25th European Young Statisticians Meeting
Vilnius, Lithuania

Book of Abstracts

July 7-10, 2026

Vilnius University
Faculty of Mathematics and Informatics
Institute of Applied Mathematics

Organisers

Local Organising Committee

Aidas Medžiūnas *Chair (EYSM 24)* aidas.medziunas@mif.vu.lt

Saulius Jokubaitis *(EYSM 22)* saulius.jokubaitis@mif.vu.lt

Viktor Skorniakov *(EYSM 21)* viktor.skorniakov@mif.vu.lt

Jurgita Markevičiūtė *(EYSM 19)* jurgita.markeviciute@mif.vu.lt

International Organising Committee and Scientific Program Committee

Marthe Elisabeth Aastveit
University of Oslo

Kalle Alaluusua
Aalto University

Danijel Aleksić
University of Belgrade

Chiara Amorino
Universitat Pompeu Fabra

Simone Bombari
Institute of Science and Technology Austria

Manuela-Simona Cojocea
University of Bucharest

David-Jacob Economides
University of Piraeus

Irène Gijbels
KU Leuven

Matteo Giordano
University of Turin

Gentiane Haesbroeck
University of Liege

Catherine Higgins
University College Dublin

Daniela Ivanković
University of Zagreb

Baalu Belay Ketema
EDF R&D/IMT

Karolina Klockmann
University of Vienna

Johann Köhne
University of Göttingen

Bartosz Majewski
AGH University of Krakow

Pedro Martín-Chávez
University of Warwick

Aidas Medžiūnas
Vilnius University

Christos Meselidis
University of the Aegean

Marcell Nagy
Budapest University of Technology and Economics

Jimmy Olsson
KTH Royal Institute of Technology

Katariina Perkonoja
University of Turku

Francisco Pina
University of Luxembourg

Hedvika Ranošová
Charles University

Cecilie Olesen Recke
University of Copenhagen

Mette Marie Skjøtt Rasmussen
Aarhus University

Katarzyna Szczerba
University of Luxembourg

Maeve Upton
University of Limerick

Jānis Valeinis
University of Latvia

Jorge Vaz Ramos Rodrigues de Cabral
University of Aveiro

Mirjana Veljović
University of Belgrade

Giorgia Zaccaria
University of Milano-Bicocca

Book of Abstracts

25th European Young Statisticians Meeting

July 7-10, Vilnius, Lithuania

Vilnius University

Faculty of Mathematics and Informatics

Institute of Applied Mathematics

2026

Contents

Interpolated Stochastic Interventions Based on Propensity Scores, Target Policies and Treatment-Specific Costs

Johan de Aguas

Integreat – Norwegian Centre for Knowledge-driven Machine Learning, Department of Mathematics, University of Oslo
johanmd@uio.no

We introduce two families of stochastic interventions with discrete treatments that connect causal modeling to cost-sensitive decision making. The interventions arise from a cost-penalized information projection of the independent product of the organic propensity scores and a reference policy, yielding closed-form Boltzmann–Gibbs couplings. The induced marginals define modified stochastic policies that interpolate smoothly, via a tilt parameter, from the organic law or from the reference law toward a product-of-experts limit when all destination costs are strictly positive. The first family recovers and extends incremental propensity score interventions, retaining identification without global positivity. For inference on the expected outcomes after these policies, we derive the efficient influence functions under a nonparametric model and construct one-step estimators. In simulations, the proposed estimators improve stability and robustness to nuisance misspecification relative to plug-in baselines. The framework can operationalize graded scientific hypotheses under realistic constraints. Because inputs are modular, analysts can sweep feasible policy spaces, prototype candidates, and align interventions with budgets and logistics before committing experimental resources.

Beyond Black-Box Predictions: Explainability-Driven Early Warning for Electricity Price Spikes in Romania

Andrei Alexandru-Victor*

* Bucharest University of Economic Studies, Romania

Price volatility in liberalized electricity markets creates persistent difficulties for traders, system operators, and regulators alike. As European countries accelerate their energy transitions, the ability to anticipate sudden price spikes has become central to both risk management practice and broader market governance. Machine learning approaches have delivered impressive predictive accuracy in this domain, yet their inherent opacity has limited their practical value — knowing that a spike is coming matters far less if you cannot explain why.

This paper examines whether explainability methods can confirm load-generation balance as the primary mechanism behind price spike formation in the Romanian day-ahead market. We assembled hourly observations spanning March through September 2024, yielding roughly 5,000 data points and 70 engineered features. Gradient boosting and ensemble tree models were trained under strict chronological ordering to avoid any form of data leakage, and two complementary interpretability techniques were applied: one characterizing global predictor importance, the other reconstructing the causal logic behind individual spike episodes.

The best-performing model accounts for 67.6% of price variance. Residual load forecast ranks as the second most influential variable in the global analysis and emerges as the dominant trigger within the early warning component — a consistent finding across both interpretability lenses. The warning system itself identifies 72.2% of high-price events, with an F1-type composite score of 0.57. Given the scarcity of established benchmarks for threshold-based spike detection, this represents a credible result. We also report, as incidental evidence rather than a core claim, that model performance deteriorates during the summer 2024 heat wave; crucially, however, balance indicators continued to flag system stress throughout, suggesting that interpretable signals can retain informational value precisely when predictive accuracy falters.

The central contribution is a validated, integrated XAI framework that bridges global and local interpretability for electricity price spike anticipation. Rather than delivering opaque probabilistic outputs, the system produces transparent, operator-readable signals — a distinction that matters considerably in high-stakes, real-time decision environments.

Causal Discovery with Tiered Background Knowledge

Christine Winther Bang[†]

[†] Department of Mathematical Sciences, University of Copenhagen

Causal graphs are useful tools for understanding complex causal structures and can help us deciding on interventions or identifying causal effects. Causal graphs are usually constructed based on expert knowledge, which may be insufficient or incorrect. Causal discovery methods construct causal graphs in a data-driven way and serve as an alternative to the classical expert-driven approach, but come with other limitations. Typically, without strong assumptions, we cannot identify all causal directions. Moreover, causal discovery methods are sensitive to errors based on e.g. statistical testing when applied to real data. In practice, we often have more causal information available than what can be obtained from the data alone, and the use of so-called background knowledge in causal discovery is a way to bridge the purely data-driven and the purely expert-driven approaches. Time structure induces a partial causal ordering of the variables, which I will refer to as tiered background knowledge. This type of background knowledge is common, e.g. for cohort data, and it improves causal discovery methods such that the output becomes more informative and reliable. In this talk, I will first explain how tiered background knowledge can be incorporated in causal discovery algorithms and show how this improves graphs estimated using finite sample data [1, 2]. Additionally, I will show in which ways tiered background knowledge is useful if we allow for unobserved confounding and consider multiple datasets with overlapping variables [4].

References

[1] Bang, C. W. and Didelez, V. 2023. Do we become wiser with time? On causal equivalence with tiered background knowledge. *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence* PMLR 216:119-129.

[2] Bang, C. W., Witte, J., Foraita, R., and Didelez, V. 2026. Improving finite sample performance of causal discovery by exploiting temporal structure. *Statistical Methods in Medical Research*.

[3] Bang, C. W. and Didelez, V. 2025. Constraint-based causal discovery with tiered background knowledge and latent variables in single or overlapping datasets. *Proceedings of the Fourth Conference on Causal Learning and Reasoning*, PMLR 275:1116-1146.

Changed Segment Detection in Functional Data via C-type Projection

Karolis Bartkus[†]

[†] Vilnius University

We consider the problem of detecting a *changed segment* (epidemic alternative) in a sequence of independent functional observations $X_1, \ldots, X_n$ taking values in $C[0,1]$. Under the null hypothesis the curves are identically distributed; under the alternative there exists an unknown interval $I^* = \{k^* + 1, \ldots, k^* + \ell^*\} \subset \{1, \ldots, n\}$ on which the distribution changes from P to $Q \neq P$ and returns to P outside it.

