EURANDOM 25th anniversary

Program | April 15th – 17th, 2024 | TU/e

Day 1: Monday April 15th

Location: MetaForum building room 11 and 12

9:30-9:40h	Welcome and Opening	
9:40-10:30h	**Keynote 1** Peter Bühlmann (*ETH Zürich)* [chair: Geurt Jongbloed]	Causality-inspired Statistical Machine Learning
10:30-11:00h	*Break*	
11:00-12:30h	**Session 1** "*Latest Theoretical Developments for Neural Networks*" [chair: Botond Szabo]	
	Sophie Langer (*University of Twente)*	The role of statistical theory in understanding deep learning
	Frans Oliehoek (*TU Delft*)	Learning and using (partial) simulators for controlling complex environments
	Johannes Schmidt-Hieber (*University of Twente*)	Statistical learning in biological neural networks
12:30-13:40h	*Lunch*	
13:40-14:30h	**Keynote 2** Hansjörg Albrecher (*University of Lausanne*) [chair: Onno Boxma]	Markov, Matrices and EURANDOM
14:30-14:50h	*Break*	
14:50-16:20h	**Session 2** "*Collective Behavior of Interacting Stochastic Systems*" [chair: Remco van der Hofstad]	
	Ellen Baake (*Bielefeld University*)	The multiple coupon collection process and Markov embedding
	Cristian Giardina (*University of Modena and Reggio Emilia*)	The multifacet Ising model on random graphs
	Silke Rolles (*TU Munich*)	The vertex-reinforced jump process
16:20-16.40h	*Break*	
16.40-17.40h	**Session 3** "*Future Workshop Formats in a Changing World*" [chair: Nelly Litvak]	
17:40-19:15h	Drinks and snacks | Connection game	

Day 2: Tuesday April 16th

Location: MetaForum building room 11 and 12

9:10-10:00h	**Keynote 3** Zbigniew Palmowski (*University of Wroclaw*) [chair: Stella Kapodistria]	Stationary states and exit times for Lévy processes with partial resetting
10:00-10:30h	*Break*	
10:30-12:00h	**Session 4** "*Statistics in the Machine Learning Era*" [chair: Johan Segers]	
	Anne Sabourin (*Université Paris Cité / CNRS*)	On regression in extreme regions
	Johannes Lederer (*Ruhr University Bochum)*	High-Dimensional Extremes
	Rémi Leluc (*École polytechnique*)	Monte Carlo methods in Machine Learning
12:00-13:10h	*Lunch*	
13:10-14:00h	**Keynote 4** Nina Gantert (*TU Munich*) [chair: Silke Rolles]	Some random walks in dynamical random environments
14:00-14:20h	*Break*	
14:20-15:50h	**Session 5** "*Memories and Perspectives*" [chair: Bernd Heidergott] • Intro presentation: Wim Senden & Frank den Hollander • Silke Rolles • Zbignieuw Palmowski • Josine Bruin • Matthias Löwe	
16:10-20:30h	Evoluon excursion and dinner	

Day 3: Wednesday April 17th

Location: MetaForum building room 11 and 12

Statistical Theory for Sound Data Analysis

(parallel track, organizers: Geurt Jongbloed, Johan Segers, Botond Szabo)

9:00-9:10h	*Welcome and Opening*	
9:10-10:40h	Gerda Claeskens (*KU Leuven*)	Valid selective inference for parameters in a model selected by the group lasso method
	Fabian Mies (*TU Delft*)	Likelihood asymptotics for stationary Gaussian arrays
	Paulo Serra (*VU Amsterdam*)	Uncertainty Quantification in Sparse Quantile Regression
10:40-11:10h	*Break*	
11:10-12:40h	Jelle Goeman (*Leiden University*)	Optimal All-Resolutions Inference using Closed Testing
	Ingrid van Keilegom (*KU Leuven*)	Copula Based Cox Proportional Hazards Models For dependent censoring
	Moritz Schauer (*University of Gothenburg*)	Causal structure learning and sampling using Markov Monte Carlo with momentum
12:40-13:50h	*Lunch*	
13:50-14:50h	Rui Castro (*Eindhoven University of Technology*)	Distribution free anomaly detection when monitoring a large number of units: a permutation/rank-based higher criticism approach
	Davy Paindaveine (*University of Brussels*)	Inference on location for noisy directional data

