

## ABSTRACTS

### World Congress in Probability and Statistics

Toronto, July 11–15, 2016

The list of abstracts will only be available on-line, not printed. Participants should download a copy to use as needed. The document will be updated periodically. This version uploaded July 7, 2016.

#### Plenary Talks

All Talks:

A	B	C	D	E	F	G	H	I	J	K	L	M
N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Posters

## ABSTRACTS OF PLENARY TALKS

- **David R. Brillinger** (University of California, Berkeley)

Plenary talk: Tukey Lecture

Tuesday 4:15 – 5:05, McLeod Auditorium

*People are different*

#### Abstract:

- (1) Some items in JWTs vita
- (2) Some reminiscences and anecdotes  
copybaras
- (3) The Labs 1960 -1964  
Where did the E in EDA come from?  
A problem JWT put to DRB in 1978  
“Measure” (?) of the “peakiness” of periodogram:
  - (a) take cepstrum (based on log periodogram)
  - (b) average in octaves, laid  $\frac{1}{2}$  to the weather 22 Jan 78
- (4) Environmental Risk Analysis  
Propagating wildfire boundaries
- (5) ‘Sunset salvo’

- **Pierre Del Moral** (INRIA)

Plenary talk: IMS Medallion Lecture

Thursday 5:10 – 6:00, McLeod Auditorium

*An introduction to mean field particle methods*

**Abstract:** In the last three decades, there has been a dramatic increase in the use of Feynman-Kac type particle methods as a powerful tool in real-world applications of Monte Carlo simulation in computational physics, population biology, computer sciences, and statistical machine learning.

The particle simulation techniques they suggest are also called resampled and diffusion Monte Carlo methods in quantum physics, genetic and evolutionary type algorithms in computer sciences, as well as Sequential Monte Carlo methods in Bayesian statistics, and particle filters in advanced signal processing.

These mean field type particle methodologies are used to approximate a flow of probability measures with an increasing level of complexity. This class of probabilistic models includes conditional distributions of signals with respect to noisy and partial observations, non absorption probabilities in Feynman-Kac-Schrödinger type models, Boltzmann-Gibbs measures, as well as conditional distributions of stochastic processes in critical regimes, including quasi-invariant measures and ground state computations.

This lecture presents a pedagogical introduction to the stochastic modeling and the theoretical analysis of these sophisticated probabilistic models.

We shall discuss the origins and the mathematical foundations of these particle stochastic methods, as well as their applications in rare event analysis, signal processing, mathematical finance and Bayesian statistical inference. We illustrate these methods through several applications including random walk confinements, particle absorption models, nonlinear filtering, stochastic optimization, combinatorial counting and directed polymer models.

- **Frank den Hollander** (Leiden University, The Netherlands)

Plenary talk: IMS Medallion Lecture

Friday 8:30 – 9:20, McLeod Auditorium

*Metastability for interacting particle systems*

**Abstract:** Metastability is the phenomenon where a system, under the influence of a stochastic dynamics, moves between different subregions of its state space on different time scales. Metastability is encountered in a wide variety of stochastic systems. The challenge is to devise realistic models and to explain the experimentally observed universality that is displayed by metastable systems, both qualitatively and quantitatively.

In statistical physics, metastability is the dynamical manifestation of a first-order phase transition. In this talk I give a brief historical account of metastability in this context. After that I describe the metastable behaviour of one particular model, namely, the Widom-Rowlinson model on a two-dimensional torus subject to a Metropolis stochastic dynamics. In this model, particles are randomly created and annihilated inside the torus as if the outside of the torus were an infinite reservoir with a given chemical potential. The particles are viewed as points carrying disks, and the energy of a particle configuration is equal to the volume of the union of the disks, called the “halo” of the configuration. Consequently, the interaction between the particles is attractive.

We are interested in the metastable behaviour at low temperature when the chemical potential is supercritical. In particular, we start with the empty torus and are interested in the first time when we reach the full torus, i.e., the torus is fully covered by disks. In order to achieve the transition from empty to full, the system needs to create a sufficiently large droplet of overlapping disks, which plays the role of a “critical droplet” that triggers the crossover. In the limit as the temperature tends

to zero, we compute the asymptotic scaling of the average crossover time, show that the crossover time divided by its average is exponentially distributed, and identify the size and the shape of the critical droplet. It turns out that the critical droplet exhibits “surface fluctuations”, which need to be understood in order to obtain a fine estimate of the crossover time.

Based on joint work with Sabine Jansen (Bochum), Roman Kotecky (Warwick) and Elena Pulvirenti (Leiden).

- **Vanessa Didelez** (Professor of Statistics and Causal Inference Leibniz Institute & Department of Mathematics, University of Bremen, Germany)

Plenary talk: IMS Medallion Lecture

Thursday 4:15 – 5:05, McLeod Auditorium

*Causal reasoning for events in continuous time*

**Abstract:** Dynamic associations among different types of events in continuous time can be represented by local independence graphs as developed by Didelez (2008). Intuitively we say that a process is locally independent of another one if its short-term prediction is not improved by using the past of the other process, similar to Granger-non-causality; the graphical representation uses nodes for processes or events and the absence of a directed edge for local independence. Important independence properties can be read off these -possibly cyclic- graphs using delta-separation (Didelez, 2006) which generalises d-separation from DAGs. In related work, Røysland (2011, 2012) showed how causal inference based on inverse probability weighting (IPW), well known for longitudinal data (Robins et al., 2000), can be extended to the continuous-time situation using a martingale approach.

In the work that I will present at the Medallion lecture (joint work with Kjetil Rysland, Odd Aalen and Theis Lange), we start by defining causal validity of local independence graphs in terms of interventions, which in the context of events in time take the form of modifications to the intensities of specific processes, e.g. a treatment process; causal validity is given if the specification of the dynamic system is rich enough to model such an intervention. (similar to what is known as ‘modularity’ for causal DAGs). We then combine the above previous developments to give graphical rules for the identifiability of the effect of such interventions via IPW; these rules can be regarded as characterising ‘unobserved confounding’. Re-weighting then simply replaces the observed intensity by the one given by the intervention of interest. For this to be meaningful, causal validity and identifiability are crucial assumptions. Our approach can be regarded as the time-continuous version of Dawid & Didelez (2010), who develop a decision theoretic approach for sequential decisions in longitudinal settings and use a graphical representation with influence diagrams that include decision nodes; specifically causal validity is analogous to the extended stability of Dawid & Didelez (2010). As an aside, we find that it is helpful to also use causal reasoning when faced with censoring as the target of inference can often be regarded as the population in which censoring is prevented, i.e. its intensity is set to zero. We apply our theoretical results to the example of cancer screening in Norway. ”

- **Arnaud Doucet** (University of Oxford)

Plenary talk: IMS Medallion Lecture

Monday 4:15 – 5:05, McLeod Auditorium

*On pseudo-marginal methods for Bayesian inference in latent variable models*

**Abstract:** The pseudo-marginal algorithm is a popular variant of the Metropolis–Hastings scheme which allows us to sample asymptotically from a target probability density when we are only able to estimate unbiasedly an un-normalized version of it. It has found numerous applications in Bayesian statistics as there are many latent variable models where the likelihood function is intractable but can be estimated unbiasedly using Monte Carlo samples. In this talk, we will first review the pseudo-marginal algorithm and show that its computational cost is for many common applications quadratic in the number of observations at each iteration. We will then present a simple modification of this methodology which can reduce very substantially this cost. A large sample analysis of this novel pseudo-marginal scheme will be presented.

- **Christina Goldschmidt** (Department of Statistics & Lady Margaret Hall, University of Oxford)  
Plenary talk: IMS Medallion Lecture

Tuesday 5:10 – 6:00, McLeod Auditorium

*Scaling limits of critical random trees and graphs*

**Abstract:** I will begin by surveying some recent developments in the theory of scaling limits for random trees and graphs. The overarching aim is to understand how random tree and graph structures on  $n$  vertices may be rescaled so that, as  $n$  tends to infinity, we obtain meaningful limits in distribution. These limits are usually formulated as random metric spaces, and are themselves fascinating mathematical objects. The prototypical example is Aldous’ Brownian continuum random tree, which is the scaling limit of (for example) a uniform random labelled tree on  $n$  vertices.

Consider first a Galton-Watson forest, that is the sequence of family trees of a collection of independent and identically distributed Galton-Watson processes. We assume that the offspring distribution is critical. It is possible to encode the forest in terms of a random walk which (thanks to the criticality condition) is centred. Assuming additionally that the offspring distribution is in the domain of attraction of a stable law with parameter  $\alpha \in (1, 2]$ , this random walk converges (after a suitable rescaling) to a spectrally positive  $\alpha$ -stable Lévy process (Brownian motion if  $\alpha = 2$ ). Foundational work of Duquesne, Le Gall and Le Jan showed that the excursions of this Lévy process above its running infimum may be interpreted as encoding a forest, composed of continuum random trees, which I will refer to as a stable forest. In the second part of the talk, I will turn to work in progress with Guillaume Conchon-Kerjan on the model of a uniform random graph with independent and identically distributed vertex-degrees. The condition for criticality in this setting is  $E[D^2] = 2E[D]$ , and we assume additionally that  $P(D = k) \sim ck^{-(\alpha+2)}$  as  $k$  tends to infinity, for some  $\alpha \in (1, 2)$ . In this situation, it turns out that the largest components have sizes on the order of  $n^{\alpha/(\alpha+1)}$ . Building on earlier work of Joseph, we show that the components have scaling limits which can be related to a stable forest via an absolute continuity relation. This gives a natural generalisation of the scaling limit for the Erdos-Renyi random graph which I obtained in collaboration with Louigi Addario-Berry and Nicolas Broutin a few years ago (extending results of Aldous), and complements recent work on random graph scaling limits of various authors including Bhamidi, Broutin, Duquesne, van der Hofstad, van Leeuwen, Riordan, Sen, M. Wang and X. Wang.

- **Martin Hairer** (University of Warwick)

Plenary talk: Plenary Lecture

Wednesday 9:25 – 10:15, McLeod Auditorium

*On the algebraic structure of renormalisation*

**Abstract:** TBA

- **Valerie Isham** (University College London)

Plenary talk: Bernoulli Lecture

Wednesday 8:30 – 9:20, McLeod Auditorium

*Applied stochastic modelling for structured physical processes*

**Abstract:** Stochastic modelling in which a mechanistic model is used to represent a physical process (albeit in highly idealised form) is an invaluable tool in a broad range of applications. The parameters of such a model are interpretable and relate directly to physical phenomena. As well as gaining insight into the process and understanding of its main drivers, the modeller can address important issues, answering ‘what if’ questions, developing appropriate control strategies and determining policy. In this talk I will discuss approaches to modelling and the need for stochasticity, together with the purposes of modelling and some of the questions that need to be asked and choices made when developing a model for a specific application. The discussion will be illustrated with examples from a variety of applications. A focus on population structure will form a common thread in discussing model development.

- **Servet Martinez** (Universidad de Chile)

Plenary talk: Lévy Lecture

Wednesday 4:15 – 5:05, McLeod Auditorium

*Quasi-stationarity in Markov chains and population dynamics*

**Abstract:** We give some of the main concepts on survival for Markov chains that are surely killed. Quasi-stationary distributions (q.s.d.) are the invariant distributions when conditioned to survive. We give some properties of q.s.d. in a global framework. We introduce some associated concepts, a main one being the  $Q$ -process which is the Markov chain of trajectories that survive forever. All these concepts are completely described for finite state Markov chains by using the Perron-Frobenius theory and they can be seen geometrically in expanding dynamical systems.

For countable Markov chains we give the existence theorem due to Ferrari, Kesten, Martinez, Picco, ensuring that if there is not entrance from infinity then exponential survival is a necessary and sufficient condition to have q.s.d.. In birth and death chains the set of q.s.d.’s the set of q.s.d., when they exist, can be a singleton or a continuum with an extremal one that is the quasi-limiting distribution and plays a fundamental role in the associated  $Q$ -process. We describe some population dynamics models which are birth and death structured and extincted almost surely which were developed jointly with Collet, Meleard and San Martin: a first one has mutation on the traits and we describe the traits conditionally to survive, a second model describe the dynamics of bacteria in a chemostat.

- **Byeong Park** (Seoul National University)

Plenary talk: Laplace Lecture

Tuesday 8:30 – 9:20, McLeod Auditorium

*In-sample density forecasting*

**Abstract:** In-sample density forecasting is a new class of forecasting methods based on structured nonparametric density models. The problem arises when the data are not observed over the full support  $S$ , but are truncated outside a proper subset of  $S$ , say  $I$ . The task is then to estimate the underlying density not only on  $I$  but also on  $S \setminus I$ . This can be accomplished in case the density restricted to  $I$  and to  $S \setminus I$  are described by the common one-dimensional nonparametric component functions in such a way that the component functions are fully identified by the density restricted to  $I$ . In this presentation, I introduce two promising structured density models and a powerful technique of estimating the component functions with sound theory. I also illustrate the approach by applying it to claims reserve prediction and to mesothelioma mortality forecasting.

- **Judith Rousseau** (CEREMADE, Université Paris Dauphine & CREST-ENSAE)

Plenary talk: Ethel Newbold Prize Lecture

Monday 5:10 – 6:00, McLeod Auditorium

*On the Bayesian measures of uncertainty in infinite dimensional models*

Coauthors: Ismael Castillo (Université Paris 6), Botond Szabo (Leiden University)

**Abstract:** Over the last 2 decades there have been many advances in the study of frequentist properties of Bayesian approaches in nonparametric or high dimensional models. Following the seminal papers of Ghosal, Ghosh and van der Vaart 2000 on posterior concentration rates, there has been a large literature on posterior concentration rates in various families of sampling and prior models. Recently, more refined properties of the posterior distribution have also been investigated, leading to a better understanding of the frequentist properties of Bayesian measures of uncertainty such as credible regions.

In this talk, I will first present the general ideas behind posterior concentration rates. Then I will describe more precisely the recent advances behind the understanding of Bayesian measures of uncertainty. In particular I will present (1) the Bernstein von Mises theorems obtained in semi and nonparametric models and (2) some general conditions to construct credible regions having optimal frequentist properties.

- **Scott Sheffield** (MIT)

Plenary talk:

Wednesday 5:10 – 6:00, McLeod Auditorium

*Random surfaces and gauge theory*

**Abstract:** I will discuss several new results (many joint with Jason Miller) related to two prominent models for random surfaces: Liouville quantum gravity and the Brownian map. I will also discuss the relationship between these subjects and the sums over random surfaces that appear in lattice gauge theory.

- **Sara van de Geer** (Seminar for Statistics, ETH Zurich)

Plenary talk: Wald Lecture 1

Tuesday 9:25 – 10:15, McLeod Auditorium

*High-dimensional statistics: a triptych (part 1)*

**Abstract:** High-dimensional statistics concerns the situation where the number of parameters  $p$  is (much) larger than the number of observations  $n$ . This is quite common nowadays, and it has led to the development of new statistical methodology. The lectures present a selected overview of mathematical theory for sparsity inducing methods.

In the first lecture we will highlight the main ingredients for proving sharp oracle inequalities for regularized empirical risk minimizers. The regularization penalty will be taken to be a norm  $\Omega$  on  $p$ -dimensional Euclidean space. Important is that the norm  $\Omega$  has a particular feature which we term the triangle property. We present as examples: the  $\ell_1$ -norm, norms generated from cones, the sorted  $\ell_1$ -norm, the nuclear norm for matrices and an extension to tensors. We then show sharp oracle inequalities for a broad class of loss functions.

The second lecture addresses the construction of asymptotic confidence intervals for parameters of interest. Here, we restrict ourselves to the linear and the graphical model. We prove asymptotic normality of de-biased estimators. We consider asymptotic lower bounds for the variance of an approximately unbiased estimator of a one-dimensional parameter as well as Le Cam type lower bounds. We ascertain the approximate unbiasedness of the de-biased estimator under sparsity conditions and show that it reaches the lower bound.

In the third lecture, we examine the null space property for sparsity inducing norms. The null space property ensures exact recovery of certain sparsity patterns and is moreover a key ingredient for oracle results. We derive this property for the Gram matrix based on  $n$  copies of a  $p$ -dimensional random variable  $X$ , where we require moment conditions for finite dimensional projections of  $X$  or the more general small ball property.

The lectures are based on joint work with Andreas Elsenner, Jana Janková, Alan Muro and Benjamin Stucky.

- **Sara van de Geer** (Seminar for Statistics, ETH Zurich)

Plenary talk: Wald Lecture 2

Thursday 9:25 – 10:15, McLeod Auditorium

*High-dimensional statistics: a triptych (part 2)*

**Abstract:** See Wald Lecture 1

- **Sara van de Geer** (Seminar for Statistics, ETH Zurich)

Plenary talk: Wald Lecture 3

Friday 9:25 – 10:15, McLeod Auditorium

*High-dimensional statistics: a triptych (part 3)*

**Abstract:** See Wald Lecture 1

- **Ruth Williams** (UC San Diego)

Plenary talk: Kolmogorov Lecture

Friday 4:15 – 5:05, McLeod Auditorium

*Reflected diffusions and (bio)chemical reaction networks*

**Abstract:** Continuous-time Markov chain models are often used to describe the stochastic dynamics of networks of reacting chemical species, especially in the growing field of systems biology. Discrete-event stochastic simulation of these models rapidly becomes computationally intensive. Consequently, more tractable diffusion approximations are commonly used in numerical computation, even for modest-sized networks. However, existing approximations (e.g., van Kampen and Langevin), do not respect the constraint that chemical concentrations are never negative.

In this talk, we propose an approximation for such Markov chains, via reflected diffusion processes, that respects the fact that concentrations of chemical species are non-negative. This fixes a difficulty with Langevin approximations that they are frequently only valid until the boundary of the positive orthant is reached. Our approximation has the added advantage that it can be written down immediately from the chemical reactions. This contrasts with the van Kampen approximation, which involves a two-stage procedure — first solving a deterministic ordinary differential equation, followed by a stochastic differential equation for fluctuations around those solutions.

An invariance principle for reflected diffusions, due to Kang and Williams, is adapted in justifying our approximation under mild assumptions. Some numerical examples illustrate the advantages of our approximation over direct simulation of the Markov chain or use of the van Kampen approximation.

This talk is based on joint work with David Anderson, Des Higham and Saul Leite.

- **Bin Yu** (UC Berkeley)

Plenary talk: Rietz Lecture

Monday 9:25 – 10:15, McLeod Auditorium

*Theory to gain insight and inform practice*

**Abstract:** Henry L. Rietz, the first president of IMS, published his book “Mathematical Statistics” in 1927. One review wrote in 1928:

“Professor Rietz has developed this theory so skillfully that the ‘workers in other fields’, provided only that they have a passing familiarity with the grammar of mathematics, can secure a satisfactory understanding of the points involved.”

In this lecture, I would like to promote the good tradition of mathematical statistics as expressed in Rietz’s book in order to gain insight and inform practice. In particular, I will recount the beginning of our theoretical study of dictionary learning (DL) as part of a multi-disciplinary project to “map a cell’s destiny” in *Drosophila* embryo. I will share insights gained regarding local identifiability of primal and dual formulations of DL. Furthermore, comparing the two formulations is leading us down the path of seeking confidence measures of the learned dictionary elements (corresponding to biologically meaningful regions in *Drosophila* embryo). Finally, I will present preliminary work using our confidence measures to identify potential knockout (or gene editing) experiments in an iterative interaction between biological and data sciences.

- **Ofer Zeitouni** (Weizmann Institute of Science & Courant Institute)

Plenary talk: Schramm Lecture

Thursday 8:30 – 9:20, McLeod Auditorium

*Extremes in logarithmically correlated fields*

**Abstract:** It was recently proved that the maximum of Gaussian logarithmically correlated fields (of which the Gaussian free field in two dimensions is a prime example) converges in distribution to a randomly shifted Gumbel distribution. I will review the link between this result and the corresponding one for branching random walks, and then will describe recent progress and challenges in obtaining similar results for certain non-Gaussian logarithmically correlated fields; a specific example of the latter is the maximum of the modulus of the characteristic polynomial of a random unitary matrix. The talk is based in part on joint works with Bramson, Ding, Paquette and Roy.

## ABSTRACTS OF ALL TALKS

A

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Louigi Addario-Berry** (McGill University)

Invited session: Random planar maps [7]

Tuesday 2:00 – 2:30, McLeod Auditorium

*Most trees are short and fat*

**Abstract:** Galton-Watson trees are a kind of planar maps so this talk is allowed in the session. Let  $T$  be any Galton-Watson tree. Write  $\text{vol}(T)$  for the volume of  $T$  (the number of nodes),  $\text{ht}(T)$  for the height of  $T$  (the greatest distance of any node from the root) and  $\text{wid}(T)$  for the width of  $T$  (the greatest number of nodes at any level). We study the relation between  $\text{vol}(T)$ ,  $\text{ht}(T)$  and  $\text{wid}(T)$ .

In the case when the offspring distribution  $p = (p_i, i \geq 0)$  has mean one and finite variance, both  $\text{ht}(T)$  and  $\text{wid}(T)$  are typically of order  $\text{vol}(T)^{1/2}$ , and have sub-Gaussian upper tails on this scale (A-B, Devroye and Janson, 2013). Heuristically, as the tail of the offspring distribution becomes heavier, the tree  $T$  becomes “shorter and bushier”. We describe a few theorems which can be viewed as justifying this heuristic. In particular, we show, under the sole assumption that  $p_1 \neq 1$ , that the random variable  $\text{ht}(T)/\text{wid}(T)$  always has sub-exponential tails and, more strongly, that the random variable  $\text{ht}(T)/\text{vol}(T)^{1/2}$  always has sub-Gaussian tails.

- **Genevera Allen** (Rice University & Baylor College of Medicine)

Invited session: Models and inference for big data [38]

Friday 11:20 – 11:50, McLeod Auditorium

*Algorithmic regularization paths: A new approach to variable selection*

Coauthors: Yue Hu (Rice University), Michael Weylandt (Rice University)

**Abstract:** Variable selection has become a cornerstone of statistical machine learning and is ubiquitously used for the analysis of high-dimensional data. While there is a robust literature on providing

guarantees for the performance of variable selection techniques, there is one setting for which there are no such guarantees and existing methods perform poorly: that of high-dimensional high-correlation (HDHC) data. Such HDHC data commonly arises from high-throughput biomedical technologies and in graph selection problems for highly connected graphs. In this paper, we develop a radically different type of variable selection method that will prove to be superior in HDHC settings. Our so-called Algorithmic Regularization Paths generate a sequence of sparse models as the iterates of an algorithm inspired by the Alternating Direction Methods of Multipliers (ADMM) algorithm for solving the Lasso. We show that our procedure naturally performs “variable selection through peer pressure” by using information from all other highly correlated variables to determine the correct sparse model. We establish basic theoretical properties of our method and draw connections to several existing regularization techniques. Extensive empirical studies demonstrate that our Algorithmic Regularization Paths yield superior performance in terms of prediction accuracy, variable selection, model selection, and computing time than all existing approaches in HDHC settings.

- **Gidi Amir** (Bar-Ilan University)

Invited session: Random walks on graphs [9]

Wednesday 3:10 – 3:40, Medical Sciences 3153

*Exponents of random walks on groups*

**Abstract:** Let  $G$  be an infinite, finitely generated group with a finite symmetric generating set  $S$ , and let  $X_n$  be a (symmetric, finitely supported) random walk on  $G$ . There are several quantities that one can associate to  $G$  and  $X_n$  that describe geometric aspects of the group and the random walk. Among the most studied are the rate of escape of the walk  $E|X_n|$ , the entropy  $H(X_n)$ , the return probabilities  $P(X_{2n} = id)$ , and the Volume growth of  $G$   $|B_n(id)|$ . Starting with the works of Kesten, many relations between the various quantities and connections to other group properties were discovered. Return probabilities have been linked to the spectral radius, amenability, isoperimetric properties and volume growth. The entropy and rate of escape to the Liouville property and Poisson Furstenberg boundary and the list is far from complete.

In this talk we will address some partial advances towards the following general question: What are the possible (joint) behaviours of the above-mentioned geometric quantities.

After describing some known bounds and some relations between these quantities, we will focus on a constructive approach and show several approaches to constructing groups and random walks with good control over some of these quantities, with an emphasis on the rate of escape and the entropy.

The talk is based on joint works with Balint Virag and with Gady Kozma, and on the work of Brioussel and Zheng.

- **Aboubacar Amiri** (Université Lille 3)

Contributed Talk, session [92]

Wednesday 10:45 – 11:15, Medical Sciences 2172

*Regression estimation by local polynomial fitting for multivariate data stream*

**Abstract:** We study a local polynomial estimator of the regression function and its derivatives. We

propose a sequential technique based on a multivariate counterpart of the stochastic approximation method for successive experiments for the local polynomial estimation problem. We present our results in a more general context by considering the weakly dependent sequence of stream data. Additionally, we provide algorithms for the practical use of the method in data stream framework.

- **Scott Armstrong** (Université Paris-Dauphine)

Invited session: Random media [5]

Monday 10:45 – 11:15, Medical Sciences 3153

*Quantitative stochastic homogenization*

Coauthors: Tuomo Kuusi (Aalto University), Jean-Christophe Mourrat (ENS Lyon)

**Abstract:** I will summarize some recent progress in the study of linear, divergence form elliptic equations with random coefficients (i.e., the random conductance model). Several years ago, Gloria, Otto and collaborators used an idea from statistical mechanics (concentration of measure) to establish strong quantitative bounds on the sizes of fluctuations of solutions, and then Nolen, Mourrat and others proved central limit theorems for Dirichlet forms and obtained the scaling limits of solutions. In all of these results, concentration inequalities were the essential tool for transferring ergodic information from the coefficients (which satisfy an iid assumption) to the solutions themselves. In this talk, I will describe an alternative approach developed in the last several years with Smart and then later with Kuusi and Mourrat. We obtain optimal quantitative estimates, central limit theorems, and the scaling limit of the correctors— under fewer assumptions, with much stronger (and optimal) stochastic integrability, and without using abstract concentration of measure. Rather, our approach is to “linearize the random dependence at the homogenized limit”: that is, we use bootstrap/renormalization arguments which reveal that certain Dirichlet forms are essentially additive quantities. Thus the complicated nonlinear structure reduces to a linear one (a sum of iid random variables) and then a complete quantitative theory can then be read off.

- **Moses Chicco Aron** (Ministry of Health, Malawi (M&E Volunteer))

Contributed Talk, session [79]

Thursday 2:35 – 3:05, Fitzgerald 139

*Modeling waiting and service time for patients: a case study of Matawale Health Centre, Zomba, Malawi*

**Abstract:** Spending more time on long queues for a basic service remains a common challenge to most developing countries, including Malawi. For health sector in particular, Out-Patient Department (OPD) experiences long queues. This puts the lives of patients at risk. However, using queuing analysis to understand the nature of the problems and efficiency of service systems, such problems can be abated. Based on a kind of service, literature proposes different possible queuing models. However, unlike using generalized assumed models proposed by literature, use of real time case study data can help in deeper understanding the particular problem model and how such a model can vary from one day to the other and also from each case to another. As such, this study uses data obtained from one urban HC for BP, Pediatric and General OPD cases to investigate an average queuing time for patients within the system. It seeks to highlight the proper queuing model by investigating the

kind of distributions functions over patients arrival time, inter-arrival time, waiting time and service time. Comparable with the standard set values by WHO, the study found that patients at this HC spend more waiting times than service times. On model investigation, different days presented different models ranging from an assumed M/M/1, M/M/2 to M/Er/2. As such, through sensitivity analysis, in general, a commonly assumed M/M/1 model failed to fit the data but rather an M/Er/2 demonstrated to fit well. An M/Er/3 model seemed to be good in terms of measuring resource utilization, proposing a need to increase medical personnel at this HC. However, an M/Er/4 showed to cause more idleness of human resources.

Keywords: health care, out-patient department, queuing model, sensitivity analysis

- **John Aston** (University of Cambridge)

Invited session: Statistical computing and complex data [36]

Monday 2:00 – 2:30, McLeod Auditorium

*Functional data for manifold neuroimaging*

Coauthors: Eardi Lila (University of Cambridge), Laura Sangalli (Politecnico di Milano)

**Abstract:** Functional data analysis (FDA) techniques are becoming more prevalent in neuroimaging. We will firstly review some of the methods used in neuroimaging based on FDA. Then, motivated by the analysis of high-dimensional neuroimaging signals located over the cortical surface, we introduce a novel Principal Component Analysis technique that can handle functional data located over a two-dimensional manifold. We approach the discretization task by means of finite element analysis and propose an efficient iterative algorithm for its resolution. We apply this to the Human Connectome Project to examine the variability in brain connectivity.

- **Rami Atar** (Technion)

Invited session: Scaling limits of stochastic networks [14]

Tuesday 3:10 – 3:40, Rosebrugh 208

*On the measure-valued Skorohod map*

**Abstract:** We describe a transformation that acts on paths with values in the space of measures over the real line, which is in a sense an infinite-dimensional analogue of the Skorokhod map. It is argued to provide a generic model for priority in queueing systems. When used to study disciplines such as earliest-deadline-first and shortest-job-first, it considerably extends known law-of-large-numbers results.

- **Siva Athreya** (Indian Statistical Institute)

Contributed Talk, session [53]

Monday 2:00 – 2:30, Medical Sciences 4279

*Subdiffusivity of a random walk among a Poisson system of moving traps on  $\mathbb{Z}$*

Coauthors: Alexander Drewitz (Universität zu Köln), Rongfeng Sun (National University of Singapore)

**Abstract:** We consider a random walk among a Poisson system of moving traps on  $\mathbb{Z}$ . In earlier work the quenched and annealed survival probabilities of this random walk have been investigated.

Here we study the path of the random walk conditioned on survival up to time  $t$  in the annealed case and show that it is subdiffusive. As a by-product, we obtain an upper bound on the number of so-called thin points of a one-dimensional random walk, as well as a bound on the total volume of the holes in the random walk's range.

**B**

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Luis A. Barboza** (Universidad de Costa Rica)  
 Contributed Talk, session [90]  
 Tuesday 11:55 – 12:25, Rosebrugh 310  
*Parameter estimation of Gaussian stationary processes using the generalized method of moments*  
 Coauthor: Frederi G. Viens (Purdue University)  
**Abstract:** We consider the class of all stationary Gaussian process with explicit parametric spectral density. Under some conditions on the autocovariance function, we defined a GMM estimator that satisfies consistency and asymptotic normality, using the Breuer-Major theorem and previous results on ergodicity. This result is applied to the joint estimation of the three parameters of a stationary Ornstein-Uhlenbeck (fOU) process driven by a fractional Brownian motion. The asymptotic normality of its GMM estimator applies for any  $H$  in  $(0,1)$  and under some restrictions on the remaining parameters. A numerical study is performed in the fOU case, to illustrate the estimator's practical performance when the number of datapoints is moderate.
- **Jean-Marc Bardet** (SAMM, University Paris 1 Panthéon-Sorbonne, France)  
 Contributed Talk, session [95]  
 Tuesday 2:00 – 2:30, Medical Sciences 4171  
*Offline and online multiple change detection for causal time series*  
 Coauthors: William Kengne (University Cergy-Pontoise, France), Olivier Wintenberger (University Pairs 6, France)  
**Abstract:** We consider time series belonging to a large class of causal models including,  $AR(\infty)$ ,  $ARCH(\infty)$ ,  $TARCH(\infty)$ ,  $ARMA-GARCH(\infty)$  processes. We first suppose that certain parameters of a fixed model change at each break-point and that there is a number  $K$  of changes,  $K$  unknown. It can be underlined that after a change-point, the model depends on previous data built with other parameters. Therefore between two successive break-points, the time series is generally not stationary. All the unknown parameters ( $K$ , the change dates and the successive parameters of the model) are estimated using a penalized contrast built from a Gaussian quasi-likelihood. Under conditions, we show that we obtain the same convergence rates that in the case of independent random variables. Numerically, we added a data-driven procedure of choice of the penalty which allow to obtain extremely convincing results. We secondly consider a sequential procedure for detecting changes in parameters of causal time series. In this framework, we construct a test statistics based on Gaussian log-likelihood and we study its asymptotic behavior. Simulations are also realized.

Joint works with William Kengne (University of Cergy, France) and Olivier Wintenberger (University Paris 6, France)

- ~~Winfried Barta~~ (George Washington University)

Contributed Talk, session [75]

Wednesday 10:45 – 11:15, Fitzgerald 139

*Loss of memory in adaptive MCMC* TALK CANCELLED

**Abstract:** Markov chain Monte Carlo (MCMC) methods are frequently used to approximately simulate high-dimensional, multi-modal probability distributions. In adaptive MCMC, the transition kernel is changed “on the fly” in the hope to speed up convergence. We study a simplified version of the equi-energy sampler called interacting tempering.

Using a coupling argument, we show that under mild and easy to verify assumptions on the target distribution (on a bounded space), the interacting tempering process rapidly forgets its starting distribution. The result applies, for example, to models from statistical physics like the Ising and Potts models and Edwards Anderson spin glasses, as well as to exponential random graph models from statistics.

Since the result holds true also in settings where the process is known to converge exponentially slowly to its limiting distribution, we argue that this has implications for the practice of convergence diagnostics.

- **Peter Bartlett** (UC Berkeley and QUT)

Invited session: Online algorithms in machine learning [33]

Friday 2:00 – 2:30, Medical Sciences 2171

*Efficient optimal strategies for adversarial linear regression*

**Abstract:** We consider linear regression games for both fixed design and adversarial design. In the fixed design case, at each round, the learner predicts a real value, the adversary reveals a label, and the learner incurs a squared error loss. The aim is to minimize the difference between the cumulative loss and that of the linear predictor that is best in hindsight. For a variety of constraints on the adversary’s labels, we show that the minimax optimal strategy is linear, with a parameter choice that is reminiscent of ordinary least squares, is easy to compute, and does not require knowledge of the constraint set. We obtain an explicit expression for the minimax regret. In the adversarial design case, we exhibit constraint sets of covariate sequences for which the same strategy is minimax optimal.

Joint work with Wouter Koolen, Alan Malek, Eiji Takimoto and Manfred Warmuth.

- **Mohsen Bayati** (Stanford University)

Invited session: Compressed sensing [23]

Thursday 2:00 – 2:30, Medical Sciences 2172

*Online decision making with high-dimensional covariates*

Coauthor: Hamsa Bastani (Stanford University)

**Abstract:** Growing availability of data has enabled decision-makers to tailor choices at the individual-level. This involves learning a model of decision rewards conditional on individual-specific covariates

or features. Recently, contextual bandits have been introduced as a framework to study these online decision making problems. However, when the space of features is high-dimensional, existing literature only considers situations where features are generated in an adversarial fashion that leads to highly conservative performance guarantees – regret bounds that scale by  $\sqrt{n}$  where  $n$  is the number of samples.

Motivated by medical decision making problems where stochastic features are more realistic, we introduce a new algorithm that relies on two sequentially updated LASSO estimators. One estimator (with a low-bias) is used when we are confident about its accuracy, otherwise a more biased (but potentially more accurate) estimator is used. We prove that our algorithm achieves a regret of order  $s^2 [\log n]^2 + s^2 [\log n] [\log p]$  where  $p$  is the dimension of the features and  $s$  is the number of relevant features. The key step in our analysis is proving a new oracle inequality that guarantees the convergence of the LASSO estimator despite the non-i.i.d. data induced by the bandit policy. We also provide a new analysis of the low-dimensional setting that improves existing bounds by a factor  $p$ .

We illustrate the practical relevance of the proposed algorithm by evaluating it on a warfarin dosing problem. A patient’s optimal warfarin dosage depends on the patient’s genetic profile and medical records; incorrect initial dosage may result in adverse consequences such as stroke or bleeding. We show that our algorithm outperforms existing bandit methods as well as physicians to correctly dose a majority of patients.

- **Andriette Bekker** (University of Pretoria)

Contributed Talk, session [74]

Tuesday 10:45 – 11:15, Fitzgerald 139

*Bivariate gamma-type distributions stemming from complex elliptical models*

Coauthors: M Arashi (Shahrood University of Technology), JT Ferreira (University of Pretoria)

**Abstract:** Ferreira et. al. (2016- Novel contributions within the complex elliptical class for performance evaluation of wireless systems. Submitted to RevStat-Statistical Journal) questioned the complex normal assumption in the MIMO environment. The authors proposed a complex central Wishart-type distribution which is based on the complex elliptical model. The objective of this paper is to study bivariate gamma-type distributions originating from the diagonal elements of the complex inverse Wishart-type. A new bivariate Nakagami-m-type results from this bivariate gamma-type model which will be used to propose alternative decision statistics. Further possible applications in the wireless communication systems arena will be outlined.

Key words: Complex inverse Wishart type; Nakagami-m distribution; signal to noise ratio; t-distribution

- **David Belius** (University of Zurich)

Invited session: Spin Glasses [4]

Monday 2:00 – 2:30, Medical Sciences 3153

*Extrema of log-correlated random fields, regularizations and branching*

Coauthors: Louis-Pierre Arguin (Baruch College, City University of New York), Paul Bourgade

(Courant Institute of Mathematics, New York University), Nicola Kistler (Frankfurt University)

**Abstract:** Log-correlated random fields arise as toy models of mean field spin glasses, and also in many other settings such as random surfaces in two dimensions (Gaussian Free Field), the statistics of the Riemann Zeta function, the characteristic polynomial of random matrices, and the study of random walk cover times. I will describe how to observe an approximate hierarchical or “branching” structure in such fields by considering appropriate multiscale regularizations. This makes it possible to apply methods developed for Branching Random Walk, where the branching structure is explicit and exact, to study the extreme values of log-correlated random fields.

- **Alexandre Belloni** (Duke University)

Invited session: High-dimensional econometrics [31]

Thursday 11:55 – 12:25, McLeod Auditorium

*Measuring context effects in Z-estimator framework*

Coauthors: Victor Chernozhukov (MIT), Roberto I. Oliveira (IMPA)

**Abstract:** We study a  $Z$ -estimators for parameters of interest associated with a stationary process. We focus on the case that the dependence can be well approximated by variable length Markov chain model which is characterized by the context tree that creates a partition of the history. We call the context effects the parameters of the model that varies with the history. We discuss examples of specific models and a general formulation that leads to uniform rates of convergence.

- **Paul Bendich** (Duke University)

Invited session: Object Oriented Data Analysis: Persistent Homology Representations [30]

Monday 11:20 – 11:50, McLeod Auditorium

*Topological features for statistics and machine learning: brain arteries and driver behavior*

**Abstract:** Topological Data Analysis (TDA) has proven itself very useful in the production of some very non-standard features that can be fed into traditional statistics and/or machine learning pipelines. A key tool is the persistence diagram, or barcode.

We go over two recent successes: an analysis of brain artery trees using the one-dimensional persistent homology of the thickening tree, and a classification of driver behavior using zero-dimensional persistent homology of motion primitives.

In the former, heightened age and sex effects were found, relative to previous non-topological analyses. In the latter, the topological features were integrated into vehicle-tracking methodology, with improved data association results.

Time permitting, we also give a warning about a current TDA weakness: although features extracted from persistence diagrams are stable to perturbation, their corresponding locations in the original data objects are often not! We give examples of this phenomenon, and discuss a few possible solutions.

- **Jérémié Bettinelli** (CNRS / École Polytechnique)

Invited session: Random planar maps [7]

Tuesday 2:35 – 3:05, McLeod Auditorium

*Brownian disks*

Coauthor: Grégory Miermont (ENS Lyon)

**Abstract:** In the same manner as the Brownian motion appears as the universal scaling limit of any mild random walk, Brownian disks are random metric spaces that appear as the universal scaling limits of mild models of random maps on the disk. These objects generalize the Brownian map of Miermont and Le Gall, which is obtained in the case of random maps of the sphere.

We will introduce the Brownian disks et give some of their remarkable properties. This talk is based on a joint work with Grégory Miermont.

- **Shankar Bhamidi** (University of North Carolina, Chapel Hill)

Invited session: Interacting particle systems and their scaling limits [10]

Thursday 2:00 – 2:30, Medical Sciences 3153

*Continuum scaling limits of critical inhomogeneous random graph models*

**Abstract:** Over the last few years a wide array of random graph models have been postulated to understand properties of empirically observed networks. Most of these models come with a parameter  $t$  (usually related to edge density) and a (model dependent) critical time  $t_c$  which specifies when a giant component emerges. There is evidence to support that for a wide class of models, under moment conditions, the nature of this emergence is universal and looks like the classical Erdos-Renyi random graph, in the sense that (a) the sizes of the maximal components in the critical scaling window scale like  $n^{2/3}$ , and (b) the structure of components in this window (rescaled by  $n^{-1/3}$ ) converge to random fractals related to the continuum random tree. We will describe a general program for proving such results. The program requires three main ingredients: (i) in the critical scaling window, components merge approximately like the multiplicative coalescent (ii) scaling exponents of susceptibility functions (including distance based susceptibility) are the same as the Erdos-Renyi random graph and (iii) macroscopic averaging of expected distances between random points in the same component in the barely subcritical regime.

- **Louis Bhim** (The University of Sydney)

Contributed Talk, session [48]

Tuesday 11:20 – 11:50, Medical Sciences 2171

*A new approach to placing hard bounds on option prices in regime-switching models*

Coauthor: Reiichiro Kawai (The University of Sydney)

**Abstract:** We approach the problem of placing hard bounds on the price of various options (including European, American and Barrier type options) in regime-switching markets, by formulating an appropriate semidefinite programming problem solvable by numerical optimization. We implement results from sum of squares polynomial representations and stochastic analysis in order to derive an optimization formulation of the problem that is solvable by available software. In particular, we utilize the analogue of the well-known Dynkin formula for regime-switching diffusions in order to place deterministic bounds on a functional that can be associated with the price function of a variety of options. We then use results of the sum of squares representations for non-negative polynomials

to relax the constraints of our problem and make it computationally tractable. The scope of application of this approach to different financial derivatives with various payoff structures and in various market models will also be discussed.

- **Ilia Binder** (University of Toronto)

Invited session: Conformally invariant processes [8]

Friday 3:10 – 3:40, Medical Sciences 3153

*Rate of convergence of the exploration process for the critical percolation.*

**Abstract:** We discuss the rate of convergence of the exploration process for the critical percolation to  $SLE_6$ . The result states that for any “reasonable” critical percolation model for which the convergence of the exploration process is established, the power rate of convergence must automatically hold. So far, the result is unconditional for the critical site percolation on hexagonal lattice and for some of its generalizations, which will be discussed at the talk. The talk is based on joint projects with Lincoln Chayes (UCLA), Helen Lei, and Larissa Richards (Toronto).

- **Horatio S. Boedihardjo** (University of Reading)

Invited session: Rough paths [13]

Wednesday 2:35 – 3:05, McLeod Auditorium

*Signature of a rough path: Uniqueness*

Coauthors: Xi Geng (Oxford), Terry Lyons (Oxford), Danyu Yang (Oxford)

**Abstract:** The signature is an exponential transform on the space of paths and plays a key role in rough path theory. The signature map has homomorphism properties analogous to those of the classical exponential map. Hambly and Lyons showed that the signature map is injective for bounded variation paths. We will discuss an extension of the injective property for rough paths. (Joint work with X. Geng, T. Lyons, D. Yang).

- **Stephane Bonhomme** (University of Chicago)

Invited session: High-dimensional econometrics [31]

Thursday 10:45 – 11:15, McLeod Auditorium

*Approximate clustering*

Coauthors: Thibaut Lamadon (University of Chicago), Elena Manresa (MIT Sloan School of Management)

**Abstract:** We characterize properties of clustering estimators in models where unobserved heterogeneity in the population can be approximated by a group structure. Clustering methods have a number of potential applications, for example difference-in-differences designs and more generally settings with low-dimensional time-varying unobservables such as nonlinear factor models. Applications also include structural economic models with latent individual heterogeneity. We study three classes of estimators:  $k$ -means clustering, single-step and two-step moments estimators, and estimators in the presence of grouped patterns of heterogeneity.

- **Omar Boukhadra** (University of Constantine 1)

Contributed Talk, session [59]

Wednesday 11:55 – 12:25, Medical Sciences 3278

*On maximal displacement of bridges in the random conductance model*

Coauthor: Nina Gantert (Fakultät für Mathematik, Technische Universität München)

**Abstract:** We study a discrete time random walk in an environment of i.i.d. non-negative conductances in  $\mathbb{Z}^d$ . We consider the maximum displacements for bridges, i.e. we condition the random walk on returning to the origin, and we prove first a normal (diffusive) behavior under some regularity assumptions: standard heat kernel decay and polynomial volume growth. Afterwards, we prove that if the heat kernel decay is such that the return probabilities are sufficiently slow, we obtain anomalous (subdiffusive) maximal displacements for bridges.

- **Cédric Boutillier** (LPMA, Université Pierre et Marie Curie)

Invited session: Integrable models in Statistical Physics [2]

Thursday 10:45 – 11:15, Medical Sciences 3153

*Integrable statistical mechanics on isoradial graph out of criticality: spanning forests and the Ising model*

Coauthors: Béatrice de Tilière (UPEC), Kilian Raschel (Université de Tours)

**Abstract:** Isoradial graphs are planar graphs, with an embedding such that every bounded face is inscribed in a circle of radius 1. Isoradiality is preserved under the star-triangle transformation of planar graphs. Since this transformation plays an important role in integrable models of statistical mechanics (via the Yang-Baxter equations), isoradial graphs are particularly well suited as a support for such integrable models.

In this talk, we will consider spanning forests and the Ising model on isoradial graphs. The weights satisfying Yang-Baxter equations for these models can be parametrized with elliptic functions. Properties of these models are related to finite difference operators on isoradial graphs (massive Laplacian, Kasteleyn matrices). We give explicit, local formulas for the inverses of these operators, which can be then used to compute correlations and the free energy in terms of the geometry of the embedding of the isoradial graph.