To make scalar scan statistics applicable, we reduce each curve by a C-type projection, i.e. a point evaluation $P_{t_0} \colon C[0,1] \to \mathbb{R}$, $P_{t_0}(f) = f(t_0)$, at a data-driven location $t_0 \in \arg\max_{t \in [0,1]} V_n(t)$, where V_n is the pointwise empirical variance of the sample. The resulting scalar series $Y_i := X_i(t_0)$, $i = 1, \ldots, n$, is then fed into three complementary epidemic scan tests sharing a common weight function $\rho_\gamma(t) = t^\gamma$, $0 \leq \gamma < 1/2$, that balances sensitivity to short and long candidate segments: a Wilcoxon antisymmetric-kernel scan [1], an L_p-type Cramér–von Mises scan and an RKHS scan based on Maximum Mean Discrepancy; both are developed in [2]. Each test admits a known asymptotic null limit (a Brownian-bridge or Wiener-sheet functional) and pre-computed Monte Carlo critical values; details are deferred to the cited works. A companion paper [3] treats the same scan family with norm and Brownian/Itô projections, providing a direct point of comparison for the C-type reduction studied here. We compare the three tests on the four functional scenarios and illustrate the methodology on Lithuanian electricity balancing prices.

References

[1] Račkauskas, A. and Wendler, M. 2020. Convergence of U-processes in Hölder spaces with application to robust detection of a changed segment. *Statistical Papers,* 61:1409–1435. doi:10.1007/s00362-020-01161-9.

[2] Bartkus, K. and Račkauskas, A. 2026. Nonparametric changed segment detection in functional data. *Nonlinear Analysis: Modelling and Control,* 31(1):194–211. doi:10.15388/namc.2026.31.44489.

[3] Bartkus, K. and Račkauskas, A. 2025. Changed segment tests via norm and Brownian projection. *Lithuanian Mathematical Journal,* 65(4):481–499. doi:10.1007/s10986-025-09699-7.

Scatter Halfspace Depth: Recent Theoretical Advances

Filip Bočinec[†], Stanislav Nagy[†], Gaëtan Louvet[‡], Germain Van Bever[‡]

[†] Charles University
[‡] Université libre de Bruxelles

Robust estimation of multivariate scatter remains a fundamental challenge in statistics. Recently, the scatter halfspace depth (sHD) has emerged as a powerful geometric tool, extending the desirable robustness properties of the classical Tukey location depth to scatter matrices. The induced scatter halfspace median provides a highly robust, non-parametric alternative to conventional scatter estimators. In this talk, we focus on selected theoretical aspects regarding the foundations and the asymptotic behavior of this estimator. First, we investigate the sHD under the broad class of α-symmetric distributions, which naturally accommodates various heavy-tailed models. Among other results, we derive concentration inequalities under Huber's contamination model and discuss its minimax optimality. Second, building upon empirical process techniques originally developed for the location halfspace median, we establish that the sample scatter halfspace median converges at the standard $\sqrt{n}$ rate to the argmax of a limiting random field.

References

[1] Bočinec, F. and Nagy, S. 2025. Location and Scatter Halfspace Median under α-Symmetric Distributions. *Journal of Nonparametric Statistics,* 1-29.

[2] Paindaveine, D. and Van Bever, G. 2018. Halfspace depths for scatter, concentration and shape matrices. *The Annals of Statistics,* 46:3276-3307.

[3] Louvet, G., Bočinec, F., Nagy, S. and Van Bever, G. 2026. Asymptotics for the scatter halfspace median. *In preparation.*

Box Confidence Depth: Simulation Based Inference with Hyper Rectangles

Elena Bortolato[†]

[†] Universitat Pompeu Fabra & Barcelona School of Economics

This work presents a novel simulation-based approach for constructing confidence regions in parametric models, which is particularly suited for generative models and situations where limited data and conventional asymptotic approximations fail to provide accurate results. The method leverages the concept of data depth and depends on creating random hyper-rectangles, i.e. boxes, in the sample space generated through simulations from the model, varying the input parameters. A probabilistic acceptance rule allows to retrieve a Depth-Confidence Distribution for the model parameters from which point estimators as well as calibrated confidence sets can be read-off. The method is designed to address cases where both the parameters and test statistics are multivariate.

References

[1] Bortolato, E. and Ventura L. 2026 Box Confidence Depth: simulation-based inference with hyper-rectangles. *Electronic Journal of Statistics, 20:942–963.*

Bayesian Binary Classification under Label Uncertainty with Network-Informed Gaussian Processes

Konstantinos Bourazas[1], Angelos Alexopoulos[1]
Konstantinos Kalogeropoulos[2], Petros Dellaportas[3,4]

[1] Department of Economics, Athens University of Economics and Business, Greece
[2] Department of Statistics, London School of Economics and Political Science, UK
[3] Department of Statistical Science, University College London, UK
[4] Department of Statistics, Athens University of Economics and Business, Greece

In this work, we address the problem of binary classification under label uncertainty in settings where both feature-based and relational data are available. We propose a Bayesian Gaussian Process classification model that combines covariate and network information through a Product of Experts (PoE) formulation, effectively capturing different modes of similarity. Furthermore, our approach accounts for uncertainty in the observed labels, and particularly missed positives, enabling robust inference. Posterior inference relies on the aGrad-z auxiliary-gradient sampler of [1]. We validate the proposed methodology through a simulation study, demonstrating improved classification accuracy and probability calibration relative to single-source alternatives.

Keywords: Bayesian Classification; Gaussian Processes; Label Uncertainty; Network.

References

[1] Titsias, M. K. and Papaspiliopoulos, O. 2018. Auxiliary gradient-based sampling algorithms. *Journal of the Royal Statistical Society: Series B,* 80:749–767.

Identifying Network Hubs with the Partial Correlation Graphical LASSO

Adam Chojecki[†], Małgorzata Bogdan[‡], Ivan Hejný[*], Bartosz Kołodziejek[†], Jonas Wallin[*]

[†] Warsaw University of Technology
[‡] University of Wrocław
[*] Lund University

Gaussian graphical models are a standard tool for learning conditional independence structures from multivariate data. The Graphical LASSO (GLASSO) is one of the most widely used methods for this task, but its penalty is applied directly to the entries of the precision matrix. As a consequence, the method is not scale-invariant, and simple rescaling of variables does affect the selected graph.

We study the Partial Correlation Graphical LASSO (PCGLASSO), a scale-invariant alternative in which sparsity is imposed on the partial correlation matrix rather than on the precision matrix itself. This formulation is natural from the perspective of Gaussian graphical models, since partial correlations directly describe conditional dependence. It is particularly well suited for detecting hub structures, where a small number of nodes have many connections.

The price for this scale-invariant formulation is a non-convex optimization problem. We develop efficient algorithms for computing the PCGLASSO estimator by alternating between optimization over the diagonal scaling matrix and the partial correlation matrix. For the latter step, we use modified coordinate-descent solvers inspired by fast GLASSO solvers.

On the theoretical side, we introduce a scale-invariant irrepresentability condition for PCGLASSO and prove model selection consistency under this condition. We show that, for hub-structured graphs, this condition is substantially weaker than the corresponding condition for GLASSO, which gives a theoretical explanation for the improved empirical performance of PCGLASSO in hub recovery.

Simulation studies and an application to gene expression data confirm that PCGLASSO can recover hub-like structures more accurately than GLASSO-based alternatives, while remaining computationally efficient. These results suggest that penalizing partial correlations provides a useful and interpretable approach to sparse graphical model estimation, especially in problems where scale effects and hub detection are central.

References

[1] Bogdan, M., Chojecki, A., Hejný, I., Kołodziejek, B. and Wallin, J. 2025. Identifying Network Hubs with the Partial Correlation Graphical LASSO. *arXiv preprint*, arXiv:2508.12258.

[2] Carter, J., Rossell, D. and Smith, J. Q. 2024. The Partial Correlation Graphical LASSO. *arXiv preprint*, arXiv:2405.01639.