Random Spatial Structures: New Developments and Perspectives

(parallel track, organizers: Cristian Giardina, Remco van der Hofstad, Silke Rolles)

9:00-9:10h	*Welcome and Opening*	
9:10-10:40h	Markus Heydenreich (*University of Augsburg*)	Prudent walk in dimension 6 and higher
	Artem Sapozhnikov (*Leipzig University*)	Brownian motions, hemiballs and Brownian interlacements
10:40-11:10h	*Break*	
11:10-12:40h	Christoph Külske (*Ruhr University Bochum*)	Extremal decomposition and branch overlaps for free states of low temperature finite-spin models on trees
	Arnaud le Ny (*Université Paris Est Créteil | CNRS*)	Metastates for Long-Range Ising Models with Random Boundary Conditions
12:40-13:50h	*Lunch*	
13:50-15:20h	Nicolas Petrelis (*Nantes Université*)	Multi-sites localization of a weakly self-avoiding walk in a heavy tailed potential
	Jan Swart (*UTIA Prague*)	Universality of the Brownian net
15:20-15:40h	*Break*	
15:40-16:25h	Maria Deijfen (*Stockholm University*)	Stable matching on the complete graph

CONTENTS

ABSTRACTS

ANNE SOUBARIN

ON REGRESSION IN EXTREME REGIONS

In the classic regression problem, the value of a real-valued random variable Y is to be predicted based on the observation of a random vector X, taking its values in $\mathbb{R}^d$ with $d\geq 1$ say.
The statistical learning problem consists in building a predictive function $\hat{f}:\mathbb{R}^d\to \mathbb{R}$ based on independent copies of the pair (X,Y) so that Y is approximated by $\hat{f}(X)$ with minimum error in the mean-squared sense. Motivated by various applications, ranging from environmental sciences to finance or insurance, special attention is paid here to the case of extreme (\textit{i.e.} very large) observations X. Because of their rarity, they contribute in a negligible manner to the (empirical) error and the predictive performance of empirical quadratic risk minimizers can be consequently very poor in extreme regions. In this paper, we develop a general framework for regression in the extremes. It is assumed that X's conditional distribution given Y belongs to a non parametric class of heavy-tailed probability distributions. It is then shown that an asymptotic notion of risk can be tailored to summarize appropriately predictive performance in extreme regions of the input space. It is also proved that minimization of an empirical and non asymptotic version of this 'extreme risk', based on a fraction of the largest observations solely, yields regression functions with good generalization capacity. In addition, numerical results providing strong empirical evidence of the relevance of the approach proposed are displayed.

ARNOUD LE NY

METASTATES FOR LONG-RANGE ISING MODELS WITH RANDOM BOUNDARY CONDITIONS

Metastates, measure-theoretic objects introduced in the early 90's by Newman and Stein to provide a framework to investigate the arising of chaotic size dependence in spin-glasses, in general are (random) measures on a product space of disorder variables and spin variables, describing the weight distribution over the possible limit Gibbs measures, whether obtained as a kind of translation averages (*Newmann-Stein metastates*) or via conditioning (*Aizenman-Wehr metatstates*). Due to the randomness of the Gibbs measures in the support of the metastate(s), in general the metastate(s) thus become measure(s) on measures on measures…

Although the notions of chaotic size dependence and metastates have been developed to describe spin glasses, there suffers from a lack of tractable non-mean-field spin-glass models and toy examples, such as ferromagnets with random boundary conditions, can provide instructive illustrations of these notions. Their marginal on the disorder variables is the disorder measure, whereas conditioned on the disorder, they provide a measure on the Gibbs measures for that disorder realisation. In this talk, we shall describe how the work of van Enter *et al* on 2d-Ising models the early 2000's, where they exhibit a dispersed metastates oscillating between the two pure extremal states, could be extend in the context of infinite-range polynomially decaying potentials in order to exhibit an example of metatstates concentrating continuously on the set of all Gibbs measures (including mixed, convex combinations, Gibbs measures) for a particular range of decays.

