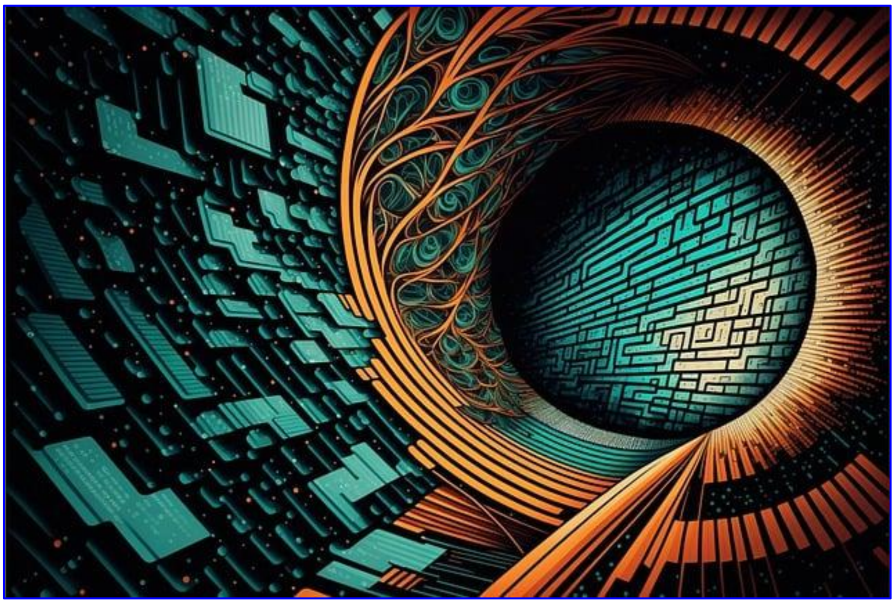

Workshop on Advances in High/Infinite-Dimensional Inference
AHIDI2024

Verona, 7-8 November 2024

Organised by
Judith Rousseau (Université Paris-Dauphine and University of Oxford)
Catia Scricciolo (University of Verona)
Lorenzo Frattarolo (University of Verona)

Venue
Aula SMT.03 - Complesso Universitario Santa Marta, Via Cantarane 24, Verona (IT)

Dipartimento di Scienze Economiche
Via Cantarane, 24 – 37129 Verona | T: +39 045 8028.095-097-236
segreteria.dse@ateneo.univr.it | www.dse.univr.it
pec: ufficio.protocollo@pec.univr.it
P. IVA 01541040232 | C.F. 93009870234

UNIVERSITÀ di VERONA | Dipartimento di SCIENZE ECONOMICHE

7th November 2024

09:45–10:00 **Welcome – Alessandro Bucciol (Head of Dept. of Economics) & Judith Rousseau**

Chair – *Judith Rousseau* (Université Paris-Dauphine and University of Oxford)

10:00–10:50 **Elisabeth Gassiat (Université Paris-Saclay)**
Fundamental limits for late change-point detection under the preferential attachment random graph model

10:50–11:40 **Vincent Rivoirard (Université Paris-Dauphine)**
PCA for point processes

11:40–12:00 Coffee break
"Corte Est" (Floor -1)

Chair – *Catia Scricciolo* (University of Verona)

12:00–12:50 **Sergios Agapiou (University of Cyprus)**
A new way for achieving Bayesian nonparametric adaptation

12:50-14:30 Lunch
"Corte Est" (Floor -1)

Chair – *Vincent Rivoirard* (Université Paris-Dauphine)

14:30–15:20 **Matey Neykov (Northwestern University)**
Convexity in the Gaussian sequence model: implications for robust mean estimation (and nonparametric regression)

15:20–16:10 **Deborah Sulem (Barcelona School of Economics and Universitat Pompeu Fabra)**
Bayesian computation for high-dimensional Gaussian graphical models

16:10–16:30 Coffee break
"Corte Est" (Floor -1)

Chair – *Lorenzo Frattarolo* (University of Verona)

