# 2019 International Conference on Kernel-Based Methods and Applications
# Talk Abstract Booklet

Xi'an Jiaotong-Liverpool University, Suzhou, China

October 10th to October 14th, 2019

# Contents

# Abstracts

## Robustness and machine learning PT

Andreas Christmann
University of Bayreuth, Germanny

**Bio** Dissertation and Habilitation at the University of Dortmund. After positions as a visiting professor at KU Leuven (Belgium) and as professor at universities in Dortmund (Germany) and Brussels (VUB, Belgium). He is a Full Professor and Chair for Stochastics at the University of Bayreuth since 2008.

www and google scholar

---

## Bridging Deep Neural Networks and Differential Equations for Image Analysis and Beyond PT

Bin Dong
Peking University, China

Deep learning continues to dominate machine learning and has been successful in computer vision, natural language processing, etc. Its impact has now expanded to many research areas in science and engineering. However, the model design of deep learning still lacks systematic guidance, and most deep models are seriously in lack of transparency andinterpretability, thus limiting the application of deep learning in some fields of science and medicine. In this talk, I will show how we can tackle this issue by presenting some of our recent work on bridging numerical differential equation and deep convolutional architecture design. We can interpret some of the popular deep CNNs in terms of numerical (stochastic) differential equations, and propose new deep architectures that can further improve the prediction accuracy of the existing networks in image classification. We also show how to design transparent deep convolutional networks to uncover hidden PDE models from observed dynamical data. Further applications of this perspective to various problems in imaging and inverse problems will be discussed.

**Bio** Prof. Bin Dong is associate director of Laboratory for Biomedical Image Analysis at Peking University, and also serves(served) as Editor for Inverse Problems and Imaging, and Applied and Computational Harmonic Analysis. His research interests include wavelet frames, theory and applications, biological and medical imaging, image processing and analysis, and deep learning. More than 40 papers and 3 monographs have been published in these areas. He has co-organised over 10 conferences/workshops and conducted 8 research grants. His research has been sponsored by the national 1000 talent plan.

www and google scholar

---

## PT Creating anisotropic meshes by combining RBFs and HDE

Patricio Farrell

Weierstrass Institute(WIAS) & Hamburg University of Technology(TUHH), Germany

Many applications heavily rely on piecewise triangular meshes to describe complex surface geometries. The quality of the meshes directly impacts numerical simulations. In practice, however, one often has to deal with several challenges to construct suitable meshes. Some regions in a naively built mesh may be overrefined, others too coarse. Additionally, the triangles may be too thin or not properly oriented. In this talk, I present a novel surface remeshing procedure which greatly improves a problematic input mesh and overcomes all of these drawbacks. By coupling surface reconstruction via radial basis functions with the higher-dimensional embedding surface remeshing technique, we can automatically generate a new anisotropic surface mesh with improved mesh quality. Moreover, we are not only able to refine or coarsen certain mesh regions but also align the triangles according to the curvature of the reconstructed surface. This yields an acceptable trade-off between computational complexity and accuracy. This is joint work with Hang Si (WIAS), Franco Dassi (Milano-Bicocca) and Tim Burg (TU Berlin).

**Bio** Patricio is a researcher at the Weierstrass Institute (WIAS) in Berlin, working on finite volume methods for semiconductors and meshfree methods. From 2017-2019, he has been a visiting professor at the Hamburg University of Technology (TUHH). He obtained his doctorate in mathematics from the University of Oxford where he studied multilevel methods for radial basis functions. Previously, he completed a Diplom at the University of Hamburg.

www

---

## Boundary effects in kernel approximation

Thomas Christiann Hangelbroek

Hawaii University, USA

**Bio**

www and arXiv and genealogy

---

# On Banach representer theorems: from RKHS to sparse kernel expansions

PT

Michael Unser

EPFL, Switzerland

Regularization addresses the ill-posedness of the training problem in machine learning or the reconstruction of a signal from a limited number of measurements. The standard strategy consists in augmenting the original cost functional by an energy that penalizes solutions with undesirable behavior. In this presentation, we present a general representer theorem that characterizes the solutions of a remarkably broad class of optimization problems in Banach spaces and helps us understand the effect of regularization. We show that the solutions fall into three categories depending on wether the underlying norm is Hilbertian, non-quadratic but strictly convex, or sparsity-promoting (and not strictly convex!). The first scenario is classical: it covers the widely-used RKHS kernel expansions of machine leaning as well a Tikhonov regularization. The second scenario is more challenging as it induces some non-linear mapping: it has a strong connection with the theory of reproducing kernel Banach spaces (RKBS). The third scenario, which is lesser known, has some interesting implication for machine learning as well. Specifically, we provide a special instance of our representer theorem (with total variation regularization) that results in sparse kernel expansions. The main difference with the standard RKHS framework is that the kernels are fewer and the data centers adaptive.