Joint work with Aernout van Enter (Groningen) and Eric O. Endo (NYU Shanghai)

ARTEM SAPOZHNIKOV

BROWNIAN MOTIONS, HEMIBALLS AND BROWNIAN INTERLACEMENTS

We discuss sharp uniform bounds on the probability that several independent Brownian motions in dimensions d>=3 jointly hit all hemiballs of a given ball and at the same time avoid a slightly smaller concentric ball. We then relate this event to the existence of multiple connected components in the restriction of the vacant set of the union of several independent Wiener sausages to a tiny ball. Finally, we talk about Brownian interlacements, which is a continuum analogue of Sznitman's random interlacements, and explain how the above estimates are used in the proof of the uniqueness of the infinite connected component of its vacant set.

Joint work with Yingxin Mu.

CHRISTOF KÜLSKE

EXTREMAL DECOMPOSITION AND BRANCH OVERLAPS FOR FREE STATES OF LOW TEMPERATURE FINITE-SPIN MODELS ON TREES

Consider the free state of low temperature Potts models, and more generally of ferromagnetic finite-spin models, on a Cayley tree. We investigate the corresponding extremal decomposition measure. We show that it is an atomless measure which is supported on uncountably many inhomogeneous extremal states.

Our proof is based on concentration bounds for branch overlaps. The method is robust, and so the results extend to the counterpart of the free state (called central state) for a range of models which have no symmetry, allowing also non-zero external field terms.

Joint work with Loren Coquille, and Arnaud Le Ny

CRISTIAN GIARDINA

THE MULTIFACET ISING MODEL ON RANDOM GRAPHS

The ferromagnetic Ising spin model is often used to model second-order phase transitions and the continuous emergence of order. We consider this model on a random graph, where the additional randomness provided by the graph gives a rich picture with a host of surprises. We identify similarities and differences between the quenched and the annealed Ising models that are obtained, respectively, by fixing or averaging over the graph realizations. We find that the annealed critical temperature is highly model-dependent, even in the case of graphs that are asymptotically equivalent (such as different versions of the simple Erdös-Rényi random graph). The quenched critical temperature is instead the same for all locally tree-like graphs. Moreover, in the presence of inhomogeneities that produce a fat-tail degree distribution, the difference between quenched and annealed becomes even more substantial, leading in some cases to different universality classes and different critical exponents. The annealed properties depend sensitively on whether the total number of edges of the underlying random graph is fixed, or is allowed to fluctuate. If time allows preliminary results on the annealed Potts model, displaying a first-order phase transition, will also be discussed.

This talk is based on several joint works with Hao Can, Sander Dommers, Claudio Giberti, Remco van der Hofstad and Maria Luisa Prioriello.

The preliminary work on Potts models also involves Neeladri Maitra and benefited from discussions with Guido Janssen.

