1-bit Sigma-Delta modulation on a circle* Olga Graf, Felix Krahmer & Sara Krause-Solberg
A29 Amphi 3	4.55-5.20	*One-Bit Compressed Sensing Using Smooth Measure of ℓ^0 Norm* Sina Alemohammad & Arash Amini
	5.45-8.00	**Poster simposio** Domaine du Haut-Carré

SampTA 2019
13th International conference on Sampling Theory and Applications
université de BORDEAUX July 8-12

Combinatorial compressed sensing with expanders

Bubacarr Bah (African Institute for Mathematical Sciences, South Africa & German Research Chair of Mathematics, Germany)

Abstract: In spite of the square-root bottleneck for doing compressed sensing with binary matrices, the computational benefits of such sparse matrices triggered a lot interest in this area dubbed combinatorial compressed sensing. This talk will start from the introduction of the l1-norm restricted isometry property, which allows for optimal sampling rates but weaker instance optimality conditions to my most recent work on model-based combinatorial compressed sensing. There will be discussion on construction of expander graphs and hence expander matrices, both deterministic and random constructions. Recent improvements in random constructions and their implications for compressed sensing will also be discussed. Algorithms motivated by linear sketching for both standard compressed sensing and model-based compressed sensing with tree-sparse and loopless overlapping group-sparse model will be presented. The current state-of-the-art with more general models (overlapping group models for groups with bounded treewith and low frequency) and more efficient algorithms using head and tail approximations with model projections done via bender's decomposition will also be presented.

Deep learning (invited session)

Chair: Misha Belkin & Mahdi Soltanolkotabi

2:55-3:20: Reconciling modern machine learning practice and the classical bias-variance trade-off

Mikhail Belkin, Daniel Hsu, Siyuan Ma & Soumik Mandal

Abstract: Breakthroughs in machine learning are rapidly changing society and science, yet fundamental understanding of learning currently lags behind. In classical machine learning, we find a model that balances under-fitting and over-fitting: complex enough to express underlying structure in data, simple enough to avoid fitting spurious patterns. In modern practice, we fit complex models like neural networks with zero training error but still obtain high accuracy on test data. Our objective is to resolve the apparent conflict between the classical understanding and modern practice. Our main result is a "double descent" risk curve that unifies the behavior in the classical and modern regimes. The mechanism underlying its emergence is posited. Our findings resolve the apparent conflict and inform how machine learning models should be designed and understood.

3.20.40-3.45: Overparameterized Nonlinear Optimization with Applications to Neural Nets

Samet Oymak

Abstract: Occam's razor is a fundamental problem-solving principal and states that one should seek the simplest possible explanation. Indeed, classical machine learning models such as (sparse) linear regression aims to find simple explanations to data by using with as few parameters as possible. On the other hand, modern techniques such as deep networks are often trained in the overparameterized regime where the model size exceeds the size of the training dataset. While this increases the risk of overfitting and the complexity of the explanation, deep networks are known to have good generalization properties. In this talk, we take a step towards resolving this paradox: We show that solution found by first order methods, such as gradient descent, has the property that it has near shortest distance to the initialization of the algorithm among all other solutions. We also advocate that shortest distance property can be a good proxy for the simplest explanation. We discuss the implications of these results on neural net training and also highlight some outstanding challenges.

SampTA 2019
13th International conference on Sampling Theory and Applications
université de BORDEAUX
July 8-12

Deep learning (invited session)

Chair: Misha Belkin & Mahdi Soltanolkotabi

3.45-4.10: General Bounds for 1-Layer ReLU approximation

Bolton R. Bailey & Matus Telgarsky

Abstract: The popularity of the *ReLU* has given rise to many neural networks which are piecewise affine. In this work, we show how a refined bound on the number of affine pieces in a single *ReLU* layer can be used to lower bound the approximation error of a *ReLU* network. We also demonstrate a method based on Rademacher complexity and random sampling to give an upper bound on the error of optimal approximations for these layers.

4.20-4.45: Generalization in deep nets: an empirical perspective

Tom Goldstein

Abstract: The power of neural networks lies in their ability to perform well on data that wasn't seen during training, a phenomena known as "generalization." Classical learning theory predicts that models generalize when they are under-parameterized, i.e., when the training set contains more samples than the number of model parameters. But strangely, neural nets generalize even when they have many more parameters than training data, and the underlying reasons for this good behavior remain elusive. Numerous rigorous attempts have been made to explain generalization, but available bounds are still quite loose, and analysis does not always lead to true understanding. The goal of this talk is to make generalization more intuitive. Using visualization methods, we discuss the mystery of generalization, the geometry of loss landscapes, and how the curse (or, rather, the blessing) of dimensionality causes optimizers to settle into minima that generalize well.

