

Colloquium 2017-2018

Spring 2018

February 13, 4:15 pm

Speaker: **Craig Guilbault** (Winsconsin - Milwaukee)

Topic: **Infinite boundary connected sums with applications to aspherical manifolds**

Abstract: A boundary connected sum $Q_1 \natural Q_2$ of n -manifolds is obtained by gluing Q_1 to Q_2 along $(n-1)$ -balls in their respective boundaries. Under mild hypotheses, this gives a well-defined operation that is commutative, associative, and has an identity element. In particular (under those hypotheses) the boundary connected sum $\natural_{i=1}^k Q_i$ of a finite collection of n -manifolds is topologically well-defined. This observation fails spectacularly when we attempt to generalize it to countable collections. In this talk I will discuss a pair of reasonable (and useful) substitutes for a well-definedness theorem for infinite boundary connected sums. An application of interest in both manifold topology and geometric group theory examines aspherical manifolds with exotic, i.e., not homeomorphic to \mathbb{R}^n , universal covers. We will describe examples different from those found in the classical papers by Davis and Davis-Januszkiewicz. Much of this work is joint with Ric Ancel and Pete Sparks.

February 15, 4:15 pm

Speaker: **Ya Su** (TAMU)

Topic: **Nonparametric Bayesian Deconvolution of a Symmetric Unimodal Density**

Abstract: We consider nonparametric measurement error density deconvolution subject to heteroscedastic measurement errors as well as symmetry about zero and shape constraints, in particular unimodality. The problem is motivated by genomics applications, where the observed data are estimated effect sizes from a regression on multiple genetic factors, as occurs in genome-wide association studies and in microarray applications. We exploit the fact that any symmetric and unimodal density can be expressed as a mixture of symmetric uniform densities, and model the mixing density using a Dirichlet process location-mixture of Gamma distributions. We do the computations within a Bayesian context, describe a simple scalable implementation that is linear in the sample size, and show that the estimate of the unknown target density is consistent. Within our application context of regression effect sizes, the target density is likely to have a large probability near zero (the near null effects) coupled with a heavy-tailed distribution (the actual effects). Simulations show that unlike standard deconvolution methods, our Constrained Bayesian method does a much better job of reconstruction of the target density. An application to a genome-wide association study to predict height shows similar results.

February 20, 4:15 pm

Speaker: **Theodore Voronov** (Manchester UK and Notre Dame)

Topic: **Supergeometry: from super de Rham theory and the Atiyah-Singer index theorem to microformal**

geometry

Abstract: Supergeometry is, roughly, the geometry associated with \mathbb{Z}_2 -graded algebra. In particular, for an odd element Q of a Lie superalgebra, the two options, $Q^2 \neq 0$ and $Q^2 = 0$, lead to “supersymmetry” and to “homological vector fields”, respectively.

The “super” notions were originally discovered as a language for describing fermions and bosons in quantum theory on an equal footing. They received their name from supersymmetric models where bosons and fermions are allowed to mix. Their mathematical roots can be traced in classical differential geometry, algebraic topology and homological algebra.

In the talk, I will introduce the basic ideas and describe some interesting results and links with other areas of mathematics. Among them: super de Rham theory and its connection with Radon transform and Gelfand's general hypergeometric equations; universal recurrence relations for super exterior powers and application to Buchstaber-Rees theory of (Frobenius) “ n -homomorphisms”; analytic proof of the Atiyah-Singer index theorem; homological vector fields as a universal language for deformation theory and bracket structures (such as homotopy Lie algebras, Lie algebroids, etc.) in mathematics and gauge systems in physics. An intriguing recent result (which started from a counterexample to a conjecture by Witten) concerns volumes of classical supermanifolds such as superspheres, super Stiefel manifolds, projective superspaces, etc. Upon some universal normalization, formulas for these “super” volumes turned out to be analytic continuations of formulas for ordinary manifolds. Another recent development is “microformal geometry”. This, roughly, is a theory that replaces ordinary maps between manifolds by certain “thick morphisms”, which induce non-linear pullbacks on functions, with remarkable properties. This is motivated by application to homotopy Poisson structures; but in general, it suggests a non-linear extension of the fundamental “algebra/geometry duality”. I hope to be able to tell about that as well.