**Bio**Michael Unser is professor and director of EPFL's Biomedical Imaging Group, Lausanne, Switzerland. His primary area of investigation is biomedical image processing. He is internationally recognized for his research contributions to sampling theory, wavelets, the use of splines for image processing, stochastic processes, and computational bioimaging. He has published over 300 journal papers on those topics. He is the author with P. Tafti of the book An introduction to sparse stochastic processes, Cambridge University Press 2014. Prof. Unser is a fellow of the IEEE (1999), an EURASIP fellow (2009), and a member of the Swiss Academy of Engineering Sciences. He is the recipient of several international prizes including four IEEE-SPS Best Paper Awards and two Technical Achievement Awards from the IEEE (2008 SPS and EMBS 2010).

www and google scholar

---

PT

## Kernel-based Reconstructions for Uncertainty Quantification

Holger Wendland
University of Bayreuth, Germany

In uncertainty quantification, an unknown quantity has to be reconstructed which depends typically on the solution of a partial differential equation. This partial differential equation itself depends on parameters, some of them are deterministic and some are random. To approximate the unknown quantity one thus has to solve the partial differential equation (usually numerically) for several instances of the parameters and then reconstruct the quantity from these simulations. As the number of parameters may be large, this becomes a high-dimensional reconstruction problem. In this talk, I will address the topic of reconstructing such unknown quantities using kernel-based reconstruction methods on sparse grids. I will give an introduction to the topic. After that, I will explain different reconstruction processes using kernel-based methods such as support vector machines and multi-scale radial basis functions and the so-called Smolyak algorithm. I will discuss techniques for deriving error estimates and show convergence results. If time permits, I will also give numerical examples. This talk is based upon joint work with R. Kempf (Bayreuth) and C. Rieger (Aachen/Bonn).

**Bio** Prof. Holger Wendland is Chair of Applied and Numerical Analysis at University of Bayreuth and serves as Editor for SIAM journal on Numerical Analysis, Mathematical Methods in the Applied Sciences, Numerical Algorithms, and Dolomite Research Notes on Approximation. Mainly research contains Numerical Analysis, Applied Mathematics, Scientific Computing, and new discretization methods for the numerical solution of partial differential equations. He has published more than 70 papers, and 3 monographs in these areas.

www and google scholar

---

PT

## The modification and applications of Prony method based on the data-driven model

Ran Zhan and **Zongmin Wu**
Shanghai University of Finance and Economics, China

This talk focuses on the modifications, generalizations and applications of Prony method based on the data-driven model. In the first part, we consider a modification of Pronys method and give its interpretation as a maximum likelihood method. With the help of an explicitly derived Jacobian matrix, we review the Levenberg-Marquardt algorithm and a new iterated gradient method (IGRA). We compare this approach with the iterative quadratic maximum likelihood (IQML). We propose two further iteration schemes based on simultaneous minimization (SIMI) approach. In the second part, a data-driven model based on a generalized Prony's method is introduced to simulate a physical problem.As a result, the nonuniform oscillations of the motion of a falling sphere in the non-Newtonian fluid is described by a jerk equation.This differential/algebraic equation is established by learning the experimental data of time vs velocity with Multiquadric (MQ) quasi-interpolation scheme, the generalized Prony method, and the regularization and variable selection method.

**Bio**

MathScinet

---

## Ten Year Development of Reproducing Kernel Banach Spaces: 2009-2019

PT

Yuesheng Xu

Old Dominion University, USA

Since the publication of the first paper "Reproducing kernel Banach spaces for machine learning" by Haizhang Zhang, Yuesheng Xu, Jun Zhang, in 10 (2009), Journal of Machine Learning Research, we have witnessed the rapid development of the theory of the reproducing kernel Banach spaces (RKBSs) and its applications. In this talk we shall review the development of RKBSs and their applications in machine learning and sampling theory in the last ten years. We shall also discuss possible future research directions of this young yet active area.

**Bio** Prof. Yuesheng Xu was "Guohua" Chair Professor at Sun Yat-sen University(2009-2017), Professor of Mathematics at Syracuse University(2003-2013) and Eberly Chair Professor at West Virginia University(2001-2003). He is currently Professor of Data Science at Old Dominion University. He served as the Managing Editor of Advances in Computational Mathematics(1999-2011) and is serving as an associate editor for 8 international Journals. His research interests include machine learning, Imaging processing, signal processing and optimization. He has published more than 150 papers in authoritative journals, such as *Menoirs of American Mathematics Society, SIAM Journal on Numerical Analysis, Numerische Mathematik, Inverse Problems, IEEE Transactions of Medical Imaging.* He also served as Vice President of the Chinese Association of Computational Mathematics(2014-2019).

www and google scholar

---

PT

# Information Geometry of Statistical Inference

Jun Zhang

University of Michigan-Ann Arbor, USA

Information Geometry is the differential geometric study of the manifold of probability models, and promises to be a unifying geometric framework for investigating statistical inference, information theory, machine learning, etc. Instead of using metric for measuring distances on such manifolds, these applications often use "divergence functions for measuring proximity of two points (that do not impose symmetry and triangular inequality), for instance Kullback-Leibler divergence, Bregman divergence, f-divergence, etc. Divergence functions are tied to generalized entropy (for instance, Tsallis entropy, Renyi entropy, phi-entropy) and cross-entropy functions widely used in machine learning and information sciences. It turns out that divergence functions enjoy pleasant geometric properties C they induce what is called "statistical structure on a manifold: a Riemannian metric together with a pair of conjugate affine connections. We investigate this geometry, and its connection to the geometry of optimal transport and also to the geometry governing Lagrange and Hamiltonian systems (geometric mechanics). We point out a universal duality principle underlying these seemingly unrelated geometries. The surprisingly rich differential geometric structure of statistical manifolds open up the intriguing possibility of unifying statistical inference, information, and machine learning using the language of geometry.