When elliptic functions degenerate to trigonometric, one recover as a particular case the results obtained by Kenyon and Boutillier & de Tilière for critical spanning trees and critical Ising model. We also relate these elliptic weights with a family of algebraic Harnack curves of genus 1.

- **Anton Bovier** (Bonn University)

Invited session: Stochastic models of cancer [17]

Wednesday 10:45 – 11:15, Medical Sciences 3154

*A stochastic individual based model for immunotherapy of cancer*

**Abstract:** I discuss an extension of a standard stochastic individual-based model in population dynamics which broadens the range of biological applications. Our primary motivation is modelling of immunotherapy of malignant tumours. In this context the different actors, T-cells, cytokines or cancer cells, are modelled as single particles (individuals) in the stochastic system. The main expansions of the model are distinguishing cancer cells by phenotype and genotype, including environment-dependent phenotypic plasticity that does not affect the genotype. Numerical simulations show that

the model can reproduce rather well experimental data make interesting and sometimes unexpected predictions for new treatment protocols.

- **Richard C. Bradley** (Indiana University (Bloomington))

Contributed Talk, session [94]

Friday 10:45 – 11:15, Medical Sciences 2172

*On mixing properties of some INAR models*

**Abstract:** In time series analysis of (nonnegative integer-valued) “count data”, one sometimes uses “integer-valued autoregressive” (INAR) models, a variant of the original autoregressive models of classical time series analysis. The INAR models of order 1 with “Poisson innovations” satisfy (with exponential mixing rate) the  $\rho^*$ -mixing condition (the stronger variant of the usual  $\rho$ -mixing condition in which the two index sets are allowed to be “interlaced” instead of being restricted to “past” and “future”). That was shown in R.C. Bradley, *Zapiski Nauchnyh Seminarov POMI* 441 (2015) 56-72, and will be explained in this talk. Earlier, S. Schweer and C.H. Weiß, *Comput. Statist. Data Anal.* 77 (2014) 267-284, had already shown that those models (as well as some other closely related ones) satisfy absolute regularity with exponential mixing rate.

- **Guy Bresler** (MIT)

Invited session: Computational versus statistical complexity in high-dimensional problems [24]

Wednesday 10:45 – 11:15, McLeod Auditorium

*Learning a tree-structured Ising model in order to make predictions*

Coauthor: Mina Karzand (MIT)

**Abstract:** We study the problem of learning a high-dimensional tree graphical model from samples such that low-order marginals are accurate. This enables making accurate predictions based on partial observations. A take-home message is that the necessary number of samples is dramatically lower than is required to learn the exact underlying tree. On the technical side, this requires reasoning about predictions computed using a completely incorrect tree. Joint work with Mina Karzand.

- **David R. Brillinger** (University of California, Berkeley)

Plenary talk: Tukey Lecture

Tuesday 4:15 – 5:05, McLeod Auditorium

*People are different*

**Abstract:**

- (1) Some items in JWTs vita
- (2) Some reminiscences and anecdotes  
capybaras
- (3) The Labs 1960 -1964

Where did the E in EDA come from?

A problem JWT put to DRB in 1978

“Measure” (?) of the “peakiness” of periodogram:

- (a) take cepstrum (based on log periodogram)
- (b) average in octaves, laid  $\frac{1}{2}$  to the weather 22 Jan 78

- (4) Environmental Risk Analysis  
Propagating wildfire boundaries
- (5) ‘Sunset salvo’

- **Victor-Emmanuel Brunel** (MIT)

Contributed Talk, session [87]

Wednesday 11:55 – 12:25, Medical Sciences 2171

*Recovering level sets of the Tukey depth*

**Abstract:** Tukey depth, aka halfspace depth, is one natural notion of depth associated with a cloud of points in a Euclidean space. When the points are random and i.i.d., one can define the population analog of Tukey depth. The superlevel sets of the population Tukey depth are convex bodies, referred to as floating bodies in some specific frameworks. We show that, for a fixed level and under some natural assumptions on the underlying probability measure, the superlevel set of the empirical Tukey depth concentrates around that of the population Tukey depth at a rate which does not depend on the ambient dimension.

- **Peter Bubenik** (University of Florida)

Invited session: Object Oriented Data Analysis: Persistent Homology Representations [30]

Monday 10:45 – 11:15, McLeod Auditorium

*An introduction to topological data analysis*

**Abstract:** In this talk I will show how ideas from the mathematical subject of topology can be used to obtain summary statistics that encode data’s geometry. In fact, these summaries may be chosen to be vectors and can thus be easily incorporated into more traditional statistical analysis. I will introduce the relevant topological constructions and apply these methods to protein data.

- **Amarjit Budhiraja** (University of North Carolina at Chapel Hill)

Contributed session: Analysis and control of stochastic network models [42]

Wednesday 11:55 – 12:25, Medical Sciences 4279

*Control of large stochastic networks with mean field interaction*

**Abstract:** We study stochastic control problems for large stochastic networks modeled as weakly interacting pure jump controlled Markov processes. Under suitable scaling limits we establish diffusion approximations that in one model setting lead to stochastic control of degenerate diffusions and in another setting lead to mean field games of Lasry and Lions. Based on joint works with Erhan Bayraktar, Asaf Cohen and Eric Friedlander.

- **Andreas Buja** (The Wharton School, University of Pennsylvania)

Invited session: Selective inference [28]

Tuesday 10:45 – 11:15, Medical Sciences 3154

*Post-selection inference and misspecification*

Coauthors: Richard Berk (University of Pennsylvania), Lawrence Brown (University of Pennsylvania)

**Abstract:** We review briefly the PoSI approach to the post-selection inference problem following

Berk et al. (AoS, 2013). We then motivate moving from a model-trusting to a model-robust framework in which one asks for valid inference under misspecification. The result is a theory in which all estimation and inference concerns statistical functionals rather than parameters in a working model. It will be shown that all sampling variability arises not only from conditional randomness of the response but also from randomness of the regressors interacting with misspecification. This type of model-robust inference can be adapted to the PoSI approach to post-selection inference.

- **Oleg Butkovsky** (Technion - Israel Institute of Technology)

Contributed Talk, session [50]

Thursday 2:35 – 3:05, Medical Sciences 4171

*Stochastic flows for a stochastic heat equation*

Coauthor: Leonid Mytnik (Technion - Israel Institute of Technology)

**Abstract:** It is well known from the results of Zvonkin-Veretennikov-Davie-Flandoli that ordinary differential equations (ODEs) regularize in the presence of noise. Even if an ODE is “very bad” and has no solutions (or has many solutions), then the addition of a random noise leads almost surely to a “nice” ODE with a unique solution. We investigate the same phenomenon for a 1D heat equation with an irregular drift

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial t^2} + b(u(t, z)) + \dot{W}(t, z), \quad t \geq 0, z \in \mathbb{R}$$

We prove that for almost all trajectories of random white noise  $W$  the perturbed heat equation has a unique solution for any initial condition (even if the original heat equation with a drift had many or no solutions). Moreover, this equation possesses a continuous flow of solutions.

C

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Mine Caglar** (Koc University)

Contributed Talk, session [50]

Thursday 3:10 – 3:40, Medical Sciences 4171

*Correlated motions on graphs*

Coauthors: Hatem Hajri (University of Bordeaux), Abdullah Harun Karakus (Koc University)

**Abstract:** We consider the stochastic differential equation

$$dX_t = 1_{\{X_t > 0\}} dB_t^+ - 1_{\{X_t \leq 0\}} dB_t^-$$

where  $B^+$  and  $B^-$  are correlated Brownian motions on  $\mathbb{R}$  satisfying  $d\langle B^+, B^- \rangle_t = \rho dt$  with  $\rho \in [0, 1)$ . The case of independent Brownian motions have already been considered in the works of Hajri, Le Jan and Raimond. We show that our equation can be cast into the perturbed Tanaka’s equation and hence it has a strong solution. In the extreme case of  $\rho = 1$ , it would be Tanaka’s equation, which does not have a strong solution. Our aim is to study the flows for which one-point motion  $(X_t)_{t \geq 0}$  follows our stochastic differential equation. Both stochastic flow of mappings and kernels are considered. We prove that the only flow solution is a coalescing flow, which is characterized by

the property that the two-points meet in finite time almost surely and then continue their motion together. Extension of flows on metric graphs is also discussed and compared with related work.

- **Daniele Cappelletti** (University of Copenhagen)

Contributed session: Chemical reaction networks [41]

Wednesday 3:10 – 3:40, Medical Sciences 2172

*Stationary distributions and graphical equilibria in Reaction Networks Theory*

Coauthors: Carsten Wiuf (University of Copenhagen), Badal Joshi (California State University, San Marcos)

**Abstract:** Chemical reaction networks can be modelled either deterministically, by means of a system of ordinary differential equations, or stochastically, by means of a continuous time Markov chain. The relationship between steady states of the deterministic model with graphical properties of the network have been intensively studied since the Seventies, with a special focus on complex balanced equilibria. However, a similar analysis for stationary distributions in the stochastic setting was not available. We introduce the notion of complex balanced stationary distribution and study the link between this notion and that of complex balanced equilibrium. We further explore more detailed relationships between graphical equilibria in the deterministic and stochastic settings.

- **Thomas Cass** (Imperial College London)

Contributed Talk, session [51]

Friday 3:10 – 3:40, Medical Sciences 3154

*A Stratonovich-to-Skorohod conversion formula for Gaussian rough paths*

Coauthor: Nengli Lim (Imperial College London and National University of Singapore)

**Abstract:** Lyons' theory of rough paths allows us to solve stochastic differential equations driven by a Gaussian processes  $X$  under certain conditions on the covariance function. The rough integral of these solutions against  $X$  again exist, and a natural question is to find a closed-form correction formula between this rough integral and the Skorohod integral of the solution. We present the results of [1] in which such a formula is proved.

This is particularly useful in applications, some of which we survey, in which one wants to compute (or estimate) the conditional expectation of the rough integral. In the case of Brownian motion our formula reduced to the classical Stratonovich-Ito conversion formula. Previous works in the literature assumes the integrand to be a smooth function of  $X$ ; our formula again recovers these results as special cases. A crucial role in the analysis is played by the tail estimates of [2].

References:

- (1) Cass, Thomas; Lim, Nengli: A Stratonovich-Skorohod integral formula for Gaussian rough paths
- (2) Cass, Thomas; Litterer, Christian; Lyons, Terry: Integrability and tail estimates for Gaussian rough differential equations. Ann. Probab. 41 (2013), no. 4, 30263050

- **ChihHao Chang** (Institute of Statistics, National University of Kaohsiung, Taiwan)

Contributed Talk, session [83]

Monday 2:35 – 3:05, Fitzgerald 139

*Asymptotic theory of conditional generalized information criterion for linear mixed-effects model selection*

Coauthors: HsinChen Huang (Institute of Statistical Science, Academia Sinica), ChingKang Ing (Institute of Statistical Science, Academia Sinica)

**Abstract:** In our research, we propose a conditional generalized information criterion (CGIC) for linear mixed-effects model selection. The CGIC includes the conditional Akaike's information criterion (CAIC) proposed by Vaida and Blanchard as a special case. In practice, CAIC is considered to be an adequate criterion for prediction purpose of mixed-effects where the criterion is an unbiased estimator of the expectation of the squared error loss function of the best linear unbiased predictor established under the linear mixed-effects model. In this talk, we propose CGIC to select fixed-effects and random-effects models simultaneously which is convinced helpful for the prediction purpose if the random-effects model could be properly selected. In general we aim for two asymptotic properties: consistency and asymptotic loss/risk efficiency, which are kernel asymptotic properties for prediction purpose of linear mixed-effects data analysis. In this talk, we establish the asymptotic theory of CGIC for the linear mixed-effects model selection under some regularity conditions.

Keywords: Variable selection, linear mixed-effects model, Conditional GIC

- **Yuan-chin Ivan Chang** (Institute of Statistical Science, Academia Sinica, Taiwan)

Contributed Talk, session [98]

Thursday 3:10 – 3:40, Rosebrugh 208

*Influential analysis for classification*

Coauthors: Bo-Shiang Ke (Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan), An Jen Chiang (Department of Obstetrics and Gynecology, Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan)

**Abstract:** The performance assessments play an essential role, when we want choose an appropriate classifier among a bunch of candidates. How to identify the influential cases in classification evaluation is an important research topic, because they will usually affect the performance measures and therefore will mislead our choices about the predictive model. Thus, these influential cases confuse not only the performance measures, but also the importance of model and its corresponding variables. In light of the robustness analysis, we adopt the concept of the influence function and local influence to study the estimate of the area under a receiver operating characteristic curve, which can help to detecting the potential influential cases. Two real data sets are used for illustration.

- **Fuqi Chen** (Western University)

Contributed Talk, session [95]

Tuesday 2:35 – 3:05, Medical Sciences 4171

*Estimation and inference for a change-point problem under a generalised Ornstein-Uhlenbeck setting*

Coauthors: Rogemar Mamon (Western University), Sévérien Nkurunziza (University of Windsor)

**Abstract:** In this talk, we focus on estimating the unknown number and locations of the change points in a generalised Ornstein-Uhlenbeck process with change points. The proposed methods for detecting the unknown locations of the change points were motivated by the least sum of squared

error (LSSE) and the maximum log-likelihood (MLL) approaches. We show further that our methods provide consistent estimators, as well as the asymptotic normality of the MLE for the drift parameters. The case where both the number and the location of the change points are unknown is investigated and an informational methodology is employed to address these issues. We also provide some numerical simulations and real financial data applications to illustrate the performance of the proposed methods.

- **Louis H. Y. Chen** (National University of Singapore)

Invited session: Stein's method [12]

Monday 10:45 – 11:15, Medical Sciences 2171

*Stein's method and Hida calculus in normal approximation for white noise functionals*

**Abstract:** Hida calculus, also known as white noise analysis, is the mathematical theory of white noise initiated by T. Hida in his 1975 Carleton Mathematical Lecture Notes. Let  $\{B(t) : t \in \mathbb{R}\}$  be a standard Brownian motion and let the white noise  $\dot{B}(t) = dB(t)/dt, t \in \mathbb{R}$ , be represented by generalized functions. By regarding the collection  $\{\dot{B}(t) : t \in \mathbb{R}\}$  as a coordinate system, Hida defined and studied generalized white noise functionals  $\varphi(\dot{B}(t) : t \in \mathbb{R})$  through their  $U$ -functionals.

In this talk we will present a connection between Stein's method and Hida calculus in normal approximation for white noise functionals. Our approach is analogous to that for the connection between Stein's method and Malliavin calculus as established by Nourdin and Peccati (*Probab. Theory Relat. Fields* **145**, 75-118, 2009). The connection between Stein's method and Hida calculus will be built on the expression of the Ornstein-Uhlenbeck operator in terms of the Hida derivatives using integration by parts techniques. This will entail an extension of the domain of the Hida derivative  $\partial_t$ , that is the  $\dot{B}(t)$ -differentiation, and a study of the regularity of  $\partial_t$ . This talk is based on joint work with Yuh-Jia Lee and Hsin-Hung Shih.

- **May-Ru Chen** (National Sun Yat-sen University)

Contributed Talk, session [60]

Thursday 11:55 – 12:25, Medical Sciences 3154

*Number of walks for the one-shoes- $N$ -doors model*

**Abstract:** Assume that a person has a house with  $N$  doors and at the beginning, he puts one pairs of walking shoes at each door. For each walk, he chooses a door at random, puts on a pair of shoes, returns after the walk to a randomly chosen door and takes off the shoes at the door. Repeat the above actions.

In this talk, we study the number of walks until the house owner discovers that no shoes are available at the door he has chosen for a further walk. By setting a good corresponding Markov chain, we can find the limiting behavior of the number of walks until no shoes as  $N$  increasing to infinity.

- **Wei-Kuo Chen** (University of Minnesota)

Invited session: Spin Glasses [4]

Monday 2:35 – 3:05, Medical Sciences 3153

*Some results on the maximum energy in the mixed  $p$ -spin models* NEW TITLE

**Abstract:** Spin glasses are disordered spin systems originated from the desire of understanding the strange magnetic behaviors of certain alloys in physics. As mathematical objects, they are often cited as examples of complex systems and have provided several fascinating structures and conjectures. In this talk, we will focus on the maximum energy in the mixed  $p$ -spin mean-field spin glass model with both Ising and spherical spin configurations. We will present their Parisi formulas. Furthermore, in the case of spherical mixed  $p$ -spin model, we will discuss results concerning the fluctuation of the maximum energy and the chaotic nature of the optimizer under small perturbations to the disorder. This talk is based on recent joint works with Antonio Auffinger and Arnab Sen.

- **Yuxin Chen** (Statistics Department, Stanford University)

Invited session: Compressed sensing [23]

Thursday 2:35 – 3:05, Medical Sciences 2172

*Solving random quadratic systems of equations is nearly as easy as solving linear systems*

Coauthor: Emmanuel Candes (Department of Statistics and of Mathematics, Stanford University)

**Abstract:** We consider the fundamental problem of solving quadratic systems of equations in  $n$  variables, where  $y_i = |\langle \mathbf{a}_i, \mathbf{x} \rangle|^2$ ,  $i = 1, \dots, m$  and  $\mathbf{x} \in \mathbb{R}^n$  is unknown. We propose a novel method, which starting with an initial guess computed by means of a spectral method, proceeds by minimizing a nonconvex functional as in the Wirtinger flow approach. There are several key distinguishing features, most notably, a distinct objective functional and novel update rules, which operate in an adaptive fashion and drop terms bearing too much influence on the search direction. These careful selection rules provide a tighter initial guess, better descent directions, and thus enhanced practical performance. On the theoretical side, we prove that for certain unstructured models of quadratic systems, our algorithms return the correct solution in linear time, i.e. in time proportional to reading the data as soon as the ratio  $m/n$  between the number of equations and unknowns exceeds a fixed numerical constant. We extend the theory to deal with noisy systems and prove that our algorithms achieve a statistical accuracy, which is nearly un-improvable. We complement our theoretical study with numerical examples showing that solving random quadratic systems is both computationally and statistically not much harder than solving linear systems of the same size.

- **Ming-Yen Cheng** (National Taiwan University)

Invited session: A celebration of the work of Peter Hall [39]

Thursday 3:10 – 3:40, McLeod Auditorium

*A simple and adaptive two-sample test in high dimensions*

**Abstract:** High-dimensional data are commonly encountered nowadays. Testing the equality of two means is a fundamental problem in the inference. But the conventional Hotelling's  $T^2$  test performs poorly or becomes inapplicable in high dimensions. Several modifications have been proposed to address this challenging issue and shown to perform well. However, they all use normal approximation to the null distributions of their test statistics, thus they all require strong regularity conditions. We study this issue thoroughly and propose an  $L^2$ -norm based test that works under milder conditions and even when there are fewer observations than the dimension. In particular, to cope with possible non-normality of the null distribution, we employ the Welch-Satterthwaite  $\chi^2$ -approximation. Simple

ratio-consistently estimators for the parameters in the approximation distribution are given. While existing tests are not, our test is adaptive to singularity or near singularity of the unknown covariance structure, which is commonly seen in high dimensions and has great impact on the shape of the null distribution. The approximate and asymptotical powers of the proposed test are also investigated. Simulation studies and real data applications show that our test has a better size controlling than a benchmark test, while the powers are comparable when their sizes are comparable.

- **Sung Chan Choi** (University of Utah)

Contributed Talk, session [54]

Monday 10:45 – 11:15, Medical Sciences 4171

*Analysis of spatial Parrondo games with modified game A*

**Abstract:** We investigate Toral’s spatial model of Parrondo’s paradox as modified by Xie et al. We give sufficient conditions for the mean profits to converge as the number of players goes to infinity. This requires verifying the ergodicity of an interacting particle system whose generator is the sum of a spin system generator and an exclusion process generator.

- **Caroline Colijn** (Imperial College London)

Invited session: Biomedical applications [20]

Wednesday 2:00 – 2:30, Medical Sciences 3154

*Cherries in branching trees: inferring the basic reproduction number*

Coauthor: Giacomo Plazzotta ()

**Abstract:** The topologies of branching trees can reveal information about the underlying branching process. This is particularly relevant where the trees in question are phylogenetic trees, and where the underlying evolutionary or branching process is not known. In particular, for evolving pathogens such as influenza or bacterial infections, genomic data are increasingly used to understand pathogen evolution. This inference is challenging, though, due to the size of the datasets and a lack of good models for how different processes shape phylogenetic trees. Using Crump-Mode-Jagers branching processes, we show that the number of cherries divided by the number of tips (the cherry-to-tip ratio) converges in distribution to a single number closely related to the basic reproduction number  $R_0$ . For finite numbers of tips, the cherry-to-tip ratio is normally distributed. We use this relationship to infer  $R_0$  from the topologies of phylogenetic trees. We also develop a tree-free algorithm to infer  $R_0$  via the number of cherries using direct comparison of sequences.

- **Andrea Collevocchio** (Monash University)

Contributed session: Reinforced processes and related fields [44]

Thursday 3:10 – 3:40, Medical Sciences 3154

*Attraction properties for general urn processes*

**Abstract:** We study attraction properties of urns composed of balls with two distinct colours and which evolve over time. This evolution may depend on the composition of the urn as well as certain other factors, external or internal depending on the history of the urn. We prove that, under mild conditions, the model localizes on one of the two colours. We extend our discussion to a system of

interacting urns, and a general class of strongly reinforced random walks. (Joint Work with Jiro Akahori, Tim Garoni and Kais Hamza)

- **Codina Cotar** (Statistical Science Department, University College London)

Contributed session: Reinforced processes and related fields [44]

Thursday 2:00 – 2:30, Medical Sciences 3154

*Edge- and vertex-reinforced random walks with super-linear reinforcement on infinite graphs*

**Abstract:** We introduce a new simple but powerful general technique for the study of edge- and vertex-reinforced processes with super-linear reinforcement, based on the use of order statistics for the number of edge, respectively of vertex, traversals. The technique relies on upper bound estimates for the number of edge traversals, proved in a different context by Cotar and Limic [Ann. Appl. Probab. (2009)] for finite graphs with edge reinforcement. We apply our new method both to edge- and to vertex-reinforced random walks with super-linear reinforcement on arbitrary infinite connected graphs of bounded degree. We stress that, unlike all previous results for processes with super-linear reinforcement, we make no other assumption on the graphs.

For edge-reinforced random walks, we complete the results of Limic and Tarres [Ann. Probab. (2007)] and we settle a conjecture of Sellke [Technical Report 94-26, Purdue University (1994)] by showing that for any reciprocally summable reinforcement weight function  $w$ , the walk traverses a random attracting edge at all large times.

For vertex-reinforced random walks, we extend results previously obtained on  $\mathbb{Z}$  by Volkov [Ann. Probab. (2001)] and by Basdevant, Schapira and Singh [Ann. Probab. (2014)], and on complete graphs by Benaim, Raimond and Schapira [ALEA (2013)]. We show that on any infinite connected graph of bounded degree, with reinforcement weight function  $w$  taken from a general class of reciprocally summable reinforcement weight functions, the walk traverses two random neighbouring attracting vertices at all large times.

(This is joint work with Debleena Thacker.)

- **Harry Crane** (Rutgers University)

Contributed Talk, session [68]

Tuesday 2:00 – 2:30, Rosebrugh 310

*Edge exchangeability: a new foundation for modeling network data*

**Abstract:** Exchangeable models for vertex labeled graphs cannot replicate the large sample behaviors of sparsity and power law degree distributions observed in many network datasets. Out of this mathematical impossibility emerges the question of how network data can be modeled in a way that reflects known empirical behaviors and respects basic statistical principles. We address this question with the new invariance principle of edge exchangeability, which unlike its vertex exchangeable counterpart admits models for networks with sparse and/or power law structure. With this, we settle a longstanding question in statistical network modeling.

Joint work with Walter Dempsey (Michigan).

D

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Kalyan Das** (National Institute of Food Technology Entrepreneurship and Management (Govt. of India, De-novo Category) Plot No.97, Sector 56, HSIIDC Industrial Estate, Kundli, Sonapat, Haryana -131028, India.)

Contributed Talk, session [72]

Thursday 3:10 – 3:40, Rosebrugh 310

*Stochastic nonlinear aspect of noise in a two predators one prey mathematical model induced cyclic oscillations*

**Abstract:** This paper deals with the stochastic stability of a three species ecosystem consisting of a prey and two competing predators. The prey and predators are being continuously harvested. The mathematical model is defined by the system of three first order nonlinear ordinary differential equations. The Local stability by using Routh-Hurwitz criteria and the global stability by using Lyapunov's function are discussed. The condition for the existence of bionomic equilibrium of the system is identified. The optimal harvesting scheme is designed using Pontryagin's maximum principle. Also the population intensities of fluctuations (variances) around the positive equilibrium due to Gaussian additive white noise are computed using Fourier transform methods. The stability of the deterministic system and chaotic nature of stochastic system are portrayed for the chosen set of parameters in MATLAB simulations.

- **Sujay Datta** (Dept. of Statistics, University of Akron, Ohio, USA)

Contributed Talk, session [88]

Tuesday 10:45 – 11:15, Medical Sciences 4279

*Bayesian network structure learning with resampling-based confidence: application to genomics*

**Abstract:** In recent years, graphical and network models have become increasingly useful in certain areas of the biomedical sciences including genomics, proteomics, genetic epidemiology and systems biology. This has been facilitated by a number of new developments in structure learning algorithms and their evolution from constraint-based to score-based and hybrid algorithms. Examples include Gaussian graphical models, multivariate Bernoulli graphical models, hypergraph models, Bayesian networks and dynamic Bayesian networks. However, until recently, not much had been done in terms of developing a 'confidence measure' for the 'point estimates' of networks provided by those algorithms. Here we introduce recently developed methods for parametric and bootstrap-based assessment of variability in the learned network structures, introducing the notion of 'confidence' to structure learning, We do it in the context of applying Bayesian networks to learn functional relationships among genes associated with differentiation potential of progenitor cells that generate muscle tissue, in order to assess the effect of aging on the heterogeneity in possible co-expression of differentiation pathways.

- **Richard Davis** (Columbia University)

IMS Presidential Address

Monday 5:10 – 6:00, Fields Institute 230

*Are we meeting the challenge?*

**Abstract:** Fifteen to twenty years ago, it was thought that the statistics profession would have to change in order to survive in a world for which technology would have a profound impact on virtually every aspect of our lives. With the ability to acquire unprecedented amount of data, who would need stochastic models and statistical inference. Many predicted that computer science and other engineering disciplines would own statistics and that eventually departments of statistics would eventually shrink or disappear entirely. This prognostication has not yet come true. I will suggest that statistics and probably have adapted well and are in fact thriving disciplines.

- **Tarlowski Dawid** (Jagiellonian University, Cracow, Poland)

Contributed Talk, session [72]

Thursday 2:35 – 3:05, Rosebrugh 310

*Nonautonomous dynamical systems in stochastic global optimization*

**Abstract:** We focus on discrete-time inhomogeneous Markov recursions of the form  $X_{t+1} = T_t(X_t, Y_t)$ , where  $Y_t$  is an independent sequence. Given the continuous objective function  $f: A \rightarrow [0, \infty)$ , where  $A$  is a metric space, we target the problem how to prove the convergence of the process  $X_t$  towards the set  $A^*$  of global minima of  $f$  (we assume  $A^* \neq \emptyset$ ). Our methodology is to view the  $X_t$  as a nonautonomous dynamical system on Borel probability measures on  $A$  and to use the Lyapunov function technique. We present general convergence theorem and provide some practical examples, including the Simulated Annealing Algorithm.

- **Hugo de la Cruz Cancino** (School of Applied Mathematic. FGV. Rio de Janeiro)

Contributed Talk, session [49]

Thursday 11:20 – 11:50, Medical Sciences 3278

*MS-Stable integrators for the strong approximation of Ito diffusion processes*

**Abstract:** This work develops a new approach for constructing strong local linearization-based methods for the approximation of Ito diffusion process solution of multidimensional non-autonomous stochastic differential equations (SDEs). Using this approach, we construct new integrators for highly stiff stochastic equations and consequently suitable for the analysis and integration of important types of systems that function in the presence of random perturbations. We analyze the mean square stability property and the rate of convergence of the proposed methods. A comparative study with other integrators in the literature is presented, showing the benefits of the proposed methods in the integration of SDEs driven, in general, by multiplicative noise.

Keywords: diffusion processes, stochastic differential equations, stability, local linearization method, multiplicative noise, approximation methods

- **Geoffrey Decrouez** (Higher School of Economics)

Contributed Talk, session [80]

Friday 11:20 – 11:50, Rosebrugh 208

*Finite sample properties of the mean occupancy counts and probabilities*

Coauthors: Quentin Paris (Higher School of Economics), Michael Grabchak (University of North Carolina)

**Abstract:** For a probability distribution  $P$  on an at most countable alphabet, we give finite sample bounds for the expected occupancy counts and probabilities. In particular, both upper and lower bounds are given in terms of the right tail  $\nu$  of the counting measure of  $P$ . Special attention is given to the case where  $\nu$  is bounded by a regularly varying function. In this case, it is shown that our general results lead to an optimal-rate control of the expected occupancy counts and probabilities with explicit constants. Our results are also put in perspective with Turing's formula and recent concentration bounds to deduce confidence regions.

- **Pierre Del Moral** (INRIA)

Plenary talk: IMS Medallion Lecture

Thursday 5:10 – 6:00, McLeod Auditorium

*An introduction to mean field particle methods*

**Abstract:** In the last three decades, there has been a dramatic increase in the use of Feynman-Kac type particle methods as a powerful tool in real-world applications of Monte Carlo simulation in computational physics, population biology, computer sciences, and statistical machine learning.

The particle simulation techniques they suggest are also called resampled and diffusion Monte Carlo methods in quantum physics, genetic and evolutionary type algorithms in computer sciences, as well as Sequential Monte Carlo methods in Bayesian statistics, and particle filters in advanced signal processing.

These mean field type particle methodologies are used to approximate a flow of probability measures with an increasing level of complexity. This class of probabilistic models includes conditional distributions of signals with respect to noisy and partial observations, non absorption probabilities in Feynman-Kac-Schrödinger type models, Boltzmann-Gibbs measures, as well as conditional distributions of stochastic processes in critical regimes, including quasi-invariant measures and ground state computations.

This lecture presents a pedagogical introduction to the stochastic modeling and the theoretical analysis of these sophisticated probabilistic models.

We shall discuss the origins and the mathematical foundations of these particle stochastic methods, as well as their applications in rare event analysis, signal processing, mathematical finance and Bayesian statistical inference. We illustrate these methods through several applications including random walk confinements, particle absorption models, nonlinear filtering, stochastic optimization, combinatorial counting and directed polymer models.

- **Aurore Delaigle** (University of Melbourne)

Contributed session: Functional data analysis [43]

Tuesday 2:35 – 3:05, Medical Sciences 3154

*New methods for analyzing partially observed functional data*

**Abstract:** We consider analysis of functional data which are only partially observed. Often in such cases, the observed fragments of curves are supported on quite different intervals, in which case

standard methods of analysis cannot be used. We propose new approaches to analysing fragments of curves observed on different intervals. The techniques we suggest involve discretising the observed fragments, and then extending them outside the interval where they were observed. Using the same approach we can construct estimators of the mean and covariance functions, and, for example, deal with functional linear regression.

- **Jean-François Delmas** (Ecole des Ponts, Univ. Paris Est)

Contributed Talk, session [60]

Thursday 11:20 – 11:50, Medical Sciences 3154

*Cost functionals for large random trees*

Coauthors: DHERSIN, Jean-Stéphane (Univ. Paris 13), SCIAUVEAU, Marion (Ecole des Ponts, Univ. Paris-Est)

**Abstract:** Additive tree functionals allows to represent the cost of many divide-and-conquer algorithms. We give a.s. asymptotics and fluctuations for such tree functionals for the Catalan model (random tree uniformly distributed among the binary ordered tree with given number of internal nodes). This relies on the natural embedding of binary trees into the Brownian excursion and then on elementary second moment computations. We recover results first given by Fill and Kapur (2005) and then by Fill and Janson (2009).

- **Frank den Hollander** (Leiden University, The Netherlands)

Plenary talk: IMS Medallion Lecture

Friday 8:30 – 9:20, McLeod Auditorium

*Metastability for interacting particle systems*

**Abstract:** Metastability is the phenomenon where a system, under the influence of a stochastic dynamics, moves between different subregions of its state space on different time scales. Metastability is encountered in a wide variety of stochastic systems. The challenge is to devise realistic models and to explain the experimentally observed universality that is displayed by metastable systems, both qualitatively and quantitatively.

In statistical physics, metastability is the dynamical manifestation of a first-order phase transition. In this talk I give a brief historical account of metastability in this context. After that I describe the metastable behaviour of one particular model, namely, the Widom-Rowlinson model on a two-dimensional torus subject to a Metropolis stochastic dynamics. In this model, particles are randomly created and annihilated inside the torus as if the outside of the torus were an infinite reservoir with a given chemical potential. The particles are viewed as points carrying disks, and the energy of a particle configuration is equal to the volume of the union of the disks, called the “halo” of the configuration. Consequently, the interaction between the particles is attractive.

We are interested in the metastable behaviour at low temperature when the chemical potential is supercritical. In particular, we start with the empty torus and are interested in the first time when we reach the full torus, i.e., the torus is fully covered by disks. In order to achieve the transition from empty to full, the system needs to create a sufficiently large droplet of overlapping disks, which plays the role of a “critical droplet” that triggers the crossover. In the limit as the temperature tends

to zero, we compute the asymptotic scaling of the average crossover time, show that the crossover time divided by its average is exponentially distributed, and identify the size and the shape of the critical droplet. It turns out that the critical droplet exhibits “surface fluctuations”, which need to be understood in order to obtain a fine estimate of the crossover time.

Based on joint work with Sabine Jansen (Bochum), Roman Kotecky (Warwick) and Elena Pulvirenti (Leiden). ”

- **Vanessa Didelez** (Professor of Statistics and Causal Inference Leibniz Institute & Department of Mathematics, University of Bremen, Germany)

Plenary talk: IMS Medallion Lecture

Thursday 4:15 – 5:05, McLeod Auditorium

*Causal reasoning for events in continuous time*

**Abstract:** Dynamic associations among different types of events in continuous time can be represented by local independence graphs as developed by Didelez (2008). Intuitively we say that a process is locally independent of another one if its short-term prediction is not improved by using the past of the other process, similar to Granger-non-causality; the graphical representation uses nodes for processes or events and the absence of a directed edge for local independence. Important independence properties can be read off these -possibly cyclic- graphs using delta-separation (Didelez, 2006) which generalises d-separation from DAGs. In related work, Røysland (2011, 2012) showed how causal inference based on inverse probability weighting (IPW), well known for longitudinal data (Robins et al., 2000), can be extended to the continuous-time situation using a martingale approach.

In the work that I will present at the Medallion lecture (joint work with Kjetil Rysland, Odd Aalen and Theis Lange), we start by defining causal validity of local independence graphs in terms of interventions, which in the context of events in time take the form of modifications to the intensities of specific processes, e.g. a treatment process; causal validity is given if the specification of the dynamic system is rich enough to model such an intervention. (similar to what is known as ‘modularity’ for causal DAGs). We then combine the above previous developments to give graphical rules for the identifiability of the effect of such interventions via IPW; these rules can be regarded as characterising ‘unobserved confounding’. Re-weighting then simply replaces the observed intensity by the one given by the intervention of interest. For this to be meaningful, causal validity and identifiability are crucial assumptions. Our approach can be regarded as the time-continuous version of Dawid & Didelez (2010), who develop a decision theoretic approach for sequential decisions in longitudinal settings and use a graphical representation with influence diagrams that include decision nodes; specifically causal validity is analogous to the extended stability of Dawid & Didelez (2010). As an aside, we find that it is helpful to also use causal reasoning when faced with censoring as the target of inference can often be regarded as the population in which censoring is prevented, i.e. its intensity is set to zero. We apply our theoretical results to the example of cancer screening in Norway. ”

- **Jian Ding** (University of Chicago)

Invited session: Random walks on graphs [9]

Wednesday 2:00 – 2:30, Medical Sciences 3153

*Random planar metrics, Gaussian free fields, simple and non-simple random walks*

**Abstract:** I will present some recent results on random planar metrics that arise from two-dimensional discrete Gaussian free fields, and discuss their connections to some simple/non-simple random walks in two dimensions.

- **Vu C. Dinh** (Fred Hutchinson Cancer Research Center)

Contributed Talk, session [76]

Friday 11:20 – 11:50, Fitzgerald 139

*Hamiltonian Monte Carlo on the space of phylogenies*

Coauthors: Arman Bilge (University of Washington), Frederick A Matsen IV (Fred Hutchinson Cancer Research Center)

**Abstract:** Evolutionary tree inference, or phylogenetics, is an essential tool for understanding biological systems from deep-time divergences to recent viral transmission. The Bayesian paradigm is now commonly used in phylogenetics to describe support for inferred tree structures or to test hypotheses that can be expressed in phylogenetic terms. However, current Bayesian phylogenetic inference algorithms are limited to about 1,000 sequences, which is much fewer than are available via modern sequencing technology.

Here we develop phylogenetic Hamiltonian Monte Carlo (HMC) as a new approach to enable phylogenetic inference on larger data sets. Because phylogenetic trees have both discrete and continuous structure, we must go beyond the current implementations of HMC on manifolds with boundary. To do so, we develop a probabilistic version of the leap-frog algorithm, which can explore tree space realized as a cubical complex. This algorithm generalizes previous algorithms by doing classical HMC in the interior of the component Euclidean spaces, but making random choices between alternative paths available at a boundary. We show that the resulting Markov chain is ergodic, and provide a proof-of-concept implementation in open-source software.

- **Arnaud Doucet** (University of Oxford)

Plenary talk: IMS Medallion Lecture

Monday 4:15 – 5:05, McLeod Auditorium

*On pseudo-marginal methods for Bayesian inference in latent variable models*

**Abstract:** The pseudo-marginal algorithm is a popular variant of the Metropolis–Hastings scheme which allows us to sample asymptotically from a target probability density when we are only able to estimate unbiasedly an un-normalized version of it. It has found numerous applications in Bayesian statistics as there are many latent variable models where the likelihood function is intractable but can be estimated unbiasedly using Monte Carlo samples. In this talk, we will first review the pseudo-marginal algorithm and show that its computational cost is for many common applications quadratic in the number of observations at each iteration. We will then present a simple modification of this methodology which can reduce very substantially this cost. A large sample analysis of this novel pseudo-marginal scheme will be presented.

- **David Driver** (University of Cambridge)

Contributed Talk, session [56]

Thursday 2:00 – 2:30, Medical Sciences 3278

*Optimisation-based representations of the FKPP equation*

Coauthor: Dr Michael Tehranchi (University of Cambridge)

**Abstract:** We study solutions  $u \equiv u(t, x)$  to an FKPP-type reaction-diffusion equations of the form

$$\frac{\partial u}{\partial t} = \mathcal{L}u + u(1 - u^k), \text{ for } (t, x) \in (0, \infty) \times \mathbb{R}^n$$

$$u(0, x) = F(x) \text{ for } x \in \mathbb{R}^n$$

for a generator  $\mathcal{L}$  (generating a Feller process for example) and  $k \in \mathbb{R} \setminus \{-1, 0\}$ . The simplest example of this is the FKPP equation arising in the study of branching Brownian motion.

Under some simple assumptions the solution of the above equation can be represented by the value function of various optimisation problems which we present. We outline how this can give probabilistic bounds on the solution to the equation. We then show how this can be applied to an example in the theory of utility maximisation.

- **Haosui Duanmu** (University of Toronto)

Contributed Talk, session [65]

Friday 11:55 – 12:25, Medical Sciences 4279

*Nonstandard analysis and its application to Markov processes*

Coauthors: Jeffery S. Rosenthal (Department of Statistics, University of Toronto), William Weiss (Department of Mathematics, University of Toronto)

**Abstract:** A time-homogeneous Markov process with a stationary probability distribution  $\pi$  will converge to  $\pi$  in an appropriate sense (i.e., will be “ergodic”), under suitable conditions (such as “irreducibility”). This phenomenon is well understood for processes in discrete time and space, and for processes in continuous time and discrete space, and for processes in discrete time and continuous space. However, for processes in continuous time and space, there are apparently no such clean results; the closest are apparently the results by Meyn and Tweedie using awkward assumptions about skeleton chains together with drift conditions.

Meanwhile, nonstandard analysis is a useful tool for providing intuitive new proofs as well as new results to all areas of mathematics, including probability and measure theory. One of the strengths of nonstandard analysis is to provide a direct passage to link discrete mathematical results to continuous mathematical results.

In this new paper, we apply nonstandard analysis to Markov Chain Theory. We give a nonstandard characterization of general Markov chains which allows us to view every Markov chain as a “discrete” process. The time line  $T$  is chosen to be  $\{0, \delta t, 2\delta t, \dots, K\}$  where  $\delta t$  is some positive infinitesimal and  $K$  is an infinite number. We then “cut” the state space into hyperfinitely many pieces with infinitesimal diameters and we look at the transition probability to every infinitesimal piece at time  $\delta t$ . We call such a process a hyperfinite Markov process. Hyperfinite Markov processes behave like Markov processes on finite state spaces with discrete time lines in many ways due to the similarity between finite sets and hyperfinite sets. For every standard Markov chain  $\{X_t\}_{t \geq 0}$ , we can construct a hyperfinite Markov chain (that is, a Markov chain with hyperfinite state space and a hyperfinite

time line)  $\{X'_t\}_{t \in T}$  that inherits most of the key properties of  $\{X_t\}_{t \geq 0}$ . Moreover, under moderate assumptions on  $\{X_t\}_{t \geq 0}$ , we can prove the Markov Chain Ergodic theorem for  $\{X'_t\}_{t \in T}$ . The proof is similar to the proof of the Ergodic theorem for Markov processes on finite state spaces with discrete time lines. Then the ergodicity of  $\{X'_t\}_{t \in T}$  will naturally imply the ergodicity of  $\{X_t\}_{t \geq 0}$ .

- **Artem Dudko** (Stony Brook University)

Contributed Talk, session [64]

Friday 2:00 – 2:30, Medical Sciences 3278

*On Kesten-Neumann-Serre measures associated to countable groups*

**Abstract:** Let  $G$  be a countable group. Fix a finite set  $S$  of elements of  $G$  and a decreasing sequence of finite index subgroups  $P_n < G$ . For each  $n$  denote by  $\Gamma_n$  the Schreier graph associated to  $S$  with the vertex set  $G/P_n$ . Grigorchuk and Zuk using results of Serre showed that the counting spectral measure  $\tau_n$  of the Markov operator of uniform random walk on  $\Gamma_n$  converges weakly to some measure  $\tau_*$  called the Kesten-Neumann-Serre (KNS) measure. I will show that  $\tau_*$  coincides with a spectral measure of a Markov operator of certain groupoid representation of  $G$ . This allows to compute momentums of  $\tau_*$  and explains the relation between the KNS measure and the Kestens measures of limits of  $\Gamma_n$ .

- **Maurice Duits** (Royal Institute of Technology (KTH))

Invited session: Integrable models in Statistical Physics [2]

Thursday 11:55 – 12:25, Medical Sciences 3153

*GlobGlobal fluctuations for non-colliding process*

**Abstract:** In this talk I will discuss a recent approach for studying the global fluctuations for a class of non-colliding processes with determinantal correlations. The method is based on recurrence relations for the associated biorthogonal family of functions. The main results are Central Limit Theorems for multi-time linear statistics. The results show the universality of the Gaussian Free Field appearing in the global fluctuations for these models. Special attention will be given to Dyson's Brownian motion with one initial point and several endpoints.

- **Claudio Durastanti** (Ruhr Universität, Bochum, Germany)

Contributed Talk, session [87]

Wednesday 10:45 – 11:15, Medical Sciences 2171

*Adaptive nonparametric estimation on the sphere*

**Abstract:** This talk is concerned with the construction of adaptive nonparametric estimates by means of wavelet thresholding techniques over the  $q$ -dimensional sphere  $\mathbb{S}^q$ . In particular, our aim is to describe global hard thresholding nonparametric estimation of functions defined on the sphere.

Our proposal is based upon a wavelet system named spherical needlets, featuring strong concentration properties in both real and harmonic domains; spherical needlets are also characterized by very useful bounds on  $L^p$ -norms. More in details, for any fixed resolution level  $j \geq 0$ , let  $\{\psi_{j,k}\}_{k=1,\dots,K_j}$  and  $\{\beta_{j,k}\}_{k=1,\dots,K_j}$  denote respectively the set of spherical needlets and the set of needlet coefficients, while  $K_j$  is the cardinality of the needlets corresponding to  $j$  (typically  $K_j \approx 2^j$ ). For  $f \in L^p(\mathbb{S}^q)$ ,

the global thresholding needlet estimator is given by

$$\hat{f}(x) = \sum_{j \geq 0} \eta_j \sum_{k=1}^{K_j} \hat{\beta}_{j,k} \psi_{j,k}(x),$$

where,  $\{\hat{\beta}_{j,k}, j \geq 0, k = 1, \dots, K_j\}$  is the set of the so-called empirical needlet coefficients, based on the set of observations of size  $n$ , while  $\eta_j$  is a suitably defined U-statistics. Observe that empirical needlet coefficients are unbiased, namely  $\mathbb{E}[\hat{\psi}_{j,k}] = \psi_{j,k}$ , for any  $j \geq 0, k = 1, \dots, K_j$ .

Needlet estimators represent a convenient choice because they attain optimal rates of convergence on a large scale of functional classes on  $\mathbb{S}^q$ : in particular, globally thresholded needlet estimators are proved to be adaptive for the  $L^p$ -risk and for the scale of Besov spaces over the sphere. Finally, we also provide a comparison between our results and the ones existing in the literature concerning local and block thresholding methods on the sphere.