16:30–17:30 **Junior researcher session**
Eric Aubinais (CNRS) *Fundamental limits of membership inference attacks on machine learning models*
Ibrahim Kaddouri (Université Paris-Saclay) *Clustering and classification risks in non-parametric Hidden Markov and I.I.D. models*
Leonardo Martins Bianco (Université Paris-Saclay) *SubSearch: Robust Estimation and Outlier Detection for Stochastic Block Models via Subgraph Search*
Yichen Zhu (Bocconi University) *Vecchia Gaussian Processes: Probabilistic Properties, Minimax Rates and Methodological Insights*

19:00–22:00 Social dinner
Ristorante Maffei
Piazza Erbe 38, 37121 Verona (VR)
Tel. +39 045 8005124

8th November 2024

Chair – *Sven Wang* (Humboldt University of Berlin)

***09:30–10:20* Kolyan Ray (Imperial College London)**
A variational Bayes approach to debiased inference in high-dimensional linear regression

***10:20–11:10* Veronika Rockova (University of Chicago Booth School of Business)**
Adaptive Bayesian predictive inference in high-dimensional regression

11:10–11:40 Coffee break
"Corte Est" (Floor -1)

Chair – *Lorenzo Frattarolo* (University of Verona)

***11:40–13:00* Junior researcher session**
Botond Szabo (Bocconi University) *Privacy constrained semiparametric inference*
Maël Duverger (Université Paris-Dauphine) *Bernstein-von Mises property for functionals of linear Hawkes process*
Francesco Pozza (Bocconi University) *Skew-symmetric approximations of posterior distributions*
Ardjen Pengel (Cambridge University) *Gaussian Approximation and Output Analysis for High-Dimensional MCMC*
Qing Wang (Ca' Foscari Venice) *Markov Switching Multiple-equation Tensor Regressions*

13:00-14:30 Lunch
"Corte Est" (Floor -1)

Chair – *TBA*

***14:30–15:25* Randolf Altmeyer (Imperial College London)**
Statistical inference for SPDEs under spatial ergodicity

***15:25–16:20* Sven Wang (Humboldt University of Berlin)**
M-estimation and statistical theory for neural operators

***16:20–16:30* Discussion and closing remarks by Judith Rousseau**

16:30–18:00 Farewell Cocktail
"Corte Est" (Floor -1)

Dipartimento di Scienze Economiche
Via Cantarane, 24 – 37129 Verona | T: +39 045 8028.095-097-236
segreteria.dse@ateneo.univr.it | www.dse.univr.it
pec: ufficio.protocollo@pec.univr.it
P. IVA 01541040232 | C.F. 93009870234

Abstracts

Elisabeth Gassiat (Université Paris-Saclay)

Title: *Fundamental limits for late change-point detection under the preferential attachment random graph model*
(Joint work with Ibrahim Kaddouri and Zacharie Naulet)

Abstract: We consider the problem of late change-point detection under the preferential attachment random graph model with time dependent attachment function. This can be formulated as a hypothesis testing problem where the null hypothesis corresponds to a preferential attachment model with a constant affine attachment parameter δ_0 and the alternative corresponds to a preferential attachment model where the affine attachment parameter changes from δ_0 to δ_1 at a moment $\tau_n = n - \Delta_n$ where $0\leq\Delta_n\leq n$ and n is the size of the graph. It was conjectured that when observing only the unlabeled graph, detection of the change is not possible for $\Delta_n=o(n^{1/2})$. In this work, we make a step towards proving the conjecture for $\Delta_n=o(n^{1/3})$. We also study change-point detection in the case where the labelled graph is observed and show that change-point detection is possible if and only if $n-\tau_n\xrightarrow[n\rightarrow+\infty]{}+\infty$, thereby exhibiting a strong difference between the two settings.