DAVY PAINDAVEINE

INFERENCE ON LOCATION FOR NOISY DIRECTIONAL DATA

We introduce parametric models for noisy directional data, in which a radial noise with magnitude σ^2 makes the observations deviate from their theoretical hyperspherical sample space, namely a hypersphere centered at θ and with radius r. We consider inference --- hypothesis testing, point estimation, and confidence zone estimation --- on the location parameter θ, in a framework where both r and σ^2 remain unspecified. We introduce several asymptotic scenarios in which the radius of the hypersphere and, most importantly, the noise magnitude may depend on the sample size n in an essentially arbitrary way. This allows us to consider very diverse cases, in which the a priori information that the data belong to a hypersphere is more and more, or on the contrary less and less, relevant. We base our investigation on Le Cam's asymptotic theory of statistical experiments and aim at a full understanding of the resulting limiting experiments. The corresponding contiguity rates, that characterize how easy/hard inference on θ is, reveal rather counter-intuitive results in some scenarios. We build locally asymptotically optimal tests and estimators, that turn out to be adaptively optimal across all asymptotic scenarios. We show that, in standard asymptotic scenarios, classical procedures that would ignore the hyperspherical a priori information are rate-consistent but do not achieve efficiency bounds, and that, in non-standard asymptotic scenarios, such classical procedures are not even rate-consistent. We investigate the finite-sample relevance of our results through Monte Carlo exercises.

ELLEN BAAKE

THE MULTIPLE COUPON COLLECTION PROCESS AND MARKOV EMBEDDING

The embedding problem of Markov transition matrices into Markov semigroups is a classic problem that regained a lot of impetus and activities in the last few years. We consider it here for the following generalisation of the classical coupon collection process: from a set $S = \{1,2,\ldots, m\}$ of distinct types of objects, a subset $K \subseteq S$ is drawn with probability p_K in every time step, $\sum_{K \subseteq S} p_K = 1$, and united with the set of types sampled so far. We obtain explicit conditions for the resulting discrete-time Markov chain to be representable as the semigroup of a continuous-time process and interpret the result in terms of the finite-sites mutation process in genetics.

This is joint work with Michael Baake.

FABIAN MIES

LIKELIHOOD ASYMPTOTICS FOR STATIONARY GAUSSIAN ARRAYS

Arrays of stationary Gaussian time series can arise naturally in econometric applications, e.g. as the discretization of continuous-time stochastic processes, or be introduced artificially to model persistency via so-called local-to-unity models, i.e. linear time series models with parameters close to a unit root. For the parametric statistical estimation of these stationary models, the spectral density plays a central role. In particular, classical results in time series analysis suggest that the Gaussian likelihood and Fisher information may be approximated in terms of the spectral density, and conditions for efficiency of the MLE have been formulated in the literature. Unfortunately, these general results do not cover arrays of time series. Our contribution is to show in which way the asymptotic likelihood theory needs to be adapted for the array case, and we demonstrate that this yields a straight-forward approach to study a broad class of processes.

As a motivating example, we investigate the estimation of the mixed fractional Brownian motion based on high-frequency observations. Our findings reveal that the achievable rates of convergence depend intricately on the size of the various components, as well as their intertemporal and crosstemporal dependence structure.

FRANS OLIEHOEK

LEARNING AND USING (PARTIAL) SIMULATORS FOR CONTROLLING COMPLEX ENVIRONMENTS

In reinforcement learning (RL), we develop techniques to learn to control complex systems, and over the last decade we have seen impressive successes ranging from beating grand masters in the game of Go, to real-world applications like chip design, power grid control, and drug design. However, nearly all applications of RL require access to an accurate and lightweight simulator from which huge numbers of trials can be sampled. In this talk, I will deal with settings where this is not the case.

Specifically, I will cover some work where good simulators are available (such as traffic control in a big city), but where simulating the full problem is too expensive. In the INFLUENCE project we have developed techniques that combine (supervised) deep learning with deep RL to improve scalability for such settings. I will also talk about some approaches, like MuZero, where such a simulator or "world model" is not available and needs to be learned, and will highlight some successes and limitations.

GERDA CLAESKENS

VALID SELECTIVE INFERENCE FOR PARAMETERS IN A MODEL SELECTED BY THE GROUP LASSO METHOD

We develop and study methods for inference after selection by group lasso estimators. The approach can be used with a wide class of distributions and loss functions, including applications with generalized linear models and quasi-ikelihood models for overdispersed count data. The use of group lasso allows for categorical or grouped covariates in the models, as well as for continuous covariates. Extra randomization is added to the optimization problem. This allows us to construct a post-selection likelihood which we show to be useful for obtaining selective inference when conditioning on the event of the selection of the grouped covariates. This likelihood provides too a selective point estimator, accounting for the selection by the group lasso. The confidence regions for the regression parameters in the selected model constructed via this method take the form of Wald-type regions and are shown to have bounded volume. The method compares favorably with other methods in simulations.