- **Richard Durrett** (Duke University)

Invited session: Statistics and networks [27]

Wednesday 3:10 – 3:40, Medical Sciences 4171

*Evolving voter models: recent results and open problems*

**Abstract:** The evolving voter model was introduced by Holme and Newman in 2006. In this model when an individual looks at the state of a neighbor with a different opinion then with probability  $p$  the connection is severed and a new neighbor is chosen (i) at random or (ii) from those with the same opinion as yours. In 2012 Durrett and seven coauthors performed an analysis of the model on sparse graphs and based on a combination of simulation of heuristics which showed that the two variants had much different behavior. Recently Basu and Sly has rigorously proved the existence of a phase transition from rapid disconnection to prolonged persistence in the dense case, but were not able to demonstrate the interesting qualitative behavior. In this talk we will review the existing results and describe new joint work with Anirban Basak on the case of thick graphs, where the degree is power  $a$  of the number of vertices with  $0 < a < 1$ .

- **Rick Durrett** (Duke University)

Invited session: Stochastic models of cancer [17]

Wednesday 11:55 – 12:25, Medical Sciences 3154

*Spatial evolutionary games with small selection coefficients*

**Abstract:** The use of evolutionary game theory biology dates to work of Maynard-Smith who used it to explain why most fights between animals were of the limited war type. In the mid 1990s Nowak and May demonstrated that a spatial distribution of players can explain the existence of altruism, which would die out in a homogeneously mixing population. Recently, evolutionary games have been used to model cancer, e.g., the fact that stromal cells may cooperate with prostate tumor cells to help cancer spread.

We use results of Cox, Durrett, and Perkins for voter model perturbations to study spatial evolutionary games when the fitness differences between strategies are small. We prove that the effect of space is equivalent to (i) changing the entries of the game matrix and (ii) replacing the replicator

ODE by a related PDE. The first idea is due to Ohtsuki and Nowak (for the pair approximation) while the second is well known in the theory of stochastic spatial processes. A remarkable aspect of our result is that the limiting PDE depends on the interaction kernel only through the values of two simple noncoalescence probabilities. Due to results of Aronson and Weinberger, and Fife and McLeod, we can analyze any  $2 \times 2$  game. However, when there are three strategies the limiting object is a system of reaction diffusion equations, so we only have results for special cases.

**E**

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Georgios Effraimidis** (University of Southern Denmark, USC)  
 Contributed Talk, session [73]  
 Wednesday 2:35 – 3:05, Fitzgerald 139  
*Measuring the effect of polygenic risk score on the aging rate*  
 Coauthors: Morgan Levine (UCLA), Eileen Crimmins (USC)  
**Abstract:** Population aging is widely recognized as a major demographic feature and challenge of the current century. The task to find genetic factors that affect the aging rate is of great importance as we can get insights into the aging process. The goal of this paper is to construct a statistical model for the study of the association between genetic factors and aging rate. By performing a genome-wide association study, we first calculate the polygenic risk score, which combines different genetic factors into a single numerical value. Next, we construct a model which allows the aging rate to depend on the polygenic risk score. We apply our methodology to mortality and genetic data from the Health and Retirement Study survey.
- **Georgios Effraimidis** (Department of Business and Economics, University of Southern Denmark)  
 Contributed Talk, session [82]  
 Thursday 11:20 – 11:50, Rosebrugh 208  
*Nonparametric identification of a time-varying frailty model*  
**Abstract:** Duration analysis is emerging as a popular technique in diverse research fields such as economics, finance, medical science and political science. In duration analysis, the Mixed Proportional Hazard model is the most common choice among practitioners for the specification of the underlying hazard rate. One major drawback of this model is that the value of the frailty term (i.e. unobserved factors) is time-invariant. This paper introduces a new model, the Mixed Random Hazard (MRH) model, which allows the frailty term to be time-varying. We provide sufficient conditions under which the new model is nonparametrically identified. Moreover, a theoretical framework is proposed for testing whether the true model is MRH. We conclude this paper with a discussion of how the arguments for the univariate MRH model can be extended to various multivariate problems.
- **Nathaniel Eldredge** (University of Northern Colorado)  
 Contributed Talk, session [64]  
 Friday 2:35 – 3:05, Medical Sciences 3278

*Strong hypercontractivity and logarithmic Sobolev inequalities on stratified complex Lie groups*

Coauthors: Leonard Gross (Cornell University), Laurent Saloff-Coste (Cornell University)

**Abstract:** The Ornstein–Uhlenbeck semigroup on Euclidean space with Gaussian measure has two key properties: hypercontractivity (an  $L^p$  improvement over time) and the logarithmic Sobolev inequality. It has long been known that these two properties remain equivalent if we work with any of a wide variety of semigroups, state spaces, and measures. If we work instead on complex Euclidean space, and restrict our attention to holomorphic functions, an improved form of hypercontractivity holds, called Janson’s strong hypercontractivity. Gross showed that in greater generality, Janson’s property is also implied by the logarithmic Sobolev inequality. In this talk, we present an extension of Gross’s result to a stratified complex Lie group equipped with a subelliptic heat kernel measure. In this setting, key assumptions of Gross’s proof are no longer satisfied, yet other features of the space and measure allow us to prove an appropriate version of the result anyway.

This talk is based on joint work with Leonard Gross and Laurent Saloff-Coste.

- **Mohamad Elmasri** (McGill University)

Contributed Talk, session [76]

Friday 11:55 – 12:25, Fitzgerald 139

*Latent score model for Host-Parasite link prediction*

Coauthor: Maxwell Farrell (McGill University)

**Abstract:** Predicting species interactions has become one of the great challenges facing modern ecologists, and particular interest is shown in understanding parasite-host interactions. We formulate this problem in a bipartite graph structure, connecting hosts on one side with parasites on the other, where edges between each host-parasite pair represent an interaction. We first construct a prior network of associations by drawing from available literature on host-parasite interactions. To predict undocumented interactions, we use a hierarchical Bayesian latent score framework for bipartite graphs where we incorporate a Markov random field type of dependence informed by host phylogeny. The addition of phylogenetic information to the model preserves many characteristics of the data, such as the sparseness of the network and the scale free degree distribution of nodes. We show that such a model can easily incorporate weighted or binary graphs, and different forms of neighbourhood structures. We test this model on subsets of two published databases, the Global Mammal Parasite Database (GMPD) and the Enhanced Infectious Diseases (EID2) database, each with thousands of documented host-parasite interactions. In addition, we extend the model by integrating a correction mechanism for false negative predictions, which proves valuable in reducing the uncertainty regarding absences of interaction in these databases. We compare the performance of different variations of our model to the standard nearest-neighbour method. We find all variations to be competitive, with some outperforming the nearest-neighbour approach, in particular when correcting for false negative proportions in the extended version.

- **Andreas Elsener** (Seminar for Statistics, ETH Zurich)

Contributed Talk, session [98]

Thursday 2:00 – 2:30, Rosebrugh 208

*Robust low-rank matrix estimation*

Coauthor: Sara van de Geer (Seminar for Statistics, ETH Zurich)

**Abstract:** The problem of matrix completion via nuclear norm penalized empirical risk minimization has attracted a lot of attention in the past years. Boosted by the Netflix Prize and other recommendation systems many estimators and theoretical results have been proposed. However, a main (desirable) property of these estimators has been neglected: robustness against outliers. Robustness is an appealing property also in this framework since it could produce estimators that are e.g. “resistent” to frauds in recommendation systems. In this work, we consider two robust estimators relying on the well-known absolute value and Huber loss functions. As it is often the case in similar contexts a comparison with an oracle that already “knows” the rank of the matrix that is estimated is of great usefulness. A non-sharp oracle inequality for the absolute value loss as well as a sharp oracle inequality for the Huber loss are shown to hold with high probability. As far as the “deterministic” part of the problem is concerned we use properties of the nuclear norm (sum of the singular values of a matrix) and describe the similarities with the  $\ell_1$ -norm penalized estimators in a regression context. On the other hand, the empirical process part is treated using techniques such as the symmetrization and contraction theorems that allow us to control the random fluctuations of the non-deterministic part of the problem. The oracle inequalities are used to derive bounds for the nuclear norm estimation error. Since in many applications a matrix is not exactly low rank but rather with some few large singular values and many small ones, we have considered also the case of an approximately low rank matrix.

- **Philip Ernst** (Rice University)

Contributed Talk, session [52]

Friday 11:20 – 11:50, Medical Sciences 4171

*On the volatility of the correlation of two independent Wiener processes*

Coauthors: Larry Shepp (deceased) (The Wharton School of the University of Pennsylvania), Abraham Wyner (The Wharton School of the University of Pennsylvania)

**Abstract:** In this talk, we resolve a longstanding open statistical problem. The problem is to analytically determine the second moment of the empirical correlation coefficient

$$\theta := \frac{\int_0^1 W_1(t)W_2(t)dt - \int_0^1 W_1(t)dt \int_0^1 W_2(t)dt}{\sqrt{\int_0^1 W_1^2(t)dt - [\int_0^1 W_1(t)dt]^2}} \sqrt{\int_0^1 W_2^2(t)dt - [\int_0^1 W_2(t)dt]^2}$$

of two independent Wiener processes,  $W_1, W_2$ . Using tools from Fredholm integral equation theory, we successfully calculate the second moment of  $\theta$  to obtain a value for the standard deviation of  $\theta$  of nearly .5. The correlation is volatile in the sense that its distribution is heavily dispersed and is frequently large in absolute value, and it is induced because each Wiener process is “self-correlated” in time. This is because a Wiener process is an integral of pure noise and thus its values at different time points are correlated. In addition to providing an explicit formula for the second moment of  $\theta$ , we offer implicit formulas for higher moments of  $\theta$ .

- **Alison Etheridge** (Oxford)

Invited session: Stochastic models of evolution [16]

Monday 3:10 – 3:40, Medical Sciences 3154

*Branching Brownian motion, mean curvature flow and the motion of hybrid zones*

**Abstract:** A hybrid zone is a narrow geographic region where two genetically distinct populations are found close together and hybridise to produce offspring of mixed ancestry. They are maintained by a balance between selection and dispersal. We discuss recent work, joint with Nic Freeman (Sheffield) and Sarah Penington (Oxford) which investigates how hybrid zones will evolve over time.

- **Stewart N. Ethier** (University of Utah)

Contributed Talk, session [55]

Friday 2:00 – 2:30, Medical Sciences 4171

*Wright–Fisher construction of the two-parameter Poisson–Dirichlet diffusion*

Coauthors: Cristina Costantini (University of Chieti-Pescara), Pierpaolo De Blasi (University of Torino and Collegio Carlo Alberto), Matteo Ruggiero (University of Torino and Collegio Carlo Alberto), Dario Spanò (University of Warwick)

**Abstract:** The two-parameter Poisson–Dirichlet diffusion, recently introduced by Petrov, extends the infinitely-many-neutral-alleles diffusion model, related to Kingman’s one-parameter Poisson–Dirichlet distribution and to certain Fleming–Viot processes. The additional parameter has been shown to regulate the clustering structure of the population, but is yet to be fully understood in the way it governs the reproductive process. Here we shed some light on these dynamics by formulating a  $K$ -allele Wright–Fisher model for a population of size  $N$ , involving a uniform parent-independent mutation pattern and a specific state-dependent immigration kernel. Suitably scaled, this process converges in distribution to a  $K$ -dimensional diffusion process as  $N \rightarrow \infty$ . Moreover, the descending order statistics of the  $K$ -dimensional diffusion converge in distribution to the two-parameter Poisson–Dirichlet diffusion as  $K \rightarrow \infty$ . The choice of the immigration kernel depends on a delicate balance between reinforcement and redistributive effects. The proof of convergence to the infinite-dimensional diffusion is nontrivial because the generators do not converge on a core. Our strategy for overcoming this complication is to prove *a priori* that in the limit there is no “loss of mass”, i.e., that, for each limit point of the finite-dimensional diffusions (after a reordering of components by size), allele frequencies sum to one.

**F**

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Adrian Falkowski** (Nicolaus Copernicus University, Toruń)

Contributed Talk, session [50]

Thursday 2:00 – 2:30, Medical Sciences 4171

*SDEs with time-dependent reflecting barriers driven by processes with bounded  $p$ -variation.*

Coauthor: Leszek Słomiński (Nicolaus Copernicus University, Toruń)

**Abstract:** We study the problems of existence, uniqueness and approximation of solutions to finite-dimensional stochastic differential equations with constraints driven by general processes with bounded  $p$ -variation,  $p \geq 1$ . More precisely, let  $f : \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $g : \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d$  be measurable functions,  $A$  be a one-dimensional process with locally bounded variation,  $Z$  be a  $d$ -dimensional process with locally bounded  $p$ -variation and  $L, U$  be  $d$ -dimensional processes such that  $L_t^i \leq U_t^i$ ,  $i = 1, \dots, d$ ,  $t \in \mathbb{R}^+$ . We consider SDEs with reflecting boundary condition of the form

$$(1) \quad X_t = X_0 + \int_0^t f(s, X_{s-}) dA_s + \int_0^t g(s, X_{s-}) dZ_s + K_t, \quad t \in \mathbb{R}^+.$$

By a solution to (1) we mean a pair  $(X, K)$ , where  $X = \{X_t\}_{t \in \mathbb{R}^+}$  is process such that  $X_t^i \in [L_t^i, U_t^i]$ ,  $i = 1, \dots, d$ ,  $t \in \mathbb{R}^+$  and  $K = \{K_t\}_{t \in \mathbb{R}^+}$ , called regulator term, is process changing values only when  $X^i$  is living on  $L^i$  or  $U^i$  for some  $i = 1, \dots, d$ . Equation (1) is called the Skorokhod SDE in analogy with the case  $L = 0$ ,  $U = \infty$  first discussed by Skorokhod [1] for a standard Brownian motion in place of  $Z$  and  $A_t = t$ ,  $t \in \mathbb{R}^+$ . Our main tool are new estimates showing Lipschitz continuity of the deterministic Skorokhod problem in  $p$ -variation norm.

[1] A. V. Skorokhod, Stochastic equations for diffusion processes in a bounded region 1,2, Theory Probab. Appl. 6 (1961) 264–274, 7 (1962) 3–23.

- **Yuguang Fan** (University of Melbourne)

Contributed Talk, session [84]

Monday 3:10 – 3:40, Rosebrugh 310

*Clustering partially observed functional data*

Coauthor: Aurore Delaigle (University of Melbourne)

**Abstract:** We study curves that are observed only on short and irregular intervals. Such fragmentary functional data are observed in practice but it poses several difficulties when applying the usual techniques designed for curves defined on the entire interval. In the fragment context, an interesting phenomenon occurs when adapting the usual functional  $K$  means clustering algorithm. Under these circumstances, we propose a novel method to cluster such fragments and overcome such problems.

- **Zhou Fan** (Stanford University)

Invited session: Inference for high-dimensional covariance structures [29]

Tuesday 3:10 – 3:40, Medical Sciences 3153

*Eigenvalue distributions of variance components estimators in high-dimensional random effects models*

**Abstract:** In quantitative genetics, estimation of variance component covariance matrices in multivariate random effects models is a common method for determining the genetic basis of variation in quantitative traits and in predicting the joint phenotypic responses of traits to selection. We study the spectra of MANOVA and other quadratic variance components estimators in such models. In the asymptotic regime where the number of traits is comparable to the number of samples, we show that the empirical eigenvalue distributions of these estimators are well-approximated by deterministic laws, characterized by systems of fixed-point equations in their Stieltjes transforms. Under a certain

global null hypothesis, we also show that the largest eigenvalue distributions of these estimators converge to Tracy-Widom limits. Our proofs of bulk and edge convergence use, respectively, tools of operator-valued free probability theory and a resolvent comparison argument using a Green function flow. We specialize our results to some simple experimental designs that arise in quantitative genetics applications.

- **Xiao Fang** (National University of Singapore)

Contributed Talk, session [70]

Friday 2:35 – 3:05, Medical Sciences 4279

*Probability approximations by Stein's method and applications*

**Abstract:** Stein's method has proved to be a powerful tool for probability approximations under dependence. In this talk, we give a brief introduction to Stein's method. We then discuss three recent applications of Stein's method including a universal central limit theorem for the number of monochromatic edges in uniformly colored graphs, steady-state diffusion approximation for a queuing model, and Poisson approximation for scan statistics.

- **Long Feng** (Rutgers University)

Contributed Talk, session [93]

Thursday 10:45 – 11:15, Rosebrugh 310

*Scaled concave penalized regression*

Coauthor: Cun-hui Zhang (Rutgers University)

**Abstract:** Over the past decade, concave penalized least-square estimator (PLSE) has been studied intensively. It is been shown that the PLSE guarantees variable selection consistency under significantly weaker conditions than the Lasso, for example, the strong irrepresentable condition on the design matrix required by the Lasso can be replaced by a sparse Riesz condition. Moreover, the concave PLSE also enjoys rate optimal error bounds in prediction and coefficients estimation. However, the error bounds for prediction and coefficients estimation in the literature still requires significant stronger condition than what Lasso requires, for example, the knowledge of the  $\ell_1$  norm of the true coefficients vector or the upper sparse eigenvalue condition. Ideally, selection, prediction and estimation property should only depend on lower sparse eigenvalue/restricted eigenvalue, it that achievable? In this paper, we give an affirmative answer to this question.

We prove that the concave PLSE matches the oracle inequalities for prediction error and  $\ell_q$  coefficients estimation error for the Lasso, with  $1 \leq q \leq 2$ , based only on the restricted eigenvalue condition, which can be viewed as nearly the weakest available condition on design matrix. Furthermore, under a uniform signal strength condition, selection consistency does not require any additional conditions for proper concave penalties such as the SCAD penalty and MCP. Our theorem applies to all the local solutions that computable by path following algorithms starting from the origin. We also developed a scaled version of concave PLSE, which jointly estimates the regression coefficients and noise level. The scaled concave PLSE also match the oracle inequalities for estimation of noise level with the scaled Lasso, and the computation cost is negligible beyond computing a continuous

solution path. All our consistency results applies to the case where the number of predictors  $p$  is much larger than the sample size  $n$ .

- **Daive Ferrari** (School of Mathematics and Statistics, University of Melbourne)

Contributed Talk, session [81]

Tuesday 2:35 – 3:05, Medical Sciences 3278

*Model selection confidence sets by likelihood ratio testing*

**Abstract:** The traditional activity of model selection aims at discovering a single model superior to all other candidate models. In the presence of pronounced noise, however, multiple models are often found to explain the same data equally well. To resolve this model selection ambiguity, we introduce model selection confidence sets (MSCSs) in the context of maximum likelihood estimation. A MSCS is defined by a list of models statistically equivalent to the true model at a user-specified level of confidence, thus extending the familiar notion of confidence intervals to the model-selection framework. We propose to construct MSCSs using the likelihood ratio test; our approach guarantees correct coverage probability of the true model when both sample size and model dimension increase. We derive conditions under which the MSCS contains all the relevant information about the true model structure. In addition, we propose natural statistics to measure importance of parameters in a principled way that accounts for the overall model uncertainty. When the overall space of feasible models is large, MSCSs is implemented by an adaptive stochastic search algorithm which samples MSCS models with high probability.

- **Johann Ferreira** (University of Pretoria)

Contributed Talk, session [74]

Tuesday 11:20 – 11:50, Fitzgerald 139

*Quadratic form of the complex singular matrix elliptical distribution in wireless communication systems*

Coauthors: Andriette Bekker (University of Pretoria), Mohammad Arashi (Sharhooth University of Technology)

**Abstract:** This paper presents some results related to the quadratic forms and their associated eigenvalues emanating from an underlying complex singular matrix elliptical distribution. These results have a computationally convenient form which allows for flexible application and calculations. These results are used to evaluate an important performance measure, namely channel capacity of spatially correlated Rayleigh distributed MIMO channels. Superior performance is observed when other underlying distributions are considered from the complex elliptical class rather than the complex matrix normal.

Keywords: Complex Wishart-type; capacity; elliptical distribution; complex singular matrix variate t-distribution

- **Alexander Fribergh** (Université de Montréal)

Invited session: Random media [5]

Monday 11:20 – 11:50, Medical Sciences 3153

*The ant in “a” labyrinth*

**Abstract:** One of the most famous open problem in random walks in random environments is to understand the behavior of a simple random walk on a critical percolation cluster, a model known as the ant in the labyrinth. I will present new results on the scaling limit for the simple random walk on the critical branching random walk in high dimension. In the light of lace expansion, we believe that the limiting behavior of this model should be universal for simple random walks on critical structures in high dimensions. This is a joint work with G. Ben Arous and M. Cabezas.

G

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Tamar Gadrich** (ORT Braude College)

Contributed Talk, session [55]

Friday 3:10 – 3:40, Medical Sciences 4171

*Asymptotic and numerical analysis of stochastic population dynamics using a finite Markov chain*

Coauthor: Guy Katriel (ORT Braude College)

**Abstract:** Deterministic discrete-time population models are widely used in the field of population ecology. Small populations are subject to stochastic effects, so it is natural to look for stochastic version of deterministic models that approach the deterministic models in the large-population limit. Population-level consequences of assumptions regarding the behavior of individual organisms (e.g. reproduction, dispersal, and competitive interactions) can be investigated using site-based models, an approach pioneered by Royama, Brännstrom and Sumpter. In site-based models density-dependence is modeled by assuming that new offspring are randomly (uniformly) distributed among a finite number of subsistence areas. An individual’s probability of survival to maturity depends on the number of other individuals sharing the same site. Those that survive to the next reproduction season produce a number of offspring according to a given fecundity distribution (e.g., a Poisson distribution).

Our study focused on a site-based model assuming scramble competition – i.e., only sole occupants of a site are able to survive to the next reproduction season. In the infinite-population limit, this stochastic converges to the Ricker model, a well-known deterministic discrete time model, which exhibits complex dynamics. We can therefore call the finite-population model a “stochastic Ricker process”. We formulate the stochastic Ricker process as an integer-valued discrete-time Markov chain (DTMC) with finite state space, whose transition probabilities are explicitly computed. The transition matrix is used to explore the limiting behavior of the model and address ecologically relevant questions such as expected time to extinction. Since the DTMC has an absorbing state, its long run behavior is described by the quasi-stationary distribution, which we compute numerically. When the population size is large, numerical computation of the quasi-stationary distribution is challenging. In this case, we applied an analytic approximation, the Gaussian Markov chain approximation, valid in large-population limit. These tools enable us to explore the behavior of the stochastic Ricker process, and compare it to that of the deterministic Ricker model.

- **Chao Gao** (Yale)

Invited session: Community detection in random graphs [26]

Thursday 11:55 – 12:25, Medical Sciences 2172

*Community detection in degree-corrected block models*

**Abstract:** Community detection is a central problem of network data analysis. Given a network, the goal of community detection is to partition the network nodes into a small number of clusters, which could often help reveal interesting structures. The present paper studies community detection in Degree-Corrected Block Models (DCBMs). We first derive asymptotic minimax risks of the problem for a misclassification proportion loss under appropriate conditions. The minimax risks are shown to depend on degree correction parameters, community sizes, and average within and between community connectivities in an intuitive and interpretable way. In addition, we propose a polynomial time algorithm to adaptively perform consistent and even asymptotically minimax optimal community detection in DCBMs.

- **Fengnan Gao** (Leiden University)

Contributed Talk, session [68]

Tuesday 2:35 – 3:05, Rosebrugh 310

*On the statistical estimation of the preferential attachment network models*

Coauthor: Aad van der Vaart (Leiden University)

**Abstract:** The Preferential attachment (PA) models are a popular way of modeling the social networks, the collaboration networks and etc. The PA model is an evolving network where new nodes keep coming in. When a new node comes in, it establishes connections with existing nodes. The random choice on the existing node is via a multi-normal distribution with probability weights based on a preferential function  $f$  on the degrees, that maps the natural numbers to the positive real line.  $f$  is assumed apriori non-decreasing, which means the nodes with high degrees are more likely to get new connections, i.e. “the rich get richer”. 1) If we pose no condition of the shape of  $f$ , we may propose an estimator on  $f$ . We show, with techniques from branching process, our estimator is consistent. 2) If we assume some parametric form for the truth  $f_0 \in \{f_\theta : \theta \in \Theta\}$ , then we show the MLE for  $\theta$  is asymptotic normal, enabling us to conduct inferences on  $f$ . 3) If  $f$  is affine ( $f(k) = k + \delta$ ), it is well known that such a model leads to a limiting power-law degree distribution with power-law exponent  $3 + \delta$ . Thus this model is of particular importance to practitioners, who look for explanations of power-law relations in real-world applications. We proposed a maximum likelihood estimator for  $\delta$  and establish an asymptotic normality result on the MLE of  $\delta$ , and henceforth solve the estimation problem for the practitioners.

- **Geneviève Gauthier** (HEC Montréal)

Contributed Talk, session [47]

Monday 10:45 – 11:15, Medical Sciences 3154

*Credit and systemic risks in the financial services sector: evidence from the 2008 global crisis*

Coauthors: Jean-Francois Bégin (HEC Montréal), Mathieu Boudreault (UQAM)

**Abstract:** The Great Recession has shaken the foundations of the financial industry and led to a

tighter solvency monitoring of both the banking and insurance industries. To this end, we develop a portfolio credit risk model that includes firm-specific Markov-switching regimes as well as individual stochastic and endogenous recovery rates. Using weekly credit default swap premiums for 35 financial firms, we analyze the credit risk of each of these companies and their statistical linkages, emphasizing especially on the 2005-2012 period. Moreover, we study the systemic risk between the banking and insurance subsectors.

- **Ezra Gayawan** (Department of Statistics, Federal University of Technology, Akure, Nigeria)

Contributed Talk, session [79]

Thursday 3:10 – 3:40, Fitzgerald 139

*Geographical variations in infant and child mortality in West Africa: a geo-additive discrete-time survival modelling*

Coauthors: Mumini I Adarabioyo (Department of Mathematics and Physical Sciences, Afe Babalola University, Ado-Ekiti, Nigeria), Dupe M Okewole (Department of Mathematical Sciences, Redeemer's University, Ede, Nigeria)

**Abstract:** This study examines the residual geographical variations in infant and child mortality and how the different categories of the risk factors account for the spatial inequality in West African countries. To this end, we pooled data for 10 of the countries extracted from Demographic and Health Surveys and used the spatial extension of discrete-time survival model to examine how the variables exact influence on infant and child mortality across space. Inference was Bayesian based on MCMC technique. The spatial component was modeled assuming Gaussian Markov Random field while Bayesian P-spline was used for smooth function. We found different geographical patterns for infant and child mortality. In the case of children under-five, demographic factors inherent to the mother and child as well as maternal status variables when accounted for, explain away a good part of the huge variations observed in the crude spatial rates. There are no evidence of significant variations, however, in infant mortality except for three neighbouring regions of Liberia and Sierra Leone. The findings can guide in evidence based allocation of scarce resources in West Africa with the aim at improving the survival chance of young children.

- **S. K. Ghoreishi** (Department of Statistics, Qom University, IR Iran)

Contributed Talk, session [88]

Tuesday 11:55 – 12:25, Medical Sciences 4279

*Shrinkage estimators properties in multi-dimensional settings*

**Abstract:** From practical point of view, the univariate empirical Bayes shrinkage estimators have found vast application in many disciplines for both homoscedastic and heteroscedastic hierarchical models. It seems there is no serious work such that it has considered the optimum asymptotic properties of multiple empirical Bayes heteroscedastic shrinkage estimators. In this paper, we focus on the empirical Bayes estimation problem for a set of multivariate normal population means via a multivariate heteroscedastic hierarchical normal models and propose a class of multivariate shrinkage estimators based on Stein's unbiased estimate of risk (SURE). The asymptotic optimality properties of our multivariate shrinkage estimators are derived. Finally, we carry out a simulation study to

compare the performance of various multivariate competing shrinkage estimators and analyze a real data set.

- **Subhashis Ghoshal** (North Carolina State University)

Invited session: Frequentist properties of Bayesian nonparametric or high dimensional models [34]

Friday 10:45 – 11:15, Medical Sciences 2171

*Bayesian methods for boundary detection in images*

Coauthor: Meng Li (Duke University)

**Abstract:** Boundary detection in images has many important applications in spatial statistics, forestry, climatology, medical sciences and other fields and may be thought of as higher dimensional generalization of change-point problems. The boundary of a  $d$ -dimensional image may be viewed as a  $d - 1$ -dimensional manifold, and in particular a smooth closed, not self-intersecting curve for 2D images. We consider a Bayesian approach to the problem using a prior indexed by the unit circle, or the unit sphere in higher dimension, typically constructed from a Gaussian process or a finite random series using trigonometric polynomials or spherical harmonics basis. For the most important case of 2D images, a very convenient prior is the squared exponential periodic Gaussian process. Its explicit eigen decomposition in terms of Bessel functions allows a convenient computational scheme and obtaining posterior contraction rate. We show that the posterior contracts at the minimax optimal rate and adapts to the unknown smoothness level of the curve. Simulation experiments show that the method is exceptionally efficient and robust against misspecification.

- **Emmanuel Gobet** (Ecole Polytechnique)

Invited session: BSDEs and their applications in finance [19]

Friday 11:55 – 12:25, Medical Sciences 3154

*Data-driven regression Monte Carlo for BSDEs*

Coauthors: Gang Liu (Ecole Polytechnique, Paris-Saclay), Jorge Zubieli (IMPA, Rio)

**Abstract:** Our goal is to solve certain dynamic programming equations (like for discrete BSDEs) associated to a given Markov chain  $X$ , using a regression-based Monte Carlo algorithm. More specifically, we assume that the model for  $X$  is not known in full detail and only a root sample  $X^1, \dots, X^M$  of such process is available. By a stratification of the space and a suitable choice of a probability measure  $\nu$ , we design a new resampling scheme that allows to compute local regressions (on basis functions) in each stratum. The combination of the stratification and the resampling allows to compute the solution to the dynamic programming equation (possibly in large dimensions) using only a relatively small set of root paths. To assess the accuracy of the algorithm, we establish non-asymptotic error estimates in  $L_2(\nu)$ . Our numerical experiments illustrate the good performance, even with  $M = 20 - 40$  root paths.

- **Christina Goldschmidt** (Department of Statistics & Lady Margaret Hall, University of Oxford)

Plenary talk: IMS Medallion Lecture

Tuesday 5:10 – 6:00, McLeod Auditorium

*Scaling limits of critical random trees and graphs*

**Abstract:** I will begin by surveying some recent developments in the theory of scaling limits for

random trees and graphs. The overarching aim is to understand how random tree and graph structures on  $n$  vertices may be rescaled so that, as  $n$  tends to infinity, we obtain meaningful limits in distribution. These limits are usually formulated as random metric spaces, and are themselves fascinating mathematical objects. The prototypical example is Aldous' Brownian continuum random tree, which is the scaling limit of (for example) a uniform random labelled tree on  $n$  vertices.

Consider first a Galton-Watson forest, that is the sequence of family trees of a collection of independent and identically distributed Galton-Watson processes. We assume that the offspring distribution is critical. It is possible to encode the forest in terms of a random walk which (thanks to the criticality condition) is centred. Assuming additionally that the offspring distribution is in the domain of attraction of a stable law with parameter  $\alpha \in (1, 2]$ , this random walk converges (after a suitable rescaling) to a spectrally positive  $\alpha$ -stable Lévy process (Brownian motion if  $\alpha = 2$ ). Foundational work of Duquesne, Le Gall and Le Jan showed that the excursions of this Lévy process above its running infimum may be interpreted as encoding a forest, composed of continuum random trees, which I will refer to as a stable forest. In the second part of the talk, I will turn to work in progress with Guillaume Conchon-Kerjan on the model of a uniform random graph with independent and identically distributed vertex-degrees. The condition for criticality in this setting is  $E[D^2] = 2E[D]$ , and we assume additionally that  $P(D = k) \sim ck^{-(\alpha+2)}$  as  $k$  tends to infinity, for some  $\alpha \in (1, 2)$ . In this situation, it turns out that the largest components have sizes on the order of  $n^{\alpha/(\alpha+1)}$ . Building on earlier work of Joseph, we show that the components have scaling limits which can be related to a stable forest via an absolute continuity relation. This gives a natural generalisation of the scaling limit for the Erdos-Renyi random graph which I obtained in collaboration with Louigi Addario-Berry and Nicolas Broutin a few years ago (extending results of Aldous), and complements recent work on random graph scaling limits of various authors including Bhamidi, Broutin, Duquesne, van der Hofstad, van Leeuwen, Riordan, Sen, M. Wang and X. Wang.

- **Larry Goldstein** (Department of Mathematics, University of Southern California)

Invited session: Stein's method [12]

Monday 11:20 – 11:50, Medical Sciences 2171

*Size biased couplings and the spectral gap for random regular graphs*

Coauthors: Nicholas Cook (Department of Statistics, Stanford University), Tobias Johnson (Department of Mathematics, NYU)

**Abstract:** Let  $\lambda$  be the second largest absolute eigenvalue of a uniform random  $d$ -regular graph on  $n$  vertices. Friedman proved a conjecture of Alon that if  $d$  is fixed independent of  $n$ , then  $\lambda = 2\sqrt{d-1} + o(1)$  with high probability. We show that  $\lambda = O(\sqrt{d})$  continues to hold with high probability as long as  $d = O(n^{2/3})$ , making progress towards a conjecture of Vu that the bound holds for all  $1 \leq d \leq n/2$ . Our work depends on new results for concentration of measure, and specifically Bennett-type tail estimates for random variables admitting certain unbounded size biased couplings.

- **Adrian Gonzalez-Casanova** (Weierstrass Institute Berlin)

Invited session: Stochastic models of evolution [16]

Monday 2:00 – 2:30, Medical Sciences 3154

*The seedbank coalescent*

Coauthors: Noemi Kurt (TU Berlin), Jochen Blath (TU Berlin)

**Abstract:** We analyse patterns of genetic variability of populations in the presence of a large seed bank with the help of a new coalescent structure called seed bank coalescent. This ancestral process appears naturally as scaling limit of the genealogy of large populations that sustain seed banks, if the seed bank size and individual dormancy times are of the same order as the active population. Mutations appear as Poisson process on the active lineages, and potentially at reduced rate also on the dormant lineages. The presence of such dormant lineages leads to qualitatively altered times to the most recent common ancestor and non-classical patterns of genetic diversity. To illustrate this we provide a Wright-Fisher model with seed bank component and mutation, motivated from recent models of microbial dormancy, whose genealogy can be described by the seed bank coalescent. Interestingly, the scaling limit of this model (forward in time) can be interpreted in terms of a stochastic delay differential equation. Our results indicate that the presence of a large seed bank considerably alters the distribution of some distance statistics, as well as the site-frequency spectrum. (Joint work with Jochen Blath, Eugenio Buzzoni, Bjarki Eldon, Noemi Kurt and Maite Wilke-Berenguer)

- **Dan Goreac** (Université Paris-Est, LAMA (UMR 8050), CNRS)

Contributed Talk, session [77]

Tuesday 2:00 – 2:30, Medical Sciences 4279

*Parameter selection in gene networks using BSVIs and control*

**Abstract:** We present a mathematical model allowing to detect the infection time in gene networks with several stable regimes. The mathematical apparatus relies on a class of switch piecewise deterministic Markov processes. The switch is governed by pure jump modes and linked to DNA bindings. The differential component follows backward stochastic dynamics (of PDMP type) and is reflected in some mode-dependent non-convex domains. We give some elements of existence and uniqueness for the solution to these backward stochastic variational inclusions (BSVI) by reducing them to a family of ordinary variational inclusions with generalized reflection in semiconvex domains. Second, by considering control-dependent drivers, we hint to some model-selection approach by embedding the (controlled) BSVI in a family of regular measures. Regularity, support and structure properties of these sets are given. The talk is based on recent work with Eduard-Paul Rotenstein.

- **Roger Grosse** (University of Toronto)

Invited session: Models and inference for big data [38]

Friday 11:55 – 12:25, McLeod Auditorium

*Optimizing neural networks using structured probabilistic models of the gradient computation*

**Abstract:** Neural networks have recently driven significant progress in machine learning applications as diverse as vision, speech, and text understanding. Despite much engineering effort to boost the computational efficiency of neural net training, most networks are still trained using variants of stochastic gradient descent. Natural gradient descent, a second-order optimization method, has the

potential to speed up training by correcting for the curvature of the loss function. Unfortunately, the exact natural gradient is impractical to compute for large networks because it requires solving a linear system involving the Fisher matrix, whose dimension may be in the millions for modern neural network architectures. The key challenge is to develop approximations to the Fisher matrix which are efficiently invertible, yet accurately reflect its structure.

The Fisher matrix is the covariance of log-likelihood derivatives with respect to the weights of the network. I will present techniques to approximate the Fisher matrix using structured probabilistic models of the computation of these derivatives. Using probabilistic modeling assumptions motivated by the structure of the computation graph and empirical analysis of the distribution over derivatives, I derive approximations to the Fisher matrix which allow for efficient approximation of the natural gradient. The resulting optimization algorithm is invariant to some common reparameterizations of neural networks, suggesting that it automatically enjoys the computational benefits of these reparameterizations. I show that this method gives significant speedups in the training of two widely used architectures: restricted Boltzmann machines and convolutional networks.

- **Ankit Gupta** (ETH Zurich)

Invited session: Mathematical systems biology [18]

Thursday 10:45 – 11:15, Medical Sciences 2171

*Stability properties of stochastic biomolecular reaction networks: analysis and applications*

Coauthor: Prof. Mustafa Khammash (ETH Zurich)

**Abstract:** The internal dynamics of a cell is generally very noisy. An important source of this noise is the intermittency of reactions among various molecular species in the cell. The role of this noise is commonly studied using stochastic models for reaction networks, where the dynamics is described by a continuous-time Markov chain whose states represent the molecular counts of various species. In this talk we will discuss how the long-term behavior of such Markov chains can be assessed using a blend of ideas from probability theory, linear algebra and optimisation theory. In particular we will describe how many biomolecular networks can be viewed as generalised birth-death networks, which leads to a simple computational framework for determining their stability properties such as ergodicity and convergence of moments. We demonstrate the wide-applicability of our framework using many examples from Systems and Synthetic Biology. We also discuss how our results can help in analysing regulatory circuits within cells and in understanding the entrainment properties of noisy biomolecular oscillators.

- **Bhupendra Gupta** (PDPM Indian Institute of Information Technology, Design & Manufacturing – Jabalpur)

Contributed Talk, session [69]

Friday 11:55 – 12:25, Medical Sciences 3278

*Random growth model of online social networks*

**Abstract:** In last years, online social networks (ONS) have gained tremendous popularity and are becoming increasingly useful for various reasons. Peoples use these networks as platform of interact, communicate and collaborate with the others. An extensive study of growth models of such network

can answer many questions related areas like socio-economic. This motivates us to formulate a random growth model for online social network websites. We design a model based on user dynamic approach influenced by suitability of the online social network website. Also we can forecast the number of the user on the bases of the present available data.

- **Ori Gurel Gurevich** (Hebrew University of Jerusalem)

Invited session: Random walks on graphs [9]

Wednesday 2:35 – 3:05, Medical Sciences 3153

*Boundaries of planar graphs*

**Abstract:** We discuss the relation between properties of the simple random walk on a bounded degree, planar graph and the geometric properties of a nice embedding of the graph in the plane (e.g. a circle packing of the graph). Specifically, we show two results:

- (1) The graph is recurrent if and only if the boundary of the embedding is a polar set (that is, Brownian motion misses it almost surely).
- (2) If the graph is one ended and transient, and it is embedded nicely in the unit disk then the boundary of the disk, i.e. the unit circle is a representation of the Poisson and Martin boundaries, via the natural extension of the embedding.

Based of joint works with Omer Angel, Martin Barlow, Asaf Nachmias and Juan Souto.

**H**

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Charles Hagwood** (Statistical Engineering Div, NIST)

Contributed Talk, session [71]

Monday 10:45 – 11:15, Rosebrugh 310

*Some analytical results for shape space computations*

Coauthors: Javier Bernal (Applied and Computational Mathematics Division , NIST), Gunay Dogan (Applied and Computational Mathematics Division , NIST)

**Abstract:** The distance between the shapes of two 2D contours is defined as the geodesic distance between two mathematical entities defined as their shapes. These shapes reside in a manifold called shape space. Usually numerical methods are applied to compute these geodesics. In this talk, a few analytical results pertaining to computing these geodesics are discussed.

- **Martin Hairer** (University of Warwick)

Plenary talk: Plenary Lecture

Wednesday 9:25 – 10:15, McLeod Auditorium

*On the algebraic structure of renormalisation*

**Abstract:** TBA

- **Martin Hairer** (Warwick)

Invited session: Rough paths [13]

Wednesday 3:10 – 3:40, McLeod Auditorium

*Discretisation of regularity structures***Abstract:** TBA

- **Christian Hansen** (University of Chicago Booth School of Business)

Invited session: High-dimensional econometrics [31]

Thursday 11:20 – 11:50, McLeod Auditorium

*Double machine learning: improved point and interval estimation of treatment and causal parameters*

Coauthors: Victor Chernozhukov (MIT), Denis Chetverikov (UCLA)

**Abstract:** Most supervised machine learning (ML) methods are explicitly designed to solve prediction problems very well. Achieving this goal does not imply that these methods automatically deliver good estimators of causal parameters. Examples of such parameters include individual regression coefficients, average treatment effects, average lifts, and demand or supply elasticities. In fact, estimates of such causal parameters obtained via naively plugging ML estimators into estimating equations for such parameters can behave very poorly, for example, by formally having inferior rates of convergence with respect to the sample size  $n$  caused by regularization bias. Fortunately, this regularization bias can be removed by solving auxiliary prediction problems via ML tools. Specifically, we can form an efficient score for the target low-dimensional parameter by combining auxiliary and main ML predictions. The efficient score may then be used to build an efficient estimator of the target parameter which typically will converge at the fastest possible  $1/\sqrt{n}$  rate and be approximately unbiased and normal, and from which valid confidence intervals for these parameters of interest may be constructed. The resulting method thus could be called a “double ML” method because it relies on estimating primary and auxiliary predictive models. Such double ML estimators achieve the fastest rates of convergence and robustness of behavior with respect to a broader class of probability distributions than naive “single” ML estimators. We illustrate the use of the proposed methods with an application to estimating the effect of 401(k) eligibility on accumulated assets.

- **Adrien Hardy** (Université de Lille)

Invited session: Random matrices [6]

Friday 10:45 – 11:15, Medical Sciences 3153

*Large complex correlated Wishart matrices: Local behavior of the eigenvalues*

Coauthors: Walid Hachem (CNRS LTCI; Télécom ParisTech), Jamal Najim (CNRS LIGM; Université Paris-Est)

**Abstract:** The aim of this talk is to provide an overview on the local behavior of the eigenvalues of complex Wishart matrices at different regimes of interest. In a few words, the Tracy-Widom law appears at the soft edges of the limiting spectrum, and processes involving Bessel and Pearcey kernels arise at the hard edge and cusp points respectively. The fluctuations for the condition number are also considered, and open problems as well.

- **Giseon Heo** (University of Alberta)

Invited session: Object Oriented Data Analysis: Persistent Homology Representations [30]

Monday 11:55 – 12:25, McLeod Auditorium

*Ensembles of multiscale representations*

Coauthors: Matthew Pietrosanu (University of Alberta), Zhichun Zhai (University of Alberta)

**Abstract:** Persistent homology (PH) is a recently-popularized branch of computational topology fitting under the umbrella of general one-dimensional scale-space theory. Included in the realm of scale-space theory are computer vision techniques, SiZer for density estimation, and LASSO for variable selection, to name a few. In this talk, we compare the scale-space representation methods of different disciplines and illustrate their applications using several real-world data sets. This research is undertaken in collaboration with Matthew Pietrosanu, Mathieu Chalifour, Steven Luoma, Yi Zhou, Rui Hu, Zhichun Zhai, Bei Jiang, and Linglong Kong.

- **Alexander E. Holroyd** (Microsoft Research)

Invited session: Random growth models [3]

Tuesday 11:20 – 11:50, McLeod Auditorium

*Random games* TALK CANCELLED

**Abstract:** Alice and Bob compete in a game of skill, making moves alternately until one or other reaches a winning position, at which the game ends. Or, perhaps neither player can force a win, in which case optimal play continues forever, and we say that the game is drawn.

What is the outcome of a typical game? That is, what happens if the game itself is chosen randomly, but is known to both players, who play optimally?

I will provide some answers (and many questions) in several settings, including trees, directed and undirected lattices, and point processes. The competitive nature of game play frequently brings out some of the subtlest and most fundamental properties of probabilistic models. We will encounter continuous and discontinuous phase transitions, hard-core models, probabilistic cellular automata, bootstrap percolation, maximum matching, and stable marriage.

Based on joint works with Riddhipratim Basu, Maria Deijfen, Irene Marcovici, James Martin and Johan Wastlund.

- **Wen-Ping Hsieh** (National Tsing Hua University)

Invited session: Analysis of next generation sequencing data for biomedicine [35]

Tuesday 11:20 – 11:50, Medical Sciences 3153

*A study on intratumor heterogeneity with two-way mixture model*

Coauthor: An-Shun Tai (National Tsing Hua University)

**Abstract:** Tumors are known to be heterogeneous in their genetic composition. Not all of the cells in the same tissue sample are carrying the same set of DNA variation. Understanding the intratumor heterogeneity leads to many therapeutic advantages. In this talk, I will introduce the computational issues involved with short-read data. A two-way mixture model is proposed to infer the size of each subclone containing the somatic copy number variants and the copy number status of each variant. The TCGA data will be used to demonstrate how we infer the evolutionary sequence of the variants.