Deep learning (invited session)

Chair: Misha Belkin & Mahdi Soltanolkotabi

4.45-5.10: Neuron birth-death dynamics accelerates gradient descent and converges asymptotically

Joan Bruna

Abstract: Neural networks with a large number of parameters admit a mean-field description, which has recently served as a theoretical explanation for the favorable training properties of "overparameterized" models. In this regime, gradient descent obeys a deterministic partial differential equation (PDE) that converges to a globally optimal solution for networks with a single hidden layer under appropriate assumptions. In this work, we propose a non-local mass transport dynamics that leads to a modified PDE with the same minimizer. We implement this non-local dynamics as a stochastic neuronal birth/death process and we prove that it accelerates the rate of convergence in the mean-field limit. We subsequently realize this PDE with two classes of numerical schemes that converge to the mean-field equation, each of which can easily be implemented for neural networks with finite numbers of parameters. We illustrate our algorithms with two models to provide intuition for the mechanism through which convergence is accelerated.

Joint work with G. Rotskoff, S. Jelassi and E. Vanden-Eijnden

Frame Thory

Chair: Ole Christensen

2.55-3.20: Banach frames and atomic decompositions in the space of bounded operators on Hilbert spaces

Peter Balazs

Abstract: The concept of frames is used extensively for the representation of signal or functions. Recently this concept ia applied more and more for the representation of operators, both in theory as well as in the application for the numerical solutions of operator equations. In this paper we first give a survey about the matrix representation of operators using frames. Then we prove that the tensor product of frames forms a Banach frame and an atomic decomposition for the space of bounded operators of Hilbert spaces.

3.20-3.45: Frames by Iterations in Shift-invariant Spaces

Alejandra Aguilera, Carlos Cabrelli, *Diana Carbajal* & Victoria Paternostro

Abstract: In this note we solve the dynamical sampling problem for a class of shift-preserving (SP) operators, acting on a finitely generated shift-invariant space (FSIS). We find conditions on the operator and on a finite set of functions in the FSIS in order that the iterations of the operator on the functions produce a frame generator set. That is, the integer translations of the frame generator set is a frame of the FSIS. In order to obtain these results, we study the structure of SP operators and obtain a generalized finite dimensional spectral theorem.

SampTA 2019 July 8-12
13th International conference on Sampling Theory and Applications
université de BORDEAUX

Frame Theory

Chair: Ole Christensen

3.45-4.10: Frame representations via suborbits of bounded operators

Ole Christensen & *Marzieh Hasannasabjaldehbakhani*

Abstract: The standard setup of dynamical sampling concerns frame properties of sequences of the form $\{T_n\varphi\}_{n=0}^{\infty}$, where T is a bounded operator on a Hilbert space H and $\varphi \in H$. In this paper we consider two generalizations of this basic idea. We first show that the class of frames that can be represented using iterations of a bounded operator increases drastically if we allow representations using just a subfamily $\{T^{\alpha(k)}\varphi\}_{k=1}^{\infty}$ of $\{T_n\varphi\}_{n=0}^{\infty}$; indeed, any linear independent frame has such a representation for a certain bounded operator T. Furthermore, we prove a number of results relating the properties of the frame and the distribution of the powers $\{\alpha(k)\}_{k=1}^{\infty}$ in $\mathbb{N}$. Finally, we show that also the condition of linear independency can be removed if we allow to consider approximate frame representations with an arbitrary small prescribed tolerance, in a sense to be made precise.

4.20-4.45: Frame Potentials and Orthogonal Vectors

Josiah Park

Abstract: An extension is given of a recent result of Glazyrin, showing that an orthonormal basis $\{e_i\}_{i=1}^{d}$ joined with the vectors $\{e_i\}_{i=1}^{d}$, where $1 \leq m < d$ minimizes the p-frame potential for $p \in [1, 2\log\frac{2m+1}{2m}/\log\frac{m+1}{m}]$ over all collections of $N = d + m$ vectors $\{x_1, \ldots, x_n\}$ in $\mathbb{S}^{d-1}$.

4.45-5.10: Sum-of-Squares Optimization and the Sparsity Structure of Equiangular Tight Frames

Dmitriy Kunisky & Afonso Bandeira

Abstract: Equiangular tight frames (ETFs) may be used to construct examples of feasible points for semidefinite programs arising in sum-of-squares (SOS) optimization. We show how generalizing the calculations in a recent work of the authors' that explored this connection also yields new bounds on the sparsity of (both real and complex) ETFs. One corollary shows that Steiner ETFs corresponding to finite projective planes are optimally sparse in the sense of achieving tightness in a matrix inequality controlling overlaps between sparsity patterns of distinct rows of the synthesis matrix. We also formulate several natural open problems concerning further generalizations of our technique.