March 15, 4:15 pm

Speaker: **Steve Marron** (UNC)

Topic: **OODA of Tree Structured Data Objects Using Persistent Homology**

Abstract: The field of Object Oriented Data Analysis has made a lot of progress on the statistical analysis of the variation in populations of complex objects. A particularly challenging example of this type is populations of tree-structured objects. Deep challenges arise, whose solutions involve a marriage of ideas from statistics, geometry, and numerical analysis, because the space of trees is strongly non-Euclidean in nature. Here these challenges are addressed using the approach of persistent homologies from topological data analysis. The benefits of this data object representation are illustrated using a real data set, where each data point is the tree of blood arteries in one person's brain. Persistent homologies gives much better results than those obtained in previous studies.

March 22, 4:15 pm

Speaker: **Stefan Steinerberger** (Yale)

Topic: **Four (confusing) Miracles in Analysis and Number Theory**

Abstract: CANCELLED

March 26, 4:15 pm

Speaker: **Jaehong Jeong** [KAUST (King Abdullah University of Science and Technology), Saudi Arabia]

Topic: **A Stochastic Generator of Global Monthly Wind Energy with Tukey g-and-h Autoregressive Processes**

Abstract: Quantifying the uncertainty of wind energy potential from climate models is a very time-consuming task and requires a considerable amount of computational resources. A statistical model trained on a small set of runs can act as a stochastic approximation of the original climate model, and be used to assess the uncertainty considerably faster than by resorting to the original climate model for additional runs. While Gaussian models have been widely employed as means to approximate climate simulations, the Gaussianity assumption is not suitable for winds at policy-relevant time scales, i.e., sub-annual. We propose a trans-Gaussian model for monthly wind speed that relies on an autoregressive structure with Tukey g-and-h transformation, a flexible new class that can separately model skewness and tail behavior. This temporal structure is integrated into a multi-step spectral framework that is able to account for global nonstationarities across land/ocean boundaries, as well as across mountain ranges. Inference can be achieved by balancing memory storage and distributed computation for a data set of 220 million points. Once fitted with as few as five runs, the statistical model can generate surrogates fast and efficiently on a simple laptop, and provide uncertainty assessments very close to those obtained from all the available climate simulations on a monthly scale. This is joint work with Yuan Yan, Stefano Castruccio, and Marc G. Genton.

Fall 2017**September 28, 4:30 pm**

Speaker: **Paul F. Velleman** (Cornell)

Topic: **Six Impossible Things Before Breakfast: Integrating Randomization Methods in the Introductory Statistics Course**

Abstract: The traditional introductory statistics course generally proceeds smoothly until the point where we have to admit to our students that the statistics they've been finding in their homework problems aren't really the answer; they are only an answer. They can believe that. Then we tell them that those answers may be random, but they aren't haphazard. In particular, if we gather the answers for all possible samples we can model them. They might accept that even though they can't see why it should be true. Then we claim to be able to estimate the parameters of those models and propose to use them for inference. Then, to top it all off, we admit that we were lying when we said the model for the mean was Normal, and that when the standard deviation is estimated (that is, almost always) or we're doing a regression, the model isn't Normal at all but only similar to the Normal.

Many students find all those results uncomfortable. They are not used to thinking that way. The Red Queen encountered by Alice may have been able to believe six impossible things before breakfast, but it is challenging to ask that of our students.

With computer technology, we can spread out these results across the first several weeks of the course to make it easier for students to understand and accept them. And then, by introducing bootstrap methods, we can carry these ideas into the discussion of inference.

I will discuss a syllabus that does just that and demonstrate some free software that supports the approach.

October 5, 4:15 pm

Speaker: **Alan Edelman** (MIT)

Topic: **Novel Computations with Random Matrix Theory and Julia**

This speaker's visit is part of the Dean's Speaker Series in Statistics and Data Science.

Abstract: Over the many years of reading random matrix papers, it has become increasingly clear that the phenomena of random matrix theory can be difficult to understand in the absence of numerical codes to illustrate the phenomena. (We wish we could require that all random matrix papers that lend themselves to computing include a numerical simulation with publicly available code.) Of course mathematics exists without numerical experiments, and all too often a numerical experiment can be seen as an unnecessary bother. On a number of occasions, however, the numerical simulations themselves have an interesting twist of their own. This talk will illustrate a few of those simulations and illustrate why in particular the Julia computing language is just perfect for these simulations. Some topics we may discuss:

1. "Free" Dice
2. Tracy Widom
3. Smallest Singular Value
4. Jacobians of Matrix Factorizations

(joint work with Bernie Wang)

October 19, 4:15 pm

Speaker: **Stefan Steinerberger** (Yale)

Topic: **Three (confusing) Miracles in Analysis and Number Theory**

Abstract: I will discuss three different topics at the intersection of Analysis and Number Theory:

1. improved versions of classical inequalities for functions on the Torus whose proof requires Number Theory,
2. mysterious interactions between the Hardy-Littlewood maximal function and transcendental number theory (I have a proof but I still don't understand what's going on) and
3. a complete mystery in an old integer sequence of Stanislaw Ulam (\$300 prize for an explanation).