- **Mingshang Hu** (Shandong University)

Invited session: BSDEs and their applications in finance [19]

Friday 10:45 – 11:15, Medical Sciences 3154

*Stochastic maximum principle for optimization with recursive utilities*

**Abstract:** We obtain the variational equations for backward stochastic differential equations in recursive stochastic optimal control problems, and then get the maximum principle which is novel. The control domain need not be convex, and the generator of the backward stochastic differential equation can contain  $z$ .

- **Schalk Human** (Department of Statistics, University of Pretoria)

Contributed Talk, session [94]

Friday 11:20 – 11:50, Medical Sciences 2172

*Generally Weighted Moving Average control charts for small shifts*

Coauthors: N. Balakrishnan (Department of Mathematics and Statistics, McMaster University), N. Chakraborty (Department of Statistics, University of Pretoria)

**Abstract:** An efficient control chart should detect any change in a process as quickly as possible. However, it is known the traditional Shewhart-type charts are inefficient at detecting small changes quickly. An alternative and more effective option is to use a time-weighted chart, which sequentially accumulates all the information over time. A class of parametric and non-parametric Generally Weighted Moving Average (GWMA) control charts is proposed. The proposed GWMA chart includes the Exponentially Weighted Moving Average (EWMA) and Shewhart-type charts as special cases. We investigate both the scenarios when the parameters are known and unknown; in the latter scenario one would first have to estimate the parameter from an in-control Phase I reference sample. Three methods for calculating the run-length distribution and the associated characteristics of the run-length distribution are investigated; this includes (i) Exact closed-form expressions, (ii) A Markov chain approach, and (iii) Monte Carlo Simulation. To aid with the implementation of the chart the necessary design parameters are provided which guarantees that the in-control ARL is equal to a specified nominal value. The performance analysis shows that the GWMA chart is better than the well-known EWMA and Shewhart charts at detecting very small to moderate changes.

- **Tobias Hurth** (University of Toronto)

Contributed Talk, session [72]

Thursday 2:00 – 2:30, Rosebrugh 310

*Dynamical systems with random switching*

**Abstract:** In this talk, we will consider some aspects of the ergodic theory for dynamical systems with random switching, also known as piecewise deterministic Markov processes. These systems arise from switching between finitely many deterministic flows at random times. In particular, we will formulate conditions under which there is a unique invariant measure that is also absolutely continuous. For a particular system involving two linear vector fields on  $\mathbb{R}^2$ , we will describe how the switching rate determines whether the density of the invariant measure is bounded. The talk is based on work with Yuri Bakhtin, Sean Lawley and Jonathan Mattingly.

I

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Valerie Isham** (University College London)

Plenary talk: Bernoulli Lecture

Wednesday 8:30 – 9:20, McLeod Auditorium

*Applied stochastic modelling for structured physical processes*

**Abstract:** Stochastic modelling in which a mechanistic model is used to represent a physical process (albeit in highly idealised form) is an invaluable tool in a broad range of applications. The parameters of such a model are interpretable and relate directly to physical phenomena. As well as gaining insight into the process and understanding of its main drivers, the modeller can address important issues, answering ‘what if’ questions, developing appropriate control strategies and determining policy. In this talk I will discuss approaches to modelling and the need for stochasticity, together with the purposes of modelling and some of the questions that need to be asked and choices made when developing a model for a specific application. The discussion will be illustrated with examples from a variety of applications. A focus on population structure will form a common thread in discussing model development.

- **Dindane Issaka** (National Institute of Statistics and Demography)

Contributed Talk, session [99]

Friday 2:35 – 3:05, Rosebrugh 208

*Impact of agricultural input subsidies on food security of rural households in Burkina Faso : by the matching method*

**Abstract:** Throughout this study, it was question of measuring the real impact of subsidy programs on food security of rural households in Burkina Faso. To do this, we used the method of propensity score matching. This was to use statistical methods to generate the best group of artificial comparison of rural households do not benefit from the subsidy during the 2015-2016 crop year compared to those who received to assess the causal effect Live from these grants. Our results reveal that subsidies for agricultural inputs have a growing positive impact on the food safety standards of recipient households. Indeed, the subsidy resulted in the 2015-2016 crop year, higher food consumption score 12 points beneficiaries of households; an improvement in their incomes over 100,377 CFA. Moreover, it also appears from this study that the causal effect of the subsidy on cereal stocks by member of beneficiary households is positive and amounted to a surplus of almost 50 kg. For improved utilization of agricultural inputs guarantee of improved food security, the subsidy measure could be combined with a system of agricultural credits. The grants would improve food security at the national level for the positive effect on agricultural production. However, this requires that the implementation of the measure subsidized agricultural inputs are available in all localities of the country in accordance with the demand of each location (in terms of quantity and time). In addition, the grant would enable improved selected seeds utilization and chemical fertilizers. Agricultural loans which rural households benefit in the last twelve would also to some extent the improvement of input utilization. Thus, a combination of these two forms of support would improve access to chemical inputs.

J

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Pierre Jacob** (Harvard University)

Invited session: Statistical computing and complex data [36]

Monday 2:35 – 3:05, McLeod Auditorium

*Couplings of particle filters*

Coauthors: Fredrik Lindsten (Uppsala University), Thomas B. Schön (Uppsala University)

**Abstract:** Particle filters provide Monte Carlo approximations of intractable quantities, such as likelihood evaluations in state-space models. In many cases, the interest does not lie in the values of the estimates themselves, but in the comparison of these values for various parameters. For instance, we might want to compare the likelihood at two parameter values. Such a comparison is facilitated by introducing positive correlations between the estimators, which is a standard variance reduction technique. In the context of particle filters, this calls for new resampling schemes. We propose coupled resampling schemes, and show how they improve the performance of finite difference estimators and pseudo-marginal algorithms for parameter inference. Furthermore, coupled resampling schemes can be embedded into debiasing algorithms (Rhee & Glynn 2014), leading to a new smoothing algorithm which is easy to parallelize and comes with accurate estimates of the associated Monte Carlo error.

- **Jana Jankova** (ETH Zurich)

Contributed Talk, session [96]

Monday 11:20 – 11:50, Fitzgerald 139

*Semi-parametric efficiency bounds and efficient estimation for high-dimensional models*

Coauthor: Sara van de Geer (ETH Zurich)

**Abstract:** Asymptotic lower bounds for estimation play a fundamental role in assessing the quality of statistical procedures. We establish semi-parametric efficiency bounds for high-dimensional models and we consider construction of estimators reaching these bounds. In particular, we derive an explicit lower bound on the variance of any estimator that is in some sense asymptotically unbiased. We show that an estimator obtained by de-sparsifying (or de-biasing) an  $\ell_1$ -penalized M-estimator is asymptotically unbiased and achieves the lower bound on the variance: thus in this sense it is asymptotically efficient. In particular, we consider the linear regression model, Gaussian graphical models and Gaussian sequence models under mild conditions.

- **Carsten Jentsch** (University of Mannheim)

Contributed Talk, session [66]

Thursday 11:55 – 12:25, Medical Sciences 4171

*Empirical characteristic functions-based estimation and distance correlation for locally stationary processes*

Coauthors: Carina Beering (Technical University Braunschweig), Anne Leucht (Technical University Braunschweig), Marco Meyer (Technical University Braunschweig)

**Abstract:** In this paper, we propose a kernel-type estimator for the local characteristic function (local CF) of locally stationary processes. Under weak moment conditions, we prove joint asymptotic normality for local empirical characteristic functions (local ECF). Precisely, for processes having a (two-sided) time-varying  $MA(\infty)$  representation, we establish a central limit theorem under the assumption of finite absolute first moments of the process. Additionally, we prove process convergence of the local ECF. We apply our asymptotic results to parameter estimation of time-varying  $\alpha$ -stable distributions. Furthermore, by extending the notion of distance correlation (dCor) of Szekely, Rizzo and Bakirov (2007, *Annals of Statistics*) to locally stationary processes, we are able to provide asymptotic theory for local empirical distance correlations. Finally, we provide a simulation study on minimum distance estimation for  $\alpha$ -stable distributions and illustrate the pairwise dependence structure over time of log returns of German stock prices via local empirical distance correlations.

- **Jiashun Jin** (Carnegie Mellon University)

Invited session: A celebration of the work of Peter Hall [39]

Thursday 2:00 – 2:30, McLeod Auditorium

*Network membership estimation by simplex vertices hunting*

Coauthors: Tracy Zheng Ke (University of Chicago), Shengming Luo (University of Science and Technology of China)

**Abstract:** Consider an undirected mixed membership network with  $n$  nodes and  $K$  communities. For each node  $i$ ,  $1 \leq i \leq n$ , we model the membership by a Probability Mass Function (PMF)  $\pi_i = (\pi_i(1), \pi_i(2), \dots, \pi_i(K))'$ , where  $\pi_i(k)$  is the probability that node  $i$  belongs to community  $k$ ,  $1 \leq k \leq K$ . We call node  $i$  “pure” if  $\pi_i$  is degenerate and “mixed” otherwise. The primary interest is to estimate  $\pi_i$ ,  $1 \leq i \leq n$ .

We model the adjacency matrix  $A$  with a Degree Corrected Mixed Membership (DCMM) model. Let  $\hat{\xi}_1, \hat{\xi}_2, \dots, \hat{\xi}_K$  be the eigenvectors of  $A$  associated with the  $K$  largest (in magnitude) eigenvalues. We define a matrix  $\hat{R} \in \mathbb{R}^{n, K-1}$  by  $\hat{R}(i, k) = \hat{\xi}_{k+1}(i) / \hat{\xi}_1(i)$ ,  $1 \leq k \leq K-1$ ,  $1 \leq i \leq n$ . The matrix can be viewed as a distorted version of its non-stochastic counterpart  $R \in \mathbb{R}^{n, K-1}$ , which is unknown but contains all information we need for the memberships.

We reveal an interesting insight: there is a simplex  $\mathcal{S}$  in  $\mathbb{R}^{K-1}$  such that row  $i$  of  $R$  corresponds to a vertex of  $\mathcal{S}$  if node  $i$  is pure, and corresponds to an interior point of  $\mathcal{S}$  otherwise. Vertices Hunting (i.e., estimating the vertices of  $\mathcal{S}$ ) is therefore the key to our problem.

We propose a new approach *Mixed-SCORE* to estimating the memberships, at the heart of which is an easy-to-use Vertices Hunting algorithm. The approach is successfully applied to 4 network data sets (a coauthorship and a citee network for statisticians, a political book network, and a football network) with encouraging results. We analyze Mixed-SCORE and derive its rate of convergence, using delicate Random Matrix Theory.

- **Oliver Jovanovski** (Leiden University)

Contributed Talk, session [54]

Monday 11:55 – 12:25, Medical Sciences 4171

*Metastability for Glauber dynamics on random graphs*

Coauthors: Frank den Hollander (Leiden University), Francesca Nardi (Technical University Eindhoven), Sander Dommers (Ruhr-Universität Bochum)

**Abstract:** Starting with a finite, random graph constructed according to the “Configuration Model” algorithm, a spin is assigned to each vertex, carrying a +1 or -1 value. Each spin interacts with a positive magnetic field, and spins at neighbouring vertices interact with each other according to a ferromagnetic pair potential. We start with the configuration where every vertex carries a -1 spin, and allow spins to flip according to Glauber dynamics. We give a description of the time it takes the system to reach the configuration with all spins being +1, in the regime when the temperature approaches 0.

**K**

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Hye-Won Kang** (University of Maryland, Baltimore County)

Contributed session: Chemical reaction networks [41]

Wednesday 2:00 – 2:30, Medical Sciences 2172

*Reduction for stochastic reaction networks with multi-scale conservation*

Coauthors: Jae Kyoung Kim (Department of Mathematical Sciences, Korea Advanced Institute of Science and Technology), Grzegorz A. Rempala (Division of Biostatistics and Mathematical Biosciences Institute, The Ohio State University)

**Abstract:** Chemical reaction networks frequently have a multi-scale property: reaction rates vary and several time scales exist in which chemical species evolve. Many works have been done on constructing multi-scale methods for stochastic chemical reaction networks to approximate temporal changes in the abundance of chemical species and to reduce network complexity. Among different approaches, stochastic models for chemical reaction networks using Markov processes and the multi-scale approximation method developed by Ball et al. (2006) and extended by Kang and Kurtz (2013) will be introduced. Then, a modified multiscale approximation method for stochastic chemical reaction networks with conservation will be investigated, when the chemical reaction network conserves linear combinations of some chemical species with quantities in different scales. The method will be applied to an example in enzyme kinetics and one generating oscillations. This is joint work with J. Kim and G.A. Rempala.

- **Bikram Karmakar** (Department of Statistics, University of Pennsylvania, PA, USA)

Contributed Talk, session [97]

Friday 2:35 – 3:05, Medical Sciences 2172

*A general statistical framework for analysis of big data*

Coauthor: Indranil Mukhopadhyay (Human Genetics Unit, Indian Statistical Institute, Kolkata, India)

**Abstract:** Information explosion creates big data, which if properly analyzed, can benefit various aspects of our lives. But inherent complexities in big data analysis might mislead inferences through standard statistical methods. We need to be particularly careful and cautious in the analysis of big

data in spite of technological advancement in computational and data management system. Towards achieving this goal, we have developed a general statistical framework with sound mathematical and statistical logic that provides guidelines to appropriate analysis. Our proposed method is based on very weak assumptions respecting the practical characteristics of data flow. This method has tremendous potential towards amenability with recent advancements of computational system that handles and manages the big data keeping in mind its complexities that separates it from standard data and the corresponding analysis. Our work will facilitate the advancements in biology, computer science, customer service, health care, statistics, social network, and any other field that can generate huge volume of data with high velocity and variety. We have shown implementation of these guidelines through some illustrations.

- **Mikolaj Kasprzak** (University of Oxford)

Contributed Talk, session [70]

Friday 2:00 – 2:30, Medical Sciences 4279

*Stein's method for functional approximation*

**Abstract:** In [1] A. Barbour used Stein's method to prove the following result. Let  $X_i$ 's be iid. with  $\mathbb{E}X_i = 0$ ,  $\mathbb{E}X_i^2 = 1$ ,  $\mathbb{E}|X_i|^3 < \infty$ ,  $W$  denote a standard Wiener process on  $[0, 1]$  and  $Y_n(\cdot) = n^{-1/2} \sum_{i=1}^n X_i \mathbb{I}_{[i/n, 1]}(\cdot) \in D[0, 1]$ . Let  $g : D[0, 1] \rightarrow \mathbb{R}$  be twice Fréchet differentiable and satisfy  $|g(w)| \leq K(1 + \|w\|_\infty^3)$  for a uniform constant  $K$ . Let its second derivative be Lipschitz with respect to the supremum norm. Then:

$$|\mathbb{E}g(Y_n) - \mathbb{E}g(W)| \leq C \|g\|_M n^{-1/2} \left( \mathbb{E}|X_1|^3 + \sqrt{\log n} \right),$$

where:

$$\|g\|_M = \sup_{w \in D} \frac{|g(w)|}{1 + \|w\|_\infty^3} + \sup_{w \in D} \frac{\|Dg(w)\|}{1 + \|w\|_\infty^2} + \sup_{w \in D} \frac{\|D^2g(w)\|}{1 + \|w\|_\infty} + \sup_{w, h \in D} \frac{\|D^2g(w+h) - D^2g(w)\|}{\|h\|_\infty}.$$

This bound was better than the one appearing in the famous approximation of Komlós, Major and Tusnády from 1975 and 1976.

We extend Barbour's result to an analogous one concerning a  $p$ -dimensional random walk with independent components approximated by a  $p$ -dimensional standard Brownian Motion for  $p > 1$  and one concerning a two-dimensional random walk with weakly dependent components:

$$Y_n(\cdot) = n^{-1/2} \sum_{i=1}^n \left( X_i \mathbb{I}_{[i/n, 1]}(\cdot), \sum_{i=n-l+1}^{2n-l} X_i \mathbb{I}_{[(i-n+l)/n, 1]}(\cdot) \right) \in D[0, 1]$$

approximated by a standard planar Brownian Motion on  $[0, 1]$ . We also derive analogous bounds on the approximation of a correlated random walk:

$$Y_n(\cdot) = n^{-1/2} \sum_{i=1}^n \left( X_i^{(1)}, \rho(\cdot) X_i^{(1)} + \sqrt{1 - \rho^2(\cdot)} X_i^{(2)} \right) \mathbb{I}_{[i/n, 1]}(\cdot) \in D[0, 1], \quad \rho : [0, 1] \rightarrow [-1, 1],$$

where  $X_i^{(1)}$ 's and  $X_i^{(2)}$ 's are all iid. with mean 0, variance 1 and a finite third moment, by a correlated Brownian Motion on  $[0, 1]$ . Similarly, Stein's method lets us obtain a bound in the approximation of

a sequence of scaled non-degenerate U-statistics by a Wiener process, as in the well-known invariance principle for U-statistics.

We provide explicit values of all the constants appearing in the above mentioned bounds. This work is part of my DPhil project supervised by Professor Gesine Reinert.

[1] A.D. Barbour. Stein’s method for diffusion approximation. *Probability Theory and Related Fields*, 84: 297-322, 1990.

- **Adrien Kassel** (ETH Zürich)

Invited session: Random growth models [3]

Tuesday 11:55 – 12:25, McLeod Auditorium

*Active spanning trees and Schramm–Loewner evolution*

**Abstract:** Active spanning trees with bending energies form a two-parameter family of growing “space-filling” discrete curves in 2D. For a particular choice of parameters, they specialize to exploration curves of Fortuin–Kasteleyn clusters which were conjectured (and have been proven in two integrable cases: the uniform spanning tree and the FK-Ising model) to converge to the Schramm–Loewner evolution (SLE) in the scaling limit. We explain why the conjecture can reasonably be extended to all active spanning trees thus providing a conjectured discrete analog of  $SLE_\kappa$  for any value of its parameter  $\kappa \geq 0$ . In the case when the underlying planar graph is itself random, this conjecture is shown to be true building on recent breakthroughs in the field of random surfaces. Joint work with David B. Wilson (Microsoft Research), and with Ewain Gwynne (MIT), Jason P. Miller (Cambridge), and David B. Wilson.

- **Adrien Kassel** (ETH Zürich)

Invited session: Integrable models in Statistical Physics [2]

Thursday 11:20 – 11:50, Medical Sciences 3153

*Integrable models related to the Laplacian*

**Abstract:** Random walks, spanning trees, loop ensembles, and the discrete Gaussian free field are basic integrable models on graphs (all related to the combinatorial Laplacian) which have served as a rich source of inspiration for the understanding of statistical physics in 2D and the underlying random conformal geometry. We explain how these models and some relations between them can be studied in the more general setup of vector bundles over graphs and simplicial complexes and discuss their relevance for the understanding of higher statistical field theories and random geometry.

- **Gursharn Kaur** (Indian Statistical Institute (Delhi))

Contributed Talk, session [63]

Monday 2:35 – 3:05, Medical Sciences 3278

*Negatively reinforced urn schemes*

Coauthor: Antar Bandyopadhyay (Indian Statistical Institute (Delhi))

**Abstract:** In this talk we will consider general negatively reinforced urn schemes with finitely many colours. We will call an urn scheme negatively reinforced, if the selection probability for a color is proportional to the weight  $w$  of the color, where  $w$  is a decreasing function. Under some assumptions on  $w$ , such as,  $w(0) < \infty$  and  $w$  is strictly decreasing differentiable function, we obtain almost sure

convergence of the random configuration of the urn for a general replacement matrix  $R$ . We show that depending on the function  $w$  and the replacement matrix  $R$ , the limit may be a constant or a random variable. For certain weight functions we show that Gaussian scaling limit of the fluctuations can also be obtained after subtracting the limiting constant.

- **Wilfrid Stephen Kendall** (University of Warwick)

Contributed Talk, session [59]

Wednesday 10:45 – 11:15, Medical Sciences 3278

*Random walks in scale-invariant random spatial networks (SIRSN)*

Coauthor: Sayan Banerjee (University of Warwick)

**Abstract:** A SIRSN obeys axioms proposed by Aldous [1] and provides random (almost surely) unique routes between specified locations in a statistically scale-invariant manner. A planar construction based on a randomized dyadic rectilinear network is established in [1]: a further construction based on Poisson line processes has now been established in [2,3] and even delivers SIRSN in dimensions 3 and higher. I will describe recent planar results concerning random walks (actually Rayleigh random flights) in such SIRSN, aimed at providing better insight into the behaviour of SIRSN routes.

- (1) Aldous, D.J. (2014). Scale-Invariant Random Spatial Networks. *Electronic J. Prob.*, 21, no. 19, 1-41.
- (2) WSK (2016). From Random Lines to Metric Spaces. *Ann. Prob.* (to appear).
- (3) Kahn, J. (2016). Improper Poisson line process as SIRSN in any dimension. *Ann. Prob.* (to appear).

- **Samer A. Kharroubi** (American University of Beirut)

Contributed Talk, session [79]

Thursday 2:00 – 2:30, Fitzgerald 139

*Valuations of EQ-5D health states: could United Kingdom results be used as informative priors for United States*

**Abstract:** Background: There is interest in the extent of using the results in one country to inform the design in another country by using the results in one country as informative priors. This kind of analysis (borrowing strength from country 1) may produce better estimation of the second country's population utility function than analysing its data separately. Methods: The data set is the US and UK EQ-5D valuation studies where a sample of 42 health states defined by the EQ-5D was valued by representative samples of the general population from each country using the time trade-off technique. We apply a nonparametric Bayesian method to improve the accuracy of predictions of the US population utility function using the results in the UK as informative priors. Results: The results suggest that drawing extra information from the UK data produces better estimation of the US population utility than analysing its data separately.

Conclusion: The implications of these results will be hugely important in countries without the same capacity to run large evaluation exercises.

- **Christian Kleiber** (Universitaet Basel)  
 Contributed Talk, session [48]  
 Tuesday 10:45 – 11:15, Medical Sciences 2171  
*On moment indeterminacy in stochastic volatility models*  
**Abstract:** The moment problem asks whether a distribution can be uniquely characterized by the sequence of its moments. Counterexamples have long been known, e.g., the lognormal and certain generalized gamma distributions. Here we show that one of the standard volatility models in mathematical finance and financial econometrics, namely the stochastic volatility (SV) model, leads to return distributions that are moment indeterminate. Perhaps somewhat unexpectedly, moment indeterminacy already arises in the classical discrete time SV model with lognormal latent volatility and independent multiplicative Gaussian noise.
- **Antti Knowles** (ETH Zurich)  
 Invited session: Random matrices [6]  
 Friday 11:20 – 11:50, Medical Sciences 3153  
*Local spectral statistics for random regular graphs*  
**Abstract:** We consider the adjacency matrix of a random regular graph, and prove, under mild growth assumptions on the degree, a local law for the Green function. Applications include eigenvector delocalization and a proof that the bulk eigenvalue statistics coincide with those of the GOE from random matrix theory. Joint work with R. Bauerschmidt, J. Huang, and H.-T. Yau.
- **Vladimir Koltchinskii** (School of Mathematics, Georgia Tech)  
 Invited session: Random matrix theory and statistics [21]  
 Thursday 2:35 – 3:05, Medical Sciences 2171  
*Concentration and normal approximation bounds for sample covariance and its spectral projections*  
 Coauthor: Karim Lounici (School of Mathematics, Georgia Tech)  
**Abstract:** New concentration and normal approximation bounds for sample covariance operators based on i.i.d. Gaussian observations and their spectral projections will be discussed. The results are proved in a dimension-free framework with so called effective rank of the true covariance operator playing the role of complexity parameter. This quantity along with the operator norm of the true covariance provide a complete characterization of the size of the operator norm deviations of empirical covariance from its true counterpart in the Gaussian case. The effective rank plays also a crucial role in the problems related to concentration and normal approximation of the spectral projections of sample covariance.
- **Gnameho Kossi** (Maastricht University)  
 Contributed Talk, session [45]  
 Monday 2:35 – 3:05, Medical Sciences 4171  
*Fourier-Hermite expansions for backward stochastic differential equations*  
 Coauthor: Prof. Antoon A. J. Pelsser (Maastricht University)  
**Abstract:** This paper deals with the numerical approximation of the class of Markovian backward stochastic differential equations (BSDEs) where the terminal condition is a functional of Brownian

motion. In this class, we show that the problem of solving a Markovian BSDE is identical to solving a countable infinite dimensional system of ordinary differential equations (ODEs). The family of ODEs belongs to the class of stiff ODEs where the associated functional is one-sided Lipschitz. On this basis, we derive a numerical algorithm for the BSDE via the standard Euler scheme with respect to the solution of the countable system of ordinary differential equations (CODEs). Finally, we provide some numerical experiments to test the performance of the algorithm.

- **Barbara Kowalczyk** (Warsaw School of Economics)

Contributed Talk, session [80]

Friday 10:45 – 11:15, Rosebrugh 208

*Item count technique based on two different treatment groups*

Coauthor: Robert Wieczorkowski (Central Statistical Office of Poland)

**Abstract:** Item Count Technique (ICT) is a method of indirect questioning that allows to protect respondents' privacy and thus helps eliciting truthful answers when dealing with vastly personal features or legally improper behaviors. In the talk we introduce a new type of ICT that uses continuous auxiliary variable and eliminates traditional division into control and treatment groups. For the new method both MM and ML (using EM algorithm) estimators are analyzed. Theoretical and practical advantages of the new approach are discussed. Presented theoretical results are supported by extensive simulation studies.

- **Kun-Lin Kuo** (Institute of Statistics, National University of Kaohsiung, Taiwan)

Contributed Talk, session [75]

Wednesday 11:55 – 12:25, Fitzgerald 139

*Pseudo-Gibbs distributions for incompatible conditional models*

**Abstract:** An incompatible model first estimates full conditional distributions from data, then forge a joint distribution from the conditionals. The conditional approach allows a higher level of flexibility because it is computationally convenient to estimate the local distribution of one variable given the remaining variables. Similar approaches have appeared in multiple imputation since 1999. However, individually determined conditional distributions are generally not coherent with any joint distribution. The pseudo-Gibbs sampling (PGS) method is often used to generate samples and use the empirical distribution to approximate joint distributions, though its properties are mostly unknown. This talk investigates the richness and varieties of the joint distributions that PGS approximates. In short, PGS produces a large number of joint distributions with shared dependence and shared marginal distributions. As an algorithm, PGS has multiple fixed points that are as similar as possible under the constraints of incompatibility; the optimal similarity is proved in terms of Kullback-Leibler information divergence. Our characterizations help the practitioners of multiple imputation to address some of their concerns. Moreover, properties of incompatible PGS provide a fresh perspective for understanding the original Gibbs sampler used in Bayesian computations.

- **Thomas G. Kurtz** (University of Wisconsin – Madison)

Contributed session: Chemical reaction networks [41]

Wednesday 2:35 – 3:05, Medical Sciences 2172

*Maximum likelihood estimation for partially observed Markov chain models.*

Coauthor: Grzegorz A. Rempala (Ohio State University)

**Abstract:** In continuous time Markov chain chemical network models, the natural parameters enter in a convenient multiplicative fashion. For a completely observed model, computing the maximum likelihood estimators for the parameters is simple and consistency and asymptotic normality is straight forward, but complete observation is an unreasonable assumption. The talk will focus on a consistency result for the partially observed maximum likelihood estimator and an (almost) central limit theorem. This work is joint with Greg Rempala.

- **Alexey Kuznetsov** (York University)

Contributed Talk, session [63]

Monday 2:00 – 2:30, Medical Sciences 3278

*Reinforced Polya urns*

**Abstract:** Consider the following toy model for neural processing in the brain: a large number of neurons are interconnected by synapses, and the brain removes connections which are seldom or never used and reinforces those which are stimulated. We introduce a class of reinforced Polya urn models which aim to describe this dynamics. Our models work as follows: at each time step  $t$ , we first choose a random subset  $A_t$  of colours (independently of the past) from  $n$  colours of balls, and then choose a colour  $i$  from the subset  $A_t$  with probability proportional to the number of balls of colour  $i$  in the urn raised to the power  $\alpha > 1$ . We are mostly interested in stability of equilibria for such models studying phase transitions in a number of examples, including when the colours are the edges of a graph. We conjecture that for any graph  $G$  and all  $\alpha$  sufficiently large, the set of stable equilibria is supported on so-called whisker-forests, which are forests whose components have diameter between 1 and 3. This talk is based on joint work with Remco van der Hofstad, Mark Holmes and Wioletta Ruszel.

**L**

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **John Lafferty** (Department of Statistics & Department of Computer Science, University of Chicago)

Invited session: Computational versus statistical complexity in high-dimensional problems [24]

Wednesday 11:20 – 11:50, McLeod Auditorium

*Statistical estimation under communication constraints*

**Abstract:** Imagine that I estimate a statistical model from data, and then want to share my model with you. But we are communicating over a resource constrained channel. By sending lots of bits, I can communicate my model accurately, with little loss in statistical risk. Sending a small number of bits will incur some excess risk. What can we say about the tradeoff between statistical risk and the communication constraints? This is a type of rate distortion and constrained minimax problem, for which we provide a sharp analysis in certain nonparametric settings. Joint work with Yuancheng Zhu.

- **Steffen Lauritzen** (University of Copenhagen)  
 Invited session: Graphical models and totally positive distributions [22]  
 Monday 11:20 – 11:50, Medical Sciences 2172  
*Total positivity and Markov structures*  
**Abstract:** Positive associations between random variables can be described in a variety of ways. One of strongest is multivariate total positivity of order two (MTP2) introduced by Karlin and Rinott (1980). The MTP2 property is stable under a number of operations, including marginalization and conditioning. The lecture investigates how the property interacts with conditional independence structures, implying, for example, that any strictly positive MTP2 distribution becomes faithful to its pairwise independence graph. In addition we shall study how this property is manifested in Gaussian, discrete, and conditional Gaussian distributions, and give a number of examples of such distributions. The lecture is based on joint work with S. Fallat, K. Sadeghi, C. Uhler, N. Wermuth, and Piotr Zwiernik; arXiv:1510.01290.
- **Steffen Lauritzen** (University of Copenhagen)  
 Invited session: Statistics and networks [27]  
 Wednesday 2:00 – 2:30, Medical Sciences 4171  
*Random networks, graphical models, and exchangeability*  
**Abstract:** We investigate connections between exponential random network models (ERGM), bidirected and undirected graphical models, and issues of invariance and exchangeability. Each point of view gives rise to its own class of parametrization of network models and we investigate how they are related to each other and to the theory of graph limits and graphons. In particular we provide insight into possibilities for formulating specific submodels which partially obey properties of the type mentioned. The lecture is based upon joint work with A. Rinaldo and K. Sadeghi.
- **Gregory F. Lawler** (University of Chicago)  
 Invited session: Conformally invariant processes [8]  
 Friday 2:00 – 2:30, Medical Sciences 3153  
*Convergence of loop-erased random walk to the Schramm-Loewner evolution in natural parametrization*  
 Coauthor: Fredrik Johansson Viklund (KTH Royal Institute of Technology)  
**Abstract:** We prove that loop-erased random walk parametrized by renormalized length converges in the lattice size scaling limit to  $SLE_2$  parametrized by the natural parametrization which is the same as the 5/4-dimensional Minkowski content.
- **Guillaume Lecué** (CNRS & ENSAE)  
 Invited session: Aggregation of estimators [25]  
 Monday 2:35 – 3:05, Medical Sciences 2171  
*Some geometric and stochastic aspects of the aggregation problem*  
 Coauthors: Mendelson (Technion), Rigollet (MIT)  
**Abstract:** In this talk, the three aggregation problems are considered:  
 (1) Model Selection aggregation, where one wants to mimic the best procedure in a given dictionary;

- (2) Convex aggregation, where one wants to mimic the best procedure in the convex hull of a given dictionary
- (3) linear aggregation, where one wants to mimic the best linear combination of elements in a dictionary.

The study of the Empirical Risk Minimization (ERM) procedure in those three problems will be central in this talk. In particular, by understanding the geometrical reasons why the ERM is suboptimal for the Model Selection aggregation problem, I will design optimal aggregation procedures for this problem. Then, I will show why ERM is optimal for the convex aggregation problem. Finally, I will show that ERM is not optimal in the linear aggregation problem unless a small ball assumption is satisfied.

This talk is based on joint works with Shahar Mendelson, Philippe Rigollet and Stéphane Gaïffas.

- **Kevin Leder** (University of Minnesota)

Invited session: Stochastic models of cancer [17]

Wednesday 11:20 – 11:50, Medical Sciences 3154

*A spatial model of cancer initiation*

**Abstract:** I will discuss a stochastic model of mutation accumulation and spread in a spatially-structured population. This situation arises in a variety of ecological and biological problems, including the process of cancer initiation from healthy tissue. Cancer arises via the accumulation of mutations to the genetic code. Although many mathematical models of cancer initiation have assumed ‘perfect mixing’ or spatial homogeneity, solid tumors often initiate from tissues with well-regulated spatial architecture and dynamics. Here, we make a thorough study of this stochastic model to investigate the impact of a spatial tissue structure on the temporal dynamics and patterns of mutation accumulation. As an application we look at the phenomena of the cancer field effect in head and neck squamous cell carcinoma (HNSCC) and its impact on treatment response. This is joint work with R. Durrett (Duke), J. Foo (Minnesota) and M. Ryser (Duke).

- **Marc Lelarge** (INRIA-ENS)

Invited session: Community detection in random graphs [26]

Thursday 11:20 – 11:50, Medical Sciences 2172

*Community detection with the non-backtracking operator*

Coauthors: Charles Bordenave (CNRS), Laurent Massoulié (MSR-INRIA)

**Abstract:** Community detection consists in identification of groups of similar items within a population. In the context of online social networks, it is a useful primitive for recommending either contacts or news items to users. We will consider a particular generative probabilistic model for the observations, namely the so-called stochastic block model and prove that the non-backtracking operator provides a significant improvement when used for spectral clustering.

- **Oleg Lepski** (Aix-Marseille Université)

Invited session: Aggregation of estimators [25]

Monday 3:10 – 3:40, Medical Sciences 2171

*Some new ideas in nonparametric estimation*

**Abstract:** In the framework of an abstract statistical model we discuss how to use the solution of one estimation problem (*Problem A*) in order to construct an estimator in another, completely different, *Problem B*. As a solution of *Problem A* we understand a data-driven selection from a given family of estimators  $\mathbf{A}(\mathfrak{H}) = \{\hat{A}_{\mathfrak{h}}, \mathfrak{h} \in \mathfrak{H}\}$  and establishing for the selected estimator so-called oracle inequality. If  $\hat{\mathfrak{h}} \in \mathfrak{H}$  is the selected parameter and  $\mathbf{B}(\mathfrak{H}) = \{\hat{B}_{\mathfrak{h}}, \mathfrak{h} \in \mathfrak{H}\}$  is an estimator's collection built in *Problem B* we suggest to use the estimator  $\hat{B}_{\hat{\mathfrak{h}}}$ . We present very general selection rule led to selector  $\hat{\mathfrak{h}}$  and find conditions under which the estimator  $\hat{B}_{\hat{\mathfrak{h}}}$  is reasonable. Our approach is illustrated by several examples related to adaptive estimation.

- **Boris Levit** (Queen's University)

Contributed Talk, session [92]

Wednesday 11:20 – 11:50, Medical Sciences 2172

*Optimal interpolation in Approximation Theory, Nonparametric Regression, and Optimal Design: a non-asymptotic approach*

**Abstract:** For some rectangular Hardy classes of analytic functions, an optimal method of interpolation has been recently found, within the framework of Optimal Recovery. It will be shown that this method of interpolation, based on the Abel-Jacobi elliptic functions, is also optimal, according to corresponding criteria of Nonparametric Regression and Optimal Design.

In a non-asymptotic setting, the maximal mean squared error of the optimal interpolant is evaluated explicitly, for all noise levels away from 0. In these results, a pivotal role is played by an interference effect, in which both the stochastic and deterministic parts of the interpolant exhibit an oscillating behavior, with the two oscillating processes mutually subduing each other.

- **Yushu Li** (Department of mathematics, University of Bergen, Norway)

Contributed Talk, session [84]

Monday 2:00 – 2:30, Rosebrugh 310

*A likelihood ratio and Markov chain based method to evaluate the density forecast*

**Abstract:** In this paper, we propose a likelihood ratio and Markov chain based method to evaluate density forecasting. This method can jointly evaluate the unconditional forecasted distribution and dependence of the outcomes. This method is an extension of the widely applied evaluation method for interval forecasting proposed by Christoffersen (1998). It is also a more refined approach than the pure contingency table based density forecasting method in Wallis (2003). Our test framework are nonparametric and purely data driven. We compare our method with goodness of fit Kolmogorov-Smirnov (KS) test and the density evaluation process proposed by Berkowitz (2001). We show that our methodology has very high power against both dependence and incorrect forecasting distributions. Moreover, the straightforwardness and ease of application of this joint test provide a high potentiality for further applications in both financial and economical areas.

- **Zeng Li** (The University of Hong Kong)

Contributed Talk, session [97]

Friday 2:00 – 2:30, Medical Sciences 2172

*Identifying the number of factors from singular values of a large sample auto-covariance matrix*

Coauthors: Qinwen Wang (University of Pennsylvania), Jianfeng Yao (The University of Hong Kong)

**Abstract:** Identifying the number of factors in a high-dimensional factor model has attracted much attention in recent years and a general solution to the problem is still lacking. A promising ratio estimator based on the singular values of the lagged auto-covariance matrix has been recently proposed in the literature and is shown to have a good performance under some specific assumption on the strength of the factors. Inspired by this ratio estimator and as a first main contribution, this paper proposes a complete theory of such sample singular values for both the factor part and the noise part under the large-dimensional scheme where the dimension and the sample size proportionally grows to infinity. In particular, we provide the exact description of the phase transition phenomenon that determines whether a factor is strong enough to be detected with the observed sample singular values. Based on these findings and as a second main contribution of the paper, we propose a new estimator of the number of factors which is strongly consistent for the detection of all significant factors (which are the only theoretically detectable ones). In particular, factors are assumed to have the minimum strength above the phase transition boundary which is of the order of a constant; they are thus not required to grow to infinity together with the dimension (as assumed in most of the existing papers on high-dimensional factor models). Empirical Monte-Carlo study as well as the analysis of stock returns data attest a very good performance of the proposed estimator. In all the tested cases, the new estimator largely outperforms the existing estimator using the same ratios of singular values.

- **Zhaoyuan Li** (The University of Hong Kong)

Contributed Talk, session [92]

Wednesday 11:55 – 12:25, Medical Sciences 2172

*Testing for heteroscedasticity in high-dimensional regressions*

Coauthor: Jianfeng Yao (The University of Hong Kong)

**Abstract:** Testing heteroscedasticity of the errors is a major challenge in high-dimensional regressions where the number of covariates is large compared to the sample size. Traditional procedures such as the White and the Breusch-Pagan tests typically suffer from low sizes and powers. This paper proposes two new test procedures based on standard OLS residuals. Using the theory of random Haar orthogonal matrices, the asymptotic normality of both test statistics is obtained under the null when the degree of freedom tends to infinity. This encompasses both the classical low-dimensional setting where the number of variables is fixed while the sample size tends to infinity, and the proportional high-dimensional setting where these dimensions grow to infinity proportionally. These procedures thus offer a wide coverage of dimensions in applications. To our best knowledge, this is the first procedures in the literature for testing heteroscedasticity which are valid for medium and high-dimensional regressions. The superiority of our proposed tests over the existing methods are demonstrated by extensive simulations and by several real data analyses as well.

- **Vlada Limic** (CNRS, Université Paris Sud)

Invited session: Interacting particle systems and their scaling limits [10]

Thursday 3:10 – 3:40, Medical Sciences 3153

*Eternal multiplicative coalescents: old look, new look*

Coauthor: David Aldous (University of California, Berkeley)

**Abstract:** An old theorem by D. Aldous and the speaker says that the marginal law of each extremal eternal multiplicative coalescent (MC) at time  $t$  equals that of the infinite vector of ordered excursion (above past minima) lengths of a certain Lévy type process  $(W^{\kappa, \tau, \vec{c}}(s) + ts, s \geq 0)$ , where  $\kappa \geq 0$  is the “scale” parameter,  $\tau$  is the “time-centering”, and  $\vec{c} \in l^3$  is a non-negative sequence of “color weights”.

A new theorem says that the two-parameter family  $(W^{\kappa, \tau, \vec{c}}(s) + ts, s \geq 0, t \in (-\infty, \infty))$  completely encodes the above mentioned extremal eternal MC  $\Pi^{\kappa, \tau, \vec{c}}$ : as  $t$  increases from  $-\infty$  to  $\infty$ , the infinite vector of ordered excursion (above past minima) lengths of  $(W^{\kappa, \tau, \vec{c}}(s) + ts, s \geq 0)$  is a realization of  $\Pi^{\kappa, \tau, \vec{c}}$ . In the Aldous standard MC setting (where  $\kappa = 1, \tau = 0, \vec{c} = (0, 0, \dots)$ ) this result was announced by I. Armendáriz at the beginning of the century, and proved recently by N. Broutin and J.-F. Marckert. The talk will explain a unified approach that applies equally well to all the entrance laws. To the best of speaker’s knowledge, the underlying continuous time/continuous mass Erdős-Rényi (CTCMER) random graph construction, based on a coupling with “simultaneous” breadth-first walks, is novel. The talk will also address its connection to the Armendáriz’ construction of the CRCMER random graph, and should end by recalling some interesting open questions.

- **Liang-Ching Lin** (Department of Statistics, National Cheng Kung University, Tainan, Taiwan.)

Contributed Talk, session [82]

Thursday 11:55 – 12:25, Rosebrugh 208

*Robust principal expectile component analysis*

**Abstract:** Principal component analysis (PCA) is a widely used dimension reduction technique especially for the high dimension data analysis. These principal components are identified by sequentially maximizing the component score variance for observations centered on the sample mean. However, in practice, one might be more interested in the variation captured by the tail characters instead of the sample mean, for example the analysis of expected shortfall. To properly capture the tail characters, principle expectile component (PEC) analysis was proposed based on an asymmetric  $L_2$  norm (Tran, Osipenko, and Hardle, 2014). Although, in order to achieve robustness against outliers, we generalize the PEC by integrating with Huber’s norm. The newly proposed method is named as principal Huber-type expectile components (PHEC). A derivative free optimization approach, particle swarm optimization (PSO), is adopted to efficiently identify the components in PHEC. Simulation studies show that the PHEC outperforms PCA and PEC in capturing the tail variation in the case of normal mixture distributions. Finally a real example is analyzed for illustration.

- **David Lipshutz** (Brown University)

Contributed session: Analysis and control of stochastic network models [42]

Wednesday 10:45 – 11:15, Medical Sciences 4279

*Differentiability of flows and sensitivity analysis of reflected diffusions*

Coauthor: Kavita Ramanan (Brown University)

**Abstract:** Differentiability of flows and sensitivity analysis are classical topics in dynamical systems. However, the analysis of these properties for constrained processes, which arise in a variety of applications, is challenging due to the discontinuous dynamics at the boundary of the domain, and is further complicated when the boundary is non-smooth. We show that the study of both differentiability of flows and sensitivities of constrained processes in convex polyhedral domains can be largely reduced to the study of directional derivatives of an associated map, called the extended Skorokhod map, and we introduce an axiomatic framework to characterize these directional derivatives. In addition, we establish pathwise differentiability of a large class of reflected diffusions in convex polyhedral domains and show that they can be described in terms of certain constrained stochastic differential equations with time-varying domains and directions of reflection.

- **Alexander Litvak** (University of Alberta)

Invited session: Random matrix theory and statistics [21]

Thursday 3:10 – 3:40, Medical Sciences 2171

*Approximating the covariance matrix with heavy tailed columns and RIP.*

**Abstract:** Let  $A$  be a matrix whose columns  $X_1, \dots, X_N$  are independent random vectors in  $\mathbb{R}^n$ . Assume that  $p$ -th moments of  $\langle X_i, a \rangle$ ,  $a \in S^{n-1}$ ,  $i \leq N$ , are uniformly bounded. For  $p > 4$  we prove that with high probability  $A$  has the Restricted Isometry Property (RIP) provided that Euclidean norms  $|X_i|$  are concentrated around  $\sqrt{n}$  and that the covariance matrix is well approximated by the empirical covariance matrix provided that  $\max_i |X_i| \leq C(nN)^{1/4}$ . We also provide estimates for RIP when  $\mathbb{E} \phi(|\langle X_i, a \rangle|) \leq 1$  for  $\phi(t) = (1/2) \exp(t^\alpha)$ , with  $\alpha \in (0, 2]$ .

Joint work with O. Guédon, A. Pajor, N. Tomczak-Jaegermann

- **Chak Hei Lo** (Durham University)

Contributed Talk, session [63]

Monday 3:10 – 3:40, Medical Sciences 3278

*Non-homogeneous random walks on a half strip*

Coauthor: Andrew R. Wade (Durham University)

**Abstract:** We study a Markov chain on  $\mathbb{R}_+ \times S$ , where  $S$  is a finite set, in which when the  $\mathbb{R}_+$ -coordinate is large, the  $S$ -coordinate of the process is approximately Markov with stationary distribution  $\pi_i$ . If  $\mu_i(x)$  is the mean drift of the  $\mathbb{R}_+$ -coordinate of the process at  $(x, i) \in \mathbb{R}_+ \times S$ , we study the case where  $\sum_i \pi_i \mu_i(x) \rightarrow 0$ , which is the critical regime for the recurrence-transience phase transition. In the Lamperti case where  $\mu_i(x) = O(1/x)$ , the recurrence classification is known, but we proved new results on existence and non-existence of moments of return times. Also, it is natural to study the generalized Lamperti case where  $\mu_i(x) = d_i + O(1/x)$ , in which we obtain a recurrence classification and existence of moments results. The generalized Lamperti case is seen to be more subtle, as the recurrence classification depends on correlation terms between the two coordinates of the process.