This is joint work with S. Pirenne (KU Leuven) and S. Panigrahi, Y. Huang (U. Michigan)

HANSJOERG ALBRECHER

MARKOV, MATRICES AND EURANDOM

After some scientific and personal reminiscences of EURANDOM over the years, I will talk about recent developments on matrix distributions and their connection to absorption times of inhomogeneous Markov processes. It will be discussed how and why such constructions are natural tools for modelling in applied probability, with a particular emphasis on non-life and life insurance applications. We also illustrate how certain extensions to the non-Markovian case involve fractional calculus and lead to matrix Mittag-Leffler distributions, which turn out to be a flexible and parsimonious class for modelling large but rare insurance loss events.

INGRID VAN KEILEGOM

COPULA BASED COX PROPORTIONAL HAZARDS MODELS FOR DEPENDENT CENSORING

Most existing copula models for dependent censoring in the literature assume that the parameter defining the copula is known. However, prior knowledge on this dependence parameter is often unavailable. In this article we propose a novel model underwhich the copula parameter does not need to be known. The model is based on a parametric copula model for the relation between the survival time (T) and the censoring time (C), whereas the marginal distributions of T and C follow a semiparametric Cox proportional hazards model and a parametric model, respectively. We show that this model is identified, and propose estimators of the nonparametric cumulative hazard and the finite-dimensional parameters. It is shown that the estimators of the model parameters and the cumulative hazard function are consistent and asymptotically normal. We also investigate the performance of the proposed method using finite-sample simulations. Finally, we apply our model and estimation procedure to a follicular cell lymphoma dataset. Supplementary materials for this article are available online.

JAN MEINDERTS SWART

UNIVERSALITY OF THE BROWNIAN NET

The Brownian web can informally be described as coalescing one-dimensional Brownian motions, started from every point in space and time. It arises as the scaling limit of nearest-neighbour coalescing random walks. The Brownian net is a closely related object that arises as the scaling limit of coalescing random walks that also branch with a small probability. It can be thought of as the "near-critical" version of the Brownian web and has some similarity to the scaling limit of near-critical two-dimensional percolation. In some sense, it can be viewed as a simpler, oriented version of the latter. The Brownian web and net are believed to be universal scaling limits for a variety of one-dimensional models. For the Brownian web, there are now several results that confirm this. In particular, thanks to the work of Belhaouari, Mountford, Sun, and Valle (2006), it is known that non-nearest-neighbour coalescing random walks also have the Brownian web as their scaling limit, provided their jump kernel has a finite "three plus epsilon" moment. The analogue result for the Brownian net has been conjectured since its introduction in 2008 but only recently have we been able to confirm this.

Joint work with Rongfeng Sun (Singapore) and Jinjiong Yu (Shanghai).

JELLE GOEMAN

OPTIMAL ALL-RESOLUTIONS INFERENCE USING CLOSED TESTING

Many modern fields of science gather data at a very fine resolution, but require inference at a higher aggregated level. For example, neuroimaging data are gathered at the level of 3 mm × 3 mm × 3 mm voxels, but the relevant biology happens at the level of cm-scale brain areas. Generally, there is not one single natural aggregation level for the inference. Too much aggregation may result in unnecessarily vague inference; more limited aggregation may result in low power. Fortunately, our recent work on closed testing has shown that any multiple testing method that is defined on one resolution level (and controls a tail probability of the number or proportion of errors there), can be naturally embedded into a closed testing method that looks at all resolution levels simultaneously. Remarkably, such additional inference is often non-trivial, and always comes for free, i.e., without loss of power at the original resolution level. The resulting all-resolutions simultaneous inference gives researchers unprecedented freedom of human-in-the-loop analysis, in which many analysis decisions may be postponed after seeing the data, without compromising error control. In this talk, I will review several closed testing embeddings of classical methods, such as Hochberg/Hommel, Higher Criticism, and Cluster-extent inference, as well as of more modern methods involving knock-offs or e-values. We motivate and explain our theory with neuroimaging examples.