**Bio** Prof. Jun Zhang is a full professor of department of mathematics and department of psychology of University of Michigan-Ann Arbor. He is the director of M3 Laboratory(Mind. Machine, Machematics) and founding editor of *Information Geometry*, associate editor of *Journal of Mathematical Psychology*. He received a BS degree in theoretical physics from Fudan University in 1985, and received a PhD in Neurobiology from University of California Berkeley. His research interests include mathematical psychology and computational neuroscience, which include neural network theory and reinforcement learning, dynamical analysis of the nervous system, computational vision, choice-reaction time model, Bayesian decision theory and game theory.

www

---

## Approximation Theory of Deep Learning

PT

Dingxuan Zhou
City University of Hong Kong, China

Deep learning has been widely applied and brought breakthroughs in speech recognition, computer vision, and many other domains. The involved deep neural network architectures and computational issues have been well studied in machine learning. But there lacks a theoretical foundation for understanding the approximation or generalization ability of deep learning models with network architectures such as deep convolutional neural networks (CNNs) with convolutional structures. The convolutional architecture gives essential differences between the deep CNNs and fully-connected deep neural networks, and the classical approximation theory of fully-connected networks developed around 30 years ago does not apply. This talk describes an approximation theory of deep CNNs. In particular, we show the universality of a deep CNN, meaning that it can be used to approximate any continuous function to an arbitrary accuracy when the depth of the neural network is large enough. Rates of approximation are provided. Our quantitative estimate, given tightly in terms of the number of free parameters to be computed, verifies the efficiency of deep CNNs in dealing with large dimensional data.

**Bio** Prof. Dingxuan Zhou is a Chair professor of department of matheamtics, City University of Hong Kong. His research interests include learning theory, wavelet analysis and approximation theory. He has published over 100 research papers and is serving on editorial board of the international journals *Advances in Computational Mathematics*, '*Aanalysis and Applications*, *Complex Analysis and Operator Theory* and *Journal of Computational Analysis and Applications*. He has co-organised over 10 international conferences and conducted more than 20 research grants.

www

---

## Numerical homogenization and kernel-based methods

PT

Lei Zhang
Shanghai Jiaotong University, China

Numerical homogenization concerns the finite dimensional approximation of the solution space of, for example, divergence form elliptic equation with $L^\infty$ coefficients which allows for nonseparable scales. Standard methods such as finite-element method with piecewise polynomial elements can perform arbitrarily badly for such problems. Interestingly, numerical homogenization can be reformulated as a Bayesian inference or kernel based problem. The corresponding numerical method precomputes $H^{-d}$ localized bases on patches of size $H\log(1/H)$, and has an optimal converge rate of $O(H)$. The localization is due to the exponential decay of the corresponding fine scale solutions with Lagrange type constraints. Furthermore, the kernel based formulation can be used, for example, in the elliptic inverse problem and image segmentation.

**Bio** Prof. Lei Zhang's research interests include include partial differential equations, numerical analysis, multiscale modeling and analysis, as well as wide applications in fields such as materials science, geophysics, and biological sciences. His research results were published in first class journals such as Comm. Pure. Appl. Math., SIAM J. Numer. Anal., SIAM MMS (Multiscale Modeling and Simulation), etc. He was the gold medalist of International Mathematical Olympiad (1993), and W.P. Carey Prize winner in applied mathematics (2007). His research has been sponsored by the national 1000 talent plan, NSFC, and Intel.

www and google scholar

---

## Learning Rates of Constrained Covariance on Reproducing Kernel Hilbert spaces

Dirong Chen

Beihang University, China

This talk considers learning rates of estimators associated with kernel cross-covariance. For kernel cross-covariance operators, we bound the summation of estimation errors of empirical singular functions in terms of the estimation error of empirical cross-covariance. It is much tighter than a result on perturbation of operators as the latter bounds only each estimation error of empirical singular function individually.

**Bio** Prof. Dirong Chen is "Lantian Scholar" at Beihang University, and "Chutian Scholar" at Wuhan Textile University. His research interests include approximation theory, information processing, and statistical learning theory. He has hosted 2 national "863" projects, 1 national "973" sub-project, and 6 National Natural Science Foundation. More than 70 papers have been published in authoritative journals.

www

---

## Recovery and Approximation of High dimensional Data Via Convex and Non-convex Minimization

Wengu Chen

Institute of Applied Physics and Computational Mathematics, China

In this talk, we consider the recovery conditions for the exact recovery of data with structures in the noiseless setting and approximation in the noisy case from incomplete information. The structure includes sparsity, the context when some prior information on the support of the signals is available. Moreover, we consider the optimality or sharpness of these sufficient conditions

**Bio** Prof. Wengu Chen is director of the Department of Basic Mathematics and Applied Mathematics at Institute of Applied Physics and Computational Mathematics. The research interests include harmonic analysis, compressed sensing, channel codec, and nonlinear dispersion equation theory and application. Prof. Chen is Doctor supervisor at University of Science and Technology of China, and reviewer of American mathematics reviews. Many papers have been published in authoritative journals.

www

---

## A high-precision thermal imaging approach based on Bayesian variation approximation for electrical device prognosis

Ning Chu

Zhejiang University, China

Remote monitoring and early warning of thermal source abnormality play more and more important roles in fire prevention for the museums and historical monuments (Notre dame de Paris e.g.), metro and electric vehicle (Tesla e.g.) etc. However, conventional thermal imaging techniques cannot obtain the accurate temperature distribution of thermal sources in the far-fields. This is due to the fact that true temperature of thermal sources, according to heat radiation model, depends on many complex factors such as background temperature, environment humidity and surface emissivity. To solve the above challenge, we propose a Bayesian deep learning approach in thermal remote imaging with hyper-resolution. And mixture Gaussian priors are employed to model the temperature distribution of thermal sources, as well as background temperature. Meanwhile, sparsity-enforcing prior of temperature gradient is also utilized for spatial hyper-resolution. Moreover, the environment humidity and surface emissivity in heat radiation model can be studied by latent variables in Bayesian Hierarchy Network, so that these two important parameters can be estimated by maximizing the entropy of variational Bayesian inference. Through this Bayesian deep learning framework (sampling-training-updating), temperature mapping of hot sources can be accurately obtained (about 0.5 degree Celsius variation) as far as 5-10 meters way through a cost-effective infra-red camera (¡100 Euros, 7 degree Celsius variation) . Even without knowing the exact environment information, proposed approach is able to learn rapidly from remote monitoring data about heat radiation parameters. Based on proposed approach, a carry-on system of remote thermal imaging system has been invented for monitor the abnormal heating in metro system in Guangzhou city China.