- **Kyle Logan** (Governor's School for Science and Technology)

Contributed Talk, session [99]

Friday 3:10 – 3:40, Rosebrugh 208

*TSER value determination for NFL African American quarterbacks*

**Abstract:** In American football, an efficient quarterback is key in scoring and winning a game. This study sought to test whether a higher quarterback efficiency value could possibly determine the chance of a probable hall of fame selection based on a TSER scale or a possible team winning record during the regular season. This study was conducted based on ten African American Quarterbacks in the NFL based on a touchdown-to-sack efficiency equation. The ten quarterbacks selected for research were from a variety of decades from the late 1970s to the early 2010s and a sixteen game season was used as well. Their efficiency mean was calculated to determine a final efficiency value for their career. The TSER equation was used to display the trends in how many total touchdowns were scored by the quarterbacks compared to how many times they were sacked. The higher values based on the data, were approximately around 8 to 16.2, Daunte Culpepper having the highest value. The overall data in this research was varied since the quarterback played for different teams and played with different styles of offense. This study could also be used for future determination in deciding a fixed salary for the players based on performance or can be a derivative in determining quarterback performance.

- **Po-Ling Loh** (UW–Madison)

Invited session: Models and inference for big data [38]

Friday 10:45 – 11:15, McLeod Auditorium

*Robust techniques in high-dimensional statistics*

**Abstract:** Robust statistics provides a powerful framework for quantifying the behavior of estimators when data are observed subject to imperfections that deviate from standard modeling assumptions. In this talk, we highlight recent work involving statistical theory for robust estimators in high dimensions, with applications to compressed sensing and graphical model estimation. Central to our analysis are (1) a basic understanding of which classes of robust loss functions to employ in order to protect against particular deviations; and (2) rigorous theoretical statements connecting particular characteristics of the loss function to robustness properties of the resulting estimator. Such ideas have been well-studied in low dimensions, but only recently brought to bear in high-dimensional contexts. Due to the fact that many attractive robust loss functions are inherently nonconvex, we leverage results in optimization theory to devise computationally efficient methods for obtaining local or global optima with provably good properties. This further elucidates some of the mostly heuristic methods employed in low-dimensional robust estimation. We conclude by discussing challenges and open questions arising in the mostly unexplored territory lying between robust and high-dimensional statistics.

- **Miles Lopes** (UC Davis, Department of Statistics)

Contributed Talk, session [98]

Thursday 2:35 – 3:05, Rosebrugh 208

*Measuring the algorithmic convergence of random forests via bootstrap extrapolation*

**Abstract:** When making predictions with a voting rule, a basic question arises: “What is the smallest number of votes needed to make a good prediction?” In the context of ensemble classifiers,

such as Random Forests or Bagging, this question represents a tradeoff between computational cost and statistical performance. Namely, by paying a larger computational price for more classifiers, the prediction error of the ensemble tends to improve and become more stable. Conversely, by using fewer classifiers and tolerating some variability in accuracy, it is possible to gain a speedup. In this paper, we propose a bootstrap method to quantify this tradeoff. To be specific, suppose the training dataset is fixed, and let the random variable  $\text{Err}_t$  denote the prediction error of a randomly generated ensemble of  $t = 1, 2, \dots$  classifiers. (The randomness of  $\text{Err}_t$  comes only from the algorithmic randomness of the ensemble.) Working under a “first order model” of Random Forests, we prove that the centered law of  $\text{Err}_t$  can be consistently approximated via our proposed method as  $t \rightarrow \infty$ . As a consequence, this result offers a guideline for choosing the smallest  $t$  needed to ensure that the algorithmic fluctuations are negligible, e.g.  $\text{var}(\text{Err}_t)$  less than a given threshold. Lastly, we explain how the technique of extrapolation can be used to substantially reduce the computational cost of resampling.

- **Wendy Lou** (University of Toronto)

Invited session: Analysis of next generation sequencing data for biomedicine [35]

Tuesday 11:55 – 12:25, Medical Sciences 3153

*Statistical methods for studying microbial compositions*

Coauthor: Anastasia Teterina (Toronto General Hospital, University of Health Network)

**Abstract:** Motivated by a study examining the characteristics of patients with chronic liver diseases, such as non-alcoholic fatty liver disease, statistical methods for analyzing data obtained from Illumina sequencing of 16S rRNA gene amplicons will be presented. Approaches for dealing with sparse data, extreme outliers, and technical variations in multiple runs, through normalization will be discussed first. Various methods, including approaches based on the median of the Hodges-Lehmann difference, permutation-based p-values, and non-parametric tests, will then be discussed to compare abundance at both OTU and genus levels across patient groups. To improve robustness, a two-stage semi-parametric approach based on ranks will be proposed. Numerical comparisons through simulations between this approach and selected common approaches will be given. The corresponding findings for the microbiome data will also be discussed, along with practical recommendations.

- **Zhan-Qian John Lu** (National Institute of Standards and Technology, Gaithersburg, MD 20899-8980)

Contributed Talk, session [71]

Monday 11:55 – 12:25, Rosebrugh 310

*Shape estimation with complex-plane Fourier representation of boundary data*

**Abstract:** Both shape and size are important in biological and medical measurements. I will discuss a statistical formulation of Fourier descriptors for 2D boundary curves and how shape functional such as perimeter and area can be estimated based on sampled boundary data via the Fourier-based smoothing spline model. In particular, I will discuss the implication of uncertainty analysis involving shape and a hypothesis testing problem involving the cell shape and size population.

- **Malwina Luczak** (Queen Mary University of London)

Invited session: Statistics and networks [27]

Wednesday 2:35 – 3:05, Medical Sciences 4171

*Extinction time for the weaker of two competing SIS epidemics*

Coauthor: Fabio Lopes (Universidad de Chile)

**Abstract:** We consider a simple stochastic model for the spread of a disease caused by two virus strains in a closed homogeneously mixing population of size  $N$ . In our model, the spread of each strain is described by the stochastic logistic SIS epidemic process in the absence of the other strain, and we assume that there is perfect cross-immunity between the two virus strains, that is, individuals infected by one strain are temporarily immune to infections by the other strain. For the case where one strain has a strictly larger basic reproductive ratio than the other, and the stronger strain on its own is supercritical (that is, its basic reproductive ratio is larger than 1), we derive precise asymptotic results for the distribution of the time when the weaker strain disappears from the population, that is, its extinction time. We further extend our results to the case where the difference between the two basic reproductive ratios may tend to 0.

In our proof, we set out a simple approach for establishing a fluid limit approximation for a sequence of Markov chains in the vicinity of a stable fixed point of the limit drift equations, valid for a time exponential in system size.

- **Terry Lyons** (University of Oxford)

Invited session: Rough paths [13]

Wednesday 2:00 – 2:30, McLeod Auditorium

*Enveloping algebras, signatures and Chinese handwriting*

**Abstract:** Sequential multimodal data can be modelled in many ways, and were traditionally modelled by using time series. However, the signature is a powerful representation, coming out of algebra and geometry, that has shown itself to be both natural, and in some circumstances, orders of magnitude more efficient. We explain some of the ideas, and an application as a core component (with deep learning) of one of the leading apps for recognising Chinese handwriting online from finger movements on the screen of your mobile phone.

M

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Bence Mélykúti** (University of Freiburg)

Contributed Talk, session [62]

Wednesday 3:10 – 3:40, Medical Sciences 4279

*Parameter estimation in a percolation model with colouring*

Coauthor: Felix Beck (University of Freiburg)

**Abstract:** In the bond percolation model on a lattice, we colour vertices with  $n_c$  colours independently at random according to Bernoulli distributions. A vertex can receive multiple colours and each of these colours is individually observable. The colours colour the entire component into which

they fall. Our goal is to estimate the  $n_c + 1$  parameters of the model: the probabilities of colouring of single vertices and the probability with which an edge is open. The input data is the configuration of colours once the complete components have been coloured, without the information which vertices were originally coloured or which edges are open.

We use a Monte Carlo method, the method of simulated moments to achieve this goal. We prove that this method is a strongly consistent estimator by proving a strong law of large numbers for the vertices' weakly dependent colour values. We evaluate the method in computer tests. The motivating application is cross-contamination rate estimation for digital PCR in lab-on-a-chip microfluidic devices.

- **Ursula U. Müller** (Texas A&M University)

Contributed Talk, session [86]

Wednesday 2:00 – 2:30, Rosebrugh 208

*Improved density estimators for incomplete and contaminated data*

**Abstract:** We will present new statistical methods to make more efficient use of data where relevant variables are not observed (missing, measured with error, surrogates). We will combine recent advances in the construction of rate optimal estimators for densities of convolutions (which converge with  $\sqrt{n}$  rate) with deconvolution techniques for dealing with measurement errors in the responses and in the covariates. The results have relevance for empirical studies in various disciplines, most notably health sciences including nutritional epidemiology, and in social/behavioral research.

- **Marloes Maathuis** (ETH Zurich)

Invited session: Model selection in non-standard high dimensional models [32]

Wednesday 2:00 – 2:30, Medical Sciences 2171

*High-dimensional consistency in score-based and hybrid structure learning*

Coauthors: Preetam Nandy (ETH Zurich), Alain Hauser (Google)

**Abstract:** The main approaches for learning Bayesian networks can be classified as constraint-based, score-based or hybrid methods. Although high-dimensional consistency results are available for the constraint-based PC algorithm, such results have been lacking for score-based and hybrid methods, and most hybrid methods are not even proved to be consistent in the classical setting where the number of variables remains fixed. We study the score-based Greedy Equivalence Search (GES) algorithm, as well as hybrid algorithms that are based on GES. We show that such hybrid algorithms can be made consistent in the classical setting by using an adaptive restriction on the search space. Moreover, we prove consistency of GES and adaptively restricted GES (ARGES) for certain sparse high-dimensional scenarios. ARGES scales well to large graphs, and our simulation studies indicate that both ARGES and GES generally outperform the PC algorithm.

- **James G. MacKinnon** (Queen's University)

Contributed Talk, session [78]

Thursday 10:45 – 11:15, Fitzgerald 139

*Difference-in-differences inference with few treated clusters*

Coauthor: Matthew D. Webb (Carleton University)

**Abstract:** Inference using difference-in-differences with clustered data requires care. Previous research has shown that, when there are few treated clusters,  $t$  tests based on a cluster-robust variance estimator (CRVE) severely over-reject, different variants of the wild cluster bootstrap can over-reject or under-reject dramatically, and procedures based on randomization show promise. We demonstrate that randomization inference (RI) procedures based on estimated coefficients, such as the one proposed by Conley and Taber (2011), fail whenever the treated clusters are atypical. We propose an RI procedure based on  $t$  statistics which fails only when the treated clusters are atypical and few in number. We also propose a bootstrap-based alternative to randomization inference, which mitigates the discrete nature of RI  $P$  values when the number of clusters is small. An empirical example demonstrates that alternative procedures can yield dramatically different inferences.

- **Neal Madras** (York University)

Contributed Talk, session [62]

Wednesday 2:00 – 2:30, Medical Sciences 4279

*Location of the adsorption transition for lattice polymers*

**Abstract:** Polymer molecules interacting with a surface undergo a simple phase transition from “adsorbed” (substantially adhered to the surface) to “desorbed” (free-floating) as the interaction strength passes through a critical value. The surface can be impenetrable (e.g. the wall of a container) or penetrable (e.g. an interface in a liquid). The configuration of a large polymer molecule is described by standard discrete undirected models: self-avoiding walks, lattice trees, and lattice animals. We consider all lattice trees, say, of a fixed large size, and weight each one according to the number of its vertices that lie in a given plane (within 3-space). In the penetrable case, I outline a new combinatorial proof that the transition occurs at a strictly positive interaction strength. In contrast, for the impenetrable case, I present a new result supporting the conjecture that the transition occurs at zero.

- **Vlad Margarit** (University of Oxford)

Contributed Talk, session [58]

Monday 11:20 – 11:50, Medical Sciences 4279

*Convergence in probability of the quantum diffusion in a random band matrix model*

**Abstract:** Random band matrices are the natural objects that interpolate between Wigner matrices and random Schrödinger operators. We consider Hermitian random band matrices  $H$  in  $d \geq 1$  dimensions. In our model, the matrix elements  $H_{xy}$ , indexed by vertices of a lattice  $x, y \in \Lambda \subset \mathbb{Z}^d$ , are independent, uniformly distributed random variables on  $\mathbb{S}^1 \subset \mathbb{C}$  if  $|x - y|$  is less than the band width  $W$ , and zero otherwise. In this model, we consider the time evolution on the lattice of a quantum particle subjected to the Hamiltonian  $H$ . I will report on recent progress in understanding this evolution, strengthening the known mode of convergence from convergence of the expectation to convergence in high probability. The main tool for proving the result is the use of combinatorial arguments emerging from certain double-rooted Feynman graphs.

This result was obtained in a project supervised by Professor Antti Knowles in ETH Zürich.

- **Greg Markowsky** (Monash University)

Contributed Talk, session [52]

Friday 10:45 – 11:15, Medical Sciences 4171

*The planar Brownian Green's function, and a probabilistic proof of the Riemann Mapping Theorem*

**Abstract:** It has been known for some time that the Green's function of a planar domain can be defined in terms of Brownian motion, and that this Green's function is conformally invariant. I will show how this conformal invariance can be generalized in order to calculate Green's function for certain stopping times more general than exit times, and how these considerations can be used to give a proof that any two simply connected domains in the plane (excluding the plane itself) are conformally invariant, also known as the Riemann Mapping Theorem.

- **Servet Martinez** (Universidad de Chile)

Plenary talk: Lévy Lecture

Wednesday 4:15 – 5:05, McLeod Auditorium

*Quasi-stationarity in Markov chains and population dynamics*

**Abstract:** We give some of the main concepts on survival for Markov chains that are surely killed. Quasi-stationary distributions (q.s.d.) are the invariant distributions when conditioned to survive. We give some properties of q.s.d. in a global framework. We introduce some associated concepts, a main one being the  $Q$ -process which is the Markov chain of trajectories that survive forever. All these concepts are completely described for finite state Markov chains by using the Perron-Frobenius theory and they can be seen geometrically in expanding dynamical systems.

For countable Markov chains we give the existence theorem due to Ferrari, Kesten, Martinez, Picco, ensuring that if there is not entrance from infinity then exponential survival is a necessary and sufficient condition to have q.s.d.. In birth and death chains the set of q.s.d.'s the set of q.s.d., when they exist, can be a singleton or a continuum with an extremal one that is the quasi-limiting distribution and plays a fundamental role in the associated  $Q$ -process. We describe some population dynamics models which are birth and death structured and extincted almost surely which were developed jointly with Collet, Meleard and San Martin: a first one has mutation on the traits and we describe the traits conditionally to survive, a second model describe the dynamics of bacteria in a chemostat.

- **Jose M. Maseda-Garcia** (Department of Applied Economics V. University of the Basque Country UPV/EHU)

Contributed Talk, session [99]

Friday 2:00 – 2:30, Rosebrugh 208

*Classification of socio-economic variables in urban regeneration*

**Abstract:** Decisions of public institutions such as Municipalities and Regional Governments concerning urban regeneration in areas with refurbishment necessities and social problems involve the analysis of simultaneous measurements of several types of variables. Our paper shows a classification of many variables that measure social, economic and habitability aspects of urban areas obtained from national and regional public databases in Spain. In a second step, a deeper analysis of relations

between variables is performed to identify dependencies of variables. Multivariate analysis techniques are proposed to understand the relationship between those variables. As a consequence, variables are selected to distinguish those that are essential for the analysis from those that are redundant or do not provide useful information. The global study of the selected variables will provide the decision maker a framework that will facilitate a ranking system of the urban areas to value the necessity of public support for regeneration. The proposed methodology based on multivariate analysis has been used to value the necessity to act in some districts with potential necessity of public support for regeneration in the Basque Country.

- **Hiroki Masuda** (Kyushu University)

Invited session: Statistics of random processes [15]

Wednesday 11:55 – 12:25, Medical Sciences 4171

*On asymptotics of multivariate non-Gaussian quasi-likelihood*

**Abstract:** We consider (semi-)parametric inference for a class of stochastic differential equation (SDE) driven by a locally stable Lévy process, focusing on multivariate setting and some computational aspects. The process is supposed to be observed at high frequency over a fixed time domain. This setting naturally gives rise to a theoretically fascinating quasi-likelihood which brings about a novel unified estimation strategy for targeting a broad spectrum of driving Lévy processes. The limit experiment is mixed normal with a clean-cut random information structure, based on which it is straightforward to make several conventional asymptotic statistical decisions. The infill-asymptotics adopted here makes the popular Gaussian quasi-likelihood useless, while instead enabling us not only to incorporate any exogenous and/or observable endogenous data into the trend and/or scale coefficients without essential difficulty, but also to sidestep most crucial assumptions on the long-term stability such as ergodicity and moment boundedness. The proposed quasi-likelihood estimator is asymptotically efficient in some special cases.

- **Frederick A. Matsen IV** (Fred Hutchinson Cancer Research Center)

Contributed Talk, session [76]

Friday 10:45 – 11:15, Fitzgerald 139

*Consistency and convergence rate of phylogenetic inference via regularization*

Coauthors: Vu Dinh (Fred Hutchinson Cancer Research Center), Lam Si Tung Ho (UCLA), Marc Suchard (UCLA)

**Abstract:** It is common in phylogenetics to have some, perhaps partial, information about the overall evolutionary tree of a group of organisms and wish to find an evolutionary tree of a specific gene for those organisms. There may not be enough information in the gene sequences alone to accurately reconstruct the correct “gene tree”. Although the gene tree may deviate from the “species tree” due to a variety of genetic processes, in the absence of evidence to the contrary it is parsimonious to assume that they agree. A common statistical approach in these situations is to develop a likelihood penalty to incorporate such additional information. Recent studies using simulation and empirical data suggest that a likelihood penalty quantifying concordance with a species tree can significantly improve the accuracy of gene tree reconstruction compared to using sequence data alone.

However, the consistency of such an approach has not yet been established, nor have convergence rates been bounded. Because phylogenetics is a non-standard inference problem, the standard theory does not apply.

In this paper, we propose a penalized maximum likelihood estimator for gene tree reconstruction, where the penalty is square of the Billera-Holmes-Vogtmann geodesic distance from the gene tree to the species tree. We prove that this method is consistent, and derive its convergence rate for estimating the discrete gene tree structure and continuous edge lengths (representing the amount of evolution that has occurred on that branch) simultaneously. We find that the regularized estimator is “adaptive fast converging”, meaning that it can reconstruct all edges of length greater than any given threshold from gene sequences of polynomial length. Our method does not require the species tree to be known exactly; in fact, our asymptotic theory holds for any such guide tree.

- **Jonathan Mattingly** (Duke University)

Invited session: SPDEs and the work of Martin Hairer [1]

Wednesday 11:55 – 12:25, Medical Sciences 3153

*Unique ergodicity of SPDEs: some too smooth and some too singular*

**Abstract:** I will give some perspective on conditions for unique ergodicity in SPDEs in different settings. I will consider systems where the forcing is too smooth in some sense and examples where it is “too” rough. I will contrast the fundamental ideas used in these two different settings to obtain rigorous results. The too rough setting will include SPDE whose local existence theory is understood through Regularity Structures as epitomized by Hairer’s work on the KPZ equation. Most of this talk is joint work with Martin Hairer.

- **Yeliz Mert Kantar** (Anadolu University)

Contributed Talk, session [67]

Monday 11:20 – 11:50, Medical Sciences 3278

*Robust estimation of the Pareto shape parameter*

Coauthor: Ibrahim ARIK (Anadolu University)

**Abstract:** The Pareto distribution have been used in many areas of sciences to model the right tail of heavy tailed distributions. Its shape parameter characterizes the tail behavior of distribution. There are a number of estimation methods for estimating shape parameter. However, the standard methods for estimating the shape parameter such as the maximum likelihood estimator (MLE) are non-robust, in the sense that it is very sensitive to outliers. In recent years, a number of robust estimation methods for distributional parameters have been proposed. In this study, we investigate the robust estimation methods for the shape parameter of Pareto distribution, which includes;  $M$ -estimators, Least Absolute Deviation (LAD), Least Median of Squares (LMS), Least Trimmed Squares, Theil Estimator, Repeated Median Estimator, Quantile Estimator, estimation for transformed location-scale family and Method of Trimmed Moments (MTM). Monte Carlo simulations show that Theil and Repeated Median estimation methods provide the desired performance against outliers and show good performance without outliers.

Keywords: Pareto Distribution, Shape Parameter, Robustness, Robust Estimator.

- **Ivana Milovic** (University of Vienna, Austria)

Contributed Talk, session [96]

Monday 10:45 – 11:15, Fitzgerald 139

*Conditional means of low-dimensional projections from high-dimensional data. Explicit error bounds.*

Coauthor: Hannes Leeb (University of Vienna, Austria)

**Abstract:** Consider a random  $d$ -vector  $Z$  that has a Lebesgue density and that is standardized so that  $\mathbb{E}Z = 0$  and  $\mathbb{E}ZZ' = I_d$ . Moreover, consider two projections defined by unit-vectors  $\alpha$  and  $\beta$ , namely a response  $y = \alpha'Z$  and an explanatory variable  $x = \beta'Z$ . Under certain regularity conditions, Leeb (2013, *The Annals of Statistics*) has shown that for most  $\beta$ 's,  $\mathbb{E}[y|x]$  is approximately linear in  $x$ , and that  $\text{Var}[y|x]$  is approximately constant in  $x$ , provided that  $d$  is large. These results imply that most simple submodels of a high-dimensional linear model are approximately weakly correct. In order to develop statistical procedures for inference with sparse linear submodels, we require a good control of error terms. But the results obtained by Leeb are asymptotic, as  $d \rightarrow \infty$  and no explicit bounds have been established. Now, for every  $\beta$  let  $\Delta(\beta, x)$  denote the worst-case deviation (in  $\alpha$ ) of  $\mathbb{E}[y|x]$  from  $\alpha'\beta x$  and let  $B_d$  be the set of all  $\beta$ 's such that  $\Delta(\beta, \beta'Z)$  is small. For a fixed  $d$ , we provide explicit lower bounds on the size of  $B_d$ , as measured with respect to the uniform distribution  $v$  on the  $d$ -unit sphere, and we show that  $v(B_d) \rightarrow 1$ , as  $d \rightarrow \infty$ , at an arbitrary polynomial rate. Additionally, we provide finite- $d$  upper bounds on  $\Delta(\beta, \beta'Z)$ , for  $\beta$ 's in  $B_d$ . Finally, we apply our findings to the class of subgaussian random variables, to obtain even better results.

- **Isabel Linda Moyo (nee Zulu)** (National University of Science and Technology)

Contributed Talk, session [48]

Tuesday 11:55 – 12:25, Medical Sciences 2171

*An empirical study on the factors influencing the performance of the financial institutions in Zimbabwe.*

**Abstract:** The purpose of the study is to learn more about the factors that influence the performance of companies in the Financial sector in Zimbabwe. The results are envisaged to benefit the individual companies striving for success in our country and also assist foreign investors in deciding on the allocation of their funds to different financial companies in the sector. Data was extracted from the companies' Balance sheets and Income Statements. Econometric and statistical models on the Return on Assets (ROA), Return on Equity (ROE) and Net Interest Margin (NIM) (independent variables) were developed where an analysis of the effects of both the Microeconomic and Macroeconomic Variables was carried out.

- **Ioane Muni Toke** (CentraleSupélec)

Invited session: Statistics of random processes [15]

Wednesday 11:20 – 11:50, Medical Sciences 4171

*Order flows in a limit order book: modelling and estimation using high-frequency aggregated data*

**Abstract:** We study a model of limit order book based on parametric point processes, its estimation using high-frequency data, and its numerical simulation.

The first step deals with data preparation and the empirical reconstruction of the order flow. We investigate tick-by-tick data provided by the Thomson-Reuters Tick History database for several stocks on three different exchanges (Paris - Euronext, London and Frankfurt - Deutsche Börse) and on a five-year span. We propose an algorithm that helps the synchronization of the trades and quotes data sources, providing enhancements to the basic procedure that, depending on the time period and the exchange, are shown to be significant. We show that the analysis of the performance of this algorithm turns out to be a forensic tool assessing the quality of the aggregated database: we are able to track through the data some significant technical changes that occurred on the studied exchanges. Importantly, we also show that the choices made when reconstructing order flows have consequences on the quantitative models that are to be calibrated afterwards on such data.

The second step proposes a parametric model for the simulation of limit order books. We assume that limit orders, market orders and cancellations are submitted according to point processes with state-dependent intensities. We propose functional forms for these intensities, formally exponentials or quadratic forms of the (log of the) studied variables, as well as parametric models for the placement of limit orders and cancellations. For cancellations, we use the concept of “priority index” to describe the selection of orders to be cancelled in the order book. Parameters of the model are estimated using likelihood maximization. We illustrate the performance of the model by providing extensive simulation results, with a comparison to empirical data and a standard Poisson reference.

- **Nina Munkholt Jakobsen** (University of Copenhagen)

Contributed Talk, session [49]

Thursday 10:45 – 11:15, Medical Sciences 3278

*Efficient estimation for jump-diffusions*

Coauthor: Michael Sørensen (University of Copenhagen)

**Abstract:** This talk concerns parametric estimation for ergodic univariate diffusion processes with finite activity jumps, given by stochastic differential equations. The processes are assumed to be observed at high frequency over an increasing time interval, with terminal sampling time going to infinity. We present a result stating that under quite general assumptions, approximate martingale estimating functions yield consistent estimators of the parameters of the process. These estimators are asymptotically normally distributed, and their asymptotic variances may be estimated consistently.

With reference to results in the literature, we conjecture on the criteria for rate optimality and efficiency of estimators of drift-jump and diffusion parameters. Conditions for rate optimality and efficiency of the approximate martingale estimating function-based estimators are discussed in special cases. These conditions are found to extend the pre-existing conditions applicable in connection with continuous diffusions, and they impose stronger requirements on the estimating functions in the presence of jumps. For example, depending on the jump dynamics of the model, it can be considerably more difficult to achieve rate optimal estimators of the diffusion parameter for a jump-diffusion, than for the corresponding continuous diffusion. Furthermore, in order to achieve efficiency,

coordinate functions of the estimating function must be able to discriminate asymptotically between observations of the process at jump times and non-jump times.

- **Kairat Mynbaev** (Kazakh-British Technical University)

Contributed Talk, session [85]

Tuesday 11:55 – 12:25, Medical Sciences 3278

*Estimation of densities in bounded and unbounded domains: the missing link*

Coauthor: Carlos Martins-Filho (Department of Economics, University of Colorado, Boulder, CO, USA and IFPRI, 2033 K Street NW Washington, DC, USA)

**Abstract:** Kernel estimation of densities in domains with boundaries is known to suffer from boundary effects. We show that in case of smooth densities the right way is to extend the density and estimate the extension. The resulting estimators in domains with boundaries have bias and variance expressed in terms of density extensions. Therefore they have the same rates as on the real line. No kernel modification near endpoints is needed and the kernels commonly used for estimation on the real line can be applied. Densities defined on the half-axis and in a unit interval are considered. The results are applied to estimation of densities that are discontinuous or have discontinuous derivatives, where they yield the same rates of convergence as for smooth densities on the real line.

**N**

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Toshio Nakata** (University of Teacher Education Fukuoka)

Contributed Talk, session [67]

Monday 11:55 – 12:25, Medical Sciences 3278

*Weak laws of large numbers for weighted independent random variables with infinite mean*

**Abstract:** We study weak laws of large numbers for weighted independent random variables with infinite mean. In particular, this paper explores the case that the decay order of the tail probability is -1. Moreover, we extend a result concerning the Pareto–Zipf distributions given by Adler (2012).

- **Kihun Nam** (Rutgers University)

Contributed Talk, session [45]

Monday 2:00 – 2:30, Medical Sciences 4171

*Fixed point framework in BSDEs: beyond the classical Banach fixed point theorem*

Coauthor: Patrick Cheridito (Princeton University)

**Abstract:** We show that BSDE and its variant can be translated into a fixed point problem in a space of random vectors. This makes it possible to employ general fixed point arguments to find solutions. For instance, Banach-type fixed point theorem can be used to derive general existence and uniqueness results for equations with Lipschitz coefficients, whereas Schauder-type fixed point arguments can be applied to non-Lipschitz equations. The approach works equally well for multidimensional as for one-dimensional equations and leads to results in several interesting cases such as equations with path-dependent coefficients, anticipating equations, McKean-Vlasov type equations and equations with coefficients of superlinear growth.

- **Ariel Neufeld** (ETH Zurich)  
 Contributed Talk, session [47]  
 Monday 11:20 – 11:50, Medical Sciences 3154  
*Nonlinear Lévy processes and robust utility maximization*  
 Coauthor: Marcel Nutz (Columbia University)  
**Abstract:** We present a tractable framework for Knightian uncertainty, the so-called nonlinear Lévy processes, and use it to formulate and solve problems of robust utility maximization. This talk is based on joint work with Marcel Nutz.
- **Philip Ngare** (University of Nairobi)  
 Contributed Talk, session [47]  
 Monday 11:55 – 12:25, Medical Sciences 3154  
*Modelling risk of financing agribusiness in an emerging market*  
**Abstract:** Our study investigates insurance-investment model where risk process can be controlled by index based insurance contract and investment in improved technology using a loan from bank. The performance criterion is either expected utility of increased production or lower agricultural credit rates for investor with both insurance and credit in his/her portfolio. We solve financial optimization problem for financial institution who is deciding between financing agribusiness for an investor with insurance or those without insurance. We assume that an investor takes credit to invest in technology to increase his/her productivity. We establish the optimal credit rates for both investors and justify low rates for an investor with insurance. Also we establish bounds for optimal insurance coverage taken by each investor depending on his or her risk preference. The structured model is aimed to solve the complexity in low uptake index-based insurance and dis-interest of financial institutions in financing small scale farmers who are major contributors to food security in developing countries.
- **Gia Bao Nguyen** (Universidad de Chile)  
 Contributed Talk, session [58]  
 Monday 11:55 – 12:25, Medical Sciences 4279  
~~*Non-intersecting Brownian bridges and the Laguerre Orthogonal Ensemble*~~ TALK CANCELLED  
 Coauthor: Daniel Remenik (Universidad de Chile)  
**Abstract:** Consider  $N$  non-intersecting Brownian bridges, all starting from 0 at time  $t = 0$  and returning to 0 at time  $t = 1$ . The  $\text{Airy}_2$  process is defined as the motion of the top path (suitably rescaled) in the large  $N$  limit. K. Johansson proved the remarkable fact that the supremum of the  $\text{Airy}_2$  process minus a parabola has the Tracy-Widom GOE distribution from random matrix theory. In this talk, I will present a result which shows that the squared maximal height of the top path in the case of finite  $N$  is distributed as the top eigenvalue of a random matrix drawn from the Laguerre Orthogonal Ensemble. This result can be thought of as a discrete version of Johansson's result, and provides an explanation of how the Tracy-Widom GOE distribution arises in the KPZ universality class. The result can also be recast in terms of the probability that the top curve of the stationary Dyson Brownian motion hits an hyperbolic cosine barrier. This is joint work with Daniel Remenik.

- **Abdou-Aziz Niang** (University of Ziguinchor)  
 Contributed Talk, session [94]  
 Friday 11:55 – 12:25, Medical Sciences 2172  
*On factor-augmented univariate forecasting*  
**Abstract:** The question we examine in this paper is the problem of improvement of univariate forecasting when the variable of interest belongs to a panel of time series which are dependant across units. We analyse in possibly nonstationary framework, to what extent forecast based on the augmented univariate process implied by a factor model can show substantial advantages in terms of expected gains, with respect to simple univariate model. Moreover we consider identical autoregressive (AR) roots over the cross sections. Analyses are done theoretically and on the basis of Monte Carlo simulations. Our results show that in the general case where non stationarity is allowed, substantial forecast error reduction of the univariate process can be achieved by simply augmenting each individual time series with its idiosyncratic factor.
  
- **Richard Nickl** (University of Cambridge)  
 Invited session: Frequentist properties of Bayesian nonparametric or high dimensional models [34]  
 Friday 11:20 – 11:50, Medical Sciences 2171  
*On Bayes solutions of some non-linear inverse problems: Diffusion models (and beyond)*  
**Abstract:** Bayesian methods are very popular in inverse problems, but little is known about rigorous (frequentist) performance guarantees for posterior based inference procedures, particularly for non-linear inverse problems where no natural conjugate priors exists. In this talk we will give some ideas on how to prove general contraction theorems for Bayes procedures in such settings, focussing on a key example where one observes discrete low-frequency samples from a continuous diffusion process, and wishes to make inference on the underlying drift and diffusion coefficient. This problem is closely related to inferring the semigroup dynamics of a second-order differential operator at discrete times, relevant in contemporary ‘data assimilation’ and filtering problems. We discuss the ingredients of the first rigorous proof of the fact that Bayesian posterior distributions solve this non-linear inverse problem in a minimax optimal way. We illustrate how the method of proof, which consists in particular of an adaptation of the usual ‘small ball estimates’ for the prior to the inverse setting, should generalise to other settings, and how the usual ‘Hellinger’ tests can be replaced by tests based on new concentration inequalities for additive functionals of Markov chains that are of independent interest.
  
- **Nynke Niezink** (University of Groningen)  
 Contributed Talk, session [69]  
 Friday 10:45 – 11:15, Medical Sciences 3278  
*Modeling the dynamics of social networks and continuous actor attributes*  
 Coauthor: Tom Snijders (University of Groningen, University of Oxford)  
**Abstract:** Social networks and the characteristics of the actors who constitute these networks are not static; they evolve interdependently over time. People may befriend others with similar

political opinions or change their own opinion based on that of their friends. The stochastic actor-oriented model is used to statistically model such dynamics. We will present an extension of this model for continuous dynamic actor characteristics. The method available until now assumed actor characteristics to be measured on an ordinal categorical scale, which yielded practical limitations for applied researchers. We now model the interdependent dynamics by a stochastic differential equation for the attribute evolution and a Markov chain model for the network evolution. Together these form a continuous-time Markov process on a state space that is the product of a continuous and discrete state space. Although the model is too complicated to calculate likelihoods or estimators in closed form, this stochastic process can be easily simulated. Therefore, we estimate model parameters using the method of moments and the Robbins-Monro algorithm for stochastic approximation. We will illustrate the proposed method by a study of the relation between friendship and obesity, analyzing body mass index as continuous dynamic actor attribute.

- **John Nolan** (American University)

Contributed Talk, session [67]

Monday 10:45 – 11:15, Medical Sciences 3278

*A measure of dependence for stable distributions*

Coauthor: Tuncay Alparslan (U.S. Department of Energy)

**Abstract:** A distance based measure of dependence is proposed for stable distributions that completely characterizes independence for a bivariate stable distribution. Properties of this measure are analyzed, and contrasted with the covariation and co-difference. A sample analog of the measure is defined and demonstrated on simulated and real data, including time series and distributions in the domain of attraction of a stable law.

- **David Nualart** (The University of Kansas)

Invited session: Limit Theorems and the Malliavin calculus [11]

Tuesday 2:00 – 2:30, Medical Sciences 2171

*Central limit theorem for symmetric integrals*

**Abstract:** The purpose of this talk is to present the convergence in distribution of  $\nu$ -symmetric integrals of functions of the fractional Brownian motion with Hurst parameter  $H$ , when  $H$  is equal to the critical value  $(2(\ell(\nu)+1))^{-1}$  and  $\ell(\nu) \geq 1$  is maximum natural number  $j$  such that the moment of order  $2j$  of the measure  $\nu$  on  $[0, 1]$  is  $1/(2j + 1)$ . This theorem characterizes previous results for  $H = \frac{1}{4}$ ,  $H = \frac{1}{6}$  and  $H = \frac{1}{10}$ , when the corresponding symmetric integrals are defined as the limit of trapezoidal, midpoint and Simpson Riemann sums, respectively. As a consequence, we establish a change-of-variable formula in law, where the correction term involves a stochastic integral with respect to an independent standard Brownian motion. The proof is based on the combination of Malliavin calculus and the classical Bernstein's big blocks/small blocks technique.

- **Victoria Nyaga** (Scientific Institute of Public Health/Universiteit Hasselt, Belgium)

Contributed Talk, session [78]

Thursday 11:55 – 12:25, Fitzgerald 139

*Beta-binomial ANOVA model for network meta-analysis of diagnostic test accuracy data*

Coauthors: Dr. Marc Arbyn (Scientific Institute of Public Health, Belgium), Prof. Marc Aerts (Universiteit Hasselt, Belgium)

**Abstract:** Network meta-analysis improves the estimation process by allowing borrowing of information between studies yielding more precise estimates especially for diagnostic tests evaluated in a small number of studies. Typical analysis assumes that the logit transformed study-specific sensitivity and specificity has approximately normal distribution with constant variance. For proportions such a model is structurally flawed because the variance depends on the underlying probability. The constant variance condition is rarely satisfied, indeed as the probability move to 0 or 1, the variance moves towards zero and is highest when the probability is 0.5. Furthermore, for sparse data, when the number of studies is small or events are rare, the assuming a normal distribution is rarely valid.

We propose the use of bivariate beta distribution for the regression analysis of correlated proportions in a multiple treatment (tests) meta-analysis. The main assumption is that the in some or all the studies, the missing test are missing at random. The open-source software Stan is used to analyze the illustration data within the Bayesian framework. p16/Ki67 immunocytochemistry was consistently identified as the most discriminatory diagnostic test in detecting cervical intraepithelial neoplasia lesions of grade 2/3 or higher in women with equivocal or mildly abnormal cytology.

Using beta distribution in regression has the following advantages that the model is exact which is particularly important in meta-analysis of sparse data which are common and provides parameters with meaningful interpretation.

- **Victoria Nyaga** (Scientific Institute of Public Health/Universiteit Hasselt, Belgium)

Contributed Talk, session [80]

Friday 11:55 – 12:25, Rosebrugh 208

*Contrast- and arm-based network meta-analysis of diagnostic test accuracy studies: a simulation study*

Coauthors: Dr. Marc Arbyn (Scientific Institute of Public Health, Belgium), Prof. Marc Aerts (Universiteit Hasselt)

**Abstract:** Network meta-analysis provide a holistic manner of combining information from related but independent studies thereby providing meta-analytic estimates simultaneously from all outcomes of interest in place of many separate comparative (two outcomes) meta-analysis. Such an approach decreases the chance of finding spurious significant treatment effects and avoids the multiplicity problems where adjustment for multiple comparisons might be needed (Bender et al 2008). Furthermore, such an integrated analysis will have more power to detect effects than disjoint comparative meta-analysis and improves the estimation process by allowing borrowing of information between studies yielding more precise estimates especially with outcomes assessed in a small number of studies. With increasing popularity of network meta-analysis for therapeutic interventions and development of statistical routines in many statistical software, literature on network meta-analysis methodology for diagnostic tests concurrently continues to grow, however, the number of simulation studies is limited on network meta-analysis methodology for diagnostic tests. To assess the influence of number of tests per study, number of studies and impact of missing data, and model misspecification we perform a

simulation study with different scenarios. The impact is measured using bias, mean squared error and coverage of the mean structure.

- **Justine Omai Nyanchoga** (Moi University Main Campus)

Contributed Talk, session [91]

Wednesday 2:35 – 3:05, Medical Sciences 3278

*Complex statistics generating line of best-fit for linear regression analysis*

Coauthors: Timothy Omwaga (Moi University), Kipkorir Bernard (Moi University)

**Abstract:** The major goal of this paper is to illustrate how the use of complex statistics has enhanced a scientific establishment of the Equation of Line of Best Fit in a scatter plot of Particular stochastic Variables which exhibit some linear correlation with extension to curvilinear system variables. Important applications of the the research are also included.

- **Justine Omai Nyanchoga** (Moi University Main Campus)

Contributed Talk, session [97]

Friday 3:10 – 3:40, Medical Sciences 2172

*Introducing the real-time normal frequency functions for analysis of dynamic data*

Coauthors: Prof. Simeon Kipkoech Mining (Moi University Main Campus), Prof. Kosgei (Moi University Main Campus)

**Abstract:** Motivated by the Wikibon Big Data Manifesto and the community’s need for better analytic tools and technologies to deal with Big Data analysis, I am pleased to share with my colleagues my discovery of the Real-Time Normal (RTN) Frequency Function as Big data analytical tool. This is mainly because the traditional Gaussian Normal Curve, which is more suited to static than to dynamic data, is straining under the massively increasing volume and exceeding variability of Big Data. The RTN frequency function, both by experimental animations and theoretical derivation, is developed from a naturally isolated dynamic system whose major state variable - that is time - is a universal latent input for all “independent” and “dependent” variables in the system. Also, it is shown that the RTN frequency function is more robust than the Gaussian Normal function. Important and diverse applications of the RTN frequency function are also identified in the paper.

**O**

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Sean O’Rourke** (University of Colorado Boulder)

Invited session: Random matrices [6]

Friday 11:55 – 12:25, Medical Sciences 3153

*Eigenvectors of random matrices*

**Abstract:** Eigenvectors of large matrices (and graphs) play an essential role in many applications. The goal of this talk is to present several properties of the eigenvectors when the matrix (or graph) is random. In particular, I will address the largest coordinate, smallest coordinate, joint distribution of several coordinates,  $\ell^p$ -norm, and amount of mass contained in a subset of coordinates. This talk is based on joint work with Van Vu and Ke Wang.

- **Andriy Olenko** (La Trobe University)  
 Contributed Talk, session [66]  
 Thursday 11:20 – 11:50, Medical Sciences 4171  
*Sampling approximation of sub-Gaussian random processes*  
**Abstract:** We generalize some classical results and obtain new truncation error upper bounds in the sampling theorem for bandlimited stochastic processes. The presentation investigates  $L_p([0, T])$  and uniform approximations of  $\varphi$ -sub-Gaussian random processes by finite time sampling sums. Explicit truncation error upper bounds are established. Some specifications of the general results for which the assumptions can be easily verified are given. Direct analytical methods are employed to obtain the results.
- **Andriy Olenko** (La Trobe University, Australia)  
 Contributed Talk, session [65]  
 Friday 11:20 – 11:50, Medical Sciences 4279  
~~*On convergence rate of functionals of random fields*~~ TALK CANCELLED  
**Abstract:** We study local functionals of homogeneous random fields with long-range dependence. The case when the summands/integrands are functionals of a long-range dependent Gaussian process is of great importance in the theory of limit theorems for sums/integrals of dependent random variables and various physics applications. Comparing with the CLT, long-range dependent summands can produce different normalizing coefficients and non-Gaussian limits.  
 Very powerful techniques and approaches based on the Malliavin calculus and Stein's method were developed by I. Nourdin and G. Peccati. However, despite recent progress in the non-central limit theory there has been remarkably little fundamental theoretical study on rates of convergence in non-central limit theorems.  
 The main result of this study is the rate of convergence in non-central limit theorems. Specifications of the main theorem are also discussed for several scenarios. In particular, special attention is paid to the Cauchy, generalized Linnik's, and local-global distinguisher random processes and fields. In contrast to the Malliavin calculus approach we use direct analysis methods to investigate the rate of convergence in the uniform metric.  
 The presentation generalizes some results of the paper Vo Anh, N. Leonenko, A. Olenko, On the rate of convergence to Rosenblatt-type distribution for long-range dependent random fields. *J. Math. Anal. Appl.* Vol. 425, No. 1, (2015), 111132.
- **Jamiu Shehu Olumoh** (American University of Nigeria, Yola, Nigeria)  
 Contributed Talk, session [83]  
 Monday 3:10 – 3:40, Fitzgerald 139  
*A Statistical model for insurgency count data*  
 Coauthors: M.O.O. Sanni (University of Ilorin, Ilorin, Nigeria), Osho O Ajayi (American University of Nigeria, Yola, Nigeria)  
**Abstract:** In insurgency count data, it is common to see some intrinsic characteristics such as a large number of zeros that denotes the days during which no victims from insurgent attacks are recorded

either because there were no attacks or attacks resulted in no causality. Outrageous number of casualties are also possible causing outliers with very heavy/long tail in any model descriptor of such data. Count-valued outcomes of this nature cannot be adequately modeled using many of the count models popularized over the years. In this work, we proposed a model which eliminates a number of the problems associated with using existing count data models to analyze data with insurgency characteristics. An exploration of the small sample and asymptotic properties of the model suggests that it is a good model for this type of data.

- **Samet Oymak** (Google Inc.)

Invited session: Compressed sensing [23]

Thursday 3:10 – 3:40, Medical Sciences 2172

*Universality laws for randomized dimension reduction*

Coauthor: Joel Tropp (California Institute of Technology)

**Abstract:** Dimension reduction is the process of embedding high-dimensional data into a lower dimensional space to facilitate its analysis. A fundamental technique for dimension reduction is to apply a random linear map to the data. It is desirable that the random map preserves the geometric features of the set (e.g. norm of the set elements). The question is how large the embedding dimension must be to ensure that dimension reduction succeeds with high probability. This work illustrates a phase transition in the behavior of the dimensionality reduction procedure as the embedding dimension increases. For several important learning tasks, the location of the associated phase transition is universal for a large class of dimension reduction maps. Furthermore, the stability and robustness properties are also universal. These results have many applications in signal processing, statistics, and randomized linear algebra.

- **Mehmet Öz** (Özyegin University, Istanbul)

Contributed Talk, session [57]

Wednesday 11:55 – 12:25, Rosebrugh 208

*Survival of branching Brownian motion in a random trap field*

Coauthors: Mine Caglar (Koc University, Istanbul), Janos Engländer (University of Colorado at Boulder)

**Abstract:** We study a branching Brownian motion  $Z$  evolving in  $\mathbb{R}^d$ , where a radially decaying field of Poissonian traps is present. Each trap is a ball with constant radius. Considering a general offspring distribution and conditioning  $Z$  on non-extinction, we find the asymptotic decay rate of the annealed probability that none of the particles of  $Z$  hits a trap. The method of proof is to use a skeleton decomposition for the underlying Galton-Watson process when supercritical and to show that the particles of finite line of descent do not contribute to the survival asymptotics. On the way, a convergence result on the conditional speed of branching Brownian motion is proved.