JOHANNES LEDERER

HIGH-DIMENSIONAL EXTREMES

Methods, theories, and algorithms for high-dimensional extremes are surprisingly scarce. We introduce a novel approach that brings the strengths of score matching, high-dimensional statistics, and convex optimization to the field. We show that this approach can fit models with thousands of parameters within minutes on a standard laptop.

JOHANNES SCHMIDT-HIEBER

STATISTICAL LEARNING IN BIOLOGICAL NEURAL NETWORKS

Compared to artificial neural networks (ANNs), the brain learns faster, generalizes better to new situations and consumes much less energy. ANNs are motivated by the functioning of the brain but differ in several crucial aspects. For instance, ANNs are deterministic while biological neural networks (BNNs) are stochastic. Moreover, it is biologically implausible that the learning of the brain is based on gradient descent. In this talk we look at biological neural networks as a statistical method for supervised learning. We relate the local updating rule of the connection parameters in BNNs to a zero-order optimization method and derive some first statistical risk bounds.

MARIA DEIJFEN

STABLE MATCHING ON THE COMPLETE GRAPH

Consider a situation where a number of objects acting to maximize their own satisfaction are to be matched. Each object ranks the other objects and a matching is then said to be stable if there is no pair of objects that would prefer to be matched to each other rather than their current partners. We consider stable matching of the vertices in the complete graph based on i.i.d. exponential edge costs. Our results concern the total cost of the matching, the typical cost and rank of an edge in the matching, and the sensitivity of the matching to small perturbations of the underlying edge costs.

MARKUS HEYDENREICH

PRUDENT WALK IN DIMENSION 6 AND HIGHER

A nearest-neighbor path of finite length on the d-dimensional Euclidean lattice is called _prudent_ if no step of the walk is in the direction of a previously visited site. Its self-avoidance constraint is thus stronger than that for (classical) self-avoiding walk. We study the uniform measure of such walks.

We prove that the prudent self-avoiding walk converges to Brownian motion under diffusive scaling if the dimension is large enough. The same result is true for weakly prudent walk in dimension $d > 5$. A challenging property of the high-dimensional prudent walk is the presence of an infinite-range self-avoidance constraint. Interestingly, as a consequence of such a strong self-avoidance constraint, the upper critical dimension of the prudent walk is five, and thus greater than that for classical self-avoiding walk.

Joint work with Lorenzo Taggi and Niccolo Torri.

MORITZ SCHAUER

CAUSAL STRUCTURE LEARNING AND SAMPLING USING MARKOV MONTE CARLO WITH MOMENTUM

In the context of inferring a Bayesian network structure from observational data, that is inferring a directed acyclic graph (DAG), we devise a non-reversible continuous-time Markov chain that targets a probability distribution over classes of observationally equivalent (Markov equivalent) DAGs. The classes are represented as completed partially directed acyclic graphs (CPDAGs). The non-reversible Markov chain relies on the operators used in Chickering's Greedy Equivalence Search (GES) and is endowed with a momentum variable, which improves mixing significantly as we show empirically. The possible target distributions include posterior distributions based on a prior and a Markov equivalent likelihood.

Joint work with Marcel Wienöbst (Universität zu Lübeck).

NICOLAS PETRELIS

MULTI-SITES LOCALIZATION OF A WEAKLY SELF-AVOIDING WALK IN A HEAVY TAILED POTENTIAL

We investigate a model of continuous-time simple random walk paths in $\Z^d$ undergoing two competing interactions: an attractive one towards the large values of a random potential, and a self-repellent one in the spirit of the well-known weakly self-avoiding random walk. We take the potential to be i.i.d.~Pareto-distributed with parameter $\alpha>d$, and we tune the strength of the interactions in such a way that they both contribute on the same scale as $t\to\infty$.