**Bio** Dr. Ning Chu received the Bachelor in information engineering from the National University of Defense Technology in 2006. He obtained the master and PhD in automatic signal, and image processing from the University of Paris Sud, France in 2010 and 2014 respectively. He then won the positions of scientific collaborator in cole Polytechnique Fdrale de Lausanne, Switzerland, and senior lecturer in Zhejiang Unviersity. His research interests mainly focus on acoustic source imaging, Bayesian deep learning in condition monitoring and inverse problem applied in super resolution imaging. He has published more than 22 peer-reveiwer journal papers, invited for lectures by top international scientific conferences, own 7 China patents and 6 software copyrights.

www

---

## Robust Functional Linear Regression in RKHS

Jun Fan

Hong Kong Baptist University, China

Functional data analysis is concerned with inherently infinity dimensional data such as curves or images. It attracts more and more attentions due to its successful applications in many areas. In this talk we consider a re-producing kernel Hilbert space (RKHS) approach to robust functional linear regression. The proposed estimator can achieve the minimax optimal rate of convergence. Despite of the infinite dimensional nature of functional data, we show that the algorithm is easily implementable.

**Bio** Dr. Jun Fan obtained Doctor degree at the City University of Hong kong in 2013. Before joining the Hong Kong Baptist University, he worked as a research associate at University of Wisconsin-Madison. His research interests include data science, statistical machine learning, information theory and applications in medical informatics. He has published many papers in these areas.

www

---

## Logarithmic Stability for Coefficients Inverse Problem of Coupled Evolution Equations

Fangfang Dou

University of Electronic Science and Technology of China, China

This talk investigates the identification of coefficients in two coupled evolution equations, especially for Schrödinger equations and wave equations, with an observation on one component of the solution. Based on the the Carleman estimate for coupled Schrödinger equations and coupled heat equations, and the F.B.I. transform, a Hölder type stability for identifying two coeffcients in the system simultaneously is obtained, with the measurements only in a nonempty open subset of the domain where the equations evolved.

**Bio** Prof. Fangfang Dou obtained doctor degree in Lanzhou University in 2009 and her research interests include inverse problems. She has published many papers and conducted several research grants.

---

## CHEBYSHEV-TYPE CUBATURE FORMULAS ON REGULAR DOMAINS

Han Feng

City University of Hong Kong

In this talk, I will present the strict Chebyshev-type cubature formula (CF) (i.e., equal weighted CF) for doubling weights on the unit sphere equipped with the usual surface Lebesgue measure and geodesic distance. Our main interest is on the minimal number of nodes required in a strict Chebyshev-type CF. Precisely, given a normalized doubling weight on unit sphere, we will establish the sharp asymptotical estimates of the minimal number of distinct nodes which admits a strict Chebyshev-type CF. If, in addition, the weight function is essentially bounded, the nodes involved can be configured well separately in some sense. The proofs of these results rely on constructing new convex partitions of the unit sphere that are regular with respect to the weight. The weighted results on the unit sphere also allow us to establish similar results on strict Chebyshev-type CFs on the unit ball and the standard simplex. This is a joint work with Professor Feng DAI.

www

---

## A new class of stochastic Bernstein polynomials and Ditzian-Totik type modulus of smoothness

**Qinjiao Gao**, Xingping Sun, Shenggang Zhang

School of Business, Dalian University of Foreign Languages, Dalian, 116044, China
Department of Mathematics, Missouri State University, Springfield, MO 65897, USA
Department of Health Statistics, Dalian Medical University, Dalian, 116044, China

We introduce a family of symmetric stochastic Bernstein polynomials based on order statistics, and establish the order of convergence in probability in terms of the second order Ditzian-Totik type modulus of smoothness on the interval $[0, 1]$, which epitomizes an optimal pointwise error estimate for the classical Bernstein polynomial approximation. Monte Carlo simulation results (presented at the end of the article) show that this new approximation scheme is efficient and robust.

---

## Distributed Learning with Minimum Error Entropy Principle

Xin Guo

The Hong Kong Polytechnic University, China

Minimum Error Entropy (MEE) principle is an important approach in Information Theoretical Learning (ITL). It is widely applied and studied in various fields for its robustness to noise. In this paper, we study a reproducing kernel-based distributed MEE algorithm, DMEE, which is designed to work with both fully supervised data and semi-supervised data. With fully supervised data, our proved learning rates equal the minimax optimal learning rates of the classical pointwise kernel-based regressions. Under the semi-supervised learning scenarios, we show that DMEE exploits unlabeled data effectively, in the sense that first, under the settings with weaker regularity assumptions, additional unlabeled data significantly improves the learning rates of DMEE. Second, with sufficient unlabeled data, labeled data can be distributed to many more computing nodes, that each node takes only O(1) labels, without spoiling the learning rates in terms of the number of labels.