**P**

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Sandra Palau Calderón** (Centro de Investigación en Matemáticas)  
 Contributed Talk, session [56]  
 Thursday 3:10 – 3:40, Medical Sciences 3278  
*Multi-type continuous-state branching processes*  
**Abstract:** In this talk we will define and study multi-type continuous-state branching processes with a countably infinite number of types. Local and global extinction will be compared. We will provide a local extinction dichotomy. At the end of this talk we will show a process that has a.s. local extinction but a positive probability of non-global extinction.
  
- **Victor Panaretos** (EPFL)  
 Contributed session: Functional data analysis [43]  
 Tuesday 2:00 – 2:30, Medical Sciences 3154  
*Smooth plus rough variation of functional data: the interplay between rank, resolution, and scale*  
**Abstract:** Functional data analyses typically proceed by smoothing, followed by functional PCA. This paradigm implicitly assumes that any roughness is due to nuisance noise. Nevertheless, relevant functional features such as time-localised or short scale variations may indeed be rough. These will be confounded with the smooth components of variation by the smoothing/PCA steps, potentially distorting the parsimony and interpretability of the analysis. We will explore how both smooth and rough variations can be recovered on the basis of discretely observed functional data. Assuming that a functional datum arises as the sum of two uncorrelated components, one smooth and one rough, we develop identifiability conditions for the estimation of the two corresponding covariance operators. We construct nonlinear estimators of the smooth and rough covariance operators and their spectra via low rank matrix completion, and establish their consistency and rates of convergence. We then use them to recover the smooth and rough components of each functional datum, effectively producing separate functional PCAs for smooth and rough variation. Based on joint work with Marie-Hélène Descary (EPFL).
  
- **Vladimir Panov** (Higher School of Economics (Moscow, Russia))  
 Contributed Talk, session [90]  
 Tuesday 11:20 – 11:50, Rosebrugh 310  
*Low-frequency estimation of continuous-time moving average Lévy processes*  
 Coauthor: Denis Belomestny (University of Duisburg-Essen (Germany) and Higher School of Economics (Moscow, Russia))  
**Abstract:** This talk is devoted to statistical inference for the integrals of the type

$$Z_t = \int_{\mathbb{R}} \mathcal{K}(t-s) dL_s, \quad t \in \mathbb{R},$$

where  $\mathcal{K} : \mathbb{R} \rightarrow \mathbb{R}_+$  is a deterministic function, and  $L_t$  is a Lévy process. We study the problem of statistical inference for the Lévy density  $\nu(\cdot)$  of the process  $L_t$  from the observations of the process  $Z_t$ .

It turns out that the methodology crucially depends on the properties of the kernel  $\mathcal{K}(\cdot)$ . For some kernels, e.g., for symmetric kernels in the form  $\mathcal{K}_\alpha(x) = (1 - \alpha x)^{1/\alpha}$ ,  $x \in (0, 1/\alpha)$  with  $\alpha \in (0, 1)$ , estimation procedure directly follows from the representation of the characteristic function, see [1].

Nevertheless, for most examples, the direct approach is not applicable. Our basic example is  $\mathcal{K}(x) = |x|^r e^{-|x|}$  with  $r = 0, 1, 2, \dots$ , which leads to many known models. For instance, if  $r = 0$  then  $Z_t$  is known as the well-balanced Ornstein-Uhlenbeck process [3].

In this research, we study the estimator for  $\nu$  based on the Mellin transform techniques. The key point in this approach is that the Mellin transform of the derivative of the function  $\Lambda(u) = \log(\mathbb{E}[e^{iuZ_t}])$  and the Mellin transform of the function  $\bar{\nu}(x) := x\nu(x)$  are related by the following formula:

$$\mathcal{M}[\Lambda'](z) = i\Gamma(z)e^{i\pi z/2} \int_{\mathbb{R}} (\mathcal{K}(x))^{1-z} dx \cdot \mathcal{M}[\bar{\nu}](1-z), \quad z \in \mathbb{C}.$$

This formula yields that the Mellin transform of  $\bar{\nu}$ , and therefore  $\nu$  itself, can be estimated by getting use of the estimator of  $\mathcal{M}[\Lambda']$ , which is based in its turn by the natural non-parametric estimator of the Fourier transform of  $Z_t$ .

As the result of this study, we construct a consistent estimator for the Lévy density  $\nu(\cdot)$ , derive the convergence rates and prove their optimality. We discuss also the possibility to apply this approach to the fractional Lévy process [2], which is recently of great interest in the stochastic community.

#### REFERENCES

- [1] Belomestny, D., and Reiss, M. (2015) *Estimation and calibration of Lévy models via Fourier methods*. Lévy matters IV. Estimation for discretely observed Lévy processes. Springer.
- [2] Marquardt, T. (2006) *Fractional Lévy processes with an application to long memory moving average processes*. Bernoulli, 12(6), 1099-1126.
- [3] Schnurr, A., and Wörner, J. (2011) *Well-balanced Lévy driven Ornstein-Uhlenbeck processes*. Stat. Risk Model. 28(4), 343-357.

- **Juan-Carlos Pardo-Fernandez** (Universidade de Vigo (Spain))

Contributed Talk, session [93]

Thursday 11:55 – 12:25, Rosebrugh 310

*Asymptotic distribution-free tests for semiparametric regressions*

Coauthors: Juan-Carlos Escanciano (Indiana University (USA)), Ingrid Van Keilegom (Université Catholique de Louvain (Belgium))

**Abstract:** This piece of research proposes a new general methodology for constructing nonparametric and semiparametric asymptotic distribution-free tests for semiparametric hypotheses in regression models for possibly dependent data coming from a strictly stationary process. Classical tests based on the difference between the estimated distributions of the restricted and unrestricted regression errors are not asymptotic distribution-free. In this article we introduce a novel transformation of this difference that leads to asymptotic distribution-free tests with well known critical values. The general methodology is illustrated with applications to testing for parametric models against non-parametric or semiparametric alternatives, and semiparametric constrained mean-variance models.

Several Monte Carlo studies and an empirical application show that the finite sample performance of the proposed tests is satisfactory in moderate sample sizes.

- **Byeong Park** (Seoul National University)

Plenary talk: Laplace Lecture

Tuesday 8:30 – 9:20, McLeod Auditorium

*In-sample density forecasting*

**Abstract:** In-sample density forecasting is a new class of forecasting methods based on structured nonparametric density models. The problem arises when the data are not observed over the full support  $S$ , but are truncated outside a proper subset of  $S$ , say  $I$ . The task is then to estimate the underlying density not only on  $I$  but also on  $S \setminus I$ . This can be accomplished in case the density restricted to  $I$  and to  $S \setminus I$  are described by the common one-dimensional nonparametric component functions in such a way that the component functions are fully identified by the density restricted to  $I$ . In this presentation, I introduce two promising structured density models and a powerful technique of estimating the component functions with sound theory. I also illustrate the approach by applying it to claims reserve prediction and to mesothelioma mortality forecasting.

- **Byeong Park** (Seoul National University)

Invited session: A celebration of the work of Peter Hall [39]

Thursday 2:35 – 3:05, McLeod Auditorium

*Smooth backfitting for errors-in-variables additive models*

**Abstract:** We study nonparametric additive regression models where covariates are contaminated by measurement errors. Based on a deconvolution technique, we construct an iterative algorithm for smooth backfitting additive component functions with the contaminated covariates. We focus on the case of ordinary smooth measurement errors to employ efficient deconvolution and present the algorithmic properties of the deconvolution smooth backfitting as well as its asymptotic properties. Interestingly, the rate of convergence of the deconvolution smooth backfitting estimator changes continuously along the smoothness of the measurement errors, but is accelerated in a certain range so that the effect of contamination is confined into a negligible magnitude in the backfitting process. In the latter case, each component function estimator achieves asymptotic normality with the oracle variance that can be obtained under the knowledge of the other components. We also present the finite sample performance of the deconvolution smooth backfitting in comparison with the naive application of standard smooth backfitting, and demonstrate that the former gives smaller mean integrated squared errors than the latter.

- **Todd Parsons** (Université Pierre et Marie Curie)

Contributed Talk, session [54]

Monday 11:20 – 11:50, Medical Sciences 4171

*Invasion in density dependent population processes*

**Abstract:** In a series of papers in the 1970s, Thomas G. Kurtz identified a class of models that he called Density Dependent Population Processes, which contained a variety of interesting examples

from biology and chemistry. Kurtz and subsequent authors derived functional laws of large numbers and central limit theorems for this class, but outside of specific examples, the problem of the establishment of a new population from very small numbers - a question of fundamental interest in ecology, epidemiology, and evolutionary biology - has remained largely unexamined. In my talk, I will present partial results for the general case, making a connection with results on permanence for dynamical systems.

- **Tim Patschkowski** (Ruhr-Universität Bochum)  
 Contributed Talk, session [85]  
 Tuesday 10:45 – 11:15, Medical Sciences 3278  
*Locally adaptive confidence bands*  
 Coauthor: Angelika Rohde (Ruhr-Universität Bochum)  
**Abstract:** We develop locally adaptive confidence bands in density estimation. In contrast to those previously proposed in the literature, their width is locally variable and shown to adapt to the unknown local smoothness except for a nowhere dense subset with respect to the underlying Banach norms (i.e. of first Baire category). On a technical level, the crucial ingredient is the identification of a stationary process as a least favorable case by means of Slepian’s comparison inequality. The asymptotic quantiles turn out to be universally those of the standard Gumbel distribution.
- **Giovanni Peccati** (University of Luxembourg)  
 Invited session: Stein’s method [12]  
 Monday 11:55 – 12:25, Medical Sciences 2171  
*Non-universal second order behaviour of the nodal sets generated by arithmetic random waves*  
**Abstract:** First studied by Rudnick and Wigman, arithmetic random waves are Gaussian Laplace eigenfunctions on the two-dimensional torus. In this talk, I will describe the asymptotic behaviour of the so-called “nodal length” (that, is the volume of the zero set) of such random objects, and show that (quite unexpectedly) it is non-central and non-universal. The non-central behaviour can be understood in terms of an underlying four-dimensional quantitative central limit theorem, and has to be studied by means of highly non-trivial arithmetic considerations. Our approach allows one to explain from a spectral standpoint the so-called “Berry cancellation phenomenon”, according to which geometric quantities related to the zeros of random waves have typically smaller fluctuations than those associated with non-zero level sets. Joint work with D. Marinucci (Rome Tor Vergata), M. Rossi (Luxembourg) and I. Wigman (King’s College, London).
- **Danilo Pellin** (University Center of Statistics for the Biomedical Sciences, Vita-Salute San Raffaele University, Via Olgettina 58, 20132 Milan, IT. Johann Bernoulli Institute, University of Groningen, Nijenborgh 9, 9747 AG Groningen, NL.)  
 Contributed Talk, session [77]  
 Tuesday 3:10 – 3:40, Medical Sciences 4279  
*A stochastic logistic growth model to investigate stem cell differentiation process*  
 Coauthors: Luca Biasco (San Raffaele Gene Therapy Institute), Clelia Di Serio (Universit Vita Salute San Raffaele), Ernst Wit (Johann Bernoulli Institute. University of Groningen)

**Abstract:** In gene therapy (GT) clinical trials, individual Hematopoietic Stem/Progenitor Cells (HSPCs) gene corrected with permanently integrating viral vectors become distinctly labelled by viral insertion sites (IS). Upon infusion into the patients, the cellular progeny of engineered HSPC will carry identical IS markings. Therefore, analyzing samples from patient’s bone marrow and peripheral blood at different timepoints after GT, it is possible to quantitatively track the population of blood cells derived from each individual HSPC clone, providing a unique experimental setting to unveil clonal properties of hematopoiesis directly in vivo in humans. In this work, a statistical model for single-clone dynamics across multiple cell types is proposed. Such a model, named Stochastic Logistic Cell Differentiation Process, is defined as a  $N$ -dimensional continuous time Markov process,  $\mathbf{X}(t) = (X_1(t), \dots, X_N(t))$ , where each component corresponds to the number of cells (counts) of specific cell type  $C_i, i = 1, \dots, N$  present in the system at time  $t$ . Given initial state,  $\mathbf{X}_0$ , process evolves according to a random sequence of events grouped in 3 main categories, cell duplication, death and differentiation from  $C_i$  to  $C_j$ . Individual cells are assumed to be independent and cells belonging to the same  $C_i$  are assumed to share the same individual cell events rates, respectively  $\alpha_i, \delta_i$  and  $\lambda_{i,j=1..N}$ . Transitions over process states space obey to a set of conditional probabilities depending on events rates and polynomial functions of  $\mathbf{X}(t)$ . Starting from the Kramers-Moyal expansion of Master Equation, which describe the evolution of  $P(\mathbf{X}; t)$  over time, is possible derive a system of ordinary differential equations for process moments,  $E[\mathbf{X}^{(m)}]$ . Inference on  $(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\lambda})$  is performed by means of a constrained generalized non-linear least square procedure, involving simultaneously first and second order process’ moments, in order to take into account for the dependence structure between process components. The method has been evaluated by means of a simulation study, which showed its reliability from at both parameters estimation and model selection levels for different sampling time distances. Finally, the model has been successfully applied to GT clinical data, supporting a novel model of human hematopoiesis.

- **Yuval Peres** (Microsoft Research)

Invited session: Online algorithms in machine learning [33]

Friday 2:35 – 3:05, Medical Sciences 2171

*Towards optimal algorithms for prediction with expert advice*

**Abstract:** We study the classical problem of prediction with expert advice in the adversarial setting. Cover (1965) gave the optimal algorithm that minimizes worst-case regret for the case of 2 experts. In this talk, I will describe the optimal algorithm, adversary and regret for the case of 3 experts. We will see that optimal algorithm for 2 and 3 experts is a probability matching algorithm (analogous to Thompson sampling) against a particular randomized adversary. Remarkably, it turns out that this algorithm is not only optimal against this adversary, but also minimax optimal against all possible adversaries. We establish a constant factor separation between the regrets achieved by the optimal algorithm and the widely used multiplicative weights algorithm. The analysis of the optimal adversary relies on delicate random walk estimates. (Joint work with Nick Gravin and Balu Sivan). I will also briefly discuss the optimal randomized algorithm when it is known that at least one of the  $k$  experts makes perfect predictions.

- **Iosif Pinelis** (Michigan Technological University)

Contributed Talk, session [81]

Tuesday 2:00 – 2:30, Medical Sciences 3278

*Optimal-order bounds on the rate of convergence to normality for maximum likelihood estimators*

**Abstract:** It is well known that, under general regularity conditions, the distribution of the maximum likelihood estimator (MLE) is asymptotically normal. Very recently, bounds of the optimal order  $O(1/\sqrt{n})$  on the closeness of the distribution of the MLE to normality in the so-called bounded Wasserstein distance were obtained, where  $n$  is the sample size. However, the corresponding bounds on the Kolmogorov distance were only of the order  $O(1/n^{1/4})$ . In this note, bounds of the optimal order  $O(1/\sqrt{n})$  on the closeness of the distribution of the MLE to normality in the Kolmogorov distance are given, as well as their nonuniform counterparts, which work better for large deviations of the MLE. These results are based on previously obtained general optimal-order bounds on the rate of convergence to normality in the multivariate delta method. The crucial observation is that, under natural conditions, the MLE can be tightly enough bracketed between two smooth enough functions of the sum of independent random vectors, which makes the delta method applicable. It appears that the nonuniform bounds for MLEs in general have no precedents in the existing literature; a special case was recently treated by Pinelis and Molzon. The results can be extended to  $M$ -estimators or, even more generally, to the estimators that are zeros of estimating functions.

- **Toniann Pitassi** (Toronto)

Invited session: Privacy and Statistics [37]

Tuesday 10:45 – 11:15, Medical Sciences 4171

*Differential privacy and adaptive data analysis*

**Abstract:** We also formalize and address the general problem of data reuse in adaptive data analysis. We show how the differential-privacy based approach given in (Dwork et al., 2014) is applicable much more broadly to adaptive data analysis. We then show that a simple approach based on description length can also be used to give guarantees of statistical validity in adaptive settings. Finally, we demonstrate that these incomparable approaches can be unified via the notion of approximate max-information that we introduce.

- **Lea Popovic** (Concordia University)

Invited session: Mathematical systems biology [18]

Thursday 11:20 – 11:50, Medical Sciences 2171

*Large deviations for two time-scale models*

Coauthor: Rohini Kumar (Wayne State University)

**Abstract:** For a number of processes in biology the appropriate stochastic modelling is done in terms of multi-scale Markov processes with fully dependent slow and fast fluctuating variables. The most common examples are diffusions, jump-diffusions and state dependent Markov chains. The large deviations behaviour for these models are of particular interest in applications. We give a proof of the large deviations principle for two time-scale jump-diffusions and show its use on a very simple example of an intracellular model on two time-scales with fully dependent slow and fast variables.

- **Busayasachee Puang-ngern** (Department of Statistics, Faculty of Science and Engineering, Macquarie University, AUSTRALIA)

Contributed Talk, session [93]

Thursday 11:20 – 11:50, Rosebrugh 310

*The estimation of semiparametric generalized linear model using the Newton-MI algorithm*

Coauthors: A/Prof Jun Ma (Department of Statistics, Faculty of Science and Engineering, Macquarie University, AUSTRALIA), A/Prof Ayse Bilgin (Department of Statistics, Faculty of Science and Engineering, Macquarie University, AUSTRALIA), Timothy Kyng (Department of Applied Finance and Actuarial Studies, Faculty of Business and Economics, Macquarie University, AUSTRALIA)

**Abstract:** Generalized linear models (GLM) are the statistical model widely used in regression analysis. However, in practice, sometimes the generalized linear models could not fit the data well enough. So the semiparametric generalized linear models (SP-GLM) was proposed to be an extension of the classical generalized linear models. The SP-GLM includes a nonparametric component to the response distribution in GLM. This nonparametric component is unspecified and will be determined by the observed data, and it enables SP-GLM better fit to the data than the traditional GLM. Nevertheless, the existing method on SP-GLM has a computational limitation especially nonefficient for large sample sizes. Particularly, it cannot handle more than 1,000 observations. We propose a new method for computing SP-GLM. That is the regression coefficient and the nonparametric component of the distribution are estimated simultaneously, subject to the constraint that the nonparametric component is a proper density function. In each iteration, the Newton-Raphson algorithm is used to update the regression coefficient. With this new regression coefficient, the canonical parameter will be estimated using the Newton-Raphson algorithm. Finally, the nonparametric components is updated via the Multiplicative Iterative algorithm (MI). This process is iterated until convergence. The results from the simulation study shown that our estimated regression coefficients were closed to the existing method. Our proposed method provide effective computations and fast convergence even for large sample size data sets.

- **Amber Puha** (CSU San Marcos)

Invited session: Scaling limits of stochastic networks [14]

Tuesday 2:00 – 2:30, Rosebrugh 208

*Analysis of processor sharing queues via relative entropy*

Coauthor: Ruth J. Williams (UCSD)

**Abstract:** Processor sharing is a mathematical idealization of round-robin scheduling algorithms commonly used in computer time-sharing. It is a fundamental example of a non-head-of-the-line service discipline. For such disciplines, it is typical that any Markov description of the system state is infinite dimensional. Due to this, measure-valued stochastic processes are becoming a key tool used in the modeling and analysis of stochastic network models operating under various non-head-of-the-line service disciplines.

In this talk, we discuss a new approach to studying the asymptotic behavior of fluid model solutions (formal functional law of large numbers limits) for critically loaded processor sharing queues. For

this, we introduce a notion of relative entropy associated with measure-valued fluid model solutions. This approach is developed with idea that similar notions involving relative entropy may be helpful for understanding the asymptotic behavior of critical fluid model solutions for stochastic networks operating under protocols naturally described by measure-valued processes.

**Q**

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Chengwei Qin** (McMaster University)

Contributed Talk, session [57]

Wednesday 11:20 – 11:50, Rosebrugh 208

*On the first passage time of degradation processes: a perspective from Birnbaum-Saunders distribution*

Coauthor: Narayanaswamy Balakrishnan (McMaster University)

**Abstract:** The Birnbaum-Saunders (BS) distribution was originally derived as a first-passage-time (FPT) model in reliability engineering based on central limit theorem. Our paper addresses a new scope to explain how does BS distribution approximate the FPT of degradation processes by constructing the intrinsic connections between them. To evaluate the approximation behaviours of BS distribution, we systematically investigate the properties of mean and variance-time-to-failure, e.g. multiple comparisons with other first-passage approximations including renewal process and Lévy Inverse subordinators in theoretical, numerical and visual ways, and analysis of the asymptotic joint distribution as well as information matrices conditioned on observing the failure times or degradation paths.

Keywords: Birnbaum-Saunders Distribution, First-Passage-Time, Renewal Process, Lévy Subordinator, Laplace Inversion, Mean-Time-To-Failure, Variance-Time-To-Failure, Censored Observations, Asymptotic Behaviours, Fisher Information.

- **Jeremy Quastel** (University of Toronto)

Invited session: SPDEs and the work of Martin Hairer [1]

Wednesday 10:45 – 11:15, Medical Sciences 3153

*Exact formulas in random growth*

**Abstract:** We describe progress on exact formulas in the KPZ universality class

**R**

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Olivier Raimond** (Université Paris Ouest Nanterre la Défense)

Contributed session: Reinforced processes and related fields [44]

Thursday 2:35 – 3:05, Medical Sciences 3154

*Vertex reinforced non-backtracking random walks*

**Abstract:** We study Vertex-Reinforced-Random-Walk (VRRW) on a complete graph with weights of the form  $w(n) = n^\alpha$ , with  $\alpha > 1$ . Unlike for the Edge-Reinforced-Random-Walk, which in this case localizes a.s. on 2 sites. Here we observe various phase transitions, and in particular

localization on a set of size  $K$  is possible, provided  $K < \frac{2\alpha-1}{\alpha-1}$ . Our proof relies on stochastic approximation techniques. We also study Vertex Reinforced Non-Backtracking Random walks, but now the localization on a set of size  $K$  is possible, provided  $K < \frac{3\alpha-1}{\alpha-1}$ . These are joint works with M. Benaim, B. Schapira and L. Le Goff.

- **Alexander Rakhlin** (University of Pennsylvania)

Invited session: Aggregation of estimators [25]

Monday 2:00 – 2:30, Medical Sciences 2171

*An aggregation procedure for nonparametric regression with convex and non-convex classes*

**Abstract:** Exact oracle inequalities for regression have been extensively studied in statistics and learning theory over the past decade. In the case of a misspecified model, the focus has been on either parametric or convex classes. We present a new estimator that steps outside of the model in the non-convex case and reduces to least squares in the convex case. To analyze the estimator for general non-parametric classes, we prove a generalized Pythagorean theorem and study the supremum of a negative-mean stochastic process (which we term the offset Rademacher complexity) via the chaining technique.

(based on work with T. Liang, K. Sridharan, and A. Tsybakov)

- **Alexander Rakhlin** (University of Pennsylvania)

Invited session: Computational versus statistical complexity in high-dimensional problems [24]

Wednesday 11:55 – 12:25, McLeod Auditorium

*How to predict when estimation is hard: algorithms for learning on graphs*

**Abstract:** We consider the problem of predicting a binary label for an individual given the information about the person and her position within a network. Such a fusion of the two sources of information naturally arises in a variety of applications, including recommendation systems, ad placement, and personalized medical care.

When formalizing this problem, one faces a computationally intractable combinatorial objective. We present an unexpected phenomenon: it is possible to develop poly-time (and statistically near-optimal) online prediction methods even when the offline problem is provably hard. These prediction methods arise in a systematic way from a new relaxation framework that has roots in the work of Cover and Blackwell. The new applications to learning on graphs are based in part on multiplicative approximation algorithms developed by the theoretical computer science community. Our approach naturally extends to the contextual multi-armed bandit setting with large sets of policies – a notoriously difficult problem, which is often encountered in real-world applications.

Joint work with K. Sridharan.

- **Kavita Ramanan** (Brown University)

Contributed session: Analysis and control of stochastic network models [42]

Wednesday 11:20 – 11:50, Medical Sciences 4279

*Interacting measure-valued processes arising in stochastic networks*

Coauthors: Reza Aghajani (Brown University), Pooja Agarwal (Brown University)

**Abstract:** We introduce a class of interacting measure-valued processes that arise as models of

large-scale stochastic networks with general service distributions operating under certain randomized routing protocols. We establish hydrodynamic limit theorems for such networks, characterize the equilibrium behavior of the hydrodynamic limits and describe some surprising properties that they reveal about network performance. We also describe how the mathematical framework introduced above may be useful for analyzing certain models arising in materials science and biology.

- **Gourab Ray** (University of Cambridge)

Invited session: Random planar maps [7]

Tuesday 3:10 – 3:40, McLeod Auditorium

*A uniformization theorem for random planar maps*

**Abstract:** We show that for infinite planar unimodular random rooted maps, many global geometric and probabilistic properties are equivalent and determined by a natural notion of average curvature, a local quantity. This dichotomy includes properties relating to amenability, conformal geometry, random walks, uniform and minimal spanning forests, and Bernoulli bond percolation. We also prove that every simply connected unimodular random rooted map is sofic, that is, a Benjamini-Schramm limit of finite maps. Joint work with Omer Angel, Tom Hutchcroft and Asaf Nachmias

- **Kolyan Ray** (Leiden University)

Contributed Talk, session [87]

Wednesday 11:20 – 11:50, Medical Sciences 2171

*Minimax theory for a class of non-linear statistical inverse problems*

Coauthor: Johannes Schmidt-Hieber (Leiden University)

**Abstract:** We study minimax estimation rates for a class of non-linear statistical inverse problems motivated by various applications, including density estimation, binary regression and spectral density estimation. We derive matching upper and lower bounds for pointwise loss using function-dependent rates that capture spatial heterogeneity of the function. The upper bound is obtained via a two-step procedure, where the first step inverts the non-linearity and is based on wavelet thresholding. The procedure is fully adaptive, both spatially and over modified Holder smoothness classes that are natural in this setting.

- **Gesine Reinert** (University of Oxford)

Invited session: Limit Theorems and the Malliavin calculus [11]

Tuesday 2:35 – 3:05, Medical Sciences 2171

*Chisquare approximations with Steins method*

Coauthors: Robert Gaunt (University of Oxford), Alastair Pickett (Chenavari Investment Managers)

**Abstract:** Chisquare approximations are central in theoretical statistics; prominent examples are Pearson's chisquare statistic Friedman's chisquare statistic. In this talk we shall derive bounds on the distance to chisquare for functions of possibly dependent random variables. Pearson's chisquare statistic Friedman's chisquare statistic are used as examples. The results will be related to Stein's method on the second Wiener chaos.

- **Franz Rembart** (Department of Statistics, University of Oxford)  
 Contributed Talk, session [60]  
 Thursday 10:45 – 11:15, Medical Sciences 3154  
*Recursive construction of CRTs and a binary embedding of the stable tree*  
 Coauthor: Matthias Winkel (winkel@stats.ox.ac.uk)  
**Abstract:** We introduce a general recursive method to construct continuum random trees (CRTs) from i.i.d. copies of a string of beads, that is, any random interval equipped with a random discrete measure. We prove the existence of these CRTs as a new application of the fixpoint method formalised in high generality by Aldous and Bandyopadhyay. We apply this recursive method to “embed” Duquesne and Le Gall’s stable tree into a binary compact CRT in a way that solves an open problem posed by Goldschmidt and Haas. We also express this binary CRT as a tree built by replacing all branch points of a stable tree by rescaled i.i.d. copies of a Ford CRT. Some of these developments are carried out in a space of  $\infty$ -marked metric spaces generalising Miermont’s notion of a  $k$ -marked metric space.
- **Yosef Rinott** (The Hebrew University and LUISS Rome)  
 Invited session: Graphical models and totally positive distributions [22]  
 Monday 10:45 – 11:15, Medical Sciences 2172  
*Total positivity and applications*  
**Abstract:** I will describe the notion of Totally Positive matrices and kernels and some of its history, and show some relations to probability, stochastic processes, and inequalities. The notion of Multivariate Total Positivity of order 2 and its relation to positive dependence of random variables and various properties will be discussed, along with the relation to graphical models.
- **Eliane R. Rodrigues** (Instituto de Matematicas - Universidad Nacional Autonoma de Mexico)  
 Contributed Talk, session [57]  
 Wednesday 10:45 – 11:15, Rosebrugh 208  
*Using a non-homogeneous Poisson model with spatial anisotropy to study ozone air pollution*  
**Abstract:** We consider a non-homogeneous Poisson model to estimate the probability that a pollutant’s concentration surpasses a given environmental threshold a certain number of times in a time interval of interest. We impose a spatial dependence on the parameters of the Poisson intensity function. That dependence is to account for the possible correlation between measurements in different monitoring sites. An anisotropic model is used. Estimation of the parameters of the model is performed using the Bayesian point of view via Markov chain Monte Carlo (MCMC) algorithms. We also consider prediction of the days in which exceedances of the threshold might occur at sites where measurements cannot be taken. This is obtained by spatial interpolation using the information provided by the sites where measurements are available. This procedure allows for estimation of the behaviour of the mean function of the non-homogeneous Poisson process associated with those sites. The models considered here are applied to ozone data obtained from the monitoring network of Mexico City. This is a joint work with Dani Gamerman, Mario H. Tarumoto and Guadalupe Tzintzun.

- **Judith Rousseau** (CEREMADE, Université Paris Dauphine & CREST-ENSAE)

Plenary talk: Ethel Newbold Prize Lecture

Monday 5:10 – 6:00, McLeod Auditorium

*On the Bayesian measures of uncertainty in infinite dimensional models*

Coauthors: Ismael Castillo (Université Paris 6), Botond Szabo (Leiden University)

**Abstract:** Over the last 2 decades there have been many advances in the study of frequentist properties of Bayesian approaches in nonparametric or high dimensional models. Following the seminal papers of Ghosal, Ghosh and van der Vaart 2000 on posterior concentration rates, there has been a large literature on posterior concentration rates in various families of sampling and prior models. Recently, more refined properties of the posterior distribution have also been investigated, leading to a better understanding of the frequentist properties of Bayesian measures of uncertainty such as credible regions.

In this talk, I will first present the general ideas behind posterior concentration rates. Then I will describe more precisely the recent advances behind the understanding of Bayesian measures of uncertainty. In particular I will present (1) the Bernstein von Mises theorems obtained in semi and nonparametric models and (2) some general conditions to construct credible regions having optimal frequentist properties.

S

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Qurat-Ul-An Sabir** (Dalhousie University, Halifax, Nova Scotia.)

Contributed Talk, session [78]

Thursday 11:20 – 11:50, Fitzgerald 139

*Using correspondence analysis and log-linear models in the study of cardiovascular disease.*

Coauthor: Tri Nguyen-Quang (Dalhousie University, Halifax, Nova Scotia.)

**Abstract:** Cardiovascular disease (CVD) is the second foremost reason of mortality within the world (WHO, 2012). This issue has been seriously alarmed by governments of developing countries (WHO, 2012). In the same token, diseases related to the heart are increasing day by day in Pakistan, a third world country, and play a role as a highest risk for human health. There are many factors contributing to the development of these diseases including poor diet, sedentary life style, high blood pressure and hypertension etc.

In this paper, we present a study of correspondence analysis and log-linear models to deal with the prediction of CVD. In order to explore the risk pattern via different factors that could be measurable for CVD, a survey has been conducted amongst affected people of different age groups, gender, and various education levels. Based on these data, we could determine which group would be at higher risk conducting to the CVD and make a conclusion on our findings. It should be noted that all participants were suffering from CVD either slightly or seriously.

Our interesting findings are that women could get a higher risk than men facing CVD. Moreover, different factors such as smoking, high cholesterol level and physical inactivity and poor diet

contribute significantly to the risk of this disease. From these findings, we also can obtain a better understanding of the data structure and better interpreting the results by combining the both approaches correspondence analysis and log-linear models. It can be concluded that correspondence analysis seems the best approach due to the fact that we can find the strong correlations between variables. That could lead to the conception of models using the inter-correlations between variables and building a good structure of data in the future.

- **Lyudmila Sakhanenko** (Michigan State University)

Contributed Talk, session [96]

Monday 11:55 – 12:25, Fitzgerald 139

*Estimation of integral curves from high angular resolution diffusion imaging (HARDI) data*

Coauthor: Owen Carmichael (Pennington Biomedical Research Center)

**Abstract:** We develop statistical methodology for a popular brain imaging technique HARDI based on the high order tensor model by Özarslan and Mareci. We investigate how uncertainty in the imaging procedure propagates through all levels of the model: signals, tensor fields, vector fields, and fibers. We construct asymptotically normal estimators of the integral curves or fibers which allow us to trace the fibers together with confidence ellipsoids. The procedure is computationally intense as it blends linear algebra concepts from high order tensors with asymptotical statistical analysis. The theoretical results are based on proving that certain empirical processes converge to a Gaussian process. We illustrate our procedure on simulated and real datasets.

This work generalizes the statistical methodology proposed for low angular resolution diffusion tensor imaging by Carmichael and Sakhanenko, to several fibers per voxel. It is also a pioneering statistical work on tractography from HARDI data. It avoids all the typical limitations of the deterministic tractography methods and it delivers the same information as probabilistic tractography methods. Our method is computationally cheap and it provides well-founded mathematical and statistical framework where diverse functionals on fibers, directions and tensors can be studied in a systematic and rigorous way.

- **Michele Salvi** (Paris-Dauphine)

Contributed Talk, session [59]

Wednesday 11:20 – 11:50, Medical Sciences 3278

*Law of large numbers for the variable range hopping model*

Coauthors: Alessandra Faggionato (La Sapienza, Roma), Nina Gantert (TU Munich)

**Abstract:** The Mott Variable Range Hopping model is considered in Physics as an accurate representation of electrical conduction in semiconductors. From the mathematical point of view, it represents a prominent example of reversible long-range random walks on random point processes, which generalize in several ways the classical random conductance model on the lattice. We ask ourselves how an external field influences the limiting velocity of the walk: So far, only very few models of biased random walks with trapping mechanisms have been rigorously studied. We will also give a precise control of the invariant measure for the process from the point of view of the particle, fundamental step towards the proof of an Einstein Relation for this model.

- **Colin Sandon** (Princeton University)

Invited session: Community detection in random graphs [26]

Thursday 10:45 – 11:15, Medical Sciences 2172

*Community detection in the stochastic block model down to and below the KS threshold*

Coauthor: Emmanuel Abbe (Princeton University)

**Abstract:** This talk describes our recent results on achieving the Kesten-Stigum (KS) threshold in the general stochastic block model with a linearized acyclic belief propagation algorithm (ABP) in quasi linear time. This proves a conjecture of Decelle et al. while extending the result to non symmetrical SBMs. ABP is further connected to a spectral algorithm on a generalized nonbacktracking operator. It is also mentioned how to cross the KS threshold with information theory when the number of communities is large enough.

- **Sriram Sankararaman** (UCLA)

Invited session: Biomedical applications [20]

Wednesday 2:35 – 3:05, Medical Sciences 3154

*Latent variable models for inferring human population admixture*

**Abstract:** Genomic analyses have revealed that admixture, a process in which populations mix and exchange genes, has been a common process through human evolution. Understanding the genomic ancestry of an admixed individual is an important step in genomic medicine, social science and evolutionary biology. To this end, we need to develop statistical models that are detailed enough to capture the biological signal while permitting efficient inference that can enable them to leverage large-scale genomic datasets.

In this talk, I will describe two inferential problems that arise in the analysis of admixed genomes and describe how latent variable models provide a natural framework for inference. The first, local ancestry inference, aims to infer the ancestry at every position along an individual's genome. I will describe a two-level factorial hidden Markov model for local ancestry inference that captures the fine-scale correlation structure across these genomes due to population genetic processes. As a result, this model is shown to be highly accurate while the nested structure enables it to scale to genomic datasets. The second problem is focussed on inferring the genomic ancestry of the parents of an admixed individual. Here, I will introduce the framework of pooled semi-Markov processes and describe efficient inference algorithms in this framework that enable us to accurately infer the genomic ancestries of parents from their offspring's genome.

Finally, I will show how these models are being applied to genomes of Mexican and Puerto Rican individuals (that are an admixture of African, European and Native American ancestral populations) to learn about genetics of complex diseases as well as to elucidate human history and mating patterns.

- **Emmanuel Schertzer** (UPMC)

Invited session: Stochastic models of evolution [16]

Monday 2:35 – 3:05, Medical Sciences 3154

*Limit theorems for the partitioning process*

Coauthors: Amaury Lambert (UPMC), Veronica Miro Pina (UPMC)

**Abstract:** In this talk, I will consider a haploid Wright–Fisher model with recombination, where each haplotype is a mosaic of its two parental chromosomes. Starting with uniformly colored and distinct chromosomes, each individual of the population at time  $t$  is a composite (or partition) of the colors originally present in the ancestral population. The partitioning process at time  $t$  is then defined as the color partition of a sampled chromosome. In particular, as time goes to infinity, it provides a description of the haplotype that will eventually fix in the population.

I will present some recent results on the partitioning process at stationarity. In particular, I will discuss the description of a typical color cluster, and a law of large numbers for the number of clusters inside a large portion of the chromosome.

- **Jang Schiltz** (University of Luxemburg, LSF)

Contributed Talk, session [83]

Monday 2:00 – 2:30, Fitzgerald 139

*Model selection in generalized finite mixture models*

**Abstract:** We present a generalization of Nagin’s finite mixture model that allows non parallel trajectories for different values of covariates. We investigate some mathematical properties of this model and illustrate its use by giving typical salary curves for the employees in the private sector in Luxembourg between 1981 and 2006, as a function of their gender, as well as of Luxembourg’s gross domestic product (GDP). Then we suggest a new criterion for model selection in finite mixture models that overcomes some of the shortcomings of existing selection methods.

- **Volker Schmidt** (Ulm University, Institute of Stochastics)

Contributed Talk, session [71]

Monday 11:20 – 11:50, Rosebrugh 310

*Stochastic 3D modeling of amorphous microstructures – a powerful tool for virtual materials testing*

**Abstract:** In this talk, we are going to introduce the concept of stochastic 3D modeling of geometrically complex (disordered) microstructures as a tool for virtual materials testing. In particular, we use models of stochastic geometry, e.g., random (marked) point processes, random graphs and random tessellations, in order to describe the 3D microstructure of particle-based materials [1, 2]. This includes applications to Lithium-ion batteries [3] as well as to fuel cells [4] and solar cells [5]. Using stochastic 3D models, one can generate a large variety of stochastically simulated microstructures with little computational effort. These virtual microstructures can be used as data basis to elucidate microstructure-property relationships. In this way, for example, effective conductivity can be expressed by microstructural characteristics, like volume fraction, tortuosity (windedness of transport paths) and constrictivity (bottleneck criterion) of the considered material phase [6].

[1] V. Schmidt, G. Gaiselmann and O. Stenzel, Stochastic 3D models for the micro-structure of advanced functional materials. In: V. Schmidt (ed.) Stochastic Geometry, Spatial Statistics and Random Fields: Models and Algorithms, Lecture Notes in Mathematics, vol. 2120, Springer, Cham 2015, 95-141.

- [2] A. Spettl, T. Brereton, Q. Duan, T. Werz, C.E. Krill, D.P. Kroese and V. Schmidt, Fitting Laguerre tessellation approximations to tomographic image data. *Philosophical Magazine* 96 (2016), 166-189.
- [3] J. Feinauer, T. Brereton, A. Spettl, M. Weber, I. Manke and V. Schmidt, Stochastic 3D modeling of the microstructure of lithium-ion battery anodes via Gaussian random fields on the sphere. *Computational Materials Science* 109 (2015), 137-146.
- [4] M. Neumann, J. Stanek, O. Pecho, L. Holzer, V. Benes and V. Schmidt, Stochastic 3D modeling of complex three-phase microstructures in SOFC-electrodes with completely connected phases. *Computational Materials Science* 118 (2016), 353-364.
- [5] O. Stenzel, C. Hirsch, V. Schmidt, T.J. Brereton, D.P. Kroese, B. Baumeier and D. Andrienko, A general framework for consistent estimation of charge transport properties via random walks in random environments. *SIAM: Multiscale Modeling and Simulation* 12 (2014), 1108-1134.
- [6] O. Stenzel, O.M. Pecho, L. Holzer, M. Neumann and V. Schmidt, Predicting effective conductivities based on geometric microstructure characteristics. *AIChE Journal* 62 (2016), 1834-1843.

- **Alexander Schnurr** (University Siegen)

Contributed Talk, session [95]

Tuesday 3:10 – 3:40, Medical Sciences 4171

*Detecting changes in the (non-linear) dependence structure between time series*

Coauthor: Herold Dehling (Ruhr University Bochum)

**Abstract:** Using so called ordinal patterns we analyze whether there is a (possibly non-linear) dependence between given time series. If we find a dependence structure, we tackle the question whether this dependence changes over time. Considering short-range or long-range dependent time-series, different techniques are used and different limit theorems are obtained. Applications include time series appearing in mathematical finance, biology, medicine and hydrology.

References:

- (1) Testing for Structural Breaks via Ordinal Pattern Dependence (with Herold Dehling). To appear in *JASA* (2016+).
- (2) An Ordinal Pattern Approach to Detect and to Model Leverage Effects and Dependence Structures Between Financial Time Series. *Stat. Papers* 55(4) (2014), 919–931.

- **Dominic Schuhmacher** (University of Goettingen)

Contributed Talk, session [69]

Friday 11:20 – 11:50, Medical Sciences 3278

*Convergence rate for the degree distribution in a dynamic network model*

Coauthor: Fabian Kueck (University of Goettingen)

**Abstract:** In the dynamic network model of Britton and Lindholm [Journal of Statistical Physics 139, No. 3, pp. 518–535, 2010], the number of nodes evolves according to a supercritical linear birth-and-death process. At the birth of a node, a random parameter is assigned that controls the

rate at which links to other nodes are established. Links break if one of their nodes dies or after an independent exponentially distributed time.

We derive a rate of convergence in total variation distance for the degree distribution of a node picked uniformly at random at time  $T$  to a mixed Poisson distribution as  $T \rightarrow \infty$ . Main ingredients of the proof are Stein's method and a new result about the age distribution of a node picked uniformly at random.

- **Rama Shanker** (Department of Statistics, Eritrea Institute of Technology, Asmara, Eritrea)

Contributed Talk, session [74]

Tuesday 11:55 – 12:25, Fitzgerald 139

*A quasi Akash distribution and its applications*

**Abstract:** In this paper, a quasi Akash distribution (QAD), of which Akash distribution introduced by Shanker (2015) is a particular case, has been proposed. Its various interesting mathematical and statistical properties including shape, moments, coefficient of variation, skewness, kurtosis, Index of dispersion, hazard rate function, mean residual life function, stochastic orderings, mean deviations, order statistics, Bonferroni and Lorenz curves, Renyi entropy measure and stress-strength reliability have been obtained and discussed. The estimation the parameters of QAD has been discussed using maximum likelihood estimation and method of moments. The QAD has been fitted to a number of data-sets from biomedical science and engineering and the goodness of fit has been compared with one parameter lifetime distributions Akash, Lindley and exponential and other two parameter lifetime distributions and the fir is found to be quite satisfactory. Therefore, QAD can be considered as an important lifetime distribution for modeling lifetime data.

Keywords: Akash distribution, Moments, Hazard rate and mean residual life functions, Stochastic orderings, Mean deviations, Order statistics, Bonferroni and Lorenz curves, Renyi entropy measure, Stress-strength reliability, Estimation of Parameters, Goodness of fit.

- **James Sharpnack** (UC Davis, Statistics Department)

Invited session: Statistical computing and complex data [36]

Monday 3:10 – 3:40, McLeod Auditorium

*Locally adaptive smoothing on graphs*

Coauthor: Ryan Tibshirani (Carnegie Mellon University)

**Abstract:** Signal processing on graphs is a framework for non-parametric function estimation that generalizes spatial signal processing to heterogeneous domains. Recently, it has been determined that the fused lasso is theoretically competitive with wavelet thresholding, which has been celebrated as the benchmark for locally adaptive smoothing procedures. We propose a generalization of the fused lasso, called graph trend filtering, which not only extends spatio-temporal structure to general graph structure, but also to higher-order smoothness. We demonstrate the flexibility of this framework, discuss fast solvers, and recent advances in our theoretical understanding.

- **Scott Sheffield** (MIT)

Plenary talk: Doob Lecture

Wednesday 5:10 – 6:00, McLeod Auditorium

*Random surfaces and gauge theory*

**Abstract:** I will discuss several new results (many joint with Jason Miller) related to two prominent models for random surfaces: Liouville quantum gravity and the Brownian map. I will also discuss the relationship between these subjects and the sums over random surfaces that appear in lattice gauge theory.

- **Hao Shen** (Columbia University)

Invited session: SPDEs and the work of Martin Hairer [1]

Wednesday 11:20 – 11:50, Medical Sciences 3153

*Stochastic quantization equations and the theory of regularity structures*

**Abstract:** Stochastic quantization is a formal procedure in which a quantum field theory measure is viewed as the stationary measure of a stochastic PDE. The examples include such equations associated with the  $\Phi_3^4$  model, the sine-Gordon model, random loops in a manifold, and so on. Hairer's theory of regularity structures provides rigorous meanings to these extremely singular equations via systematic renormalization procedures. We will review the ideas of Hairer's theory, and explain the algebraic and analytic aspects in the renormalization method recently developed by Hairer and his collaborators.