Explainable AI in Financial Fraud Detection: Evidence from the IEEE-CIS Fraud Dataset

Delia Diaconu[†], Daniel Traian Pele[‡]

[†] Bucharest University of Economic Studies, Romania
[‡] Bucharest University of Economic Studies, IDA Institute Digital Assets, Institute for Economic Forecasting, Romanian Academy

Financial fraud detection systems increasingly rely on advanced machine learning models, whose decisions should remain easily interpretable, in order to make transparent decisions and to align with regulatory requirements. In this study, we explore explainable AI methods for fraud detection, using the IEEE-CIS Fraud detection database. Thus, we compare the results obtained using tabular models (XGBoost, CatBoost) with those generated by graph-based models (GraphSAGE) and observe that although gradient boosting models have a better performance - measured in terms of ROC-AUC -, graph-based models capture relational structures, which cannot be described by tabular models. We propose a multi-level explainability framework, which combines methods of explaining model variables (SHAP values) with graph analysis methods, such as edge and node ablation. The results highlight a trade-off between discriminatory power and the interpretation of interactions between network nodes. Thus, although tabular models are better at identifying risk at feature level, graph-based models allow for a better interpretation of transactional fraud patterns. These results demonstrate that explainable AI in fraud detection should not be based on a single mpdel, but it should combine several types of models that capture both the relational structure and the predictive power at feature level.

Meanimile regression: a general class of conditional functionals

Dieter Debrauwer, Irène Gijbels and Gerda Claeskens

KU Leuven, Belgium

Statistical modeling fundamentally seeks to describe the behavior of random variables through various functionals. While classical measures such as the mean and median are defined via symmetric loss functions, more flexible tools, including quantiles, expectiles, and extremiles, have emerged to better capture risk and tail behavior. These diverse measures can be unified under the meanimile framework (Debrauwer et al., 2025), where functionals are defined as solutions to optimization problems involving a specific loss function and a distributional weighting function.

When assessing real-world risks, variables of interest are rarely isolated; they are dynamically driven by external covariates (e.g., varying economic climates, different patient groups, or climatological settings). Therefore, instead of looking into an unconditional functional, we must evaluate the conditional setting. While conditional risk measures like quantile regression are classical topics, the literature on general conditional risk functionals is far more limited, with some very recent developments, such as regression extremiles. Existing state-of-the-art research predominantly focuses on the square loss function.

We provide a comprehensive review of the state-of-the-art literature on conditional (risk) functionals and cast it into the unified conditional meanimile framework. Building upon this, we significantly contribute to methodologies for such conditional functionals. Specifically, we extend recent extremile regression techniques to accommodate arbitrary distributional weighting functions, both in linear and nonparametric regression settings. In addition, we propose a general estimation framework for conditional meanimiles. The finite-sample performance of these proposed methods is evaluated through extensive simulation studies. The practical relevance of this regression framework is demonstrated through various real-world applications

References

[1] Debrauwer, D., Gijbels, I., & Herrmann, K. 2025. On a general class of functionals: statistical inference and application to risk measures. *Electronic Journal of Statistics*, 19(1), 2456–2510.

On the Asymptotic Behaviour of the Haezendonck-Goovaerts Risk Measure for Sums with Consistently Varying Increments

Mantas Dirma[†], Jonas Šiaulys[†], Neda Nakliuda[†], Luca Zanardelli[†]

[†] Vilnius University

Let $\xi_1, \ldots, \xi_n$, $n \in \mathbb{N}$, be possibly dependent and not necessarily identically distributed real-valued random variables (r.v.s) from the class $\mathcal{C}$ of consistently varying distributions (a subclass of heavy-tailed distributions). In practical applications in the fields of finance or insurance, such r.v.s and their sums, denoted by S_n^{ξ}, may represent potential losses incurred by an investor or an insurance company. In such a context, the level of risk taken is often assessed using risk measures – functionals from a set of r.v.s to the real numbers. The presentation will focus on one of them – the so-called Haezendonck-Goovaerts (HG) risk measure (first introduced in 1982, [1]), which is defined using a nonnegative and convex on the interval $[0, \infty)$ function φ, satisfying $\varphi(0) = 0$, $\varphi(1) = 1$ and $\varphi(\infty) = \infty$. In the case when φ is a power function with an exponent $\varkappa \geqslant 1$ (i.e., $\varphi(t) = t^{\varkappa}$) and the parameter q, which represents the confidence level, tends to 1 the asymptotic estimates of the HG measure for the sums S_n^{ξ} will be explored. The obtained results extend the findings in [2] and [3].

References

[1] Haezendonck J. and Goovaerts M. J. 1982. A new premium calculation principle based on Orlitz norms, *Insur. Math. Econ.*, 1:41–53.

[2] Tang Q. and Yang F. 2012. On the Haezendonck-Goovaerts risk measure for extreme risks, *Insur. Math. Econ.*, 50:217– 227.

[3] Tang Q. and Yang F. 2014, Extreme value analysis of the Haezendonck-Goovaerts risk measure with a general Young function, *Insur. Math. Econ.*, 59:311–320.

The *post-hoc* Test for Local Dependence

Bartłomiej Gibas[†], Bogdan Ćmiel[†]

[†] AGH University of Krakow, Krakow, Poland

Independence plays a fundamental role in probability theory and mathematical statistics, making the problem of testing it an important and widely studied topic. Many testing procedures have been proposed, but most of them focus on the problem from a global perspective, leading only to a decision on whether the hypothesis of independence should be rejected. In applications, however, it is often important not only to detect dependence, but also to identify where it occurs and how strong it is. A graphical presentation of the results is therefore also valuable, as it can make the structure of dependence easier to interpret.

The aim of our work is to propose a method that addresses these aspects simultaneously. Based on copula theory, we introduce a new procedure for testing both global and local statistical independence using the quantile dependence function introduced in [2, 4]. The proposed method can be viewed as a continuation and extension of the approach developed in [1]. Instead of reducing the testing procedure to a comparison of the test statistic with a single critical value, we introduce critical surfaces constructed so that, under independence, the probability of exceeding the surface is locally controlled at the same level. This makes it possible to identify local deviations from independence and assess their statistical significance, while preserving the overall significance level of the test.

For the resulting test, we establish a consistency theorem and investigate finite-sample performance through simulation studies, including empirical power comparisons with commonly used independence tests. We also introduce a Monte Carlo method for determining the critical surfaces. Overall, the procedure provides both a formal statistical test and a graphical tool for detecting and localizing dependence structures.

References

[1] Ćmiel, B. and Ledwina, T. 2026. Detecting dependence structure: visualization and inference. *Statistical Papers,* 67(57).

[2] Ćmiel, B. and Gibas, B. 2026. The post-hoc test for local dependence. *arXiv preprint* arXiv:2512.20280.

[3] Ledwina, T. 2014. Dependence function for bivariate cdf's. *arXiv preprint* arXiv:1405.2200.

[4] Ledwina, T. 2015. Visualizing association structure in bivariate copulas using new dependence function. In: *Stochastic Models, Statistics and Their Applications,* Springer Proceedings in Mathematics & Statistics, 122:19–27. Springer, Cham.

Beyond Means: Empirical Likelihood Confidence Bands for Quantile Comparisons

Janis Gredzens[†]

[†] University of Latvia

Comparing two populations through their quantile functions provides a clearer picture of distributional differences than classical summary measures do - quantile functions capture shifts in location, spread, and tail behaviour simultaneously. Empirical likelihood (EL) offers a natural inferential framework for such comparisons, yielding asymmetric, range-respecting, and Bartlett-correctable confidence regions without making parametric assumptions, a property that bootstrap alternatives lack. Additionally to the two-sample case, we also consider an EL-based approach to quantile ANOVA, extending these ideas to multiple group comparisons within the same unified framework. While the theoretical basis of EL-based quantile inference is well documented, accessible and unified solutions for practical applications remain scarce. In this talk, we present our ongoing work on an R package that consolidates EL methodology for both two-sample and multi-sample quantile comparisons.