Vincent Rivoirard (Université Paris-Dauphine)

Title: *PCA for point processes*

Abstract: In this talk, we introduce a novel statistical framework for the analysis of replicated point processes that allows for the study of point pattern variability at a population level. By treating point process realizations as random measures, we adopt a functional analysis perspective and propose a form of functional Principal Component Analysis (fPCA) for point processes. The originality of our method is to base our analysis on the cumulative mass functions of the random measures which gives us a direct and interpretable analysis. Key theoretical contributions include establishing a Karhunen-Loève expansion for the random measures and a Mercer Theorem for covariance measures. We establish convergence in a strong sense, and introduce the concept of principal measures, which can be seen as latent processes governing the dynamics of the observed point patterns. We propose an easy-to-implement estimation strategy of eigenelements for which parametric rates are achieved. We characterize the solutions of our approach to Poisson and Hawkes processes and validate our methodology via simulations and diverse applications in seismology, single-cell biology and neurosiences, demonstrating its versatility and effectiveness.

Joint work with Franck Picard, Angelina Roche and Victor Panaretos

Sergios Agapiou (University of Cyprus)

Title: *A new way for achieving Bayesian nonparametric adaptation*

Abstract: We will consider Bayesian nonparametric settings with functional unknowns and we will be interested in evaluating the asymptotic performance of the posterior in the infinitely informative data limit, in terms of rates of contraction. We will be especially interested in priors which are adaptive to the smoothness of the unknown function. In the last decade, certain hierarchical and empirical Bayes procedures based on Gaussian process priors, have been shown to achieve adaptation to spatially homogenous smoothness. However, we have recently shown that Gaussian priors are suboptimal for spatially inhomogeneous unknowns, that is, functions which are smooth in some areas and rough or even discontinuous in other areas of their domain. In contrast, we have shown that (similar) hierarchical and empirical Bayes procedures based on Laplace (series) priors, achieve adaptation to both homogeneously and inhomogeneously smooth functions. All of these procedures involve the tuning of a hyperparameter of the Gaussian or Laplace prior.
After briefly reviewing the above results, we will present a new strategy for adaptation to smoothness based on heavy-tailed priors. We will illustrate it in a variety of nonparametric settings, showing in particular that adaptive rates of contraction in the minimax sense (up to logarithmic factors) are achieved without tuning of any hyperparameters and for both homogeneously and inhomogeneously smooth unknowns. We will also present numerical simulations corroborating the theory.

The main part of the talk is joint work with Ismaël Castillo.

Matey Neykov (Northwestern University)

Title: *Convexity in the Gaussian sequence model: implications for robust mean estimation (and nonparametric regression)*

Abstract: In this talk we will discuss the exact minimax rate of a Gaussian sequence model under convex constraints, purely in terms of the local geometry of the given constraint set. We will argue that the local entropy of the set K determines the rate via a certain equation. We then connect this result to the information theoretic limitations of robust mean estimation with sub-Gaussian noise (and time permitting nonparametric regression).

Deborah Sulem (Barcelona School of Economics and Universitat Pompeu Fabra)

Title: *Bayesian computation for high-dimensional Gaussian graphical models*

Abstract: Gaussian graphical models are widely used to analyse the inter-dependence structure among variables such as gene expression and brain fMRI data. Under this model, conditional independence statements are encoded by zero entries in the model's precision matrix. However, the computational demands of estimating a high-dimensional precision matrix have limited the scope of applications when the number of observed variables is large. This work introduces a scalable, interpretable, and fully Bayesian method for estimating precision matrices in high-dimensional settings. Our method capitalises on the relationship between the conditional dependence structure and a linear regression model and decomposes the high-dimensional estimation problem via row-wise computations. Lastly, this approach enables us to parallelise some computations of posterior conditional distributions and fosters an efficient exploration of the network structure.

Kolyan Ray (Imperial College London)

Title: *A variational Bayes approach to debiased inference in high-dimensional linear regression*

Abstract: We consider statistical inference for a single coordinate of a high-dimensional parameter in sparse linear regression. It is well-known that high-dimensional procedures such as the LASSO can provide biased estimators for this problem and thus require debiasing. We propose a scalable variational Bayes method for this problem based on assigning a mean-field approximation to the nuisance coordinates and carefully modelling the conditional distribution of the target given the nuisance. We investigate the numerical performance of our algorithm and establish accompanying theoretical guarantees for estimation and uncertainty quantification.