- **José Eliud Silva Urrutia** (Universidad Anahuac México)

Contributed Talk, session [73]

Wednesday 3:10 – 3:40, Fitzgerald 139

*Measuring the induced smoothness in the application of the Heligman and Pollard (HP) model with a smoothness index*

Coauthor: Adrian Ovin (Universidad Anahuac México)

**Abstract:** The purpose of this research is measuring the induced smoothness of the estimated mortality curve coming from the application of the Heligman and Pollard (HP) model. With this, we make an approximation to this model using a non-parametric perspective proposal, where instead of estimating the eight parameters, provide initial values for each of them and use nonlinear optimization, a similar mortality curve is estimated with a smoothness index controlled by the analyst. Additionally, not only a point estimate is provided, but also an interval estimate. The main idea is to smooth the observed data in such way that its trend is as closely as possible as to the curve produced by the HP model through the use of loss functions. To illustrate the measurements, 43 estimated and published mortality series were used. We infer the value of the induced smoothness and therefore the smoothing parameters that approximates to such trend. In the examples, we focused on age 12 and older, where the results of the estimates are presented and compared. A better fit of the male over female mortality is observed and this implies different values of the smoothing index and the smoothing parameters obtained.

- **Florian Simatos** (ISAE SUPAERO)

Invited session: Scaling limits of stochastic networks [14]

Tuesday 2:35 – 3:05, Rosebrugh 208

*Scaling limits for a limit order book model using the regenerative characterization of Lévy trees*

Coauthors: Peter Lakner (Leonard N. Stern School of Business, New York University), Josh Reed (Leonard N. Stern School of Business, New York University)

**Abstract:** We consider the following Markovian dynamic on point processes: at constant rate and with equal probability, either the rightmost atom of the current configuration is removed, or a new atom is added at a random distance from the rightmost atom. Interpreting atoms as limit buy orders, this process was introduced by Lakner et al. to model a one-sided limit order book.

We consider this model in the regime where the total number of orders converges to a reflected Brownian motion, and complement the results of Lakner et al. by showing that, in the case where the mean displacement at which a new order is added is positive, the measure-valued process describing the whole limit order book converges to a simple functional of this reflected Brownian motion.

Our approach leverages an unexpected connection with Lévy trees. More precisely, the cornerstone of our approach is the regenerative characterization of Lévy trees due to Weill, which provides an elegant proof strategy which we unfold.

- **David Sivakoff** (The Ohio State University)

Contributed Talk, session [62]

Wednesday 2:35 – 3:05, Medical Sciences 4279

*Sharp and gradual phase transitions in bootstrap percolation*

**Abstract:** We consider bootstrap percolation on graphs with local and non-local connectivity, and observe a range of behavior in the types of phase transitions possible. We observe sharp, gradual and hybrid phase transitions depending on the relative scaling of local versus non-local connectivity. Furthermore, in the sharp scaling regime, we observe two distinct phase transitions with the same order parameter. Based on joint works with Janko Gravner.

- **Erik Slivken** (University of California Davis)

Contributed Talk, session [61]

Tuesday 10:45 – 11:15, Rosebrugh 208

*Local limit of the fixed point forest*

Coauthors: Tobias Johnson (NYU), Anne Schilling (University of California Davis)

**Abstract:** Consider the following partial “sorting algorithm” on permutations: take the first entry of the permutation in one-line notation and insert it into the slot of its own value. Continue until the first entry is 1. This process imposes a forest structure on the set of all permutations of size  $n$ , where the roots are the permutations starting with 1 and the leaves are derangements. Viewing the process in the opposite direction towards the leaves, one picks a fixed point and moves it to the beginning. Despite its simplicity, this “fixed point forest” exhibits a rich structure. We consider the fixed point forest in the limit  $n \rightarrow \infty$  and show using Stein’s method that at a random permutation the local structure weakly converges to a tree in terms of independent Poisson point processes. We also show that the distribution of the length of the longest path to a leaf converges to the geometric distribution with mean  $e - 1$ , and the length of the shortest path converges to the Poisson distribution with mean 1.

- **Allan Sly** (University of California, Berkeley)

Invited session: Random media [5]

Monday 11:55 – 12:25, Medical Sciences 3153

*First passage percolation on rotationally invariant fields*

Coauthors: Riddhipratim Basu (Stanford University), Vladas Sidoravicius (NYU Shanghai)

**Abstract:** I will discuss new results on continuum models of first passage percolation which are rotationally invariant. For such models we give a multi-scale argument showing that the variance grows as  $O(n^{1-\epsilon})$ .

- **Yongsheng Song** (Academy of Mathematics and Systems Science, CAS)

Invited session: BSDEs and their applications in finance [19]

Friday 11:20 – 11:50, Medical Sciences 3154

*Backward SDEs driven by G-Brownian motion*

**Abstract:** In this paper, we study the backward stochastic differential equations driven by a  $G$ -Brownian motion  $(B_t)_{t \geq 0}$  in the following form:

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds + \int_t^T g(s, Y_s) ds - \int_t^T Z_s dB_s - (K_T - K_t),$$

where  $K$  is a decreasing  $G$ -martingale. Under Lipschitz conditions of  $f$  and  $g$  in  $Y$  and  $Z$ , the existence and uniqueness of the solution  $(Y, Z, K)$  of the above BSDE in the  $G$ -framework is proved.

- **Michael Sørensen** (Department of Mathematical Sciences, University of Copenhagen)

Contributed Talk, session [90]

Tuesday 10:45 – 11:15, Rosebrugh 310

*Likelihood estimation for stochastic differential equations with random effects*

**Abstract:** bbs sampler that uses the method for simulation of diffusion bridges developed in Bladt and Sørensen (2014) and Finch, Bladt and Sørensen (2016). The method is illustrated by an application to an Ornstein-Uhlenbeck process with random effects in the drift and in the diffusion coefficient. An EM-algorithm is briefly considered. The lecture is based on joint work with Mogens Bladt and Fernando Baltazar Larios.

References:

- Bladt, M. and Sørensen, M. (2014): Simple simulation of diffusion bridges with application to likelihood inference for diffusions. *Bernoulli*, 20, 645-675.
- Bladt, M., Finch, S. and Sørensen, M. (2016): Simulation of multivariate diffusion bridges. *J. Roy. Statist. Soc., Ser. B*, 78, 343-369.

- **Peter Spreij** (Universiteit van Amsterdam)

Contributed Talk, session [53]

Monday 3:10 – 3:40, Medical Sciences 4279

*Large deviations of Markov-modulated diffusion processes with rapid switching*

Coauthors: Gang Huang (Universiteit van Amsterdam), Michel Mandjes (Universiteit van Amsterdam)

**Abstract:** We study small noise asymptotics of Markov-modulated diffusion processes in the regime that the modulating Markov chain is rapidly switching. We prove the joint sample-path large deviations principle for the Markov-modulated diffusion process and the occupation measure of the Markov chain (which evidently also yields the large deviations principle for each of them separately by applying the contraction principle). The structure of the proof is such that we first prove exponential tightness, and then establish a local large deviations principle (where the latter part is split into proving the corresponding upper bound and lower bound).

- **Karthik Sridharan** (Cornell University)

Invited session: Online algorithms in machine learning [33]

Friday 3:10 – 3:40, Medical Sciences 2171

*On the equivalence of martingale tail bounds and deterministic regret bounds*

Coauthor: Alexander Rakhlin (University of Pennsylvania)

**Abstract:** We will see an equivalence of (i) deterministic pathwise statements appearing in the online learning literature (termed “regret bounds”), (ii) high-probability tail bounds for the supremum of a collection of martingales (of a specific form arising from uniform laws of large numbers for martingales), and (iii) in-expectation bounds for the supremum.

This equivalence provides us a two way street. First, by virtue of the equivalence, we will see how one can use deterministic regret bounds for the online mirror descent algorithm with an adaptive step size to easily prove exponential tail bounds for norms of Banach space valued martingales. Second, using the rich literature of tail bounds for martingales, (especially matrix valued) we will be able to conclude existence of new online learning algorithms.

I will also briefly talk about extending these results beyond the linear structure of the Banach spaces. Specifically, we define a notion of martingale type for general classes of real-valued functions and show its equivalence (up to a logarithmic factor) to various sequential complexities of the class (in particular, the sequential Rademacher complexity and its offset version). For classes with the general martingale type 2, we exhibit a finer notion of variation that allows partial adaptation to the function indexing the martingale.

- **Thomas Steinke** (Harvard University)

Invited session: Privacy and Statistics [37]

Tuesday 11:55 – 12:25, Medical Sciences 4171

*Robust traceability from trace amounts*

Coauthors: Cynthia Dwork (Microsoft), Adam Smith (Pennsylvania State University), Jonathan Ullman (Northeastern University), Salil Vadhan (Harvard University)

**Abstract:** The privacy risks inherent in the release of a large number of summary statistics were illustrated by Homer et al. (PLoS Genetics, 2008), who considered the case of 1-way marginals of SNP allele frequencies obtained in a genome-wide association study: Given a large number of minor allele frequencies from a case group of individuals diagnosed with a particular disease, together

with the genomic data of a single target individual and statistics from a sizable reference dataset independently drawn from the same population, an attacker can determine with high confidence whether or not the target is in the case group.

In this work we describe and analyze a simple attack that succeeds even if the summary statistics are significantly distorted, whether due to measurement error or noise intentionally introduced to protect privacy.

Our attack only requires that the vector of distorted summary statistics is close to the vector of true marginals in  $\ell_1$  norm. Moreover, the reference pool required by previous attacks can be replaced by a single sample drawn from the underlying population.

The new attack, which is not specific to genomics and which handles Gaussian as well as Bernoulli data, significantly generalizes recent lower bounds on the noise needed to ensure differential privacy (Bun, Ullman, and Vadhan, STOC 2014; Steinke and Ullman, 2015), obviating the need for the attacker to control the exact distribution of the data.

- **Weijie Su** (The University of Pennsylvania)

Invited session: Privacy and Statistics [37]

Tuesday 11:20 – 11:50, Medical Sciences 4171

*Private false discovery rate control*

Coauthors: Cynthia Dwork (Microsoft Research), Li Zhang (Google Inc.)

**Abstract:** We provide the first differentially private algorithms for controlling the false discovery rate (FDR) in multiple hypothesis testing, with essentially no loss in power under certain conditions. Our general approach is to adapt a well-known variant of the famous Benjamini-Hochberg procedure (BHq), making each step differentially private. This destroys the classical proof of FDR control. To prove FDR control of our method we

- (1) Develop a new proof of the original (non-private) BHq algorithm and its robust variants – a proof requiring milder assumptions on the dependence structure of  $p$ -values than those previously shown in the vast literature on this topic – and
- (2) Relate the FDR control properties of the differentially private version to the control properties of the non-private version.

We also present a low-distortion “one-shot” differentially private primitive for “top  $k$ ” problems, *e.g.*, “Which are the  $k$  most popular hobbies?” (which we apply to: “Which hypotheses have the  $k$  most significant  $p$ -values?”), and use it to get a faster privacy-preserving instantiation of our general approach at little cost in accuracy. The proof of privacy for the one-shot top  $k$  algorithm introduces a new technique of independent interest.

- **Eliran Subag** (Weizmann Institute)

Invited session: Spin Glasses [4]

Monday 3:10 – 3:40, Medical Sciences 3153

*Critical points and the Gibbs measure of pure spherical spin glasses*

Coauthor: Ofer Zeitouni (Weizmann Institute)

**Abstract:** Recently, sharp results concerning the critical points of the energy landscape of pure

$p$ -spin spherical spin glasses have been obtained by means of moment computations and a proof of a certain invariance property. I will describe those and explain how they can be boosted by an investigation of the behavior around the critical points to obtain a rather complete geometric description for the Gibbs measure at low enough temperature. The latter allows, in particular, to conclude the absence of temperature chaos and calculate the free energy asymptotics with high accuracy. The talk is based on joint work with Ofer Zeitouni.

- **David Suda** (University of Malta)

Contributed Talk, session [75]

Wednesday 11:20 – 11:50, Fitzgerald 139

*Stability of sampling proposals for diffusions over large time intervals*

**Abstract:** The theme of importance sampling for diffusion processes has been tackled several times in literature. We generally want to impute values for a target diffusion process between two observations. One of the motivations for this is that of estimating the transition density evaluated at these points, particularly within the context of Bayesian inference. It would be preferable if our sampling proposals remained stable also when observations are sparse. We shall look at the stability (or lack thereof) of sampling proposals on classes of target diffusions as the time  $T$  between two observations increases. Since we derive our main results within a continuous time context, we focus on the continuous time aspect of importance sampling, with a discussion of its links to the discrete time context. We shall use the Kullback Leibler (KL) divergence to measure the “closeness” of two diffusion measures and to define stability, the latter occurring when the KL divergence converges as  $T$  increases. The main results proved are that, for a large class of target diffusions, the proposal by Durham and Gallant (2002) is unstable, while for certain conditions on the target diffusion, the proposal by Delyon and Hu (2006) is stable. We also discuss the relevance of these results with respect to empirical outputs for varying  $T$  of the two sampling proposals when applied to some examples of target diffusions.

- **Cornelia J. Swanepoel** (Department of Statistics, North-West University, Potchefstroom, South Africa)

Contributed Talk, session [86]

Wednesday 2:35 – 3:05, Rosebrugh 208

*Multiple imputation in the presence of a detection limit, with applications: an empirical approach*

Coauthor: Shawn C Liebenberg (Department of statistical consultation services, North-West University, Potchefstroom, South Africa)

**Abstract:** Missing measurements that are reported to be below a fixed, known detection limit, is a regular occurrence especially in the environmental sciences. Such censored data are often ignored or “guessed” because measurements were made which were incorrectly reported, usually to be zero or to be equal to the detection limit. However, reliable estimates of the population parameters are required to perform statistical analysis. It becomes a complex task to perform when a large number of observations are below the detection limit. Rigorous robust estimation procedures are then needed. This study focuses on density estimation in such scenarios by imputing data in a

sensible way to replace the censored data below the detection limit. First, the maximum likelihood procedure of Cohen (1959) and several variants thereof, are applied to the uncensored data above the detection limit, to estimate the parameters of the underlying density function. Estimation of this density function is then attempted by using the completed imputed data set. Various boundary kernel density estimators are applied comparatively. More specifically, in this study three different Log-normal distributions will be considered to be estimated. The above-mentioned methods are implemented in combination with four new multiple imputation procedures, to assess which of these nonparametric methods are most effective in imputing data to replace the censored values. Several kernel density estimators are fitted to the complete filled-in data set. Comparative measures are applied to establish which combination of strategies are the best to estimate the underlying density function in the presence of a detection limit. The results of an extensive Monte Carlo simulation study are presented and conclusions and recommendations are made.

- **Jan W. H. Swanepoel** (Department of Statistics, North-West University, Potchefstroom, South Africa)

Contributed Talk, session [84]

Monday 2:35 – 3:05, Rosebrugh 310

*Smooth copula-based estimation of conditional density function and conditional hazard rate function*

Coauthors: Nol Veraverbeke (University of Hasselt, Belgium), Paul Janssen (University of Hasselt, Belgium)

**Abstract:** Some recent papers deal with smooth nonparametric estimators for copula functions and copula derivatives. These results lead to copula-based Bernstein estimators for conditional distribution functions and related functionals such as regression and quantile functions. See Janssen, Swanepoel, Veraverbeke (2012, 2014, 2016) and the references in these papers.

The focus in the present paper is on new copula-based smooth Bernstein estimators for the conditional density and the related conditional hazard rate function.

The classical nonparametric kernel estimator for the conditional density function goes back to Rosenblatt (1969). It is the ratio of the joint and the marginal kernel density estimator. Our approach avoids going through separate density estimation of numerator and denominator. The drawbacks of the kernel-based ratio estimator motivated Faugeras (2009) to consider a product-type estimator of a marginal density and a copula density estimator. The above estimators, however, are not smooth. Our estimator is defined as a smoother of the copula-based Bernstein estimator of the conditional distribution function. In a similar spirit, the conditional hazard rate estimator is not obtained from the ratio of the conditional density and conditional survival function, but as a smoother of the conditional cumulative hazard function estimator. The latter is based on Bernstein estimation. We establish asymptotic properties of bias and variance and discuss the asymptotic mean squared error in terms of the smoothing parameters. We also obtain the asymptotic normality of the new estimators. In a simulation study we show the good performance of the new estimator in comparison with some recent proposals of Faugeras (2009).

References

- Faugeras, O.P. (2009). A quantile-copula approach to conditional density estimation. *J. Multivariate Analysis* 100, 2083-2099.
- Janssen, P., Swanepoel, J. and Veraverbeke, N. (2012). Large sample behavior of the Bernstein copula estimator. *J. Statist. Planning Inference*, 142, 1189-1197.
- Janssen, P., Swanepoel, J. and Veraverbeke, N. (2014). A note on the asymptotic behavior of the Bernstein estimator of a copula density. *J. Multivariate Analysis* 124, 480-487.
- Janssen, P., Swanepoel, J. and Veraverbeke, N. (2016). Bernstein estimation for a copula derivative with application to conditional distribution and regression functionals. *TEST* (in print).
- Rosenblatt, M. (1969). Conditional probability and regression estimators. In: *Multivariate Analysis II, Proc. Second International Symposium, Dayton, Ohio, 1968*. Academic Press, New York, pp. 25-31.

- **Jan M. Swart** (Institute of Information Theory and Automation of the ASCR (UTIA))

Invited session: Interacting particle systems and their scaling limits [10]

Thursday 2:35 – 3:05, Medical Sciences 3153

*The Stigler-Luckock model for the evolution of an order book*

**Abstract:** In 1964, G.J. Stigler introduced a stochastic model for the evolution of an order book on a stock market. This model was independently rediscovered and generalized by H. Luckock in 2003. In his formulation, traders place buy and sell limit orders of unit size according to independent Poisson processes with possibly different intensities. Newly arriving buy (sell) orders are either immediately matched to the best available matching sell (buy) order or stay in the order book until a matching order arrives.

Although the model is a huge simplification -and even distortion- of reality, it gives insight in what goes on in real trading. Mathematically, the model is interesting because it appears to exhibit self-organized criticality, similar to the well-known Bak-Sneppen model and a handful of related models. Assuming stationarity, Luckock was able to calculate the critical point explicitly. Proving the existence of stationary states as postulated by Luckock is still a mathematical open problem. In my talk, I will show how progress can be made by adding market orders to the model, which gives rise to a new, rigorous interpretation of the critical point of Luckock.

- **Botond Szabo** (Leiden University)

Invited session: Frequentist properties of Bayesian nonparametric or high dimensional models [34]

Friday 11:55 – 12:25, Medical Sciences 2171

*How many needles in the haystack? Adaptive inference and uncertainty quantification for the horseshoe*

Coauthors: Stephanie van der Pas (Leiden University), Aad van der Vaart (Leiden University)

**Abstract:** In our work we investigate the frequentist properties of the hierarchical Bayes and the maximum marginal likelihood empirical Bayes methods in the sparse multivariate mean model with unknown sparsity level. We consider the popular horseshoe prior introduced in Carvalho, Polson,

and Scott (2010) and show that both adaptive Bayesian techniques lead to rate optimal posterior contraction without using any information on the sparsity level.

Furthermore, we also investigate the frequentist coverage properties of Bayesian credible sets resulting from the horseshoe prior both in the non-adaptive and adaptive setting. We show that the credible sets have good frequentist coverage and optimal size for appropriate choice of the tuning parameter (using information about the sparsity level). In case this information is not available the construction of adaptive and honest confidence sets is not possible, hence we have to introduce some additional restriction. We show that under a non-restrictive assumption both the (slightly modified) hierarchical and empirical Bayes credible sets have (almost) rate adaptive size and good coverage. Finally, we also investigate the coverage properties of marginal credible sets and show that large and small signal components will be covered with probability tending to one, while medium sized signal components will fall outside of the credible interval.

- **Zbigniew Szkutnik** (AGH University of Science and Technology, Krakow, Poland)

Contributed Talk, session [85]

Tuesday 11:20 – 11:50, Medical Sciences 3278

*Confidence bands in stereological problems*

Coauthor: Jakub Wojdyła (AGH University of Science and Technology, Krakow, Poland)

**Abstract:** Construction of confidence bands for a density function of directly observed i.i.d. data started in 1973 with a seminal paper by Bickel and Rosenblatt. In the last decade, asymptotic nonparametric confidence bands have been constructed in some inverse problems, like density deconvolution, inverse regression with a convolution operator and regression with errors in variables, with the pioneering 2007 contribution of Bissantz, Dumbgen, Holzman and Munk. There seems to be, however, no such construction for practically important inverse problems of stereology. We partially fill this gap by constructing a kernel-type estimator for the density of squared radii in the stereological Wicksell’s problem, along with corresponding asymptotic uniform confidence bands and an automatic bandwidth selection method. Let us recall that the Wicksell’s problem of stereology consists in unfolding the distribution of random radii of balls randomly placed in an opaque medium and only observed as circles on a random plane slice through the medium, and that the density of the observed circles radii is related to the density of balls radii via an Abel integral equation. Following earlier contributions, we construct asymptotic confidence bands that are based on strong approximations and on a limit theorem for the supremum of a stationary Gaussian process. The performance of the new procedures is also investigated in a simulation experiment and demonstrated with some real astronomical data related to M62 globular cluster and obtained in an observation process described with a very similar Abel equation, as discussed recently by Sen and Woodroffe (2012).

**T**

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Laleh Tafakori** (University of Melbourne)

Contributed Talk, session [77]

Tuesday 2:35 – 3:05, Medical Sciences 4279

*A model for cell proliferation in a developing organism*

Coauthor: Peter Taylor (University of Melbourne)

**Abstract:** In mathematical biology, there is much interest in building continuum models by scaling discrete agent-based models, governed by local stochastic rules.

We shall discuss a particular example: a model for the process that leads to Hirschprung’s disease, a potentially-fatal condition in which the enteric nervous system of a new-born child does not extend all the way through the intestine and colon, as well as various other stochastic models for foetal tissue growth.

We start with a simple discrete-state Markov chain model proposed by Hywood in 2012 for the location of the neural crest cells that make up the enteric nervous system, and consider a variety of limiting regimes that lead to partial differential equation models that describe the dynamics of crest cell density as the whole gut grows.

The initial discrete-state model has properties that are reminiscent of the totally asymmetric exclusion process with a variable-size lattice. When a neural crest cell proliferates, the size of the whole lattice increases by one and a cell from the underlying domain is left behind the proliferating cell. Since tissue growth is a significant cellular transport mechanism during embryonic growth, it has an indispensable role in the derivation of the correct partial differential equation description.

- **Tomoko Takemura** (Nara Women’s University)

Contributed Talk, session [52]

Friday 11:55 – 12:25, Medical Sciences 4171

*Exponent of Lévy processes corresponding to inverse local time for harmonic transformed diffusion processes*

Coauthor: Matsuyo Tomisaki (Nara Women’s University)

**Abstract:** We are concerned with inverse local time at the regular end point for harmonic transform of one dimensional diffusion process, and consider the corresponding exponent, entrance law and excursion law. On the exponent of inverse local time, there are some interesting works due to R. M. Blumenthal and R. K. Gettoor [1], M. Fukushima and H. Tanaka [2], K. Itô and H. P. McKean [3], etc. In [3], K. Itô and H. P. McKean showed that the Lévy measure density corresponding to the inverse local time at the regular end point for a *recurrent* one dimensional diffusion process is represented as the Laplace transform of the spectral measure corresponding to the diffusion process, where the absorbing boundary condition is posed at the end point. In our previous result [6], employing their representation theorem and our results [4] and [5], we showed that the Lévy measure density corresponding to the inverse local time at the regular end point for a recurrent harmonic transformed process is represented as the Laplace transform of the spectral measure corresponding to the original diffusion process, where the absorbing boundary condition is posed at the end point if it is regular. Here we show that a representation theorem due to K. Itô and H. P. McKean is

available for a *transient* one dimensional diffusion process, and deduce a representation theorem of the Lévy measure density corresponding to the inverse local time for a *transient* harmonic transformed process. Furthermore we show a relation between exponents of inverse local time in [1] and [2], and relations between entrance laws or excursion laws of original diffusion process and its harmonic transform. So far we have treated harmonic transform of minimal processes. However it is possible to consider harmonic transform of non-minimal processes. We present a new consideration for harmonic transform of non-minimal processes.

- (1) R. M. Blumenthal and R. K. Gettoor (1968), Markov Processes and Potential Theory, Academic Press, New York and London.
- (2) M. Fukushima and H. Tanaka (2005), Poisson point processes attached to symmetric diffusions, *Ann. I. H. Poincaré*, **41** 419–459.
- (3) K. Itô and H. P. McKean, Jr. (1974), Diffusion Processes and their Sample Paths, Springer-Verlag, New York.
- (4) T. Takemura (2010), State of boundaries for harmonic transforms of one-dimensional generalized diffusion processes, *Annual Reports of Graduate School of Humanities and Sciences, Nara Women's University*, vol. 25, 285-294.
- (5) T. Takemura and M. Tomisaki (2011),  $h$ -transform of one dimensional generalized diffusion operators, *Kyusyu J.*
- (6) T. Takemura and M. Tomisaki (2013), Lévy measure density corresponding to inverse local time, *Publ. RIMS Kyoto Univ.* **49**, 563–599.
- (7) T. Takemura and M. Tomisaki (2016), Exponent of inverse local time for harmonic transformed process *Annual Reports of Graduate School of Humanities and Sciences, Nara Women's University*.

- **Murad S. Taqqu** (Boston University)

Contributed Talk, session [66]

Thursday 10:45 – 11:15, Medical Sciences 4171

*Weak dependence*

Coauthor: Shuyang Bai (Boston University)

**Abstract:** There are all kinds of weak dependence. For example, strong mixing. Short-range dependence is also a form of weak dependence. It occurs in the context of processes that are subordinated to the Gaussian. Is a short-range dependent process strong mixing if the underlying Gaussian process is long-range dependent? We show that this is not necessarily the case. This is joint work with Shuyang Bai.

- **Jonathan Taylor** (Stanford)

Invited session: Selective inference [28]

Tuesday 11:20 – 11:50, Medical Sciences 3154

*Selective inference in linear regression*

**Abstract:** We consider inference after model selection in linear regression problems, specifically after fitting the LASSO. A classical approach to this problem is data splitting, using some randomly

chosen portion of the data to choose the model and the remaining data for inference in the form of confidence intervals and hypothesis tests. Viewing this problem in the framework of selective inference, conditional on a selection event, we describe other randomized algorithms with similar guarantees to data splitting, at least in the parametric setting. Time permitting, we describe analogous results for statistical functionals obeying a CLT in the classical fixed dimensional setting and inference after choosing a tuning parameter by cross-validation.

- **Kasia Taylor** (University of Warwick)

Contributed Talk, session [51]

Friday 2:00 – 2:30, Medical Sciences 3154

*Approximate likelihood construction for differential equations driven by rough paths.*

Coauthor: Anastasia Papavasiliou (University of Warwick)

**Abstract:** In our presentation we introduce an approximate likelihood construction for the type of differential equations

$$(1) \quad dY_t = a(Y_t; \theta)dt + b(Y_t; \theta)dX_t, \quad Y_0 = y_0, \quad t \leq T,$$

where  $X \in G\Omega_p(\mathbb{R}^m)$  is the realisation of a random geometric  $p$ -rough path and  $\theta \in \Theta$ , where  $\Theta$  is the parameter space.

First, we develop a framework for performing statistical inference for the differential equation

$$(2) \quad dY_t^{\mathcal{D}} = a(Y_t^{\mathcal{D}}; \theta)dt + b(Y_t^{\mathcal{D}}; \theta)dX_t^{\mathcal{D}}, \quad Y_0 = y_0, \quad t \leq T,$$

where  $X^{\mathcal{D}}$  is a realisation of a random piecewise linear path in  $\mathbb{R}^m$  corresponding to partition  $\mathcal{D}$  of  $[0, T]$ . Under the assumption that we know the distribution of  $X^{\mathcal{D}}$  we construct the likelihood of discrete observations of  $Y^{\mathcal{D}}$  on the grid  $\mathcal{D}$ . The main idea is to use the observations to explicitly construct the map that maps a finite parametrization of  $Y^{\mathcal{D}}$  to a finite parametrization of  $X^{\mathcal{D}}$ .

Next, we consider equation (1) and  $\pi_n(X)$  the sequence of piecewise linear paths with nested grids  $\mathcal{D}(n)$ , such that  $d_p(\pi_n(X), X) \rightarrow 0$  as  $n \rightarrow \infty$  which in the  $p$ -variation topology implies that  $d_p(Y(n), Y) \rightarrow 0$  as  $n \rightarrow \infty$  where  $Y(n)$  is solution of (2) driven by  $\pi_n(X)$ . We use the likelihood constructed before to construct an approximate likelihood of observing a realisation of (1) on grid  $\mathcal{D}(n)$  for some fixed  $n$ . The main idea behind the construction is to replace the model (1) that produces the data by (2), which is tractable and converges to (1) for  $n \rightarrow \infty$ . We show that an appropriate distance between the likelihood for discrete observations on a grid  $\mathcal{D}(n)$  of the corresponding process  $Y(n)$  and of the limiting process  $Y$  disappears, as  $n \rightarrow \infty$ . We give conditions necessary for the methodology to work.

We demonstrate how the method works in the simple case of a discretely observed Ornstein-Uhlenbeck model driven by fractional Brownian motion and also we construct the limiting likelihood for a general multivariate diffusion process.

- **Christoph Thäle** (Faculty of Mathematics at Ruhr University Bochum)

Invited session: Limit Theorems and the Malliavin calculus [11]

Tuesday 3:10 – 3:40, Medical Sciences 2171

*Malliavin-Stein method and spatial random systems*

**Abstract:** The Malliavin-Stein method for normal approximation of non-linear functionals of Poisson random measures was introduced in a seminal paper of Peccati, Solé, Taqqu and Utzet. If the kernels of the chaos decomposition of the functional under consideration is available, it allows to derive explicit error bounds by computing the so-called contraction norms of these kernels. On the other hand, if the chaos decomposition cannot be computed, a second-order Poincaré inequality can be used as a substitute that also leads to rates of convergence in the central limit theorem. In my talk I will discuss the applicability of the Malliavin-Stein method to a representative collection of problems that arise in stochastic geometry. Examples include the random geometric graph, the radial spanning tree or the Poisson hyperplane tessellation. The results are based on joint works with G. Last, M.D. Penrose, M. Reitzner and M. Schulte.

- **Alexander Tikhomirov** (Komi Science Center Ural Division of RAS, Russia)

Contributed Talk, session [58]

Monday 10:45 – 11:15, Medical Sciences 4279

*Local semicircle and Marchenko – Pastur laws for random matrices*

Coauthors: Goetze Friedrich (Bielefeld University, Germany), Naumov Alexej (Moscow State University, Moscow)

**Abstract:** We consider a random symmetric matrix  $\mathbf{X} = [X_{jk}]_{j,k=1}^n$  with upper triangular entries being independent identically distributed random variables with mean zero and unit variance. We additionally suppose that  $\mathbb{E}|X_{11}|^{4+\delta} =: \mu_{4+\delta} < C$  for some  $\delta > 0$  and some absolute constant  $C$ . Under these conditions we show that the typical Kolmogorov distance between the empirical spectral distribution function of eigenvalues of  $n^{-1/2}\mathbf{X}$  and Wigner’s semicircle law is of order  $1/n$  up to some logarithmic correction factor. As a direct consequence of this result we establish that the semicircle law holds on a short scale. Furthermore, we show for this finite moment ensemble rigidity of eigenvalues and delocalization properties of the eigenvectors. Similar results are obtained for sample covariance matrices.

- **Tin Lam Toh** (National Institute of Education, Nanyang Technological University, Singapore)

Contributed Talk, session [49]

Thursday 11:55 – 12:25, Medical Sciences 3278

*Henstock approach to stochastic integral: An innovative approach*

**Abstract:** In this talk, the author will discuss his work for the past decade on Henstock approach to stochastic integral - Ito integral and the stochastic properties, including Ito Lemma and Integration-by-Parts formula, and a characterization of Henstock-Ito integrable processes, and work has begun on studying Stratonovich integral using Henstock’s approach. This approach, compared to the classical approach of using simple functions to define the stochastic integral, is more intuitive and less technically involved. Pedagogically, it could be easier for undergraduate or beginning postgraduate mathematics students to comprehend, and could make stochastic integration theory more accessible to more students.

- **Cristina Tone** (University of Louisville)

Contributed Talk, session [65]

Friday 10:45 – 11:15, Medical Sciences 4279

*A central limit theorem for non-stationary strongly mixing random fields*

Coauthor: Richard C. Bradley (Indiana University)

**Abstract:** We introduce a central limit theorem for sequences of random fields that satisfy a Lindeberg condition and uniformly satisfy both strong mixing and an upper bound less than 1 on  $\rho'(\cdot, 1)$ , in the absence of stationarity. There is no requirement of either a mixing rate assumption or the existence of moments of order higher than two. The additional assumption of a uniform upper bound less than 1 for  $\rho'(\cdot, 1)$  cannot simply be deleted altogether from the theorem, even in the case of strict stationarity. For the case  $d = 1$ , that can be seen from any (finite-variance) strictly stationary, strongly mixing counterexample to the CLT such that the rate of growth of the variances of the partial sums is at least linear; for several such examples, see, e.g., [1] Theorem 10.25 and Chapters 30–33. Our main theorem extends certain central limit theorems of Peligrad [2] involving “arrays of random sequences”.

- (1) Bradley, R.C., *Introduction to Strong Mixing Conditions*, vol. 1, 2, & 3. Kendrick Press, Heber City (Utah) (2007);
- (2) Peligrad, M., *On the asymptotic normality of sequences of weak dependent random variables*, J. Theor. Probab. 9, 703715 (1996).

- **Mathias Trabs** (University of Hamburg)

Journal paper session [40]: Bernoulli Paper

Friday 2:00 – 2:45, McLeod Auditorium

*Adaptive quantile estimation in deconvolution with unknown error distribution*

Coauthors: Itai Dattner (University of Haifa), Markus Reiß (Humboldt-Universität zu Berlin)

**Abstract:** Quantile estimation in deconvolution problems is studied comprehensively. In particular, the more realistic setup of unknown error distributions is covered. Our plug-in method is based on a deconvolution density estimator and is minimax optimal under minimal and natural conditions. This closes an important gap in the literature. Optimal adaptive estimation is obtained by a data-driven bandwidth choice. As a side result, we obtain optimal rates for the plug-in estimation of distribution functions with unknown error distributions. The method is applied to a real data example.

- **A. Alexandre Trindade** (Texas Tech University)

Contributed Talk, session [82]

Thursday 10:45 – 11:15, Rosebrugh 208

*An empirical saddlepoint approximation based method for smoothing survival functions under right censoring*

**Abstract:** The Kaplan-Meier (KM) estimator is a commonly used non-parametric procedure for estimating survival functions. However, KM only defines the approximate probability of observed failure times, and may not deliver a proper density function if the largest observation is right censored. In addition, existing smoothing methods based on KM also assume that the largest observation is

not censored. To alleviate these issues, we devise a method for smoothing KM survival functions based on an empirical saddlepoint approximation. The method inverts the moment generating function (MGF) defined through a Riemann-Stieltjes integral with respect to the KM approximation to the failure time cumulative distribution function and exponential right-tail completion. Using tools from the theory of empirical processes, uniform consistency, weak, and strong convergence results are established for this modified version of the empirical MGF based on KM weights. The performance of the methodology is examined in simulation studies, which demonstrates that the proposed empirical saddlepoint approximation method is faster and more accurate than existing methods for smoothing survival functions.

- **Alexandre Tsybakov** (CREST-ENSAE)

Invited session: Model selection in non-standard high dimensional models [32]

Wednesday 3:10 – 3:40, Medical Sciences 2171

*Slope meets Lasso: Improved oracle bounds and optimality*

Coauthors: Pierre C. Bellec (CREST-ENSAE), Guillaume Lecue (CREST-ENSAE)

**Abstract:** We show that two polynomial time methods, a Lasso estimator with adaptively chosen tuning parameter and a Slope estimator, adaptively achieve the exact minimax prediction and  $\ell_2$  estimation rate  $(s/n) \log(p/s)$  in high-dimensional linear regression on the class of  $s$ -sparse target vectors in  $R^p$ . This is done under the Restricted Eigenvalue (RE) condition for the Lasso and under a slightly more constraining assumption on the design for the Slope. The main results have the form of sharp oracle inequalities accounting for the model misspecification error. The minimax optimal bounds are also obtained for the  $\ell_q$  estimation errors with  $1 \leq q \leq 2$  when the model is well-specified. The results are non-asymptotic, and hold both in probability and in expectation. The assumptions that we impose on the design are satisfied with high probability for a large class of random matrices with independent and possibly anisotropically distributed rows. We give a comparative analysis of conditions, under which oracle bounds for the Lasso and Slope estimators can be obtained. In particular, we show that several known conditions, such as the RE condition and the sparse eigenvalue condition are equivalent if the  $\ell_2$ -norms of regressors are uniformly bounded.

- **Amanda Turner** (Lancaster University)

Invited session: Random growth models [3]

Tuesday 10:45 – 11:15, McLeod Auditorium

*Scaling limits of Laplacian random growth models*

Coauthors: Alan Sola (Stockholm University), Fredrik Viklund (KTH Royal Institute of Technology)

**Abstract:** The idea of using conformal mappings to represent randomly growing clusters has been around for almost 20 years. Examples include the Hastings-Levitov models for planar random growth, which cover physically occurring processes such as diffusion-limited aggregation (DLA), dielectric breakdown and the Eden model for biological cell growth, and more recently Miller and Sheffield's Quantum Loewner Evolution (QLE). In this talk we will discuss ongoing work on a natural variation of the Hastings-Levitov family. We will show that for certain parameter values, under a sufficiently

weak regularisation, the resulting cluster can be shown to converge to a randomly oriented one-dimensional slit. This contrasts with earlier work in which scaling limits, under sufficiently strong regularisations, are shown to be deterministically growing disks.

This is based on work in progress with Alan Sola (Stockholm) and Fredrik Viklund (KTH).

U

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Masayuki Uchida** (Osaka University)

Contributed Talk, session [88]

Tuesday 11:20 – 11:50, Medical Sciences 4279

*Adaptive Bayes estimators for small diffusion processes based on sampled data*

**Abstract:** We deal with an estimation problem of both drift and volatility parameters for diffusion processes defined by the stochastic differential equations with small dispersion parameters from discrete observations. We call such diffusion processes “small diffusion processes”. Applying both adaptive maximum likelihood type estimation and adaptive Bayes type estimation for ergodic diffusion processes to small diffusion processes, we propose the adaptive Bayes type estimators based on the quasi likelihood functions for small diffusion processes. Moreover, we prove that the adaptive Bayes type estimators have asymptotic normality and convergence of moments by using the Ibragimov-Has’minskii-Kutoyants program together with the polynomial type large deviation inequality for the statistical random field. An example and simulation results are also given. This is a joint work with Ryosuke Nomura.

- **Caroline Uhler** (MIT)

Invited session: Graphical models and totally positive distributions [22]

Monday 11:55 – 12:25, Medical Sciences 2172

*Totally positive exponential families and graphical models*

Coauthors: S. L. Lauritzen (University of Copenhagen), Piotr Zwiernik (Universitat Pompeu Fabra)

**Abstract:** We discuss properties of exponential families that are multivariate totally positive of order two (MTP2). Such distributions appear in the context of positive dependence, ferromagnetism in the Ising model, and various latent models. We show that maximum likelihood estimation for MTP2 exponential families is a convex problem. Hence, if the MLE exists, it is unique. For quadratic exponential families, such as Ising models or Gaussian graphical models, we show that MTP2 implies sparsity of the underlying graph without the need of a tuning parameter. Moreover, we show that the MLE always exists even in the high-dimensional setting. These properties make MTP2 constraints interesting as an alternative to methods for learning sparse graphical models such as the graphical lasso. The lecture is based on joint work with S. L. Lauritzen and P. Zwiernik.

V

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Jean Vaillancourt** (ENAP)

Contributed Talk, session [61]

Tuesday 11:55 – 12:25, Rosebrugh 208

*Winner takes all*

Coauthor: Bruno Rémillard (HEC)

**Abstract:** Combining losing games into a winning game is a century old dream. An illustration of what can be done is the so-called Parrondo’s paradox in the Physics literature. This “paradox” is extended to regime switching random walks in random environments. The paradoxical behavior of the resulting random walk is explained by the effect of the random environment. Full characterization of the asymptotic behavior is achieved in terms of the dimensions of some random subspaces occurring in Oseledec’s ergodic theorem. Basically, these models exhibit the same dichotomy, in their asymptotic behavior, as simple random walks, in terms of transience and recurrence.

A wide range of regime switching strategies, motivated by gaming applications, render the resulting Markov chain neither reversible nor even irreducible, in several of the most interesting special cases under consideration here. Previous results cannot be applied in general here as a result. This is the case in particular for many periodic strategies which exhibit this seemingly paradoxical behavior.

- **Sara van de Geer** (Seminar for Statistics, ETH Zurich)

Plenary talk: Wald Lecture 1

Tuesday 9:25 – 10:15, McLeod Auditorium

*High-dimensional statistics: a triptych (part 1)*

**Abstract:** High-dimensional statistics concerns the situation where the number of parameters  $p$  is (much) larger than the number of observations  $n$ . This is quite common nowadays, and it has led to the development of new statistical methodology. The lectures present a selected overview of mathematical theory for sparsity inducing methods.

In the first lecture we will highlight the main ingredients for proving sharp oracle inequalities for regularized empirical risk minimizers. The regularization penalty will be taken to be a norm  $\Omega$  on  $p$ -dimensional Euclidean space. Important is that the norm  $\Omega$  has a particular feature which we term the triangle property. We present as examples: the  $\ell_1$ -norm, norms generated from cones, the sorted  $\ell_1$ -norm, the nuclear norm for matrices and an extension to tensors. We then show sharp oracle inequalities for a broad class of loss functions.

The second lecture addresses the construction of asymptotic confidence intervals for parameters of interest. Here, we restrict ourselves to the linear and the graphical model. We prove asymptotic normality of de-biased estimators. We consider asymptotic lower bounds for the variance of an approximately unbiased estimator of a one-dimensional parameter as well as Le Cam type lower bounds. We ascertain the approximate unbiasedness of the de-biased estimator under sparsity conditions and show that it reaches the lower bound.

In the third lecture, we examine the null space property for sparsity inducing norms. The null space property ensures exact recovery of certain sparsity patterns and is moreover a key ingredient

for oracle results. We derive this property for the Gram matrix based on  $n$  copies of a  $p$ -dimensional random variable  $X$ , where we require moment conditions for finite dimensional projections of  $X$  or the more general small ball property.

The lectures are based on joint work with Andreas Elsener, Jana Janková, Alan Muro and Benjamin Stucky.

- **Sara van de Geer** (Seminar for Statistics, ETH Zurich)  
Plenary talk: Wald Lecture 2  
Thursday 9:25 – 10:15, McLeod Auditorium  
*High-dimensional statistics: a triptych (part 2)*  
**Abstract:** See Wald Lecture 1
- **Sara van de Geer** (Seminar for Statistics, ETH Zurich)  
Plenary talk: Wald Lecture 3  
Friday 9:25 – 10:15, McLeod Auditorium  
*High-dimensional statistics: a triptych (part 3)*  
**Abstract:** See Wald Lecture 1
- **Ramon Van Handel** (Princeton University)  
Invited session: Random matrix theory and statistics [21]  
Thursday 2:00 – 2:30, Medical Sciences 2171  
*Random matrices and Gaussian processes*  
**Abstract:** The study of norms of inhomogeneous random matrices is intimately connected to the geometry of random processes. I will describe some recent results and conjectures that motivate the development of new ideas at the intersection of these areas.
- **Anand N. Vidyashankar** (George Mason University, Fairfax, VA)  
Contributed Talk, session [56]  
Thursday 2:35 – 3:05, Medical Sciences 3278  
*Local and global Gibbs conditioning principles for branching processes*  
Coauthor: Peter E. Ney (University of Wisconsin, Madison, WI)  
**Abstract:** Branching processes with and without immigration appear in a variety of scientific contexts. A probabilistic description of the evolution of such processes, based on current information, is important. Specifically, let  $\{X_n : n \geq 1\}$  denote a branching process with immigration initiated by a single ancestor and  $R_n = X_n^{-1}X_{n+1}$ ; let  $m$  denote the mean of the offspring distribution and assume that  $m > 1$ . We study the asymptotic behavior of

$$P(X_{r_n} = (\cdot) | X_n \in (cv_n dv_n))$$

where  $v_n = o(m^n)$ ,  $v_n \nearrow \infty$ ,  $0 \leq c < d < \infty$ , and  $r_n$  grows at various rates. We establish that under the classical  $Z \log Z$  assumption, there exists a critical time  $t_n^* < n$  (referred to as the divergence time) such that

$$\lim_{n \rightarrow \infty} P(X_{t_n^*} = j | X_n \in (cv_n dv_n)) = \pi_j, \quad j = 1, 2, \dots,$$

where  $\{\pi_j : j \geq 1\}$  is a proper probability distribution. In a sense to be made precise, the conditioned process is “quiescent” up to  $t_n^*$  and evolves like a branching process after  $t_n^*$ . Turning to the behavior of the last generation, we show that under an exponential moment condition

$$P(\xi_{n,1} = k_1, \xi_{n,2} = k_2, \dots, \xi_{n,l_n} = k_{l_n} | X_n \in (cv_n, dv_n), R_n > a), l_n \leq cv_n,$$

where  $\xi_{n,r}$  represents the number of offspring produced by the  $r^{\text{th}}$  parent in  $n^{\text{th}}$  generation  $k_1, k_2, \dots$  are integers,  $0 < c < d < \infty$ ,  $a > m$  and  $v_n = o(m^n)$ , the above law behaves like a product of “twisted offspring distributions” which is reminiscent of the classical Gibbs conditional principle for i.i.d. random variables.