Our main results are (1) the identification of the logarithmic asymptotics of the partition function of the model in terms of a random variational formula, and, (2) the identification of the path behaviour that gives the overwhelming contribution to the partition function for $\alpha>2d$.

NINA GANTERT

SOME RANDOM WALKS IN DYNAMICAL RANDOM ENVIRONMENTS

We explain some models and questions about random walks in dynamical random environments as well as possible applications. In particular, we consider directed random walk on (the backbone of) an oriented site percolation cluster which can serve as a model for ancestry lines. We also present recent results on biased random walks on dynamical percolation.

The talk is based on joint work with Matthias Birkner, Jiri Cerny and Andrej Depperschmidt, and on joint work with Sebastian Andres, Perla Sousi and Dominik Schmid.

PAULO SERRA

UNCERTAINTY QUANTIFICATION IN SPARSE QUANTILE REGRESSION

In statistics we often want to discover (or account for) structure in observed data, and dimension plays a crucial role in this task. For instance, high-dimensional data sometimes live in a lower dimensional space and sparse models are a popular way to account for this.

Sparse quantile regression combined with appropriate penalties produces sparse, robust estimators. In this talk I will share some results pertaining to advantages of sparse quantile regression over mean-based estimators. Performance of these estimators is studied under a very general model, allowing for correlation in the data, asymmetric distributions, sub-gaussian but also heavier data distributions. I'll present local (also minimax) rates for prediction and estimation, and share the current state of the results on uncertainty quantification.

PETER BÜHLMANN

CAUSALITY-INSPIRED STATISTICAL MACHINE LEARNING

Reliable, robust and interpretable machine learning is a big emerging theme in data science and statistics, complementing the development of pure black box prediction algorithms. New connections between distributional robustness, external validity and causality provide methodological paths for improving the reliability and understanding of machine learning algorithms, with wide-ranging prospects for various applications.

RÉMI LELUC

MONTE CARLO METHODS IN MACHINE LEARNING

Across a breadth of research areas, whether in Bayesian inference, reinforcement learning or variational inference, the need for accurate and efficient computation of integrals arises, making Monte Carlo methods one of the fundamental problems of statistical and machine learning research. This talk focuses on Monte Carlo integration, both from a theoretical and practical perspective, where the core idea is to use randomness to solve deterministic numerical problems. We will go through various control variates techniques for Monte Carlo integration in the context of machine learning problems. The study is based on mathematical tools coming from probability theory and statistics aiming to understand the behavior of certain existing algorithms and to design new ones with thorough analysis of the integration error.

RUI CASTRO

DISTRIBUTION FREE ANOMALY DETECTION WHEN MONITORING A LARGE NUMBER OF UNITS: A PERMUTATION/RANK-BASED HIGHER CRITICISM APPROACH

Anomaly detection when monitoring a large number of units is essential in a variety of applications, ranging from epidemiological studies to monitoring of complex systems. In this work we take a distribution-free stance and introduce a variant of the higher criticism test that does not require knowledge of the null distributions for proper calibration. This results in an exact test in finite samples. Our rank-based approach is also suitable when each unit is associated with a set of observations of potentially very different nature. We provide an asymptotic test power characterization and show detection is characterized by a phase-transition reminiscent of that of detection of sparse heteroskedastic Gaussian mixtures. Our analysis requires a refined understanding of the distribution of the ranks under the presence of anomalies, and in particular of the rank-induced dependencies. Within the exponential family and a family of convolutional models, we analytically quantify the asymptotic performance of our methodology and the performance of a suitable oracle, and show the difference is small for many common models. As the proposed test itself does not rely on asymptotic approximations it typically performs better than popular variants of higher criticism relying on such approximations. Finally, we show the applicability of the methodology through an analysis of quality control data of a pharmaceutical manufacturing process.