Joint work with I. Castillo, A. L'Huillier, L. Travis

Veronika Rockova (Imperial College London)

Title: *Adaptive Bayesian predictive inference in high-dimensional regression*

Abstract: Bayesian predictive inference provides a coherent description of entire predictive uncertainty through predictive distributions. We examine several widely used sparsity priors from the predictive (as opposed to estimation) inference viewpoint. To start, we investigate predictive distributions in the context of a high-dimensional Gaussian observation with a known variance but an unknown sparse mean under the Kullback-Leibler loss. First, we show that LASSO (Laplace) priors are incapable of achieving rate-optimal predictive distributions. However, deploying the Laplace prior inside the Spike-and-Slab framework (e.g. with the Spike-and-Slab LASSO prior), rate-minimax performance can be attained with properly tuned parameters (depending on the sparsity level sn). We highlight the discrepancy between prior calibration for the purpose of prediction and estimation. Going further, we investigate popular hierarchical priors which are known to attain adaptive rate-minimax performance for estimation. Whether or not they are rate-minimax also for predictive inference has, until now, been unclear. We answer affirmatively by showing that hierarchical Spike-and-Slab priors are adaptive and attain the minimax rate without the knowledge of sn. This is the first rate-adaptive result in the literature on predictive density estimation in sparse setups. Building on the sparse normal-means model, we extend our adaptive rate findings to the case of sparse high-dimensional regression with Spike-and-Slab priors. All of these results underscore benefits of fully Bayesian predictive inference.

Randolf Altmeyer (Imperial College London)

Title: *Statistical inference for SPDEs under spatial ergodicity*

Abstract: In this talk, we consider two different estimation problems for semilinear SPDEs on a fixed time interval, based on spatial asymptotics for functionals of the process.
In the first part we focus on estimating a parametric diffusivity. We observe continuously in time spatial averages of the process, centred at different locations and relative to a known compactly supported point spread kernel function. The process is therefore only partially observed (in space) and the resulting vector-valued observation process is non-Markovian. We verify the LAN (local asymptotic normality) property of the likelihood process as the number of measurements increases and the support of the point spread functions decreases. The proof is inspired by the classical Kalman-Bucy filter.
In the second part, we consider the Bayesian estimation of the reaction function in the asymptotic regime, where the diffusivity level and the noise level of the SPDE tend to zero in a realistic coupling. The analysis of the estimation error requires new concentration results for spatial averages of transformation of the SPDE, which combine the Clark-Ocone formula with bounds on the marginal densities.

Sven Wang (Humboldt University of Berlin)

Title: *M-estimation and statistical theory for neural operators*

Abstract: In this talk, we discuss convergence rates for neural network-based operator surrogates, which approximate smooth maps between infinite-dimensional Hilbert spaces. Such surrogates have a wide range of applications and can be used in uncertainty quantification and parameter estimation problems in fields such as classical mechanics, fluid mechanics, electrodynamics, earth sciences etc. Here, the operator input typically represents unknown parameters partial differential equation (PDE) describing the underlying physics. The operator output is the corresponding PDE solution. Our analysis demonstrates that, under suitable smoothness assumptions, the empirical risk minimizer for specific neural network architectures can overcome the curse of dimensionality both in terms of required network parameters and the input-output pairs needed for training.

Abstracts
Junior Researcher Sessions

7th November 2024

Eric Aubinais (CNRS)

Title : *Fundamental Limits of Membership Inference Attacks on Machine Learning Models*

Abstract : Membership inference attacks (MIA) can reveal whether a particular data point was part of the training dataset, potentially exposing sensitive information about individuals. This article provides theoretical guarantees by exploring the fundamental statistical limitations associated with MIAs on machine learning models at large. More precisely, we first derive the statistical quantity that governs the effectiveness and success of such attacks. We then theoretically prove that in a non-linear regression setting with overfitting algorithms, attacks may have a high probability of success. Finally, we investigate several situations for which we provide bounds on this quantity of interest.Interestingly, our findings indicate that discretizing the data might enhance the algorithm's security. Specifically, it is demonstrated to be limited by a constant, which quantifies the diversity of the underlying data distribution. We illustrate those results through two simple simulations.