• **Vladimir Vinogradov** (Ohio University)

Contributed Talk, session [55]

Friday 2:35 – 3:05, Medical Sciences 4171

*On properties of Poisson-Tweedie mixtures*

Coauthor: Richard B. Paris (University of Abertay Dundee)

**Abstract:** We derive a number of new local limit theorems for specific members of the three-parameter class of Poisson-Tweedie mixtures. This proper subclass of the family of Poisson mixtures corresponds to the case where the mixing measure is generated by a member of the power-variance family of univariate distributions with non-negative support. It contains back-shifted negative binomial, Neyman type A, Pólya-Aeppli and Poisson-inverse Gaussian distributions as important special cases. It is known that for a specific value of the “power” parameter associated with the corresponding members of the power-variance family, such mixtures comprise an additive exponential dispersion model as well as a factorial dispersion model which are characterized by the variance and the dispersion functions, respectively. In the case where the former structure is considered, we derive closed-form expressions for the related variance functions in terms of the exponential tilting invariants and particular special functions, and also investigate their continuity properties with respect to the “power” parameter. We construct asymptotic expansions in the local limit theorems in the cases where a specific one-parameter class of these mixtures is approximated by a member of the power-variance family with the same value of the “power” parameter as well as when it converges to a Poisson distribution. Some of the latter results are closely related to the Poisson law of thin numbers and binomial thinning. In the special case which pertains to Pólya-Aeppli distributions, we illustrate such a Poisson convergence result via the stochastic evolution of the cluster structure of branching populations. We also present new refined saddlepoint-type approximations for these two-parameter classes of non-negative integer-valued distributions relating them to some earlier work on the exact asymptotics of the probabilities of large deviations for both polar types of the mechanism of their formation. The techniques employed combine probabilistic methods with our new analytic results on subtle properties of particular Wright special functions and their approximations.

W

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Yizao Wang** (University of Cincinnati, USA)

Journal paper session [40]: SPA Paper

Friday 2:50 – 3:40, McLeod Auditorium

*Limit laws for maximal standardized increment of a random walk*

Coauthor: Zakhar Kabluchko (University of Münster)

**Abstract:** We investigate the limit laws for the maximal standardized increment of a random walk  $S_n$ , defined as

$$\max_{0 \leq i < j \leq n} \frac{S_j - S_i}{\sqrt{j - i}}.$$

We assume that the jumps are formed by i.i.d. random variables, the distribution of which has finite Laplace transform. The case that the jumps are Gaussian has been addressed by Siegmund and Venkatraman (1995). For general distributions, our results reveal a more subtle picture: the limit law being always Gumbel, the normalization sequence depends on the distribution through their Laplace transform. In particular, we distinguish 4 different cases.

Joint work with Zakhar Kabluchko (University of Münster)

- **Larry Wasserman** (Carnegie Mellon)

Invited session: Selective inference [28]

Tuesday 11:55 – 12:25, Medical Sciences 3154

*Regression without assumptions*

Coauthors: Max G'Sell (CMU), Jing Lei (CMU)

**Abstract:** Most methods for high dimensional inference rely on strong assumptions. I will discuss distribution-free methods for high-dimensional regression.

- **Marten H. Wegkamp** (Cornell University)

Invited session: Model selection in non-standard high dimensional models [32]

Wednesday 2:35 – 3:05, Medical Sciences 2171

*Adaptive estimation of the copula correlation matrix in factor models*

Coauthor: Yue Zhao (McGill University)

**Abstract:** We study the adaptive estimation of copula correlation matrix  $\Sigma$  for elliptical copulas. In this context, the correlations are connected to Kendall's tau through a sine function transformation. Hence, a natural estimate for  $\Sigma$  is the plug-in estimator  $\widehat{\Sigma}$  with Kendall's tau statistic. We first obtain a sharp bound for the operator norm of  $\widehat{\Sigma} - \Sigma$ . Then, we study a factor model for  $\Sigma$ , for which we propose a refined estimator  $\widetilde{\Sigma}$  by fitting a low-rank matrix plus a diagonal matrix to  $\widehat{\Sigma}$  using least squares with a nuclear norm penalty on the low-rank matrix. The bound on the operator norm  $\widehat{\Sigma} - \Sigma$  serves to scale the penalty term, and we obtain finite sample oracle inequalities for  $\widetilde{\Sigma}$ . We provide data-driven versions for our estimation procedure and performance bounds.

- **Martin Weidner** (Imperial College London)

Contributed Talk, session [51]

Friday 2:35 – 3:05, Medical Sciences 3154

*Rough differential equations on manifolds and Hörmander's theorem*

Coauthor: Thomas Cass (Imperial College London)

**Abstract:** We consider differential equations driven by weakly geometric rough paths along vector fields on a smooth manifold. For this purpose we revisit the behaviour of rough differential equations under a change of coordinates and we give a purely combinatorial proof of the respective chain rule using the Grossman-Larson algebra of labelled trees. We also discuss those aspects of Malliavin calculus that are different for manifold valued random variables in comparison to vector space valued ones. Under the assumptions that the driving noise belongs to a certain class of Gaussian processes (which includes fractional Brownian motion with Hurst parameter  $H > 1/4$ ) and that the vector fields satisfy Hörmander's condition we show that the law of the solution of the rough differential equation has a smooth density. This generalises an existing result for vector space valued vector fields.

- **Ruth Williams** (UC San Diego)

Plenary talk: Kolmogorov Lecture

Friday 4:15 – 5:05, McLeod Auditorium

*Reflected diffusions and (bio)chemical reaction networks*

**Abstract:** Continuous-time Markov chain models are often used to describe the stochastic dynamics of networks of reacting chemical species, especially in the growing field of systems biology. Discrete-event stochastic simulation of these models rapidly becomes computationally intensive. Consequently, more tractable diffusion approximations are commonly used in numerical computation, even for modest-sized networks. However, existing approximations (e.g., van Kampen and Langevin), do not respect the constraint that chemical concentrations are never negative.

In this talk, we propose an approximation for such Markov chains, via reflected diffusion processes, that respects the fact that concentrations of chemical species are non-negative. This fixes a difficulty with Langevin approximations that they are frequently only valid until the boundary of the positive orthant is reached. Our approximation has the added advantage that it can be written down immediately from the chemical reactions. This contrasts with the van Kampen approximation, which involves a two-stage procedure — first solving a deterministic ordinary differential equation, followed by a stochastic differential equation for fluctuations around those solutions.

An invariance principle for reflected diffusions, due to Kang and Williams, is adapted in justifying our approximation under mild assumptions. Some numerical examples illustrate the advantages of our approximation over direct simulation of the Markov chain or use of the van Kampen approximation.

This talk is based on joint work with David Anderson, Des Higham and Saul Leite.

- **Carsten Wiuf** (University of Copenhagen)

Invited session: Mathematical systems biology [18]

Thursday 11:55 – 12:25, Medical Sciences 2171

*Where does the Lyapunov Function comes from? NEW TITLE*

**Abstract:** The relationship between stationary distributions for stochastic models of reaction systems and Lyapunov functions for their deterministic counterparts is discussed. Specifically, the

well-known Lyapunov function of reaction network theory is derived as a scaling limit of the non-equilibrium potential of the stationary distribution of stochastically modeled complex balanced systems. This result is extended to general birthdeath models and demonstrate via example that similar scaling limits can yield Lyapunov functions even for models that are not complex balanced, and may even have multiple equilibria. Some general comments about the generality of the result are offered.

- **Ruoyu Wu** (University of North Carolina at Chapel Hill)

Contributed Talk, session [53]

Monday 2:35 – 3:05, Medical Sciences 4279

*Large deviations for particle approximations of a nonlinear heat equation by Brownian motions with killing*

Coauthors: Amarjit Budhiraja (University of North Carolina at Chapel Hill), Wai-Tong (Louis) Fan (University of Wisconsin-Madison)

**Abstract:** Particle approximations for certain nonlinear (and nonlocal) heat equations are studied using a system of Brownian motions with killing. The system is described by a collection of i.i.d. Brownian particles where each particle is killed independently at a rate determined by the empirical sub-probability measure of the states of the alive particles. A large deviation principle (LDP) for such sub-probability measure-valued processes is established. As a byproduct, a convenient variational representation for expectations of nonnegative functionals of Brownian motions along with an i.i.d. sequence of random variables is established. Proofs of the LDP rely on this variational representation and weak convergence arguments.

X

Y

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Dong Yan** (Leiden University)

Contributed Talk, session [89]

Monday 2:35 – 3:05, Medical Sciences 2172

*Bayesian inverse problems with partial observations*

Coauthor: Aad van der Vaart (Leiden University)

**Abstract:** We study a nonparametric Bayesian approach to linear inverse problems with discrete observations only available at design points, as shown in the following model

$$Y_i = Af(x_i) + \varepsilon_i,$$

where  $A$  is a linear operator with unbounded inverse,  $\{x_i\}_{i=1,\dots,n}$  are deterministic design points and  $\varepsilon_i$  are standard Gaussian noise. The singular value decomposition of the operator  $A$  categorizes the ill-posedness into mildly and extremely ill-posed cases, depending on whether the decay of eigenvalues  $\{a_i\}$  of  $A$  is polynomial or exponential. Then, we use discrete Fourier transform to convert the previous model to a truncated sequence model which is closely related to the white noise sequence

model. For both situations of ill-posedness, we consider a class of prior distributions on the truncated sequence model, indexed by a parameter characterizing smoothness, and show that the corresponding posterior distributions contract around the true parameter at a rate depending on the smoothness of the true parameter, the smoothness of the prior, and the ill-posedness. Correct combinations of these variables lead to the optimal minimax rate. The frequentist coverage of credible regions is shown to be dependent on the combination of the prior, the true parameter and the ill-posedness as well. Over-smoothing priors lead to zero coverage, while under-smoothing priors give (highly) conservative results. The results for both ill-posed cases are illustrated by numerical examples.

- **Jun Yang** (University of Toronto)

Contributed Talk, session [89]

Monday 2:00 – 2:30, Medical Sciences 2172

*Meta-Bayesian analysis*

Coauthor: Daniel M. Roy (University of Toronto)

**Abstract:** The “optimality” of the Bayesian approach to inference does not hold when the model is misspecified: there is no setting of the parameters believed to describe the data at hand, and so, strictly speaking, there can be no subjective prior. As essentially every statistical model is misspecified, this raises the question: what is a prior? We formalize the problem of choosing a (surrogate) prior as a Bayesian decision theory task, and develop theory and algorithms for choosing optimal surrogate priors. The resulting theory, which we call meta-Bayesian analysis, gives (optimal surrogate) priors a pragmatic interpretation: relative to one’s actual subjective beliefs, they lead to the best inference possible using the misspecified model. In this paper, we discuss some early results on meta-Bayesian analysis, which have some surprising consequences. For example, in violation of tradition Bayesian tenets, the optimal surrogate prior may depend on the loss function, on the number of data points you plan to observe, and on the number of predictions you expect to make. Also, for the i.i.d Bernoulli model versus any belief on binary sequence, under some conditions, the marginal prior belief on the limiting frequency of 1’s is asymptotically optimal when the number of predictions goes to infinity.

- **Mingan Yang** (San Diego State University)

Contributed Talk, session [89]

Monday 3:10 – 3:40, Medical Sciences 2172

*Bayesian variable selection for mixed effect models with nonignorable dropout*

**Abstract:** In this article, we use Bayesian nonparametric approach for joint variable selection of both mixed and random effects in presence of nonignorable dropout. We integrate both shrinkage prior, which is expressed as a scale mixture of normal distributions, and the mixed G-prior in our approach. By this way, we greatly improve efficiency and accuracy. In particular, we show that our approach greatly decreases bias in estimation and selection for nonignorable missing data. A stochastic search Gibbs sampler is implemented for variable selection. We further illustrate the proposed approach using simulated data and a real example.

- **Fang Yao** (University of Toronto)

Contributed session: Functional data analysis [43]

Tuesday 3:10 – 3:40, Medical Sciences 3154

*Mixture models and densities for functional data*

Coauthors: Zhenhua Lin (University of Toronto), Hans G. Mueller (University of California, Davis)

**Abstract:** We propose a novel perspective to represent infinite-dimensional functional data as mixtures, where each realization of the underlying stochastic process is represented by a finite number of basis functions. In the proposed mixture representation, the number of basis functions that constitutes the mixture may be arbitrarily large, where the number of included mixture components is specifically adapted for each random trajectory. We show that within this framework, a probability density can be well defined under mild regularity conditions, without the need for finite truncation or approximation, where genuine probability densities do not exist. This is due to the low small ball probabilities and as a consequence only approximative solutions to the density problem can be found in the usually considered  $L_2$  space. Unlike traditional functional principal component analysis that uses a common number of components to represent all trajectories, individual trajectories possess a trajectory-specific dimension that may be treated as a latent random variable. We establish a notion of consistency for estimating the functional mixture density and introduce an algorithm for fitting the functional mixture model based on a modified expectation-maximization algorithm. Simulations confirm that in comparison to traditional functional principal component analysis the proposed mixture modelling achieves similar or better data recovery while using fewer components on average. The practical merits of functional mixture modelling are demonstrated in an analysis of egg-laying trajectories for medflies.

- **Murat Yazici** (JFORCE Information Technologies Inc.)

Contributed Talk, session [91]

Wednesday 2:00 – 2:30, Medical Sciences 3278

*A least squares ratio (LSR) approach to fuzzy linear regression analysis*

**Abstract:** The Ordinary Least Squares (OLS) approach to Fuzzy Linear Regression Analysis which includes a doubly linear adaptive fuzzy regression model is one of the used methods for fitting a regression model and forecasting. The doubly linear adaptive fuzzy regression model based on two linear models: a core regression model and a spread regression model. It aims to minimize the Euclidean distance between the observed dependent fuzzy values and the obtained dependent fuzzy values. This paper includes a regression method called Least Squares Ratio (LSR) approach to fuzzy linear regression, and comparison of OLS and LSR according to Mean Absolute Error (MAE). In this study, symmetric triangular fuzzy numbers are used, the observed symmetric triangular fuzzy numbers are shown as  $\tilde{y}_i$ , and the obtained symmetric triangular fuzzy numbers are shown as  $\hat{y}_i$ .  $\tilde{y}_i$  and  $\hat{y}_i$  indicates the observed centers and spreads, and  $\tilde{y}_i$  and  $\hat{y}_i$  indicates the obtained centers and spreads, respectively.

- **Gokhan Yildirim** (York University)

Contributed Talk, session [61]

Tuesday 11:20 – 11:50, Rosebrugh 208

*On the longest monotone and alternating subsequences of pattern-avoiding permutations*

Coauthor: Neal Madras (York University)

**Abstract:** Let  $S_k$  be the set of all permutations  $\tau = \tau_1\tau_2\cdots\tau_k$  on  $\{1, 2, \dots, k\}$ . For  $\tau \in S_k$ , we say that a string of  $k$  distinct integers  $g_1g_2\cdots g_k$  forms the pattern  $\tau$  if for each  $i = 1, \dots, k$ ,  $g_i$  is the  $\tau_i$ -th smallest element of  $\{g_1, \dots, g_k\}$ . For example, 694 forms the pattern 231; 1845 forms the pattern 1423.

Let  $n$  be an integer with  $n \geq k$ . For  $\tau \in S_k$ , we say that a permutation  $\sigma \in S_n$  contains the pattern  $\tau$  if some  $k$ -element subsequence  $\sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_k}$  of  $\sigma$  forms the pattern  $\tau$ . We say that  $\sigma$  avoids the pattern  $\tau$  if  $\sigma$  does not contain  $\tau$ . The set of all permutations in  $S_n$  avoiding  $\tau$  is denoted by  $S_n(\tau)$ . For any set of patterns  $T$ , we write  $S_n(T)$  to denote the set of permutations in  $S_n$  avoiding all the patterns in  $T$ , that is,  $S_n(T) = \bigcap_{\tau \in T} S_n(\tau)$ .

We determine the exact distributions for the longest monotone and alternating subsequences of pattern-avoiding permutations from  $S_n(\tau_1, \tau_2)$  equipped with the uniform probability measure where  $\tau_1, \tau_2 \in S_3$ . We also present some results related to the pattern-avoiding permutations for monotone patterns,  $\tau = 12\cdots k$ , and patterns of the form  $\tau = k\tau_2\cdots\tau_k$ . Typically, the longest monotone subsequences have expected length proportional to  $n$  for  $S_n(T)$  rather than the  $\sqrt{n}$  behaviour that holds for all of  $S_n$ .

*This is a joint work with Neal Madras.*

- **Nakahiro Yoshida** (University of Tokyo)

Invited session: Statistics of random processes [15]

Wednesday 10:45 – 11:15, Medical Sciences 4171

*Point processes and limit order book modeling*

**Abstract:** The latest financial statistics is directed toward analysis of ultra high frequency phenomena by modeling in a more and more precise time-scale. There is no Brownian motion driving the system, differently from the standard framework, since the central limit theorem is not effective at this level of description. Recently point processes are attracting attention in modeling limit order book (LOB). The aim of this talk is to give a theoretical basis of statistical inference for point processes.

The quasi likelihood analysis (QLA) is a systematic analysis of the quasi likelihood random field and the associated estimators. The QLA features a large deviation technique that provides more precise tail probability estimates of the random field and estimators than those limit distributions give. Such estimates are required in the theory of prediction and model selection.

The point process regression model can express asynchronicity of observations and microstructure, besides nonstationarity and self-exciting/self-correcting effects as well as exogenous effects. A non-ergodic QLA is established when the intensities of the point processes become large. A related topic is a nonparametric estimation of the covariation between latent intensity processes based on counting data.

QLA can be developed also in long term asymptotics. Then establishing ergodicity of point processes becomes an issue.

Non-ergodic QLA can be regarded as a collection of short time scale ergodic QLAs. Local ergodic modeling of LOB is the first step toward global non-ergodic descriptions. AIC is applied to selection of LOB models.

### References

- [1] Yoshida, N. (2011) *Polynomial type large deviation inequalities and quasi-likelihood analysis for stochastic differential equations*, Annals of the Institute of Statistical Mathematics 63, 431–479
- [2] Ogihara, T, Yoshida, N. (2015) *Quasi likelihood analysis of point processes for ultra high frequency data*, arXiv:1512.01619
- [3] Clinet, S., Yoshida, N. (2015) *Statistical Inference for Ergodic Point Processes and Limit Order Book* arXiv:1512.01899
- [4] Kimura, A., Yoshida, N.: Estimation of correlation between latent processes. submitted (2015)
- [5] Muni Toke, I., Yoshida, N. (2016) *Modelling intensities of order flows in a limit order book* arXiv:1602.03944

- **Bin Yu** (UC Berkeley)

Plenary talk: Rietz Lecture

Monday 9:25 – 10:15, McLeod Auditorium

*Theory to gain insight and inform practice*

**Abstract:** Henry L. Rietz, the first president of IMS, published his book “Mathematical Statistics” in 1927. One review wrote in 1928:

“Professor Rietz has developed this theory so skillfully that the ‘workers in other fields’, provided only that they have a passing familiarity with the grammar of mathematics, can secure a satisfactory understanding of the points involved.”

In this lecture, I would like to promote the good tradition of mathematical statistics as expressed in Rietz’s book in order to gain insight and inform practice. In particular, I will recount the beginning of our theoretical study of dictionary learning (DL) as part of a multi-disciplinary project to “map a cell’s destiny” in *Drosophila* embryo. I will share insights gained regarding local identifiability of primal and dual formulations of DL. Furthermore, comparing the two formulations is leading us down the path of seeking confidence measures of the learned dictionary elements (corresponding to biologically meaningful regions in *Drosophila* embryo). Finally, I will present preliminary work using our confidence measures to identify potential knockout (or gene editing) experiments in an iterative interaction between biological and data sciences.

- **Shinsheng Robert Yuan** (Institute of Statistical Sciences, Academia Sinica)

Invited session: Analysis of next generation sequencing data for biomedicine [35]

Tuesday 10:45 – 11:15, Medical Sciences 3153

*Amplicon-based analysis suite for regional sequencing data*

Coauthors: Yu-cheng Li (Department of Engineering, National Taiwan University), Ker-Chau Li (Institute of Statistical Sciences, Academia Sinica)

**Abstract:** With increasing advancement of sequencing technology, more sequencing data are available. The next-generation sequencing allows us to sequence with higher throughput, cheaper cost and shorter experimental time. The new sequencing technology is often called targeted sequencing or amplicon-sequencing. The primers were designed to target the regions of interests in the genome and their reads often share the same starting and end position. This is different from whole genome sequencing whose the start and end position of reads are randomly distributed. Despite the small differences of the design, it has a great impact on many analyses such as variant callings, mutation frequency estimation, copy number estimation and so on.

**Z**

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Ofer Zeitouni** (Weizmann Institute of Science & Courant Institute)

Plenary talk: Schramm Lecture

Thursday 8:30 – 9:20, McLeod Auditorium

*Extremes in logarithmically correlated fields*

**Abstract:** It was recently proved that the maximum of Gaussian logarithmically correlated fields (of which the Gaussian free field in two dimensions is a prime example) converges in distribution to a randomly shifted Gumbel distribution. I will review the link between this result and the corresponding one for branching random walks, and then will describe recent progress and challenges in obtaining similar results for certain non-Gaussian logarithmically correlated fields; a specific example of the latter is the maximum of the modulus of the characteristic polynomial of a random unitary matrix. The talk is based in part on joint works with Bramson, Ding, Paquette and Roy.

- **Dapeng Zhan** (Michigan State University)

Invited session: Conformally invariant processes [8]

Friday 2:35 – 3:05, Medical Sciences 3153

*Green's functions for SLE curves*

Coauthor: Mohammad Rezaei (Michigan State University)

**Abstract:** Let  $\gamma$  be a Schramm-Loewner evolution (SLE) curve with parameter  $\kappa \in (0, 8)$  in a domain  $D$ . Let  $z_1, \dots, z_n$  be distinct points in  $D$ . The Green's function for  $\gamma$  valued at  $(z_1, \dots, z_n)$  is defined by

$$G(z_1, \dots, z_n) = \lim_{r_1, \dots, r_n \downarrow 0} \prod_{j=1}^n r_j^{d-2} \mathbb{P}[\text{dist}(\gamma, z_j) < r_j, 1 \leq j \leq n],$$

where  $d = 1 + \frac{\kappa}{8}$  is the Hausdorff dimension of  $\gamma$ , provided that the limit converges. We prove that the Green's function for any number of points exist and are Hölder continuous. We also get the convergence rates and sharp bounds for the Green's functions in terms of some simple functions. The talk is based on a joint project with Mohammad Rezaei.

- **Anru Zhang** (University of Wisconsin-Madison)

Invited session: Inference for high-dimensional covariance structures [29]

Tuesday 2:35 – 3:05, Medical Sciences 3153

*Rate-optimal perturbation bounds for singular subspaces with applications to high-dimensional statistics*

**Abstract:** Perturbation bounds for singular spaces, in particular Wedin’s  $\sin \Theta$  theorem, are a fundamental tool in many fields including high-dimensional statistics, machine learning, and applied mathematics. In this paper, we establish separate perturbation bounds, measured in both spectral and Frobenius  $\sin \Theta$  distances, for the left and right singular subspaces. Lower bounds, which show that the individual perturbation bounds are rate-optimal, are also given.

The new perturbation bounds are applicable to a wide range of problems. In this paper, we consider in detail applications to low-rank matrix denoising and singular space estimation, high-dimensional clustering, and canonical correlation analysis (CCA). In particular, separate matching upper and lower bounds are obtained for estimating the left and right singular spaces. To the best of our knowledge, this is the first result that gives different optimal rates for the left and right singular spaces under the same perturbation. In addition to these problems, applications to other high-dimensional problems such as community detection in bipartite networks, multidimensional scaling, and cross-covariance matrix estimation are also discussed.

- **Cun-Hui Zhang** (Rutgers University)

Invited session: Inference for high-dimensional covariance structures [29]

Tuesday 2:00 – 2:30, Medical Sciences 3153

*Incoherent tensor norms and their applications in higher order tensor completion*

Coauthor: Ming Yuan (University of Wisconsin-Madison)

**Abstract:** In this paper, we investigate the sample size requirement for a general class of nuclear norm minimization methods for higher order tensor completion. We introduce a class of tensor norms by allowing for different levels of coherence, which allows us to leverage the incoherence of a tensor. In particular, we show that a  $k$ th order tensor of rank  $r$  and dimension  $d \times \dots \times d$  can be recovered perfectly from as few as  $O((r^{(k-1)/2}d^{3/2} + r^{k-1}d)(\log(d))^2)$  uniformly sampled entries through an appropriate incoherent nuclear norm minimization. Our results demonstrate some key differences between completing a matrix and a higher order tensor: They not only point to potential room for improvement over the usual nuclear norm minimization but also highlight the importance of explicitly accounting for incoherence, when dealing with higher order tensors.

- **Jiwei Zhao** (State University of New York at Buffalo)

Contributed Talk, session [86]

Wednesday 3:10 – 3:40, Rosebrugh 208

~~*Variable selection with nonignorable missing data*~~ TALK CANCELLED

**Abstract:** Variable selection methods are well developed for a completely observed data set in the past two decades. In the presence of missing values, those methods need to be tailored to different missing data mechanisms. In this paper, we focus on a flexible and generally applicable missing data mechanism, which contains both ignorable and nonignorable missing data mechanism assumptions. We show how the regularization approach for variable selection can be adapted to the situation under

this missing data mechanism. The computational and theoretical properties for variable selection consistency are established. The proposed method is further illustrated by comprehensive simulation studies, for both low and high dimensional settings.

- **Chen Zhou** (Bank of The Netherlands and Erasmus University Rotterdam)

Contributed Talk, session [73]

Wednesday 2:00 – 2:30, Fitzgerald 139

*Nonparametric trend in extreme value indices*

Coauthor: Laurens de Haan (Erasmus University Rotterdam)

**Abstract:** Let  $\{X_i\}_{i=1,2,\dots,n}$  be independent random variables. Assume that the distribution function of each  $X_i$ ,  $F_i$ , is in the domain of attraction of some extreme value distribution. The extreme value index corresponding to  $F_i$  is denoted as  $\gamma(i/n)$ , where  $\gamma(s)$  is a continuous positive function defined on  $[0, 1]$ . Under appropriate conditions on these distribution functions we can estimate  $\gamma(s)$  locally. A global estimator of the integrated  $\gamma(s)$ -function can also be established. This allows for accurate estimation and testing regarding the  $\gamma(s)$  function. The asymptotic normality for both local and global estimators are shown. The main tool in the proof is the extension of the tail empirical process: we derive asymptotic properties of the moments of the error terms in tail empirical processes.

- **Quan Zhou** (Baylor College of Medicine)

Contributed Talk, session [45]

Monday 3:10 – 3:40, Medical Sciences 4171

*The value of foresight*

Coauthors: Philip Ernst (Rice University), L. C. G. Rogers (University of Cambridge)

**Abstract:** Suppose you have one unit of stock, currently worth 1, which you must sell before time  $T$ . The Optional Sampling Theorem tells us that whatever stopping time we choose to sell, the expected discounted value we get when we sell will be 1. Suppose however that we are able to see a units of time into the future, and base our stopping rule on that; we should be able to do better than expected value 1. But how much better can we do? And how would we exploit the additional information? The optimal solution to this problem will never be found, but in this paper we establish remarkably close bounds on the value of the problem, and we derive a fairly simple exercise rule that manages to extract most of the value of foresight.

- **Shuheng Zhou** (University of Michigan)

Contributed Talk, session [91]

Wednesday 3:10 – 3:40, Medical Sciences 3278

*Errors-in-variables models with dependent measurements* ADDED TALK

**Abstract:** I will discuss an errors-in-variables model where the covariates in the data matrix are contaminated with random noise. This model is significantly different from those analyzed in the literature in the sense that we allow the measurement error for each covariate to be dependent across observations. Such error structures appear in the science literature, for example, when modeling the trial-to-trial fluctuations in response strength shared across a set of neurons. We provide theory,

real-data examples and simulation evidence showing that we can recover the regression coefficients for a class of errors-in-variables problems. This is joint work with Mark Rudelson.

- **Xiang Zhou** (University of Michigan)

Invited session: Biomedical applications [20]

Wednesday 3:10 – 3:40, Medical Sciences 3154

*Variance component estimation with summary statistics in genome-wide association studies*

**Abstract:** Linear mixed models (LMMs) are among the most commonly used tools for genetic association studies. However, the standard method for estimating variance components in LMMs – the restricted maximum likelihood estimation method (REML) – suffers from several important drawbacks: REML requires individual-level genotypes and phenotypes, is computationally slow, and produces biased estimates in case control studies. To remedy these drawbacks, we present an alternative framework for variance component estimation based on summary statistics, which we refer to as MQS. MQS is based on the method of moments (MoM) and the minimal norm quadratic unbiased estimation (MINQUE) criterion, and brings two seemingly unrelated methods – the renowned Haseman-Elston (HE) regression and the recent LD score regression (LDSC) – into the same unified statistical framework. With this new framework, we provide an alternative but mathematically equivalent form of HE that allows for the use of summary statistics. We provide an exact estimation form of LDSC to yield unbiased and statistically more efficient estimates. Our framework also provides asymptotic variance forms for both HE and LDSC, avoiding the need of commonly used jackknife re-sampling procedures that rely on individual independence or block-wise SNP independence assumptions. A key feature of our method is that it can effectively use a small random subset of individuals for computation while still producing estimates that are almost as accurate as if the full data were used. As a result, our method produces unbiased and statistically efficient estimates with calibrated standard errors, requires only summary statistics, while it is computationally efficient for large data sets. Using simulations and applications to 33 phenotypes from 7 real data sets, we illustrate the benefits of our method for estimating and partitioning SNP heritability. Our method is implemented in the GEMMA software package, freely available at [www.xzlab.org/software.html](http://www.xzlab.org/software.html).

## ABSTRACTS OF POSTERS

A	B	C	D	E	F	G	H	I	J	K	L	M	Plenary Talks
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Posters

- **Helton Graziadei** (Department of Statistics - University of Sao Paulo)

Poster session

Thursday 10:15 – 4:15, McLeod Auditorium Lobby

*A Bayesian counterpart to the marginal homogeneity test on contingency tables*

Coauthor: Lus Gustavo Esteves (Department of Statistics - University of Sao Paulo)

**Abstract:** Matched sample studies have become popular in a wide range of areas especially those

dealing with categorical variables in several occasions. For instance, suppose that for a group of individuals, the quality of the vision in each eye is classified into four categories. In this context, it is usual to investigate whether the qualities of left and right eyes are the same or, equivalently, if the marginal distributions of the qualities of vision are the same, the so-called Marginal Homogeneity (MH) hypothesis. Classical approaches to the problem of testing MH rely on the asymptotic distribution of the test statistics which may yield imprecise results in certain situations. To overcome these limitations, we develop the Full Bayesian Significance Test (FBST) for MH in two-dimensional contingency tables. The FBST is a procedure that has some important features such as: (a) it does not rely on asymptotic distributions (b) it does not depend on the elimination of nuisance parameters (c) it is consistent in simultaneous tests. Furthermore, we calculate p-values and compare them with the FBST. To summarize, we propose a coherent measure of evidence to test MH and compare it with classical approaches to the problem.

- **Masaaki Imaizumi** (University of Tokyo)

Poster session

Friday 10:15 – 4:15, McLeod Auditorium Lobby

*Regression with infinite dimensional spaces by reproducing kernel Hilbert space approach*

**Abstract:** We develop and analyze an estimation method for a nonlinear regression function with functional covariates and responses. Such a problem mainly appears in the field of functional data analysis and it has various application topics. Estimating the nonlinear regression function with infinite dimensional inputs and outputs is still a developing problem. We propose an estimation method based on the theory of the reproducing kernel Hilbert space (RKHS). We consider an RKHS equipped with a certain kernel, which represents some smoothness and additional conditions for restricting complexity of the regression functions. We present an analysis of entropy bound and convergence rate for the estimator. Given some conditions on the settings of the regression problem, we show that the convergence rate of our estimator depends on regularities of operators and arguments.

- **Young Kyung Lee** (Kangwon National University)

Poster session

Saturday 10:15 – 4:15, McLeod Auditorium Lobby

*Operational time and in-sample density forecasting*

Coauthors: Enno Mammen (Heidelberg University), Byeong U. Park (Seoul National University)

**Abstract:** In this paper we consider a new structural model for in-sample density forecasting. In-sample density forecasting is to estimate a density function on a region where the density is observed and then re-use the estimated density while estimating the density on a region it is not observed. Our structural assumption is that the density is a product of one-dimensional functions with one function sitting on the scale of a transformed space of observations. The transformation involves another unknown one-dimensional function, so that our model is formulated via a known smooth function of three underlying unknown one-dimensional functions. We present an innovative way of estimating the one-dimensional functions and show that all the estimators of the three components achieve the optimal one-dimensional rate of convergence. We illustrate how one can use our approach

by analyzing two real datasets from insurance and epidemiology. We also investigate the finite sample performance of the method via a simulation study.

- **Hiroshi Matsuzoe** (Nagoya Institute of Technology)

Poster session

Sunday 10:15 – 4:15, McLeod Auditorium Lobby

*Generalized expectations for deformed exponential families*

**Abstract:** A deformed exponential family is a generalization of exponential family. Such a statistical model was introduced in non-extensive statistical physics, and it is applied to anomalous statistical phenomena such as economics, disaster statistics, etc. Since a deformed exponential family includes long tail distributions or fat tail distributions, expectations or variances of random variables may not be defined for this statistical model.

In this presentation, we discuss generalizations of expectation functionals. From an arguments of information geometry, we can find that it is natural to generalize the notion of expectation for non-exponential type probability distributions. In addition, we apply generalized expectations for construction of divergence functions. As a consequence, we elucidate relations among KL-, alpha-, and beta-divergences from viewpoints of generalized expectations and unbiasedness of estimating functions.

- **Debopriya Mukherjee** (Indian Institute of Science Education and Research, Thiruvananthapuram)

Poster session

Monday 10:15 – 4:15, McLeod Auditorium Lobby

*Stochastic analysis of parabolic integro-differential equations perturbed by Lévy noise*

**Abstract:** In this presentation, global existence and uniqueness of strong solutions for the abstract parabolic integro-differential equations influenced by Lévy noise will be established exploiting local monotonicity arguments. The abstract theory will then be applied to specific problems related to Viscoelastic fluid flows, such as Oldroyd B fluid and Maxwell fluid.

- **Jasdeep Pannu** (Lamar University)

Poster session

Tuesday 10:15 – 4:15, McLeod Auditorium Lobby

*Robust functional weighted LAD-adaptive group LASSO*

Coauthor: Nedret Billor (Auburn University)

**Abstract:** Functional data has become increasingly frequent and important in diverse fields of sciences, engineering, and humanities, in the last two decades. Data recorded densely over time is called functional data. Variable selection is an important problem in functional regression analysis, just as in ordinary regression analysis. In this study, our focus is on a functional linear regression model with functional predictors and scalar responses where functional coefficients of functional predictors are to be estimated. Since a functional dataset has issues of high dimensionality and multicollinearity inherently, estimation of functional parameters is complicated. Further, the presence of outlying curves adds extra complication to the estimation of the functional regression coefficients. The objective of this research is to develop a robust functional variable selection method based on  $L^1$  regularization.

Recently, a method called functional LAD-groupLASSO (LAD-gLASSO) has been developed which uses the combination of the well known robust loss function least absolute deviation (LAD) and penalty function group LASSO, where the functional parameters are estimated and selected through the minimization of the sum of the absolute value of the errors and penalizing the parameter functions. However, it is well-known that the LAD based method is only resistant to the outlier in the response variable, but not resistant to the outliers in the explanatory variables (leverage points). To address this problem we propose a new method called functional Weighted LAD-Adaptive groupLASSO (WLAD-agLASSO). This method uses WLAD function to downweight the leverage points and Adaptive groupLASSO penalty criterion to assign different weights to different coefficients to penalize them differently.

- **Heewon Park** (Yamaguchi University)

Poster session

Wednesday 10:15 – 4:15, McLeod Auditorium Lobby

*Interaction based feature selection for driver gene selection via copy number driven expression levels*  
Coauthors: Atushi Niida (Health Intelligence Center, The University of Tokyo), Seiya Imoto (Health Intelligence Center, The University of Tokyo), Satoru Miyano (Human Genome Center, The University of Tokyo)

**Abstract:** Much research on  $L^1$ -type regularization is currently underway for high dimensional data analysis. Especially, the  $L^1$ -type regularization methods have been widely used to reveal the complex system of cancer based on genomic data analysis, because the genomic data is a typical high dimensional data. Although the  $L^1$ -type approaches have been widely used in bioinformatics area, the existing studies apply the method without consideration of biological knowledge. We develop statistical strategies for cancer driver gene selection incorporating causal interaction of the comprehensive genome-scale information. To incorporate interaction between expression levels and copy number variation, which is a crucial symptom to reveal heterogeneous system of cancer, we propose a novel  $L^1$ -type regularization in line with the adaptive  $L^1$ -type approaches. Numerical studies show that the proposed method is effective for high-dimensional regression modeling, especially for feature selection.

- **Joonha Park** (University of Michigan)

Poster session

Thursday 10:15 – 4:15, McLeod Auditorium Lobby

*Particle filter in high dimension via intermediate resampling and its application to analysis of the Measles epidemic in the UK, 1949-1964*

Coauthor: Edward Ionides (University of Michigan)

**Abstract:** Particle filter has formed the basis of most inference methods for Partially Observed Markov Processes (POMP). However, it is well known that the performance of particle filter deteriorates as the underlying state dimension increases. This so called ‘curse-of-dimensionality’ has been the main limiting factor in large scale inference. A novel particle filter method that performs well in high dimension is presented. This method is a statistically consistent algorithm with the

‘plug-and-play’ property, which offers key practical value from the modelling perspective. Both theory and experiment demonstrate that the method scales much better than standard methods. This method was applied to a stochastic model of the Measles dynamics in the UK in the years 1949-1964. The new particle filter made it possible to analyze epidemic coupling between multiple cities from spatio-temporal data with proper likelihood-based statistical method.

- **Cesar Alberto Rosales-Alcantar** (Universidad de Sonora)

Poster session

Friday 10:15 – 4:15, McLeod Auditorium Lobby

*Comparison between Ito and Stratonovich calculus: which is better to use in mathematical modelling?*

Coauthor: Oscar Vega-Amaya (Universidad de Sonora)

**Abstract:** The principal aim of this poster is to present a nice comparison between the most famous stochastic calculus: Itô and Stratonovich. This aim will be realized studying the conditions for extinction of the Stochastic Malthusian Model using both calculus.

When we consider this model with density-independent growth, the solution using the Stratonovich calculus give the same results of the extinction behaviour than it deterministic version. But, when this model is solved with the Itô calculus, significant differences appears in the results of the extinction behaviour.

In the order to get an answer of the reason about this differences, we proceed to give a proof in terms of the interpretation of the parameters. Also, we sketch the proof when the growth factor in the model present conditions of density-dependence. To finally this poster, we present a similar analysis for a stochastic harvesting model.

- **Merveille Koissi Savi** (Laboratory of Biomathematics and Forest Estimation)

Poster session

Saturday 10:15 – 4:15, McLeod Auditorium Lobby

*Toward a revisiting of permutation methods when the exact probability value is used: case of permutation of residuals in one way ANOVA framework*

**Abstract:** Permutation tests constitute the best alternative to the traditional Analysis of Variance (ANOVA) when it fails in the fulfilment of parametrical assumptions. Three models were considered as the best permutation methods after different simulation studies. These permutation methods are residuals permutation under full, reduced and modified model. All these three methods, a part from the residuals that is common for them, they basically share the same procedure of probability value computation. This probability value computation generally leads to the inflated behaviour of test. The current paper addresses (1) the issue of the implementation of exact probability computation in these three permutation methods and (2) assessing the relative performance of these three residuals permutation. The first part of this document clarifies the concept of permutation and the different procedures available in the literature as well as their implementation in the free software R. The objectives (1) and (2) were reached through Monte Carlo simulation study. A total of 198 simulations were run under the unique scenario of balanced and homoscedastic design. For each simulation 1000 data sets were generated and 999 time permutations were done on each data set residual. The

outcome of these simulations showed that, when the exact p-value is used, the behaviour of the residuals permutation tests changes. And when data follow lognormal distribution it is advisable the use of permutation of residuals under reduced model.

- **Bizhan Shabankhani** (Dep of Biostatistics, Faculty of Health, Mazandaran University of Medical Sciences. Sari, IRAN)

Poster session

Sunday 10:15 – 4:15, McLeod Auditorium Lobby

*Survival factors in patients with end-stage renal disease in Mazandaran province, Iran*

Coauthors: Bizhan Shabankhani (Dept of Biostatistics, Faculty of Health, mazandaran university of medical sciences. Sari, IRAN), Anoushirvan Kazemnejad (Department of Biostatistics, Faculty of Medical sciences, Tarbiat Modares University, Tehran, Iran)

**Abstract:** Survival analysis for patients with end-stage renal disease and factors influencing their survival is crucial due to the increase in the number of these patients along with their high mortality rate. This study aimed to analyse the survival rate of patients in north of Iran undergoing hemodialysis and to assess factors influencing their survival.

A historical cohort study was conducted on 500 patients on maintenance hemodialysis in 3 hospitals of 2 cities (Sari and Babol) in Mazandaran province during a 6-year period from 2007 to 2013. The Cox regression analysis was used to assess the impact of sex, age, education, smoking habit, primary cause of kidney failure, living with family, cardiovascular diseases, weight, age at diagnosis, and age at initiating hemodialysis on survival of the patients.

The median survival time for the 500 hemodialysis patients was 108 months. Death occurred in 174 patients (34.8%). History of smoking, age, being unemployed, being illiterate, and renal cyst, congenital diseases, and unspecified diseases as the cause of kidney failure were the associated factors with survival of the patients. The 1-, 2-, 3-, 5-, 10-, and 12-year survival for these patients was estimated to be 84%, 77%, 71%, 58%, 43%, and 33%, respectively.

This study showed a high level of mortality and poor survival prognosis for patient undergoing maintenance hemodialysis. History of smoking, age, being unemployed, being illiterate, and renal cyst, congenital diseases, and unspecified conditions as the cause of kidney failure were the associated factors with survival of these patients.

- **Takaaki Shimura** (The Institute of Statistical Mathematics)

Poster session

Monday 10:15 – 4:15, McLeod Auditorium Lobby

*Variety of distributions with asymptotically logarithmic truncated mean*

Coauthor: Toshio Nakata (University of Teacher Education Fukuoka)

**Abstract:** The tail behaviors of the one-sided Cauchy distribution and the Peter and Paul distribution in the St. Petersburg game are different. Nevertheless, their truncated means are asymptotically equal (logarithmic). We discuss the variety of distributions with same asymptotic truncated mean.

- **Issah Nazif Suleiman** (Gazi University)

Poster session

Tuesday 10:15 – 4:15, McLeod Auditorium Lobby

*Nonparametric Bayesian approach to change point detection*

Coauthor: ()

**Abstract:** This research proposes a new Nonparametric Bayesian Change point detection approach which in contrast to the Markov approach of Chib (1998) uses the Dirichlet process prior to allow an integrative transition of probability from the posterior distribution. Although the Bayesian non-parametric technique on the mixture does not serve as an automated tool for the selection of the number of components in the finite mixture. Many research has been done proposing the prior for flexible estimation of the mean function, using some basis function (splines or wavelets) (Denison et al 2002). We realize most of these research assumes constant residual density within a scale factor which introduces heteroscedasticity in the process. The Bayesian nonparametric mixture shows a misspecification model properly which has been explained further in the methodology. This research shows the principal step-by-step algorithm using nonparametric Bayesian technique with the Dirichlet process prior defined on the distribution to the detection of change point. This technique is rooted from model selection which will allow us determining both the number and location of change points. The proposed approach is motivated by methods from cluster analysis. This approach can be further extended in the multivariate change point detection which will be studied in the near future.

- **Ferede Asena Tilahun** (Hawassa University)

Poster session

Wednesday 10:15 – 4:15, McLeod Auditorium Lobby

*Bayesian semi-parametric regression analysis of childhood malnutrition in Gamo Gofa zone:-the socio-economic impact of child undernutrition.*

Coauthors: Derbachew Asfaw Teni, Arba Minch University, Email: dasfaw469@gmail.com (Hawassa University), Derbachew Asfaw Teni (Arba Minch University)

**Abstract:** Major progress has been made over the last decades in reducing the prevalence of malnutrition amongst children less than 5 years of age in developing countries. However, approximately 27% of children under the age of 5 in these countries are still malnourished. This work focuses on the childhood malnutrition in in Gamo Gofa Zone, Ethiopia. This study examined the association between demographic and socioeconomic determinants and the malnutrition problem in children less than 5 years of age using Data obtained from both rural and urban sampled surveys conducted in both Dita and Mirab Abaya Weredas from December 1 to January 5, 2013. The study on the Child undernutrition and underweight prevalence in Gamo Gofa has allowed us to quantify the negative impacts of child undernutrition in both social and economic terms. Today, more than 4 out of every 7 children in Ethiopia are stunted. The result revealed that as many as 75% of all cases of child undernutrition and its related pathologies go untreated. It is also observed that about 35% of the health costs associated with undernutrition occur before the child turns 1 year-old. Generally, The results of the analysis show that place of residence, employment status of mother, employment status of partners, age of the child, educational status of mothers, diarrhea, household economic level and

source of drinking water were found to be the most important determinants of health/nutritional status of children. The study revealed that socio-economic, demographic and health and environmental variables have significant effect on the nutritional and health status of children in Ethiopia. The study revealed that children from employed mothers are at a higher risk of health problem and malnutrition.

Keyword: Bayesian Models, Childhood Malnutrition, Gamo Gofa Zone, Ethiopia