Based on past and ongoing work with Ivo Stoepker and Ery Arias-Castro

SILKE ROLLES

THE VERTEX-REINFORCED JUMP PROCESS

The area of random walks with reinforcement has been evolving since the late 1980s and seen growing interest in recent years. The talk will focus on the vertex-reinforced jump process (vrjp) which was invented by Wendelin Werner around 2000 when EURANDOM was in its early days. Davis and Volkov wrote the first two papers on the subject around the time when Stas Volkov was a postdoc at EURANDOM. In 2015, Sabot and Tarres discovered that the discrete time process associated to vrjp is a mixture of reversible Markov chains with a mixing measure which can be described in terms of a on-linear hyperbolic supersymmetric sigma model introduced by Zirnbauer in 1991. This connection is the reason that the vrjp has attracted attention from both the probability and the mathematical physics community.

In the talk, I will highlight some of the known results. This includes a recent transience result for vrjp with long-range jumps on the d-dimensional integer lattice, obtained jointly with Margherita Disertori and the former EURANDOM postdoc Franz Merkl.

SOPHIE LANGER

THE ROLE OF STATISTICAL THEORY IN UNDERSTANDING DEEP LEARNING

In recent years, there has been a surge of interest across different research areas to improve the theoretical understanding of deep learning. A very promising approach is the statistical one, which interprets deep learning as a nonlinear or nonparametric generalization of existing statistical models. For instance, a simple fully connected neural network is equivalent to a recursive generalized linear model with a hierarchical structure. Given this connection, many papers in recent years derived convergence rates of neural networks in a nonparametric regression or classification setting. Nevertheless, phenomena like overparameterization seem to contradict the statistical principle of bias-variance trade-off. Therefore, deep learning cannot only be explained by existing techniques of mathematical statistics but also requires a radical overthinking. In this talk we will explore both, the importance of statistics for the understanding of deep learning, as well as its limitations, i.e., the necessity to connect with other research areas.

ZBIGNIEW PALMOWSKI

STATIONARY STATES AND EXIT TIMES FOR LÉVY PROCESSES WITH PARTIAL RESETTING

In this talk we consider a d-dimensional stochastic process X which arises from a L\'evy process Y by partial resetting, that is the position of the process X at a Poisson moment equals c times its position right before the moment, and it develops as Y between these two consecutive moments, $c \in (0, 1)$.

We focus on Y being a strictly α-stable process with $\alpha\in (0,2]$ having a transition density. We analyze properties of the transition density p of the process X.

We establish a series representation of p. We prove its convergence as time goes to infinity (ergodicity), and we show that the limit ρ_{Y} (density of the ergodic measure) can be expressed by means of the transition density of the process Y starting from zero, which results in closed concise formulae for its moments. We show that the process X reaches a non-equilibrium stationary state. Furthermore, we check that p satisfies the Fokker-Planck equation, and we confirm the harmonicity of ρ_{Y} with respect to the adjoint generator.

In detail, we discuss the following cases: Brownian motion, isotropic and d-cylindrical α-stable processes for $\alpha \in (0,2)$, and α-stable subordinator for $\alpha\in (0,1)$. We find the asymptotic behavior of $p(t;x,y)$ as $t\to +\infty$ while (t,y) stays in a certain space-time region.

For Brownian motion, we discover a phase transition, that is a change of the asymptotic behavior of $p(t;0,y)$ with respect to $\rho_{Y}(y)$.

In the generated case of $Y_t=t$, our process X becomes an additive-increase and multiplicative-decrease (aka growth-collapse) process that grows linearly in time and that experiences downward jumps at Poisson epochs that are proportional to its present position. For this process, and also for its reflected versions, we solve one- and two-sided exit problems that concern the identification of the laws of exit times from fixed intervals and half-lines.

This process is a special case of piecewise deterministic Markov processes which will be discussed during the talk as well.