Ibrahim Kaddouri (Université Paris-Saclay)

Title: *Clustering and classification risks in non-parametric Hidden Markov and I.I.D models*

Abstract: We provide an in-depth analysis of the Bayes risk of clustering and the Bayes risk of classification in the context of Hidden Markov and i.i.d. Models. In both settings, we identify the situations where the two risks are comparable and those where the associated minimizers coincide, as well as the key quantity determining the difficulty of both tasks. Then, leveraging the nonparametric identifiability of HMMs, we control the excess risk of a plug-in clustering procedure. Simulations illustrate our findings.

Leonardo Martins Bianco (Université Paris-Saclay)

Title: *SubSearch: Robust Estimation and Outlier Detection for Stochastic Block Models via Subgraph Search*

Abstract: Community detection is a fundamental task in graph analysis, with methods often relying on fitting models like the Stochastic Block Model (SBM) to observed networks. While many algorithms can accurately estimate SBM parameters when the input graph is a perfect sample from the model, real-world graphs rarely conform to such idealized assumptions. Therefore, robust algorithms are crucial—ones that can recover model parameters even when the data deviates from the assumed distribution. We propose SubSearch, an algorithm for robustly estimating SBM parameters by exploring the space of subgraphs in search of one that closely aligns with the model's assumptions. Our approach also functions as an outlier detection method, properly identifying nodes responsible for the graph's deviation from the model and going beyond simple techniques like pruning high-degree nodes. Extensive experiments on both synthetic and real-world datasets demonstrate the effectiveness of our method.

Yichen Zhu (Bocconi University)

Title: *Vecchia Gaussian Processes: Probabilistic Properties, Minimax Rates and Methodological Insights*

Abstract: Gaussian Processes (GPs) are widely used to model dependency in spatial statistics and machine learning, yet the exact computation suffers an intractable time complexity of $O(n^3)$. Vecchia approximation allows scalable Bayesian inference of Gaussian processes in $O(n)$ time by introducing sparsity in the spatial dependency structure that is characterized by a directed acyclic graph (DAG). Despite the popularity in practice, it is still unclear how to choose the DAG structures and there is still theoretical guarantees in nonparametric settings. In this paper, we systematically study the Vecchia GPs as standalone stochastic processes and uncover important probabilistic properties and statistical results in methodology and theory. For probability properties, we prove the conditional distributions Mat\'{e}rn GPs, as well as Vecchia approximations of Mat\'{e}rn GPs, behave like polynomials. This allows us to prove series of results regarding the small ball probabilities and RKHSs of Vecchia GPs. For statistical methodology, we provide principled guideline to set parent sets as norming sets with fixed cardinality and provide detailed algorithms fulfilling such guidelines. For statistical theory, we prove the posterior contraction rates of applying Vecchia GPs to regression problems, where minimax optimality is achieved by optimally tuned Gaussian processes via either oracle rescaling or hierarchical Bayesian methods. Our theory and methodology are demonstrated with numerical studies, where we also provide efficient realization of our methods in C++ with R interfaces.

8^{th} November 2024

Botond Szabo (Bocconi University)

Title: *Privacy constrained semiparametric inference*

Abstract: For semi-parametric problems differential private estimators are typically constructed in a case-by-case basis. In this work we develop a privacy constrained semi-parametric plug-in approach, which can be used in general, over a collection of semi-parametric problems. We derive minimax lower and matching upper bounds for this approach and provide an adaptive procedure in case of irregular (atomic) functionals. Joint work with Lukas Steinberger (Vienna) and Thibault Randrianarisoa (Toronto, Vector Institute).