**Bio** Dr. Xin Guo got Ph.D at the City University of Hong Kong in 2011 and his research interest is statistical learning theory and many papers have been published in this area.

www

---

## Lower bounds of eigenvalues for Steklov eigenvalue problems

Qin Li

Beijing Technology and Business University, China

For the eigenvalue problem of the Steklov differential operator, by following Liu's approach, an algorithm utilizing the conforming finite element method (FEM) is proposed to provide guaranteed lower bounds for the eigenvalues. The proposed method requires the a priori error estimation for FEM solution to nonhomogeneous Neumann problems, which is solved by constructing the hypercircle for the corresponding FEM spaces and boundary conditions. As an application of proposed eigenvalue bounds for the Steklov operator, the optimal and explicit bound for the constant in the trace theorem is evaluated for several domains. Furthermore, Kansa's method is used to solve Steklov eigenvalue problems and it can also provide the lower bounds of the exact eigenvalues.

**Bio** Prof. Qin Li got Ph.D in Computational Mathematics at Academy of Mathematics and Systems Science Chinese Academy of Sciences in 2013. Her research interests include numerical methods for partial differential equations, finite element methods, and fast algorithms based on eigenvalues. The research results have been published in authoritative journals.

www and google scholar

---

## Kernel functions and their fast evaluations

Yingzhou Li
Duke University, USA

Kernel matrices are popular in machine learning and scientific computing, but they are limited by their quadratic complexity in both construction and storage. It is well-known that as one varies the kernel parameter, e.g., the width parameter in radial basis function kernels, the kernel matrix changes from a smooth low-rank kernel to a diagonally-dominant and then fully-diagonal kernel. Low-rank approximation methods have been widely-studied, mostly in the first case, to reduce the memory storage and the cost of computing matrix-vector products. Here, we use ideas from scientific computing to propose an extension of these methods to situations where the matrix is not well-approximated by a low-rank matrix. In particular, we construct an efficient block low-rank approximation methodwhich we call the Block Basis Factorizationand we show that it has O(n) complexity in both time and memory. Our method works for a wide range of kernel parameters, extending the domain of applicability of low-rank approximation methods, and our empirical results demonstrate the stability (small standard deviation in error) and superiority over current state-of-art kernel approximation algorithms.

**Bio** Dr. Yingzhou Li's research interests include hierarchical matrix and its applications, fast algorithms for elliptic PDEs, Fourier integral equations and high-frequency waves, distributed-memory algorithm and software, fast evaluation of kernel matrices from high dimensions data, 3D reconstruction problems, computational problems in density functional theory, and computational problems in many-body Schrödinger equation. The related results were published in authoritative journals.

www and google scholar

---

## Machine Learning in Reproducing Kernel Banach Spaces of Continuous Functions

Rongrong Lin
Sun Yat-sen University, China

Kernel methods have been broadly applied to machine learning and other questions of approximating an unknown continuous function from its finite sample data. To ensure arbitrary accuracy of such an approximation, various denseness conditions are imposed on the selected kernels. For this purpose, the existing work explores the universality of positive definite kernels in the Banach space $C_0(X)$ of all continuous functions vanishing at infinity on a locally compact Hausdorff space X. In other words, the denseness of reproducing kernel Hilbert spaces was considered in $C_0(X)$. In this talk, we will directly treat $C_0(X)$ as a reproducing kernel Banach space since point evaluation functionals on such a fundamental Banach space are meaningful and bounded. More importantly, we see that the reproducing kernels of the reproducing kernel Banach space $C_0(X)$ are closely related to universal kernels, characteristic kernels, and $c_0$-universal kernels in the literature. Concrete examples of positive definite or non-positive definite reproducing kernels are given. Finally, we discuss the representer theorem for the learning schemes in $C_0(X)$.

**Bio** Dr. Rongrong Lin got Ph.D at the Sun Yat-Sen University in 2017 and has been engaging in the research of kernel methods in machine learning and time-frequency analysis. He has published many papers have been published in these areas and conducted two research grants from Natural Science Foundation of China and Fundamental Research Funds for the Central Universities.

www google scholar

---

## A data dependent stencil selection algorithm for RBF-FD

Leevan Ling

Hong Kong Baptist University, China

In this paper, a new RBF-generalized finite difference stencil selection algorithm is proposed. For classical local radial basis function (RBF) approximation, the key for success is to choose suitable stencils for the sake of numerical stability and accuracy. For functions with rapid variations or discontinuities, the commonly used stencil selection approaches may cause numerical instability and oscillatory solutions. Therefore, to circumvent such limitation, we propose a newdata-dependent stencil selection algorithm. Numerical examples are provided to demonstrate that the proposed stencil selection algorithm outperforms the traditional nearest neighborhood stencil selection approaches.

**Bio** Prof. Leevan Ling's research interests include numerical analysis, partial differential equations, meshfree methods, adaptive greedy algorithm, and inverse problems. More than 60 papers and 10 monographs have been published in these areas. He is serving as the editorial board for Engineering Analysis with Boundary Elements, and Advances in Applied Mathematics and Mechanics.

www and google scholar

---

## The choice of the shape parameter in the collocation method

Lintian Luh

Providence University, Republic of China

In this talk, a concrete method of of choosing optimally the shape parameter in the smooth RBFs multiquadrics and inverse multiquadrics for solving PDEs is presented. We use the newly born MN curve theory to choose it and achieve unprecedentedly high accuracy. The problem of ill-conditioning is overcome by the arbitrarily precise computation software Mathematica. This is a first attempt of combining the MN curve theory and numerical PDEs. From this starting point, fruitful future research may be foreseen.