SampTA 2019
13th International conference on Sampling Theory and Applications
université de BORDEAUX
July 8-12

Frame Theory

Chair: Ole Christensen

5.10-5.35: Compactly Supported Tensor Product Complex Tight Framelets with Directionality

Xiaosheng Zhuang & Bin Han

Abstract: Construction of directional compactly supported tensor product complex tight framelets $cptTP - \mathbb{C}TF6$ is discussed. The construction algorithms employ optimization techniques and put extensive emphasis on frequency response and spatial localization of the underlying one-dimensional tight framelet filter banks. A concrete example of $cptTP - \mathbb{C}TF6$ is provided. Numerical experiments show that such constructed $cptTP - \mathbb{C}TF6$ have good performance for applications such as image denoising.

Phase Retrieval

Chair: Joseph Lakey

2.55-3.20: Phase Estimation from Noisy Data with Gaps

Yitong Huang, Clark Bowman, Olivia Walch & Daniel Forger

Abstract: Determining the phase of a rhythm embedded in a time series is a key step in understanding many oscillatory systems. While existing approaches such as harmonic regression and cross-correlation are effective even when some data are missing, we show that they can produce biased estimates of phase when missing data are consecutive (i.e., there is a gap). We propose a simple modification of the least-squares approach, Gap Orthogonalized Accelerated Least Squares (GOALS), which addresses this issue with a negligible increase in computational expense. We test GOALS against other approaches on a synthetic dataset and on a real-world dataset of activity recorded by an Apple Watch, showing in both cases that GOALS is effective at recovering phase estimates from noisy data with gaps.

3.20-3.45: Phase retrieval from local correlation measurements with fixed shift length

Oleh Melnyk, Frank Filbir & Felix Krahmer

Abstract: Driven by ptychography, we consider an extension of the phase retrieval from local correlation measurements with shifts of length one to any fixed shift length. As a result, we provide an algorithm and recovery guarantees for extended model.

3.45-4.10: Ill-conditionedness of discrete Gabor phase retrieval and a possible remedy

Matthias Wellershoff & Rima Alaifari

Abstract: In light of recent work on continuous Gabor phase retrieval, we analyse discrete Gabor phase retrieval problems and note that under realistic decay assumptions on the window functions, the stability constants increase significantly in the space dimension. When using discretisations of the Gaussian as windows, we are in fact able to show that the stability constants grow at least exponentially as the dimension of the space increases. At the same time, we observe that the adversarial examples, which we construct to estimate the stability constants, all contain long modes of silence. This suggests that one should try to reconstruct signals up to so-called semi-global phase factors and not up to a global phase factor as is the canon in the literature. This observation is further corroborated by a stability result for discrete Gabor phase retrieval which we have proven recently.

Quantization

Chair: Ozgur Yilmaz

4.20-4.45: Higher order 1-bit Sigma-Delta modulation on a circle

Olga Graf, *Felix Krahmer* & Sara Krause-Solberg

Abstract: Manifold models in data analysis and signal processing have become more prominent in recent years. In this paper, we will look at one of the main tasks of modern signal processing, namely, at analog-to- digital (A/D) conversion in connection with a simple manifold model (circle). We will focus on Sigma-Delta modulation which is a popular method for A/D conversion of bandlimited signals that employs coarse quantization coupled with oversampling. Classical Sigma-Delta schemes would provide mismatches and large errors at the initialization point if the signal to be converted is defined on a circle. In this paper, our goal is to get around these problems for higher order Sigma-Delta schemes. Our results show how to design an update for the second and third order schemes based on the reconstruction error analysis such that for the updated scheme the reconstruction error is improved.

4.45-5.10: One-Bit Compressed Sensing Using Smooth Measure of ℓ^0 Norm

Sina Alemohammad & *Arash Amini*

Abstract: Quantization of signals and parameters happens in all digital data acquisition devices. It is commonly regarded as a non-ideality of the system, and shall be taken into account when designing or analyzing a system. The topic of one-bit compressed sensing studies the effect of quantization in the extreme case where the samples are quantized with only one bit, i.e., the sign bit. The recovery of a sparse signal based on one-bit measurements is widely accomplished via thresholding methods or variants of ℓ^1-minimization techniques. In this paper, we introduce a recovery method arising from smoothing directly the ℓ^0 pseudo-norm. While we numerically verify the superior performance of the proposed method compared to the state-of-the-art techniques in our simulations, we briefly discuss the convergence analysis of this method.