Maël Duverger (Universitè Paris-Dauphine)

Title: *Bernstein-von Mises property for functionals of linear Hawkes process*

Abstract: Hawkes processes are point processes used to model self-exciting phenomena. In the linear multivariate case, where the number of components K is fixed, the Hawkes process is defined by a set of parameters composed of K positive real numbers and K^2 non negative functions (called excitation functions). There have been some recent results on Bayesian estimation of these parameters, in particular concentration rates for the posterior distribution have been obtained by Donnet et al. (2020) and Sulem et al. (2024). In this ongoing work, based on these recent results, we aim to derive a Bernstein-von Mises property for some smooth functionals of the parameters. Broadly speaking, Bernstein-von Mises property states that the posterior distribution behaves asymptotically as a gaussian distribution, centered at an efficient estimator. Castillo and Rousseau proved in 2015 a general theorem for obtaining such a property in semiparametric models, which we intend to apply in this Hawkes processes framework.*

Francesco Pozza (Bocconi University)

Title: *Skew-symmetric approximations of posterior distributions*

Abstract: Routinely implemented deterministic approximations of posterior distributions, e.g., from Laplace, variational Bayes, and expectation propagation, typically rely on symmetric approximation densities, often assumed to be Gaussian. This choice facilitates optimization and inference, but typically compromises the quality of the overall approximation. In fact, even in simple parametric models, the posterior distribution often exhibits asymmetries that lead to bias and reduced accuracy when symmetric approximations are considered. For this reason, recent research has moved toward more flexible approximations that account for skewness. However, current solutions are model-specific and lack a general supporting theory. To fill this gap, we derive, through a novel treatment of a third-order version of the Laplace method, a skewed perturbation of the classical Gaussian Laplace approximation that is similarly tractable to its Gaussian counterpart, but provably more accurate. In particular, by obtaining a skewed version of the celebrated Bernstein-von Mises theorem, we show that this new approximation converges to the posterior distribution in probability, and under the total variation norm, an order of magnitude faster than the classical Gaussian limit. These results are further extended with a general and provably optimal strategy to improve the accuracy of any symmetric approximation to a generic posterior distribution by means of a simple perturbation scheme. Crucially, such a novel correction is derived without additional optimization steps, and yields a similarly tractable approximation within the class of skew-symmetric densities that provably improves the finite-sample accuracy of the original symmetric approximation.

Ardjen Pengel (University of Cambridge)

Title: *Gaussian Approximation and Output Analysis for High-Dimensional MCM***C**

Abstract: The widespread use of Markov Chain Monte Carlo (MCMC) methods for high-dimensional applications has motivated research into the scalability of these algorithms with respect to the dimension of the problem. Despite this, numerous problems concerning output analysis in high-dimensional settings have remained unaddressed. We present novel quantitative Gaussian approximation results for a broad range of MCMC algorithms. Notably, we ana- lyse the dependency of the obtained approximation errors on the dimension of both the target distribution and the feature space. We demonstrate how these Gaussian approxim- ations can be applied in output analysis. This includes determining the simulation effort required to guarantee Markov chain central limit theorems and consistent variance estima- tion in high-dimensional settings. We give quantitative convergence bounds for termination criteria and show that the termination time of a wide class of MCMC algorithms scales polynomially in dimension while ensuring a desired level of precision. Our results offer guid- ance to practitioners for obtaining appropriate standard errors and deciding the minimum simulation effort of MCMC algorithms in both multivariate and high-dimensional settings. Furthermore, we present related ongoing research on uncertainty quantification for high-dimensional stochastic gradient descent.

Qing Wang (Ca' Foscari University)

Title: *Markov Switching Multiple-equation Tensor Regressions*

Abstract: We propose a new flexible tensor model for multiple-equation regression that accounts for latent regime changes. The model allows for dynamic coefficients and multi-dimensional covariates that vary across equations. We assume the coefficients are driven by a common hidden Markov process that addresses structural breaks to enhance the model flexibility and preserve parsimony. We introduce a new Soft PARAFAC hierarchical prior to achieve dimensionality reduction while preserving the structural information of the covariate tensor. The proposed prior includes a new multi-way shrinking effect to address over-parametrization issues. We developed theoretical results to help hyperparameter choice. An efficient MCMC algorithm based on random scan Gibbs and back-fitting strategy is developed to achieve better computational scalability of the posterior sampling. The validity of the MCMC algorithm is demonstrated theoretically, and its computational efficiency is studied using numerical experiments in different parameter settings. The effectiveness of the model framework is illustrated using two original real data analyses. The proposed model exhibits superior performance when compared to the current benchmark, Lasso regression.