**Bio** Prof. Lintian Lun's research interests include the theory and applications of radial basis functions, in particular, the selection rule for the shape parameter. As an independent author, he has published many papers in these areas.

www

---

## The Finite Point Method and Its Applications to Diffusion Problems

Guixia Lv and Longjun Shen

Institute of Applied Physics and Computational Mathematics, China

In the last few decades, to avoid difficulties in mesh generation for complex problems, meshless methods have emerged and received intensive attention in computational scientific and engineering fields. In various meshless methods, one of the key issues is how to approximate a function on scattered points. To deal with this issue, one kind of methods employ much more points than unknowns, which often have good robustness, however, result in large discrete stencils reducing computational efficiency; another kind of methods employ just adequate points in terms of requirement of accuracy, which are simple and straightforward to implement, but often suffer from instability.

In the present paper, we consider the finite point method (FPM) based on directional differences on 2D scattered point distributions. It can be viewed as a generalization to the classical finite difference method. Distinguished from most of meshless methods often involving dozens of neighboring points, the FPM needs only five neighbors of the point under consideration, and gives explicit numerical formulae for approximations to directional differentials with expected accuracy. By virtue of explicit expressions, solvability conditions of numerical derivatives are rigorously derived as general guiding principles for selecting neighboring points avoiding singularity in computing derivatives, the estimation of the discriminant function for the solvability of numerical differentials is given, and the estimations of elements in the directional difference coefficient matrix are presented, which exclude the existence of singularity.

Applying the FPM to diffusion problems, we design a new algorithm to select neighboring points on 2D scattered point distributions, which has good stability in numerical computation, and well reflects variations of gradients of physical quantities. By this, we propose a new scheme for solving diffusion problems with nonlinear conductivities, which has advantages as follows: it leads to minimal stencils, gives coefficients of the resulted scheme explicitly avoiding solving matrix equations, and is first-order consistent to the corresponding diffusion equation. Finally, we present several numerical examples with different computational domains and different point distributions. The results show the good performance in accuracy and efficiency of the proposed approaches.

**Bio**

---

## Kernel-based Basyesian Optimization of Multiple Objectives

Michael McCourt
SIGOPT, USA

Kernels provide an effective tool for approximating scattered data in arbitrary dimensions; they appear in the functional analysis, spatial statistics and machine learning communities. In this talk, we will review how kernel approximations appear in the context of Gaussian processes/random fields, and how such approximations give rise to the Bayesian optimization framework: a sample-efficient strategy for black-box optimization. We will then introduce our adaptation of this strategy for identifying the Pareto frontier of multiple competing objectives, as well as variations which enable users to interactively focus on the most desired components of the frontier in a more sample-efficient fashion. This will be demonstrated on examples from deep learning and materials science.

**Bio** Michael is a member of the research engineering team at SigOpt, with interests in kernel-based approximation theory, Bayesian optimization, spatial statistics, and matrix computations. At SigOpt, he applies his expertise in leading the development of an enterprise-grade Bayesian optimization platform and facilitating collaborations with academic partners. Prior to joining SigOpt, he spent time in the math and computer science division at Argonne National Laboratory and was a visiting assistant professor at the University of Colorado-Denver and Illinois Institute of Technology. Mike holds a Ph.D. in applied math from Cornell.

www

---

## Distances and divergences between Gaussian measures in reproducing kernel Hilbert spaces: from Information Geometry to Optimal Transport

Ha Quang Minh

REIN, Japan

Distances and divergences between probability measures play a crucial role in s-tatistics, machine learning, and many other fields in science and engineering. This talk will survey some of the recent results in the generalization of the distances and divergences between Gaussian measures on Euclidean space to the generally infinite-dimensional setting of reproducing kernel Hilbert spaces (RKHS). Examples include Fisher-Rao distance, Kullback-Leibler divergence, Maximum Mean Discrepancy, and Wasserstein distance. A great advantage of the RKHS setting is that many quantities can obtained in closed forms via the corresponding kernel Gram matrices. We will also discuss recent formulations that unify many seemingly different, unrelated distances and divergences in a common framework. The mathematical exposition will be illustrated by numerical experiments.

**Bio** Prof. Ha Quang Minh is unit leader at the RIKEN center for advanced intelligence project, Tokyo. The research interests include mathematical foundations and algorithmic developments in machine learning, AI, computer vision, and image and signal processing, and problems in applied and computational functional analysis, and applied and computational differential geometry. Many papers have been published in these areas.

www and google scholar

---

## On the regularized functional regression

Prof. Sergei V. Pereverzyev

Johann Radon Institute For Computational and Applied Mathematics, (RICAM), Austria

Functional Regression (FR) involves data consisting of a sample of functions taken from some population. Most work in FR is based on a variant of the functional linear model first introduced by Ramsay and Dalzell in 1991. A more general form of polynomial functional regression has been introduced only quite recently by Yao and Mller (2010), with quadratic functional regression as the most prominent case. A crucial issue in constructing FR models is the need to combine information both across and within observed functions, which Ramsay and Silverman (1997) called replication and regularization, respectively. In this talk we are going to present a general approach for the analysis of regularized polynomial functional regression of arbitrary order and indicate the possibility for using here a technique that has been recently developed in the context of kernel supervised learning. Finally we are briefly discuss the application of FR in stenosis detection.

Joint research with S. Pereverzyev Jr. (Uni. Med. Innsbruck), A. Pilipenko (IMATH, Kiev) and V.Yu. Semenov (DELTA SPE, Kiev) supported by the consortium of Horizon-2020 project AMMODIT and the Austrian National Science Foundation (FWF).

**Bio** Prof. Sergei Pereverzyev's research interests include inverse and ill-posed problems, functional analysis, approximation theory, and complexity theory. More than 50 journal papers and 3 monographs have been published in these areas. He is senior fellow inverse problems and mathematical imaging.