Using the Phase-Type Distribution to Conduct Inference in Population Genetics

Elisabeth Sommer James[†], Simon Boitard[‡], Asger Hobolth[†], Raphael Leblois[‡]

[†] Aarhus University
[‡] CBGP, INRAE, CIRAD, IRD, Institut Agro, Univ Montpellier

Phase-type distributions are a useful framework for population genetics: many key quantities can be shown to be phase-type distributed, such as time to most recent common ancestor or the external or internal branch length of a coalescent tree. Furthermore, [1] devised a sampling formula, allowing the probability of a genetic sample in a time homogeneous coalescent model under the infinite sites assumption to be calculated. We provide an efficient recursive implementation of this sampling formula and investigate its suitability to infer parameters such as the coalescent and migration rates using genetic data.

References

[1] Hobolth, A., Boitard, S., Futschik, A., and Leblois, R. 2025. A matrix-analytical sampling formula for time-homogeneous coalescent processes under the infinite sites mutation model. *Theoretical Population Biology,* 163:62-79.

[2] Hobolth, A., Rivas-González, I., Bladt, M. and Futschik, A., 2024. Phase-type distributions in mathematical population genetics: An emerging framework. *Theoretical Population Biology,* 157:14-32.

Sharp Oracle Inequalities for Covariate Selection via the AIC

Georg Köstenberger[†]

[†] University of Vienna

Prediction is a core task in time series analysis, and the choice of the forecasting model is of fundamental importance. To address this problem (among other things), Akaike introduced the AIC and FPE, and demonstrated their significant usefulness for prediction in two landmark papers. In subsequent seminal works, Shibata developed a notion of asymptotic efficiency and showed that both AIC and FPE are optimal, setting the stage for decades-long developments and research in this area and beyond.

Most of the literature on the usage of AIC for prediction focuses on the case of nested models. However, there is no fundamental information theoretic reason for this restriction, as the AIC is essentially an optimal estimator of the prediction error, which is based on minimizing the empirical Kullback-Leibler divergence between any two models. This point of view suggests that the AIC (and its variants) should be able to select an (optimal) forecasting model from any set of candidate models. In this work, we establish sharp, finite-sample oracle inequalities for the AIC in the non-nested case, subject only to a very general notion of weak dependence. This establishes a unversality property of the AIC, in the sense that it can not just be used to compare between nested models, but rather between arbitrary models. Our framework contains many prominent dynamical systems such as random walks on the regular group, functionals of iterated random systems, functionals of (augmented) Garch models of any order, functionals of (Banach space valued) linear processes, possibly infinite memory Markov chains, dynamical systems arising from SDEs, and many more. The proofs require a new notion of uniform integrability in the Wiener algebra, a new uniform version of the Wiener-Levy theorem and a uniform Baxter-type inequality, which may be of independent interest.

Data Driven Modeling of Multiple Interest Rates with Generalized Vasicek-type Models

Pauliina Ilmonen[a], Milla Laurikkala[a], Kostiantyn Ralchenko[bc], Tommi Sottinen[b] and Lauri Viitasaari[d]

[a]Aalto University School of Science, Department of Mathematics and Systems Analysis,
[b]University of Vaasa, School of Technology and Innovations,
[c]Taras Shevchenko National University of Kyiv, Department of Probability, Statistics and Actuarial Mathematics,
[d]Aalto University School of Business, Department of Information and Service Management

The Vasicek model is a commonly used interest rate model, and there exist many extensions and generalizations of it. However, most generalizations of the model are either univariate or assume the noise process to be Gaussian, or both. In this article, we study a generalized multivariate Vasicek model that allows simultaneous modeling of multiple interest rates while making minimal assumptions. In the model, we only assume that the noise process has stationary increments with a suitably decaying autocovariance structure. We provide estimators for the unknown parameters and prove their consistencies. A continuous-time algebraic Riccati equation based approach is used in estimation. We also derive limiting distributions for each estimator and provide theoretical examples. Furthermore, the model is tested empirically with both simulated data and real data.

Construction of Optimal Tests for Symmetry on the Torus and their Quantitative Error Bounds

Andreas Anastasiou[†], Christophe Ley[‡], Sophia Loizidou[‡]

[†] University of Cyprus
[‡] University of Luxembourg

Several complex real-world data can be viewed as points on the hyper-torus, which is the cartesian product of circles. Over the past few years, this has motivated new proposals of distributions on the torus, both (pointwise) symmetric and sine-skewed asymmetric. In practice, it is relevant to know whether one should use the simpler symmetric models or the more convoluted yet more general asymmetric ones. So far, only parametric likelihood ratio tests have been defined to distinguish between a symmetric density and its sine-skewed counterpart. A new semi-parametric test is presented, a test which is valid not only under a given parametric hypothesis but also under a very broad class of symmetric distributions. A description of its construction and asymptotic properties under the null and alternative hypotheses will be presented. Using Stein's method, bounds for the rate of convergence of the test statistic are derived, and finite sample behavior (through Monte Carlo simulations) will be given, as well as an application of the test on protein data.

Collective Outlier Detection and Enumeration with Conformalized Closed Testing

Chiara Gaia Magnani[†], Matteo Sesia[‡], Aldo Solari[§]

[†] Bocconi University
[‡] University of Southern California
[§] Ca' Foscari University of Venice

In this talk I will present ACODE, a flexible distribution-free method for collective outlier detection and enumeration, designed for situations in which the presence of outliers can be detected powerfully even though their precise identification may be challenging due to the sparsity, weakness, or elusiveness of their signals. This method builds upon recent developments in conformal inference and integrates classical ideas from other areas, including multiple testing, locally most powerful and adaptive rank tests, and non-parametric large-sample asymptotics. The key innovation lies in developing a principled and effective approach for automatically choosing the most appropriate machine learning classifier and two-sample testing procedure for a given data set. The performance of our method is investigated through extensive empirical demonstrations, including an analysis of the LHCO high-energy particle collision data set.

References

[1] Magnani, C. G., Sesia, M. and Solari, A. 2026. Collective Outlier Detection and Enumeration with Conformalized Closed Testing. *arXiv:2308.05534*.

Investigation of Correlations Between Modules and Clinical Traits using Sparse Principal Component Analysis

Markéta Barać Makarová[†]

[†] Masaryk University

High-dimensional omics datasets are frequently partitioned into clusters, or modules, to improve biological interpretability and reduce analytical complexity. These modules are commonly summarised by representative vectors, such as eigengenes, which are then correlated with a range of clinical traits and phenotypic variables. Despite their widespread use, such associations are often interpreted only at a superficial level, while systematic comparison of module behaviour across patient groups remains limited.

To address these challenges, we propose an extended analytical pipeline based on sparse principal component analysis (sPCA) [1]. By exploiting the inherent feature selection properties of sPCA, the framework enables identification of key molecular drivers within modules, facilitates comparison between biological or clinical groups, and improves the interpretability of module–trait relationships. The proposed approach provides a more comprehensive strategy for downstream analysis of omics modules and their associated clinical phenotypes, supporting deeper biological insight.

References

[1] Zou, H., Hastie, T., & Tibshirani, R. 2006. Sparse Principal Component Analysis. *Journal of Computational and Graphical Statistics*, 15(2), 265–286. https://doi.org/10.1198/106186006X113430

Surrogate Model Comparison for Sensitivity Analysis and Uncertainty Quantification of Neutral Beam Injection Heating of Nuclear Fusion Plasmas

Ioannis Mavrogiannis, David Zarzoso, Mitra Foularidad

Aix Marseille Univ, CNRS, Centrale Med, M2P2, Marseille, France

Surrogate models can be used to approximate the statistical behaviour of complex, non-linear numerical code outputs, under the influence of model parameter uncertainty. Such an uncertainty quantification (UQ) analysis can provide further confidence in the code results, while referring to a more realistic scenario when the parameters do not have a fixed value. Surrogates also facilitate the accurate estimation of global, variance-based sensitivity indices (GSA), with a smaller number of samples that would otherwise be required through traditional Monte Carlo methods. The physical system considered concerns the heating of a plasma in a tokamak machine, with neutral beam injection (NBI). Nuclear fusion is the physical process that fuels the stars like our sun. On Earth, machines such as tokamaks and stellarators use strong magnetic fields to confine plasma. The plasma in a tokamak can reach temperatures of approximately 100 million degrees Celsius, enabling nuclear fusion reactions. Auxiliary heating methods are essential in order for a tokamak plasma to reach such extreme temperatures, with such a method being NBI. Neutral particles with energy larger than the plasma bypass the strong magnetic field confining the plasma, subsequently ionize within the plasma, and then transfer energy to the plasma ions and electrons through Coulomb collisions. Such auxiliary heating systems are necessary for reaching the extreme plasma temperatures at which fusion reactions occur.