November 16, 4:15 pm

Speaker: **Slawomir Solecki** (Cornell)

Topic: **Logic and homogeneity of the pseudoarc**

Abstract: Fraïssé theory is a method of classical Model Theory of producing canonical limits of certain families of finite structures. For example, the random graph is the Fraïssé limit of the family of finite graphs. It turns out that this method can be dualized, with the dualization producing projective Fraïssé theory, and applied to the study of compact metric spaces. The pseudoarc is a remarkable compact connected space; it is the generic, in a precise sense, compact connected subset of the plane or the Hilbert cube. I will explain the connection between the pseudoarc and projective Fraïssé limits.

November 30, 4:15 pm

Speaker: **Dong Xia** (Columbia U.)

Topic: **Quantum State Tomography via Structured Density Matrix Estimation.**

Abstract: The density matrices are positively semi-definite Hermitian matrices of unit trace that describe the state of a quantum system. Quantum state tomography (QST) refers to the estimation of an unknown density matrix through specifically designed measurements on identically prepared copies of quantum systems. The dimension of the associated density matrix grows exponentially with the size of quantum system. This talk is on the efficient QST when the underlying density matrix possesses structural constraints.

The first part is on the low rank structure, which has been popular in the community of quantum physicists. We develop minimax lower bounds on error rates of estimation of low rank density matrices, and introduce several estimators showing that these minimax lower bounds can be attained up to logarithmic terms. These bounds are established over all the Schatten norms and quantum Kullback-Leibler divergence. This is based on a series of work with Vladimir Koltchinskii.

The second part is built upon decomposable graphical models for quantum multi-qubits system. The goal is to reduce the sample complexity required for quantum state tomography, one of the central obstacles in large scale quantum computing and quantum communication. By considering the decomposable graphical models, we show that the sample complexity is allowed to grow linearly with the system size and exponentially with only the maximum clique size. This is based on a joint work with Ming Yuan.

December 4, 4:40 pm

Speaker: **Xuening Zhu** (Penn State U.)

Topic: **Network Vector Autoregression**

Abstract: We consider here a large-scale social network with a continuous response observed for each node at equally spaced time points. The responses from different nodes constitute an ultra-high dimensional vector, whose time series dynamic is to be investigated. In addition, the network structure is also taken into consideration, for which we propose a network vector autoregressive (NAR) model. The NAR model assumes each node's response at a given time point as a linear combination of (a) its

previous value, (b) the average of its connected neighbors, © a set of node-specific covariates, and (d) an independent noise. The corresponding coefficients are referred to as the momentum effect, the network effect, and the nodal effect respectively. Conditions for strict stationarity of the NAR models are obtained. In order to estimate the NAR model, an ordinary least squares type estimator is developed, and its asymptotic properties are investigated. We further illustrate the usefulness of the NAR model through a number of interesting potential applications. Simulation studies and an empirical example are presented.

December 7, 4:15 pm

Speaker: **Ziwei Zhu** (Princeton U.)

Topic: **Estimation of principal eigenspaces with decentralized and incomplete data**

Abstract: Modern data sets are often decentralized; they are generated and stored in multiple sources across which the communication is constrained by bandwidth or privacy. Besides, the data quality often suffers from incompleteness. This talk focuses on estimation of principal eigenspaces of covariance matrices when data are decentralized and incomplete. We first introduce and analyze a distributed algorithm that aggregates multiple principal eigenspaces through averaging the corresponding projection matrices. When the number of data splits is not large, this algorithm is shown to achieve the same statistical efficiency as the full-sample oracle. We then consider the presence of missing values. We show that the minimax optimal rate of estimating the principal eigenspace has a phase transition with respect to the observation probability, and this rate can be achieved by the principal eigenspace of an entry-wise weighted covariance matrix.

From:

<https://www2.math.binghamton.edu/> - **Department of Mathematics and Statistics, Binghamton University**

Permanent link:

<https://www2.math.binghamton.edu/p/seminars/colloquium/y2017-2018>

Last update: **2020/01/28 18:30**