Personal website

---

## Kernel-based Regression on Massive Data

Lei Shi

Fudan University, China

Kernel methods are attractive in data analysis as they can model nonlinear similarities between observations and provide means to rich representations, both of which are useful for the regression problems in general domains. Despite their popularity, they suffer from two primary inherent drawbacks. One drawback is the positive definiteness requirement of the kernel functions, which greatly restricts their applications to some real data analysis. The other drawback is their poor scalability in massive data scenarios. In this talk, we aim to address these two problems by considering the distributed and Nyström subsampling approaches for coefficient-based regularized regression. Distributed learning and Nyström subsampling are two effective approaches to analyze big data, which serve as standard tools for reducing computational complexity in machine learning problems where massive data sets are involved. Coefficient-based regularized regression can provide a simple paradigm for designing indefinite kernel methods. We show that combinations of these two schemes with coefficient-based regularized regression are not only computationally efficient but also statistically consistent with a mini-max optimal rates of convergence.

**Bio** Prof. Lei Shi got the Ph.D. in Applied Mathematics jointly awarded by City University of Hong Kong and University of Science and Technology of China in 2010. His research interests include learning theory and approximation theory. Many papers have been published in these areas.

www

---

## Monte Carlo Methods for $L_q$ $(2 \leq q \leq \infty)$ Approximation on periodic Sobolev spaces with mixed smoothness

Heping Wang

Capital Normal University, China

In this talk we consider multivariate approximation of compact embeddings of periodic Sobolev spaces of dominating mixed smoothness into the $L_q$, $2 < q \leq \infty$ spaces. Such Sobolev spaces are the reproducing kernel Hilbert spaces and are widely used in numerical analysis and approximation theory. We use linear Monte Carlo methods that use arbitrary linear information. We construct linear Monte Carlo methods and obtain explicit-in-dimension upper estimates. These estimates catch up with the rate of convergence.

www

---

## Approximating signals from random sampling in a reproducing kernel subspace

Jun Xian

Sun Yat-sen University, China

This paper studies the problem of approximating a signal in a reproducing kernel subspace from its random samples. We employ the iterative frame algorithm to solve this problem. The approximation error is derived. Due to randomness of the samples, the magnitude of the approximation error is a probabilistic estimate and the associated probability tends to one when the sampling size is of order $|\Omega| \log |\Omega|$ with $\Omega$ being the sampling region. We also obtain the approximation error when samples are corrupted by noise. Numerical analysis is conducted to illustrate the effectiveness of the proposed algorithm.

**Bio** Professor in applied harmonic analysis, deputy dean of school of mathematics, Sun Yat-sen University.

www1 and Personal website2

---

## Minimal Sample Subspace Leaning

Zhenyue Zhang, **Yuqing Xia**
National University of Singapore

Subspace learning is a challenging and complicated task in machine learning. This talk will introduce our current work on minimal subspaces of finite samples, based on which we build a primary frame for subspace learning. We discussed the existence and conditional uniqueness of the introduced minimal subspaces and provided equivalent description via rank sum of sample segmentations. According to the equivalent description, learning minimal subspaces can be modeled as an optimization problem based on rank-restricted self-expressiveness method. Theoretically, the proposed optimization can retrieve minimal subspaces when the underlying ones are not heavily intersected. The solution of the optimization can be achieve via a manifold conjugate gradient algorithm, which learns subspaces from finite samples.

arXiv

---

## Convergence Analysis of Kernel Based Large-margin Learning

Daohong Xiang
Zhejiang Normal University, China

In this talk we study binary classification problem associated with a family of loss functions called large-margin unified machines (LUM), which offers a natural bridge between distribution-based likelihood approaches and margin-based approaches. It also can overcome the so-called data piling issue of support vector machine in the high-dimension and low-sample size setting. To the best our knowledge, a systematic error analysis is still lacking for LUMs from the learning theory perspective. We establish some new comparison theorems for all LUM loss functions which play a key role in the further error analysis of large-margin learning algorithms. Quantitative convergence analysis has been carried out for these algorithms by means of a novel application of projection operators to overcome the technical difficulty. The rates are explicitly derived under priori conditions on approximation and capacity of the reproducing kernel Hilbert space.

**Bio** Prof. Daohong Xiang's research interests include statistical learning theory, robust statistics, and deep learning. She has publised many papers and monographs in these areas and conducted 5 research grants.

www

---

## Positive Definite Tensor Kernels for Scattered Data Interpolations

Qi Ye

South China Normal University, China

Kernel methods are attractive in data analysis as they can model nonlinear similarities between observations and provide means to rich representations, both of which are useful for the regression problems in general domains. Despite their popularity, they suffer from two primary inherent drawbacks. One drawback is the positive definiteness requirement of the kernel functions, which greatly restricts their applications to some real data analysis. The other drawback is their poor scalability in massive data scenarios. In this talk, we aim to address these two problems by considering the distributed and Nyström subsampling approaches for coefficient-based regularized regression. Distributed learning and Nyström subsampling are two effective approaches to analyze big data, which serve as standard tools for reducing computational complexity in machine learning problems where massive data sets are involved. Coefficient-based regularized regression can provide a simple paradigm for designing indefinite kernel methods. We show that combinations of these two schemes with coefficient-based regularized regression are not only computationally efficient but also statistically consistent with a mini-max optimal rates of convergence.