In the present work, uncertainties in the thermal plasma equilibrium and ionization cross-sections are propagated through the code (Toroidal Accelerated PArticle Simulator), obtaining quantities of interest (QoIs) that describe the beam attenuation in the plasma, the heating power density transferred to the plasma ions and electrons, and the NBI shine-through losses, which account for the amount of neutral particles that did not ionize and instead damaged the wall of the machine. For each QoI, we also further investigate the ability of different surrogate models in approximating UQ and GSA metrics. Specifically, we consider the performance of different surrogate models, trained under different sampling methods and sample sizes. For each surrogate, an uncertainty range is obtained for its accuracy, the estimated statistical moments and Sobol' indices.

References

[1] Mavrogiannis, I., Zarzoso, D., & Fouladirad, M. 2026. Uncertainty quantification and sensitivity analysis of the energy and particle sources from hydrogen neutral beam injection in a JET-like fusion plasma. *Nucl. Fusion* **66** 062001

Multifractal Models for Taylor's Law

Dominik Mihalčić[†], Danijel Grahovac[†], Nikolai N. Leonenko[‡]

[†] University of Osijek
[‡] Cardiff University

Taylor's law describes a power-law relationship between the variance and the mean of population densities. Although it is a widely verified phenomenon, its underlying mechanism is still not fully understood. We propose a sampling scheme based on random measures, where observations are obtained by aggregating contributions over groups which are determined by an intrinsic notion of size. Different scaling properties of the underlying random measure correspond to different exponents in Taylor's law: completely random measures lead to exponent 1, whereas self-similar random measures lead to exponent 2. Intermediate values may arise by allowing a more general form of scale invariance, such as multifractality. We illustrate the approach on several datasets exhibiting Taylor's law. The empirical analysis suggests that the scaling relation extends beyond second-order moments, consistent with the behavior of multifractal random measures.

References

[1] Taylor, L. R. 1961. Aggregation, variance and the mean. *Nature,* 189:732-735.

[2] Rhodes, R. and Vargas, V. 2010. Multidimensional multifractal random measures. *Electronic Journal of Probability,* 15(9):241-258.

[3] Eisler, Z. and Bartos, I. and Kertész, J. 2008. Fluctuation scaling in complex systems: Taylor's law and beyond. *Advances in Physics,* 57(1):89-142.

[4] Saitoh, T. and Cohen, J. E. 2024. Quantifying factors that explain the slopes of the temporal Taylor's law of Hokkaido vole populations. *Population Ecology,* 66:125-142.

Large Deviations for MCMC Sampling

Federica Milinanni[†], Pierre Nyquist[‡]

[†] Northwestern University
[‡] Chalmers & Gothenburg University

Sampling algorithms from the class of Markov chain Monte Carlo (MCMC) methods are widely used across scientific disciplines. Good performance measures are essential to analyse these methods, to compare different MCMC algorithms, and to tune parameters within a given method. Common tools that are used for analysing convergence properties of MCMC algorithms are, e.g., mixing times, spectral gap and functional inequalities (e.g., Poincaré, log-Sobolev). A further, rather novel, approach consists in the use of large deviations theory to study the convergence of empirical measures of MCMC chains. At the heart of large deviations theory is the large deviation principle, which allows us to describe the rate of convergence of the empirical measures through a so-called rate function. In this talk we will consider Markov chains generated via MCMC methods of Metropolis-Hastings type for sampling from a target distribution on a Polish space. We will state a large deviation principle for the corresponding empirical measure, show examples of algorithms from this class for which the theorem applies, and illustrate how the result can be used to tune algorithms' parameters.

References

[1] Milinanni, F., and Nyquist, P. 2024. A large deviation principle for the empirical measures of Metropolis–Hastings chains. *Stochastic Processes and their Applications,* 170, 104293.

[2] Milinanni, F., and Nyquist, P. 2026. Large deviations for Independent Metropolis Hastings and Metropolis-adjusted Langevin algorithm. To appear in *Bernoulli*.

[3] Milinanni, F. 2024. Large deviation-based tuning schemes for Metropolis-Hastings algorithms. arXiv preprint arXiv:2409.20337.

A shared-parameter joint modeling framework for predicting dynamic conditional treatment benefits from observational data

Pedro Miranda-Afonso[†], Hao Liu[‡], Michele Molinari[§], Dimitris Rizopoulos[†]

[†] Erasmus University Medical Center
[‡] University of Pittsburgh Medical Center
[§] Houston Methodist Hospital

Whenever treatment or intervention resources are scarce, they should be allocated to individuals who are expected to derive the greatest benefit when the resource becomes available. One possible allocation criterion is the conditional treatment benefit, defined as the difference between an individual patient's predicted outcome with and without treatment, conditional on their longitudinal history up to the decision time. Estimating this quantity from observational data is challenging because treatment decisions are typically time dependent and guided by evolving prognostic information, which is affected by the same longitudinal processes as future event risk.

To address this, we propose a Bayesian shared-parameter joint modeling framework for multiple longitudinal markers and time-to-event outcomes that can dynamically estimate conditional event risk under alternative treatment strategies. The framework links marker-specific mixed-effects submodels to pre- and post-treatment relative-risk submodels through shared random effects and flexible association functions, such as the marker's rate of change and its standardized cumulative effect. Within a potential outcomes framework, we define the assumptions required to identify these patient-specific treatment effects from observational data. The model estimates survival probabilities in two scenarios: immediate treatment and continued follow-up without treatment. Individualized treatment effects can then be expressed as differences in expected event-free times or event-free probabilities over a clinically relevant horizon. The model is available in the CRAN `R` package `JMbayes2`, with full posterior sampling implemented in `C++` to maximize computational efficiency.

As an application, we analyzed data from 7,471 patients with hepatocellular carcinoma listed for liver transplantation in the US Scientific Registry for Transplant Recipients between 2012 and 2022, of whom 4,786 received a liver transplant. The longitudinal processes included serum α-fetoprotein, tumor burden score, and model for end-stage liver disease score.

Our general joint modeling framework can be applied to the dynamic prediction of individualized causal effects in settings with multiple longitudinal markers, time-dependent treatment, and time-to-event outcomes. Unlike other approaches, it allows these effects to be estimated without explicitly specifying the treatment assignment or censoring mechanisms.

Basis Function Embedding for Predicting Spatio-Temporal Climate Data

Brian O'Sullivan[†], Barry Coonan[†]

[†] Climate Services Division, Met Éireann

Gridded climate data, typically generated by interpolating spatially irregular station observations, are essential for many meteorological and climatological applications. Geospatial methods can effectively capture the spatio-temporal patterns exhibited by climate parameters, including kriging and generalized additive models. However, these approaches are often computationally intensive or depend on restrictive modelling assumptions.

Machine learning methods offer greater flexibility, but require high-quality covariates that are unavailable in many practical applications. Moreover, capturing spatial patterns remains challenging with few spatial predictors (e.g., longitude and latitude). There are methods that incorporate spatial structure directly, such as convolutional neural networks, although they are not readily applicable to spatially irregular station observations.

An alternative approach is DeepKriging, a deep learning method that represents the spatial domain through basis functions centered at spatial knot points [1]. These basis functions serve as additional predictors, increasing the spatial information available and improving the representation of spatial patterns. However, current implementations of DeepKriging remain inflexible, relying on regularly spaced knot points and requiring basis functions to be manually specified prior to model training. Additionally, the neural networks used in DeepKriging typically require large amounts of training data.