**Bio** Prof. Qi Ye is director of the Laboratory for Machine Learning and Computational Optimization at South China Normal University. His research interests include approximation theory in data analysis and machine learning, kernel-based approximation method, support vector machine, neural network, nonsmooth analysis, and convex analysis. The related results were published in authoritative journals such as Numerische Mathematik, Applied and Computational Harmonic Analysis, Journal of Mathematical Analysis and Applications, etc. His research has been sponsored by the national 1000 talent plan.

www and google scholar

---

## Reproducing kernel Banach spaces for multi-task learning

Haizhang Zhang

Sun Yat-sen University, China

Many applications of machine learning involve multiple tasks. Currently, the dominant methodology in handling a multi-task problem is to convert it into a series of single task problems. This is inefficient and does not make use of relation and knowledge between tasks. Vector-valued reproducing kernel spaces are the foundation for developing kernel methods for multi-task learning. We present our recent work on the theory and construction of vector-valued reproducing kernel Banach spaces. These spaces are useful for pursuing sparsity and estimating learning rates of multi-task learning algorithms. This talk is based on joint work with Liangzhi Chen, Rongrong Lin, Guohui Song, and Jun Zhang.

**Bio** Prof. Zhanghai Zhang's research interests include applicable and computational harmonic analysis, sampling theory, machine learning and reproducing kernel spaces, and mathematical foundation for time-frequency analysis, especially for empirical mode decomposition. He has published more than 34 papers and conducted 4 research grants.

www and MathScinet

---

## Information-based complexity of oscillatory integrals

Shun Zhang

Anhui University, China

We study optimal quadrature formulas for arbitrary weighted integrals and integrands from the Sobolev space $H^1([0,1])$. We obtain general formulas for the worst case error depending on the nodes $x_j$. A particular case is the computation of Fourier coefficients, where the oscillatory weight is given by $\rho_k(x) = \exp(-2\pi ikx)$. Here we study the question whether equidistant nodes are optimal or not. We prove that this depends on $n$ and $k$: equidistant nodes are *optimal* if $n \geq 2.7|k|+1$ but might be suboptimal for small $n$. As a comparison, we also get numerically the equidistance constants for large $|k|$ and $n$, respectively, for the average case error on uniformly sampling.

**Bio** Prof. Shun Zhang's research interests include information security, intelligent information processing, and structural optimization algorithm design and complexity analysis. He has published more than 10 papers in these areas and conducted 5 research grants.

Personal Webstie

---

## An accurate and efficient numerical method for nonlocal problems

Wei Zhao

City University of Hong Kong, China

Nonlocal models are based upon integro-differential equations, thereby avoiding s-patial derivatives, which are not defined at discontinuities, such as jumps, cracks, and any other singularities. As a result, these models have been widely used for fracture and anisotropic problems. In this talk, models arising from nonlocal diffusion and peridynamics are studied to obtain their numerical solutions accurately and efficiently. For this purpose, an effective numerical method, namely localized radial basis functions-based pseudo-spectral method (LRBF-PSM), is proposed for the spatial discretization, which inherits full advantages of the meshless nature, for instance, offering high accuracy and efficiency in complex geometries for both u-niform and non-uniform discretizations, extending the applicability of the method to multi-dimensions easily, etc. In addition, the LRBF-PSM combined with an exponential time-splitting technique is presented for the time-dependent peridynamic equation. Numerical results show that the proposed approach is a powerful tool for solving various nonlocal problems.

**Bio** Dr. Wei Zhao got the Ph.D. at Max Planck Institute for Dynamics of Complex Technical Systems in 2018. Her research interests include nonlocal theory and applications, radial basis functions, fast algorithms, and computational fluid dynamics with chemical reactions. The related results have been published in Applied Mathematical Modelling, Applied Mathematics and Computation, Computers & Mathematics with Applications, etc.

---

## Polynomial convergence order of stochastic Bernstein approximation

**Xuan Zhou** and Zongmin Wu

Fudan University

Recently, authors of (Wu et al., 2013) studied a Bernstein polynomial approximation scheme based on stochastic sampling, and obtained a sixth order moment estimate for the underlying random variable in terms of the modulus of continuity. In the current paper, we employ a new technique, and establish estimates for all the even order moments. our work gives a strong indication that the probabilistic convergence rate of the stochastic Bernstein approximation is exponential with respect to the modulus of continuity, which we leave as a conjecture at the end of the paper.

---

## Directional Tensor Product Complex Tight Framelets: Construction and Application

Xiaosheng Zhuang
City University of Hong Kong, China

Directional multiscale representation systems play an important role in both theory and applications. In theory, directional systems from "non-tensor-product" approach such as curvelets and shearlets have been proved to provide (nearly) optimal approximate rate for cartoon-like functions. In applications such as image/video denoising/inpainting, directional systems have been shown to outperform classical "tensor product" real-value wavelet/framelet systems. In this talks, we will focus on the construction and applications of a special type of directional multiscale representation systems (directional framelets) based on the "tensor product" approach, which we called "Tensor Product Complex Tight Framelets (TPCTF). By allowing complex-value framelets, we show that such directional TPCTFs can be easily built with band-limited framelet generators. By studying the frequency separation property of filters and employing optimization techniques, we show that such directional TPCTFs can be constructed to be with compactly supported framelet generators. Applications of both band-limited and compactly supported directional TPCTFs in image/video processing such as denoising and inpainting are demonstrated.

**Bio** Prof. Xiaosheng Zhuang is associate head of the Department of Mathematics at the City University of Hong Kong. His research interests include applied and computational harmonic analysis, sparse approximation and directional multiscale representation systems, compressed sensing, image/signal processing, and machine learning and pattern recognition. More than 40 papers have been published in these areas.

www and google scholar

---

# Author Index