We present several extensions to DeepKriging, including the generation of knot points via constrained Delaunay triangulation, improving the spatial expressiveness of the model. Furthermore, we include the basis functions within the neural network architecture, enabling them to be fine-tuned during model training. We also show how basis function representation can improve a range of machine learning methods for spatio-temporal gridding, providing a simple yet flexible means of incorporating spatial dependence in methods such as Random Forests or XGBoost. These lightweight methods are well suited to sparsely sampled climate parameters like relative humidity, reducing the need for extensive covariate data. These methods are applied to precipitation and relative humidity in Ireland, demonstrating improved performance over the original DeepKriging approach and several state-of-the-art geostatistical and machine learning gridding methods.

References

[1] Chen, W. et al. 2024. DeepKriging: Spatially dependent deep neural networks for spatial prediction. *Statistica Sinica,* 34:291-311.

Moment-type Estimators for the Dirichlet Distribution

Ioannis Oikonomidis[†], Samis Trevezas[†]

[†] National and Kapodistrian University of Athens

This study develops new closed-form estimators for the shape parameters of the Dirichlet distribution, whose maximum likelihood estimators cannot be explicitly derived. The proposed approach extends the score-adjusted estimation method from the beta to the Dirichlet family [1]. Two moment-based estimators and one score-adjusted estimator, denoted by M1, M2, and SA, respectively, are explicitely derived, along with closed-form expressions for their asymptotic variance–covariance matrices. The main results of the study are the following: (i) M2 is shown to universally dominate M1 in terms of asymptotic variance; (ii) SA is shown to dominate M2 under specific conditions; and (iii) general counterexamples demonstrate that no universal dominance between SA and M2 holds in the general case. Small-sample properties and comparisons are studied numerically. To facilitate the application of these estimators, the `joker` R package is developed and made publicly available.

References

[1] Tamae, H., Irie, K. and Kubokawa, T. 2020. A score-adjusted approach to closed-form estimators for the gamma and beta distributions. *Japanese Journal of Statistics and Data Science,* 3(2):543-561.

Estimating True Event-time Distributions from Rounded Circular Records

Chiara Passamonti[†], Marco Di Marzio[‡], Stefania Fensore[‡], Charles C. Taylor[§]

[†] University of Luxembourg
[‡] University of Chieti-Pescara
[§] University of Leeds

In many applied fields, the analysis of clock-time data is affected by measurement error, most notably rounding preference, i.e. the human tendency to report times using "attractive" values. Such mechanisms generate artificial temporal spikes, distort the underlying diurnal structure, and may substantially bias inference if ignored. Although rounding phenomena are increasingly recognized across several disciplines, they remain relatively underexplored, especially in circular statistics (see [1]).

We consider the problem of estimating the distribution of actual event times when only rounded observations are available. Our methodology proceeds in two stages. First, rounding probabilities are estimated through a regression framework based on the empirical distribution of "minutes past the hour". These probabilities are then incorporated into a stochastic matrix representation of the rounding mechanism and used within a penalized inversion procedure to recover the full 24-hour event-time distribution.

The methodology is illustrated using Boston crime data, where pronounced reporting artifacts and substantial heterogeneity in temporal precision emerge across offence categories.

References

[1] Jammalamadaka, S. R., Ghosh, K., and Akiri, S. 2024. Testing uniformity on the circle using spacings when data are rounded. *Statistics & Probability Letters,* 210, 110-114.

Testing Independence of High-Dimensional Vectors in the Presence of Missing Data

Bojana Milošević[‡], Jelena Radojević[†‡]

[†] University of Belgrade, Faculty of Mathematics
[‡] University of Belgrade, Faculty of Civil Engineering

Testing independence in high-dimensional settings is a challenging problem that remains relatively underexplored in the presence of missing data. We develop a distance correlation-based framework for testing independence between high-dimensional random vectors with incomplete observations. Assuming data are missing completely at random (MCAR), we derive the asymptotic null distribution of the proposed test statistics. These statistics are applicable in regimes where the dimensionality is comparable to or exceeds the sample size. Their finite-sample performance is evaluated through simulations under a range of dependence structures, including nonlinear relationships. We further assess robustness to violations of the MCAR assumption.

References

[1] Cuparić, M., Milošević, B., Radojević, J. 2026. Testing independence in the presence of missing data: high-dimensional case. *arXiv preprint arXiv:2604.22980*.

[2] Aleksić, D., Cuparić, M., Milošević, B. 2023. Non-degenerate U-statistics for data missing completely at random with application to testing independence. Stat, 12(1), e634.

Hedging Unpriced Risks

Roope Rihtamo[†]

[†] University of Turku

Unpriced risks are systematic sources of common variation that carry zero risk premia. Exposure to such risks increases portfolio variance without increasing expected returns and therefore lowers Sharpe ratios relative to the mean-variance efficient frontier. This paper proposes a portfolio-level method for identifying and hedging unpriced risk exposures directly in the space of portfolio weights. For a given target allocation, such as an extreme-decile long-short characteristic portfolio, the method constructs hedge allocations whose returns are maximally correlated with the target return but are constrained to have zero expected excess return. The target allocation is then tilted against these hedges, reducing variance while preserving expected returns. Under the population model, this operation improves the Sharpe ratio whenever the target portfolio has nonzero exposure to unpriced risks. Empirically, the results indicate that a large share of the variance in conventional characteristic-based pricing factors is unpriced. Hedging this variation substantially increases Sharpe ratios, primarily through large reductions in out-of-sample volatility.

Hierarchical Bayesian Recovery of Sparse Sources from Multichannel Signals

Robin Rohlén

Imperial College London

The recovery of latent sparse spike trains from multichannel observations is a convolutive blind source separation problem central to many applications in neural signal processing, where the goal is to decompose recordings into individual neural firing patterns. Existing approaches rely on time-domain signal extension followed by independent component analysis, providing point estimates but no uncertainty quantification. We propose a hierarchical Bayesian framework formulated directly in convolutional space. Each latent source is modelled as a Bernoulli-Gaussian sparse representation convolved with a channel-specific finite impulse response, with Student-t residuals for robustness to heavy-tailed noise. Inference proceeds via variational EM: a mean-field E-step yields closed-form posteriors over spike presence and amplitude, while the M-step solves a ridge-regularised deconvolution problem in the frequency domain via preconditioned conjugate gradient. We evaluate the method on simulated and experimental high-density surface electromyography data, demonstrating spike train estimates with negligible baseline noise compared to state-of-the-art methods and reduced memory requirements.

From Control to Machine Learning: Generalization bounds for hypothesis classes of stable models

Dániel Rácz[†,‡]

[†] Institute for Computer Science and Control (SZTAKI), Budapest, Hungary
[‡] Eötvös Loránd University, Budapest, Hungary

Recent high-performing machine learning models often rely on architectural or optimization heuristics whose statistical role remains unclear. In this talk, I present some results where stability concepts from control theory lead to useful insights for studying generalization.

First we consider (non-selective) deep state-space models (deep SSMs), a family underlying several sequence-to-sequence architectures. Here stability-induced system norms from control theory, such as $\mathcal{H}_2$ and ℓ_1-type norms, upper bound the Rademacher complexity of stable SSM layers. Together with a composition tool called Rademacher contraction, this yields PAC bounds for multilayer deep SSMs that are independent of sequence length.

I will also briefly discuss a related PAC bound for stable linear parameter-varying (LPV) systems and explain why this viewpoint is relevant for Mamba-like selective SSMs, whose selective block can be formulated as a discrete-time LPV system.

Bootstrapping not under the null?

Alexis Derumigny[†], Miltiadis Galanis[‡], Wieger Schipper[†], Aad van der Vaart[†]

[†] Delft University of Technology
[‡] National and Kapodistrian University of Athens

We propose a bootstrap testing framework for a general class of hypothesis tests, which allows resampling under the null hypothesis as well as other forms of bootstrapping. We identify combinations of resampling schemes and bootstrap statistics for which the resulting tests are asymptotically exact and consistent against fixed alternatives. We show that in these cases the limiting local power functions are the same for the different resampling schemes. We also show that certain naive bootstrap schemes do not work. To demonstrate its versatility, we apply the framework to several examples: independence tests, tests on the coefficients in linear regression models, goodness-of-fit tests for general parametric models and for semi-parametric copula models. Simulation results confirm the asymptotic results and suggest that in smaller samples non-traditional bootstrap schemes may have advantages. This bootstrap-based hypothesis testing framework is implemented in the R package BootstrapTests.

References

[1] Derumigny, A., Galanis, M., Schipper, W., and van der Vaart, A. 2025. Bootstrapping not under the null? *arXiv preprint arXiv:2512.10546*.

Statistical Inference for L-Estimators

Emīls Siliņš[†]

[†] University of Latvia, Faculty of Science and Technology

Define the sample L-estimator as follows:

$$L_n = \sum_{i=1}^{n} a_i X_{(i)} = \frac{1}{m} \sum_{i=1}^{n} J\left(\frac{i}{n+1}\right) X_{(i)}, \tag{1}$$

where $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ denote the order statistics of the random sample $X_1, \ldots, X_n$, and $J(\cdot)$ is a weight function. L-estimators, and in particular the trimmed mean, are among the most commonly used robust estimators in statistical inference. For example, K. K. Yuen introduced a two-sample t-test based on trimmed means, and R. R. Wilcox developed an ANOVA-type test for trimmed means [1]. Later, M. Delesa-Vēļiņa established empirical likelihood methods for trimmed mean in both two-sample settings [2] and multiple-sample settings [3].

In this work [4], we unify these developments for the general L-estimator and conduct a simulation comparing the performance of the trimmed mean and the general smoothly trimmed mean [5] in terms of coverage accuracy and empirical type I error. Furthermore, we explore covariate adjustment methods for L-estimators based on the conditional empirical likelihood [6].

References

[1] Wilcox, Rand R. 2012 Introduction to Robust Estimation and Hypothesis Testing. Netherlands: Elsevier Science.

[2] Delesa-Vēļiņa, M., Valeinis, J. and Luta, G. 2021. Comparing two independent populations using a test based on empirical likelihood and trimmed means. *Lithuanian Mathematical Journal,* 61(2):199-216.

[3] Velina, M., Valeinis, J., Greco, L. and Luta, G. 2016. Empirical likelihood-based ANOVA for trimmed means. *International journal of environmental research and public health,* 13(10):953.

[4] Silins, E. and Valeinis, J. 2026. Empirical likelihood inference for L-estimators. *Communications in Statistics-Theory and Methods*, 1-34.

[5] Kresse, E., Silins, E. and Valeinis, J. 2024. Empirical likelihood for generalized smoothly trimmed mean. *arXiv preprint arXiv:2409.05631.*

[6] Owen, A. 1991. Empirical Likelihood for Linear Models. *The Annals of Statistics,* 19(4):1725–1747.

The Asymptotic Distribution of Gromov-Wasserstein based Barycenters

Florian Steinkamp[†], Luis Alberto Rodriguez[†], Axel Munk[†,‡]

[†] University of Göttingen, Germany
[‡] Cluster of Excellence "Multiscale Bioimaging: from Molecular Machines to Networks of Excitable Cells" (MBExC),University Medical Center, Göttingen, Germany

We introduce statistical theory for the matching of finitely many objects, represented as metric measure spaces (mm-spaces). The approach is based on the second lower bound (SLB) of the Gromov-Wasserstein distance and thus is able to identify deviations in the distributions of the (pairwise) distances within each mm-space. In order to discriminate the mm-spaces, we introduce a surrogate of the SLB barycenter which can be expressed explicitly in terms of the distance distributions of each object. When comparing K mm-spaces using n samples of each space, the resulting statistic can be calculated efficiently in only $O(K \cdot n^2)$ basic calculation operations. We derive the asymptotic distribution and finite-sample bounds of the proposed test statistic, enabling a variety of tools for statistical inference, specifically an asymptotic test for pose-invariant object discrimination and a classification method (based on SLB barycenter representation). These methods are investigated in simulations and applied to the structural comparison of protein domains.

Estimation of the invariant measure of a multidimensional diffusion from noisy observations

Raphaël Maillet[†], Grégoire Szymanski[‡]

[†] Université Paris-Dauphine
[‡] University of Luxembourg

We introduce a new approach for estimating the invariant density of a multidimensional diffusion when dealing with high-frequency observations blurred by independent noises. We consider the intermediate regime, where observations occur at discrete time instances $k\Delta_n$ for $k = 0, \dots, n$, under the conditions $\Delta_n \to 0$ and $n\Delta_n \to \infty$.

Our methodology involves the construction of a kernel density estimator that uses a pre-averaging technique to proficiently remove noise from the data while preserving the analytical characteristics of the underlying signal and its asymptotic properties. The rate of convergence of our estimator depends on both the anisotropic regularity of the density and the intensity of the noise.

We establish conditions on the intensity of the noise that ensure the recovery of convergence rates similar to those achievable without any noise. Furthermore, we prove a Bernstein concentration inequality for our estimator, from which we derive an adaptive procedure for the kernel bandwidth selection.

Marginal Likelihood for Intractable Hidden Markov Models

Daniele Tancini[†], Riccardo Rastelli[‡], Francesco Bartolucci[†]

[†] University of Perugia
[‡] University College Dublin

Hidden Markov models are important in many research areas. However, for some of these models the latent process distribution involves an intractable normalizing constant, especially in spatial and spatio-temporal applications. Under the Bayesian framework, these models result in the so-called doubly intractable posterior distributions. In order to estimate these models, different procedures can be considered, such as the exchange algorithm [1] and the noisy Markov chain Monte Carlo (MCMC) proposed in [2]; see [3] for a review.

An additional layer of intractability is involved when the focus is on estimating the marginal likelihood (ML). Different methods have been proposed for estimating the ML and the Bayes factor of these types of models; see in particular [4, 5]. These methods require extra simulations beyond the primary posterior simulation, leading to significant computational challenges. In this work, we propose a method to approximate the ML of intractable hidden Markov models, which uses only the unnormalized posterior densities of the sampled parameter values, without requiring any extra simulations beyond the primary posterior sampling. The method relies on the reciprocal importance sampling [6] integrated with the exchange algorithm [1], providing an approximation of the ML directly from the MCMC output.

References

[1] Murray, I., Ghahramani, Z., and MacKay, D. 2012. MCMC for doubly-intractable distributions. In *Proc. 22nd Conference of Uncertainty in Artificial Intelligence*, pp. 359–366.

[2] Alquier, P., Friel, N., Everitt, R., and Boland, A. 2016. Noisy Monte Carlo: Convergence of Markov chains with approximate transition kernels. *Statistics and Computing*, 26:29–47.

[3] Park, J. and Haran, M. 2018. Bayesian inference in the presence of intractable normalizing functions. *Journal of the American Statistical Association*, 113:1372–1390.

[4] Friel, N. 2013. Evidence and Bayes factor estimation for Gibbs random fields. *Journal of Computational and Graphical Statistics*, 22:518–532.

[5] Everitt, R. G., Johansen, A. M., Rowing, E., and Evdemon-Hogan, M. 2017. Bayesian model comparison with un-normalised likelihoods. *Statistics and Computing*, 27:403–422.

[6] Gelfand, A. E. and Dey, D. K. 1994. Bayesian model choice: asymptotics and exact calculations. *Journal of the Royal Statistical Society: Series B*, 56:501–514.

Non-Parametric Bayesian Estimation of Distance Distributions from ENDOR Data

Alina Tschirpke[†], Yvo Pokern[‡], Lucca Sielaff[†§]

[†] University of Göttingen
[‡] University College London
[§] Max Planck Institute for Multidisciplinary Sciences Göttingen

Electron-Nuclear Double Resonance (ENDOR) spectroscopy can be used to measure intramolecular distances. Due to conformational flexibility, the population of intramolecular distances in a chemical sample is described by a probability distribution whose estimation is equivalent to solving an ill-posed inverse problem with a non-negativity constraint. We employ a non-parametric Bayesian approach using a Gibbs sampler, a type of Markov chain Monte Carlo method. Replacing Tikhonov regularization, the standard method used by experimentalists, with this Bayesian approach provides natural quantification of uncertainty via posterior credible regions. The Gibbs sampler can be extended to estimate additional physical parameters and incorporate prior information on them thanks to the Bayesian approach. Also, different methodologies of measuring ENDOR spectra can be accommodated by a simple adjustment of the forward problem. The dependency between the distance distribution and the regularisation parameter which deteriorates Gibbs sampler performance is addressed by a non-centring trick which is shown to apply even in the non-negativity-constrained case. The talk is about work in progress.

References

[1] Edwards, T. H. and Stoll, S. 2016. A Bayesian approach to quantifying uncertainty from experimental noise in DEER spectroscopy. *Journal of Magnetic Resonance,* 270:87-97.

Causal Inference using Staged Event Trees

Gherardo Varando[†]

[†] Department of Statistics and Operational Research, University of Valencia

Causal inference seeks to estimate causal effects from data by combining statistical models, assumptions, and observed information. In this work, we present a framework for causal inference and effect estimation using staged event tree models, a class of probabilistic graphical models capable of representing complex, asymmetric, and context-specific dependencies among categorical random variables [1, 4].

First, we introduce staged event tree models and how they can be used to encode causal information. Building upon this structure, we show how staged event trees can be leveraged for transparent treatment effect estimation. We characterize classical causal inference estimators within this framework to calculate average and conditional treatment effects [2]. Staged trees offer a distinct advantage by allowing practitioners to visually and explicitly describe when standard causal assumptions, such as positivity, hold, thereby enhancing their interpretability in real-world settings.

Furthermore, to address scenarios where the true data-generating causal model is uncertain, we propose a fully Bayesian approach that jointly learns the staged tree structure and causal effects [3]. This method incorporates flexible prior distributions to encourage parsimony and promote interpretable dependence patterns [3].

All methodologies are implemented within the open-source R package `stagedtrees`, providing an accessible toolkit for structural learning and causal inference [4].

References

[1] Leonelli, M. and Varando, G. 2023. Context-Specific Causal Discovery for Categorical Data Using Staged Trees. *Proceedings of Machine Learning Research,* 206:8871-8888.

[2] Varando, G., Leonelli, M., Cerdà-Bautista, J., Sitokonstantinou, V. and Camps-Valls, G. 2025. Staged Event Trees for Transparent Treatment Effect Estimation. *arXiv preprint arXiv:2509.26265.*

[3] Cremaschi, A., Leonelli, M. and Varando, G. 2025. Bayesian Causal Effect Estimation for Categorical Data using Staged Tree Models. *arXiv preprint arXiv:2511.03399.*

[4] Carli, F., Leonelli, M., Riccomagno, E. and Varando, G. 2022. The R Package stagedtrees for Structural Learning of Stratified Staged Trees. *Journal of Statistical Software,* 102(6):1-30.

Calibrated Debiased Machine Learning for Parameters without Mixed Bias Property

Willem Weyens[†], Máté Kormos[†], Oliver Dukes[†] and Stijn Vansteelandt[†]

[†] Ghent University

Debiased machine learning (DML) has become a highly active area of modern statistics. Broadly speaking, one begins with a parameter that summarizes information relevant to a scientific question of interest. DML then provides a framework for constructing both estimators and confidence intervals for that parameter by combining modern predictive methods with ideas from semiparametric statistics. The key principle is to separate prediction from inference: machine learning methods are used to estimate complex nuisance structure, while orthogonalization techniques remove the leading bias induced by these estimators, thereby allowing reliable uncertainty quantification.

Conceptually, the estimator is constructed to be approximately orthogonal to nuisance estimation errors, much like projecting away unwanted directions in a high-dimensional space.

Current state-of-the-art methods require certain convergence rate conditions on the nuisance estimators used in the procedure. Although these assumptions are relatively mild, they may fail in practice. Recently, calibrated debiased machine learning procedures introduced by Van der Laan, Luedtke and Carone [1] substantially weakened these conditions, leading to significant improvements in empirical performance.

In this talk, we present an extension of their framework to a parameter not covered by their theory. The parameter is motivated by the partially linear model $Y = \beta A + g(X) + \epsilon$, where Y is an outcome, A is a treatment or action, β is an unknown constant, g is an unknown function, and ϵ satisfies $E[\epsilon \mid A, X] = 0$.

Under this model, together with standard causal assumptions, β summarizes the average effect of an intervention on A (or, when A is continuous, the effect of a shift intervention). It can be defined in a model-free way as $\beta_0 := \frac{E[(Y,A|X)]}{E[(A|X)]}$ and learned using DML [2]. By extending techniques from the calibrated DML framework, we construct an estimator for β_0 that remains $\sqrt{n}$-consistent under weaker L_2-rate conditions on the nuisance learners.

References

[1] Van der Laan, L., Luedtke, A. and Carone, M. 2025. Doubly robust inference via calibration. *arXiv preprint arXiv:2411.02771.*

[2] Vansteelandt, S. and Dukes, O. 2022. Assumption-lean Inference for Generalised Linear Model Parameters. *Journal of the Royal Statistical Society Series B: Statistical Methodology,* 84:657-685.

Goodness-of-fit testing for point processes

Gesine Reinert[†], Dominic Schuhmacher[‡], Leoni Carla Wirth[†]

[†] University of Oxford
[‡] University of Göttingen

Point processes provide a flexible framework for modeling spatial events, such as disease cases in an epidemiological study or cells in biological tissue. They are especially useful because they can capture how events influence each other's locations, leading to patterns like clustering or repulsion.

A central statistical challenge is the goodness-of-fit assessment for such models. In this talk, I will present ongoing work on such a goodness-of-fit test for spatial point processes based on a kernelized Stein discrepancy (KSD). Unlike many existing approaches, the proposed framework is designed for the relevant setting in which only a single realization of the process is observed, rendering standard i.i.d. based methods inapplicable. The proposed methodology is sensitive to both interaction structure and inhomogeneity, and can therefore identify discrepancies between the fitted model and the observed spatial pattern, including those driven by clustering, repulsion, or incorrect specification of the intensity function. This talk is based on joint work with Dominic Schuhmacher (Göttingen) and Gesine Reinert (Oxford).

Meta-Learning Theory-Informed Inductive Biases using Deep Kernel Gaussian Processes

Bahti Zakirov[†], Gašper Tkačik[†]

[†] Institute of Science and Technology, Austria

Normative and task-driven theories offer powerful top-down explanations for biological systems, yet the goals of quantitatively arbitrating between competing theories, and utilizing them as inductive biases to improve data-driven fits of real biological datasets are prohibitively laborious, and often impossible. To this end, we introduce a Bayesian meta-learning framework designed to automatically convert raw functional predictions from normative theories into tractable probabilistic models. We employ adaptive deep kernel Gaussian processes, meta-learning a kernel on synthetic data generated from a normative theory. This Theory-Informed Kernel specifies a probabilistic model representing the predictions of a given theoretical model: usable for both fitting data and rigorously validating the theory. As a demonstration, we apply our framework to the early visual system, using efficient coding as our normative theory. We show improved response prediction accuracy in ex vivo recordings of mouse retinal ganglion cells stimulated by natural scenes compared to conventional data-driven baselines, while providing accurate uncertainty estimates and interpretable representations. Using exact Bayesian model selection, we also show that our informed kernel can accurately infer the degree of theory-match from data, confirming faithful encapsulation of theory structure. This work provides a more general, scalable, and automated approach for integrating theoretical knowledge into data-driven scientific inquiry in neuroscience and beyond.


VILNIAUS UNIVERSITETAS
•1579•
UNIVERSITAS VILNENSIS